Signal Processing 144 (2018) 214–225

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Performance analysis of the deficient length augmented CLMS algorithm for second order noncircular complex signals

Yili Xia[a,*], Scott C. Douglas[b], Danilo P. Mandic[c]

[a] School of Information Science and Engineering, Southeast University, Nanjing 210096, China
[b] Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275-0338, USA
[c] Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 22 June 2017
Revised 8 October 2017
Accepted 15 October 2017
Available online 16 October 2017

Keywords:
Widely linear model
Deficient length adaptive filter
Augmented complex LMS (ACLMS)
Second order noncircularity (improperness)
Mean square analysis
Approximate uncorrelating transform (AUT)

ABSTRACT

The augmented complex LMS (ACLMS) algorithm deals with second order noncircular (improper) input signals, based on widely linear modelling and the use of full second order statistical information. In current analyses of ACLMS, it is implicitly or explicitly assumed that the length of the adaptive filter is equal to that of the unknown system's impulse response (optimal model order). In many applications, however, the length of the adaptive filter is smaller than required, the so called deficient length case, which renders the analysis for a 'sufficient length' ACLMS inadequate. To this end, we examine the statistical behaviour of the ACLMS algorithm in undermodelling situations. Exact expressions are developed to completely characterise both the transient and steady-state mean and mean square performances of the deficient length ACLMS for general second order noncircular Gaussian input signals. This is achieved using the recently introduced approximate uncorrelating transform (AUT), in order to jointly diagonalise the covariance and pseudo-covariance matrices with a single singular value decomposition (SVD), which both simplifies the analysis and enables a link between the degree of input noncircularity and the steady state mean square error (MSE) performance of the deficient length ACLMS. Simulations in system identification settings support the analysis.

1. Introduction

The complex least mean square (CLMS) adaptive filtering algorithm has found application in numerous areas, including noise cancellation, system identification and prediction [1]. Standard learning algorithms in the complex domain, $\mathbb{C}$, are normally considered as generic extensions of the corresponding algorithms in the real domain, $\mathbb{R}$, a consequence of a restricted use of the available second order statistics analysis which is narrowed down to only the $M \times M$ covariance matrix, $\mathbf{R_x} = E[\mathbf{x}(k)\mathbf{x}^H(k)]$, where $\mathbf{x}(k) \in \mathbb{C}^{M\times 1}$ [2,3]. However, recent results in the so-called *augmented complex statistics* [4–7] show that the covariance matrix $\mathbf{R_x}$ cannot completely describe the second order behaviour of general complex signals, and another second order moment, the $M \times M$ pseudo-covariance matrix, $\mathbf{P_x} = E[\mathbf{x}(k)\mathbf{x}^T(k)]$, should also be taken into account, especially when processing second order noncircular (improper) signals. For second order circular signals, the pseudo-covariance matrix $\mathbf{P_x} = \mathbf{0}$, and the standard adaptive filtering algorithms (based on the covariance only) are adequate. For noncircular signals, $\mathbf{P_x} \neq \mathbf{0}$, and complex-valued adaptive filtering algorithms should be designed based on the $2M \times 1$ so-called augmented input vector $\mathbf{z}(k) = [\mathbf{x}^T(k), \mathbf{x}^H(k)]^T$, for which the $2M \times 2M$ augmented covariance matrix is given by [4–9]

$$\mathbf{R_z} = E[\mathbf{z}(k)\mathbf{z}^H(k)] = \begin{bmatrix} E[\mathbf{x}(k)\mathbf{x}^H(k)] & E[\mathbf{x}(k)\mathbf{x}^T(k)] \\ E[\mathbf{x}^*(k)\mathbf{x}^H(k)] & E[\mathbf{x}^*(k)\mathbf{x}^T(k)] \end{bmatrix} = \begin{bmatrix} \mathbf{R_x} & \mathbf{P_x} \\ \mathbf{P_x^*} & \mathbf{R_x^*} \end{bmatrix} \quad (1)$$

The augmented complex statistics have opened the possibility to design LMS-type adaptive filtering algorithms based on the widely linear model [4,6,7,10], leading to the so called augmented CLMS (ACLMS) [11,12], which is suitable for processing both second order circular and noncircular signals. This advantage over the conventional CLMS algorithm has led to its applications in signal processing [9,13,14], communications [15], and power systems [16,17]. In most current analyses, it is implicitly or explicitly assumed that the filter length is equal to the length of the unknown system impulse response, an optimal modelling situation [18–24], which is often not the case in practical settings. Therefore, the results for the 'sufficient length' ACLMS do not directly apply to the 'deficient length' situations; these arise in scenarios where a system with an

* Corresponding author.
E-mail addresses: yili.xia06@gmail.com (Y. Xia), douglas@engr.smu.edu (S.C. Douglas), d.mandic@imperial.ac.uk (D.P. Mandic).

https://doi.org/10.1016/j.sigpro.2017.10.021

unknown but long impulse response is modelled with an adaptive filter of a relatively shorter length. On the other hand, it is important for practical purposes to understand and quantify the statistical behaviour of the deficient length ACLMS, a subject of this work. This analysis could allow answers to more detailed performance analysis questions in specific contexts, such as the identification of a filter order that would allow a specific performance criterion such as an achievable minimum MSE to be obtained.

The first attempt to address accurate performance analysis of adaptive filters in the undermodelling situations is established in [25], where a deficient length LMS in $\mathbb{R}$ has been considered for correlated Gaussian data. An insufficient length transform domain LMS, which employs a real-valued discrete cosine transformation (DCT) to speed up the convergence, has been discussed in [26]. This work has been implicitly extended into the complex domain $\mathbb{C}$ in [27] in the sense that a complex-valued discrete Fourier transformation (DFT) has been utilised. However, from the perspective of the augmented complex statistics in $\mathbb{C}$, the analysis in [27] is still rather limited. On one hand, it directly extends the analyses in $\mathbb{R}$ [25,26] by considering a strictly linear undermodelling problem. On the other hand, the assumption of signal properness (second order circularity) has been implicitly imposed on the complex-valued data after the DFT operation, which has been justified based on central limit theorem in [28].

Therefore, in this paper, we develop a rigourous analysis of both the transient and steady-state performances of the deficient length ACLMS, which employs the most general widely linear framework in $\mathbb{C}$ for the processing of complex-valued data in terms of the full second order statistics. For rigour, this is achieved for second order correlated noncircular Gaussian inputs; the corresponding results for the case of doubly white[1] Gaussian inputs are readily obtained as a special case. In order to both simplify the analysis and explicitly link the degree of noncircularity of the input with steady state performance of the deficient length ACLMS, we employ the recently introduced approximate uncorrelating transform (AUT), which jointly diagonalises the covariance and pseudo-covariance matrices with a single singular value decomposition (SVD) [29]. The analysis shows that for correlated second order noncircular inputs, the deficient length ACLMS exhibits a distinctive statistical behaviour compared to noncircular doubly white Gaussian inputs. Simulations in the system identification setting validate the theoretical results.

2. Performance Analysis

Consider a second order noncircular (improper) desired signal $d(k)$ generated by the following widely linear model [4,6,30]

$$d(k) = \mathbf{h}_N^{oH}\mathbf{x}_N(k) + \mathbf{g}_N^{oH}\mathbf{x}_N^*(k) + q(k) \tag{2}$$

where $\mathbf{h}_N^o = [h_1^o, h_2^o, \ldots, h_N^o]^T$ and $\mathbf{g}_N^o = [g_1^o, g_2^o, \ldots, g_N^o]^T$ are respectively the standard and conjugate unknown (optimal) system impulse response vectors of length N, $\mathbf{x}_N(k) = [x(k), x(k-1), \ldots, x(k-N+1)]^T$ is the $N \times 1$ tap input vector to the unknown system at the time instant k, $q(k)$ is stationary, zero-mean, independent noise, while $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ are respectively the transpose, complex conjugate and Hermitian operators. The widely linear model in (2) has provided modelling advantages in the representation of complex-valued baseband communication signals at the receiver side which arise in real-valued modulation schemes, e.g., amplitude-shift keying (ASK), binary phase-shift keying (BPSK) and pulse amplitude modulation (PAM) [15,22,31,32]. Improper baseband communication signals represented by (2) can also arise due to amplitude and phase imbalances between their in-phase and quadrature (I/Q) components [33–35]. The utility of the model in (2) has also been demonstrated in unbalanced three-phase power systems [16,17]. For generality, we further assume that the input $x(k)$ is a correlated second order noncircular Gaussian process, i.e. the off-diagonal elements in both $\mathbf{R_x}$ and $\mathbf{P_x}$ do exist. Note that in (2), performance analysis of the deficient length ACLMS is implicitly performed in a stationary environment, where the unknown system coefficients $\mathbf{h}_N^o$ and $\mathbf{g}_N^o$ are time invariant. However, in practice, nonstationary environments may appear, where $\mathbf{h}_N^o$ and $\mathbf{g}_N^o$ exhibit time variations, which for modelling purposes typically assumes a random walk process. The purpose of the tracking analysis of an adaptive filter is to study its ability to track such time variations, and we refer to [36–38] for more detail along this direction.

In order to estimate the the unknown system impulse response vectors in (2), the ACLMS algorithm updates its weight vectors according to [11]

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e^*(k)\mathbf{x}(k) \tag{3}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \mu e^*(k)\mathbf{x}^*(k) \tag{4}$$

where μ is the step-size and the output error $e(k)$ is given by

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

with $\mathbf{x}(k) = [x(k), x(k-1), \ldots, x(k-M+1)]^T$ being the input signal vector and $\mathbf{h}(k)$ and $\mathbf{g}(k)$ the $M \times 1$ weight vectors of the widely linear adaptive filter. In the deficient length case, $M < N$.

For compactness of the analysis, we shall employ the following form of the desired signal in (2)

$$d(k) = \mathbf{w}_{2N}^{oH}\mathbf{z}_{2N}(k) + q(k) \tag{6}$$

where $\mathbf{w}_{2N}^o = [\mathbf{h}_N^{oT}, \mathbf{g}_N^{oT}]^T$ and $\mathbf{z}_{2N}(k) = [\mathbf{x}_N^T(k), \mathbf{x}_N^H(k)]^T$ are respectively the augmented system impulse response vector and the augmented system input vector.

The ACLMS update in (3) and (4) can now be compactly expressed as

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

where $\mathbf{w}(k) = [\mathbf{h}^T(k), \mathbf{g}^T(k)]^T$ and $\mathbf{z}(k) = [\mathbf{x}^T(k), \mathbf{x}^H(k)]^T$ are respectively the $2M \times 1$ augmented weight and input vectors of ACLMS, and

$$e(k) = d(k) - \mathbf{w}^H(k)\mathbf{z}(k) \tag{8}$$

Since for the deficient length case $M < N$, it is convenient to rearrange $\mathbf{w}_{2N}^o$ and $\mathbf{z}_{2N}(k)$ so as to reflect the $2M \times 1$ 'modelled', $\mathbf{w}^o$, and the $2(N-M) \times 1$ 'unmodelled', $\bar{\mathbf{w}}^o$, parts of the system in the form

$$\mathbf{w}_{2N}^o = \begin{bmatrix} \mathbf{w}^o \\ \bar{\mathbf{w}}^o \end{bmatrix}, \quad \mathbf{z}_{2N}(k) = \begin{bmatrix} \mathbf{z}(k) \\ \bar{\mathbf{z}}(k) \end{bmatrix} \tag{9}$$

where

$$\begin{aligned} \mathbf{w}^o &= [\mathbf{h}^{oT}, \mathbf{g}^{oT}]^T \\ &= [h_1^o, h_2^o, \ldots, h_M^o, g_1^o, g_2^o, \ldots, g_M^o]^T \\ \bar{\mathbf{w}}^o &= [\bar{\mathbf{h}}^{oT}, \bar{\mathbf{g}}^{oT}]^T \\ &= [h_{M+1}^o, h_{M+2}^o, \ldots, h_N^o, g_{M+1}^o, g_{M+2}^o, \ldots, g_N^o]^T \\ \bar{\mathbf{z}}(k) &= [\bar{\mathbf{x}}^T(k), \bar{\mathbf{x}}^H(k)]^T \\ &= [x(k-M), x(k-M-1), \ldots, x(k-N+1), \\ &\quad\; x^*(k-M), x^*(k-M-1), \ldots, x^*(k-N+1)]^T \end{aligned}$$

Upon introducing the $2M \times 1$ augmented weight error vector

$$\widetilde{\mathbf{w}}(k) = \mathbf{w}(k) - \mathbf{w}^o \tag{10}$$

[1] For a doubly white input vector $\mathbf{x}(k)$, we have $\mathbf{R_x} = \sigma_x^2\mathbf{I}_M$ and $\mathbf{P_x} = \rho_x\mathbf{I}_M$, where $\sigma_x^2 = E[x(k)x^*(k)]$ and $\rho_x = E[x^2(k)]$ are respectively the covariance and pseudo-covariance of $\mathbf{x}(k)$, and $\mathbf{I}_M$ is an $M \times M$ identity matrix [6,7].

from (6), (9) and (10) and after some mathematical manipulations, the output error in (8) can be rewritten as

$$e(k) = \underbrace{\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)}_{\text{unmodelled part of } d(k)} - \underbrace{\widetilde{\mathbf{w}}^{H}(k)\mathbf{z}(k)}_{\text{modelled part of } (k)} + q(k) \tag{11}$$

while for its conjugate we have

$$e^*(k) = \bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k) - \widetilde{\mathbf{w}}^{T}(k)\mathbf{z}^*(k) + q^*(k) \tag{12}$$

Observe that the output error $e(k)$ in (11) is now governed by both the weight error part (second term) and the undermodelling part (first term). From (7) and (12), the recursion for the update of the augmented weight error vector $\widetilde{\mathbf{w}}(k)$ becomes

$$\begin{aligned}\widetilde{\mathbf{w}}(k+1) &= \widetilde{\mathbf{w}}(k) + \mu\big(\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k) - \widetilde{\mathbf{w}}^{T}(k)\mathbf{z}^*(k) + q^*(k)\big)\mathbf{z}(k)\\ &= \big(\mathbf{I}_{2M} - \mu\mathbf{z}(k)\mathbf{z}^H(k)\big)\widetilde{\mathbf{w}}(k)\\ &\quad + \mu\big(\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k)\big)\mathbf{z}(k) + \mu q^*(k)\mathbf{z}(k)\end{aligned} \tag{13}$$

where $\mathbf{I}_{2M}$ is a $2M \times 2M$ identity matrix.

2.1. Convergence in the mean

The mean behaviour of the augmented weight error vector of deficient length ACLMS can now be determined by applying the statistical expectation operator $E[\cdot]$ to both sides of (13) and using the standard independence assumptions, that is, the noise $q(k)$ is statistically independent of any other variable in the ACLMS and $\widetilde{\mathbf{w}}(k)$ is statistically independent of the filter input $\mathbf{z}(k)$ [39,40], to yield

$$E[\widetilde{\mathbf{w}}(k+1)] = [\mathbf{I}_{2M} - \mu\mathbf{R}_{\mathbf{z}}]E[\widetilde{\mathbf{w}}(k)] + \mu\mathbf{b} \tag{14}$$

where $\mathbf{R}_{\mathbf{z}}$ is the augmented covariance matrix of the input data to the ACLMS, as defined in (1), and the $2M \times 1$ vector $\mathbf{b}$ is defined as

$$\mathbf{b} = E\Big[\big(\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k)\big)\mathbf{z}(k)\Big] \tag{15}$$

which reflects the unmodelled part of the system being modelled, that is, the scalar product $\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k)$, although its dimensionality $2M \times 1$ is subject to the modelled part of the system by the ACLMS adaptive filter, that is, the augmented input vector $\mathbf{z}(k)$. Since as defined, $\mathbf{z}(k) = [\mathbf{x}^T(k), \mathbf{x}^H(k)]^T$, we further have $\mathbf{b} = [\mathbf{b}_1^T, \mathbf{b}_2^T]^T$, where $\mathbf{b}_1 = [b_{1,1}, b_{1,2}, \ldots, b_{1,M}]^T$ and $\mathbf{b}_2 = [b_{2,1}, b_{2,2}, \ldots, b_{2,M}]^T$ can be expressed as

$$\mathbf{b}_1 = E\Big[\big(\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k)\big)\mathbf{x}(k)\Big] = E\Big[\big(\bar{\mathbf{h}}^{oT}\bar{\mathbf{x}}^*(k) + \bar{\mathbf{g}}^{oT}\bar{\mathbf{x}}(k)\big)\mathbf{x}(k)\Big] \tag{16}$$

$$\mathbf{b}_2 = E\Big[\big(\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k)\big)\mathbf{x}^*(k)\Big] = E\Big[\big(\bar{\mathbf{h}}^{oT}\bar{\mathbf{x}}^*(k) + \bar{\mathbf{g}}^{oT}\bar{\mathbf{x}}(k)\big)\mathbf{x}^*(k)\Big] \tag{17}$$

The ith components of $\mathbf{b}_1$ and $\mathbf{b}_2$ can be derived as

$$\begin{aligned}b_{1,i} &= E\Big[\sum_{j=M+1}^{N}\big(h_j^o x^*(k-j+1) + g_j^o x(k-j+1)\big)x(k-i+1)\Big]\\ &= \sum_{j=M+1}^{N}\big(h_j^o r(j-i) + g_j^o p(j-i)\big)\end{aligned} \tag{18}$$

$$\begin{aligned}b_{2,i} &= E\Big[\sum_{j=M+1}^{N}\big(h_j^o x^*(k-j+1) + g_j^o x(k-j+1)\big)x^*(k-i+1)\Big]\\ &= \sum_{j=M+1}^{N}\big(h_j^o p^*(j-i) + g_j^o r^*(j-i)\big)\end{aligned} \tag{19}$$

where $r(j-i) = E[x(k-i)x^*(k-j)]$ and $p(j-i) = E[x(k-i)x(k-j)]$ are respectively the covariance and pseudo-covariance coefficient at a discrete lag $(j-i)$.

Recall that $\lambda_{\max} \le \mathrm{tr}(\mathbf{R}_{\mathbf{z}})$, where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}_{\mathbf{z}}$ and $\mathrm{tr}(\cdot)$ is the trace operator. Hence, according to (14), the convergence of the deficient length ACLMS in the mean is guaranteed if the step-size μ satisfies

$$0 < \mu < \frac{2}{\mathrm{tr}(\mathbf{R}_{\mathbf{z}})} \tag{20}$$

Remark 1. From the structure of the augmented covariance matrix $\mathbf{R}_{\mathbf{z}}$ in (1), we have $\mathrm{tr}(\mathbf{R}_{\mathbf{z}}) = 2\mathrm{tr}(\mathbf{R}_{\mathbf{x}})$, and hence the step-size bound for deficient length ACLMS can be found from the mean analysis in (20), as

$$0 < \mu < \frac{1}{\mathrm{tr}(\mathbf{R}_{\mathbf{x}})} \tag{21}$$

Observe that this form is identical to that for mean stability of the sufficient length ACLMS in [18,19], due to the identical transition matrix $[\mathbf{I}_{2M} - \mu\mathbf{R}_{\mathbf{z}}]$ within the recursion of the weight error vector $E[\widetilde{\mathbf{w}}(k)]$ in (14). It is also interesting to observe that the upper bound on the step-size for the deficient length ACLMS is half of that for the real-valued deficient length LMS [25].

At the steady state, from (14), for the deficient length weight error vector we have

$$E[\widetilde{\mathbf{w}}(\infty)] = \mathbf{R}_{\mathbf{z}}^{-1}\mathbf{b} \tag{22}$$

so that, from (10), the steady-state value for the augmented weight vector of ACLMS becomes

$$E[\mathbf{w}(\infty)] = \mathbf{w}^o + \mathbf{R}_{\mathbf{z}}^{-1}\mathbf{b} \tag{23}$$

Remark 2. Eq. (23) indicates that a correlated or pseudo-correlated complex-valued input to the deficient length ACLMS results in biased estimation of the optimal weight vector, $\mathbf{w}^o$, which is reflected in the term $\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{b}$ in (23). The value of this bias depends upon the degree of undermodelling (M relative to N), the values of the weight coefficients within $\mathbf{w}^o$, as well as the full second order statistics of the input signal, that is, the covariance and pseudo-covariance coefficients contained in the augmented covariance matrix $\mathbf{R}_{\mathbf{z}}$ and the vector $\mathbf{b}$.

When the input signal $x(k)$ is doubly white (DW), both the covariance, $r(j-i)$, and the pseudo-covariance, $p(j-i)$, coefficients in (18) and (19) vanish, since $j \ge M+1$ and $i = 1, 2, \ldots, M$. Thus,

$$\mathbf{b}_{\mathrm{DW}} = \mathbf{0} \quad \text{and} \quad E[\mathbf{w}_{\mathrm{DW}}(\infty)] = \mathbf{w}^o \tag{24}$$

Remark 3. For doubly white inputs, the ACLMS converges in the mean to the first $2M$ weight coefficients within $\mathbf{w}_{2N}^o$ in (9) in an unbiased manner.

2.2. Convergence in the mean square

We now consider the mean square error (MSE) performance of the deficient length ACLMS. From (5), the MSE of the ACLMS, denoted by $J(k)$, can be defined as

$$J(k) = E[|e(k)|^2] = E[e(k)e^*(k)] \tag{25}$$

Using the independence assumptions stated in Section 2.1, the MSE in (25) can be evaluated as

$$\begin{aligned}J(k) &= E[\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)\bar{\mathbf{z}}^H(k)\bar{\mathbf{w}}^o] - E[(\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k))\mathbf{z}^H(k)\widetilde{\mathbf{w}}(k)]\\ &\quad - E[\big(\bar{\mathbf{w}}^{oT}\bar{\mathbf{z}}^*(k)\big)\mathbf{z}^T(k)\widetilde{\mathbf{w}}^*(k)] + E[\widetilde{\mathbf{w}}^H(k)\mathbf{z}(k)\mathbf{z}^H(k)\widetilde{\mathbf{w}}(k)] + \sigma_q^2\\ &= \bar{\mathbf{w}}^{oH}\mathbf{R}_{\bar{\mathbf{z}}}\bar{\mathbf{w}}^o - 2\Re\big[\mathbf{b}^H E[\widetilde{\mathbf{w}}(k)]\big] + \mathrm{tr}\big(\mathbf{R}_{\mathbf{z}}\mathbf{K}(k)\big) + \sigma_q^2\end{aligned} \tag{26}$$

where $\Re[\cdot]$ is the 'real part' operator, $\mathbf{R}_{\bar{\mathbf{z}}} = E[\bar{\mathbf{z}}(k)\bar{\mathbf{z}}^H(k)]$ is the $2(N-M) \times 2(N-M)$ augmented covariance matrix of the part of system input which is not considered by the deficient length ACLMS, and $\mathbf{K}(k) = E[\widetilde{\mathbf{w}}(k)\widetilde{\mathbf{w}}^H(k)]$ is the $2M \times 2M$ augmented weight error covariance matrix.

The mean square analysis rests upon the second order properties of the weight error vector $E[\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^H(k)]$ in (26). To this end, we first apply the Hermitian operator $(\cdot)^H$ to both sides of (13), to yield

$$\tilde{\mathbf{w}}^H(k+1) = \tilde{\mathbf{w}}^H(k)\left(\mathbf{I}_{2M} - \mu\mathbf{z}(k)\mathbf{z}^H(k)\right) + \mu\left(\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)\right)\mathbf{z}^H(k) + \mu q(k)\mathbf{z}^H(k) \tag{27}$$

Upon multiplying both sides of (13) by $\tilde{\mathbf{w}}^H(k+1)$ and taking the statistical expectation, the mean square evolution of the augmented weight error vector $\tilde{\mathbf{w}}(k)$ becomes

$$\begin{aligned}\mathbf{K}(k+1) = {} & \mathbf{K}(k) - \mu\mathbf{R}_\mathbf{z}\mathbf{K}(k) - \mu\mathbf{K}(k)\mathbf{R}_\mathbf{z} + \mu^2\mathbf{R}_\mathbf{z}\mathbf{K}(k)\mathbf{R}_\mathbf{z} \\ & +\mu^2\mathbf{P}_\mathbf{z}\mathbf{K}^T(k)\mathbf{P}_\mathbf{z}^* + \mu^2\mathbf{R}_\mathbf{z}\mathrm{tr}\left(\mathbf{R}_\mathbf{z}\mathbf{K}(k)\right) + \mu E[\tilde{\mathbf{w}}(k)]\mathbf{b}^H \\ & +\mu\mathbf{b}E[\tilde{\mathbf{w}}^H(k)] - \mu^2\left(\mathbf{T}(k)+\mathbf{T}^H(k)\right) \\ & +\mu^2\mathbf{D} + \mu^2\sigma_q^2\mathbf{R}_\mathbf{z}\end{aligned} \tag{28}$$

where

$$\mathbf{P}_\mathbf{z} = E[\mathbf{z}(k)\mathbf{z}^T(k)] = \begin{bmatrix} E[\mathbf{x}(k)\mathbf{x}^T(k)] & E[\mathbf{x}(k)\mathbf{x}^H(k)] \\ E[\mathbf{x}^*(k)\mathbf{x}^T(k)] & E[\mathbf{x}^*(k)\mathbf{x}^H(k)]\end{bmatrix} = \begin{bmatrix}\mathbf{P}_\mathbf{x} & \mathbf{R}_\mathbf{x} \\ \mathbf{R}_\mathbf{x}^* & \mathbf{P}_\mathbf{x}^*\end{bmatrix} \tag{29}$$

is the $2M\times 2M$ augmented pseudo-covariance matrix [21], and

$$\mathbf{T}(k) = E[\left(\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)\right)\mathbf{z}(k)\mathbf{z}^H(k)\tilde{\mathbf{w}}(k)\mathbf{z}^H(k)] \tag{30}$$

$$\mathbf{D} = E[|\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)|^2\mathbf{z}(k)\mathbf{z}^H(k)] \tag{31}$$

The detailed expressions of the $2M\times 2M$ matrices $\mathbf{T}(k)$ and $\mathbf{D}$ are given respectively in Appendix A and Appendix B. Note that expressions (26) and (28), for which the convergence of $E[\tilde{\mathbf{w}}(k)]$ is guaranteed by (14), now completely describe the MSE convergence behaviour of the ACLMS algorithm for $M<N$, for both correlated and pseudo-correlated Gaussian input data. The impact of the degree of undermodelling on the mean-square evolution of ACLMS, in the form of $\mathbf{b}$, $\mathbf{T}(k)$ and $\mathbf{D}$, can also be observed.

Remark 4. The MSE behaviour of the deficient length ACLMS is different from that of its sufficient length counterpart, discussed in [18,21], which is mainly due to the influence of the terms $\mathbf{T}(k)$ and $\mathbf{D}$ in (28), as well as the dependence of the mean behaviour on $E[\tilde{\mathbf{w}}(k)]$ in both (26) and (28), all arising from the system undermodelling.

From the above general analysis, the MSE behaviour of ACLMS for the special case of doubly white (DW) Gaussian input can be deduced by noting that in this case, the vector $\mathbf{b}$ in (15) vanishes, and (26) simplifies into

$$J_{\mathrm{DW}}(k) = \bar{\mathbf{w}}^{oH}\{\mathbf{R}_{\bar{\mathbf{z}}}\}_{\mathrm{DW}}\bar{\mathbf{w}}^o + \mathrm{tr}\left(\mathbf{R}_\mathbf{z}\mathbf{K}(k)\right)_{\mathrm{DW}} + \sigma_q^2 \tag{32}$$

For doubly white Gaussian input data, we have

$$\{\mathbf{R}_{\bar{\mathbf{z}}}\}_{\mathrm{DW}} = \begin{bmatrix}\sigma_x^2\mathbf{I}_{N-M} & \rho_x\mathbf{I}_{N-M} \\ \rho_x^*\mathbf{I}_{N-M} & \sigma_x^2\mathbf{I}_{N-M}\end{bmatrix} \tag{33}$$

and

$$\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}} = \begin{bmatrix}\sigma_x^2\mathbf{I}_M & \rho_x\mathbf{I}_M \\ \rho_x^*\mathbf{I}_M & \sigma_x^2\mathbf{I}_M\end{bmatrix} \tag{34}$$

so that the first term on the right hand side (RHS) of (32) can be further decomposed as

$$\begin{aligned}\bar{\mathbf{w}}^{oH}\{\mathbf{R}_{\bar{\mathbf{z}}}\}_{\mathrm{DW}}\bar{\mathbf{w}}^o &= [\bar{\mathbf{h}}^{oH}, \bar{\mathbf{g}}^{oH}]\begin{bmatrix}\sigma_x^2\mathbf{I}_{N-M} & \rho_x\mathbf{I}_{N-M} \\ \rho_x^*\mathbf{I}_{N-M} & \sigma_x^2\mathbf{I}_{N-M}\end{bmatrix}\begin{bmatrix}\bar{\mathbf{h}}^o \\ \bar{\mathbf{g}}^o\end{bmatrix} \\ &= \sigma_x^2\left(\|\bar{\mathbf{h}}^o\|_2^2 + \|\bar{\mathbf{g}}^o\|_2^2\right) + 2\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o]\end{aligned} \tag{35}$$

In the same spirit, for doubly white Gaussian input data, the second term on the RHS of (32) can be decomposed as

$$\mathrm{tr}\left(\mathbf{R}_\mathbf{z}\mathbf{K}(k)\right)_{\mathrm{DW}} = \sigma_x^2\left(\|\tilde{\mathbf{h}}_{\mathrm{DW}}(k)\|_2^2 + \|\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\|_2^2\right) + 2\Re\left[\rho_x\tilde{\mathbf{h}}_{\mathrm{DW}}^H(k)\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\right] \tag{36}$$

Therefore, the MSE in (32) now becomes

$$\begin{aligned}J_{\mathrm{DW}}(k) = {} & \sigma_x^2\left(\|\bar{\mathbf{h}}^o\|_2^2 + \|\bar{\mathbf{g}}^o\|_2^2\right) + 2\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o] \\ & +\sigma_x^2\left(\|\tilde{\mathbf{h}}_{\mathrm{DW}}(k)\|_2^2 + \|\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\|_2^2\right) \\ & +2\Re\left[\rho_x\tilde{\mathbf{h}}_{\mathrm{DW}}^H(k)\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\right] + \sigma_q^2\end{aligned} \tag{37}$$

To establish the evolution of the variance of the weight error vector, $\mathbf{K}_{\mathrm{DW}}(k)$, note that according to Appendix B, for doubly white Gaussian input data, the matrix $\mathbf{D}$ in (31) can be simplified as

$$\mathbf{D}_{\mathrm{DW}} = \begin{bmatrix} d_1\mathbf{I}_M & d_2\mathbf{I}_M \\ d_2^*\mathbf{I}_M & d_1\mathbf{I}_M\end{bmatrix} \tag{38}$$

where

$$d_1 = \sigma_x^4\left(\|\bar{\mathbf{h}}^o\|_2^2 + \|\bar{\mathbf{g}}^o\|_2^2\right) + 2\sigma_x^2\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o] \tag{39}$$

$$d_2 = \sigma_x^2\rho_x\left(\|\bar{\mathbf{h}}^o\|_2^2 + \|\bar{\mathbf{g}}^o\|_2^2\right) + 2\rho_x\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o] \tag{40}$$

and also $\mathbf{b}_{\mathrm{DW}} = \mathbf{0}$, $\mathbf{T}_{\mathrm{DW}}(k) = \mathbf{0}$. Therefore, the expression for $\mathbf{K}_{\mathrm{DW}}(k)$ in (28) can be simplified as

$$\begin{aligned}\mathbf{K}_{\mathrm{DW}}(k+1) = {} & \mathbf{K}_{\mathrm{DW}}(k) - \mu\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}}\mathbf{K}_{\mathrm{DW}}(k) - \mu\mathbf{K}_{\mathrm{DW}}(k)\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}} \\ & +\mu^2\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}}\mathbf{K}_{\mathrm{DW}}(k)\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}} \\ & +\mu^2\{\mathbf{P}_\mathbf{z}\}_{\mathrm{DW}}\mathbf{K}_{\mathrm{DW}}^T(k)\{\mathbf{P}_\mathbf{z}^*\}_{\mathrm{DW}} \\ & +\mu^2\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}}\mathrm{tr}\left(\mathbf{R}_\mathbf{z}\mathbf{K}(k)\right)_{\mathrm{DW}} \\ & +\mu^2\mathbf{D}_{\mathrm{DW}} + \mu^2\sigma_q^2\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}}\end{aligned} \tag{41}$$

where $\{\mathbf{R}_\mathbf{z}\}_{\mathrm{DW}}$, $\mathrm{tr}\left(\mathbf{R}_\mathbf{z}\mathbf{K}(k)\right)_{\mathrm{DW}}$ and $\mathbf{D}_{\mathrm{DW}}$ are given respectively in (34), (36) and (38), and according to (29),

$$\{\mathbf{P}_\mathbf{z}\}_{\mathrm{DW}} = \begin{bmatrix}\rho_x\mathbf{I}_M & \sigma_x^2\mathbf{I}_M \\ \sigma_x^2\mathbf{I}_M & \rho_x^*\mathbf{I}_M\end{bmatrix} \tag{42}$$

Within the above expression for $\mathbf{K}_{\mathrm{DW}}(k)$, of particular interest are the two $M\times M$ diagonal sub-matrices, $\mathbf{H}_{\mathrm{DW}}(k) = E[\tilde{\mathbf{h}}_{\mathrm{DW}}(k)\tilde{\mathbf{h}}_{\mathrm{DW}}^H(k)]$ and $\mathbf{G}_{\mathrm{DW}}(k) = E[\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\tilde{\mathbf{g}}_{\mathrm{DW}}^H(k)]$, since they respectively quantify the mean square evolution of the standard weight error vector and the conjugate weight error vector within the ACLMS. These can be derived from (41) as

$$\begin{aligned}\mathbf{H}_{\mathrm{DW}}(k+1) = {} & (1 - 2\mu\sigma_x^2 + \mu^2\sigma_x^4)\mathbf{H}_{\mathrm{DW}}(k) \\ & +\mu^2|\rho_x|^2\left(\mathbf{H}_{\mathrm{DW}}^T(k) + \mathbf{G}_{\mathrm{DW}}(k)\right) + \mu^2\sigma_x^4\mathbf{G}_{\mathrm{DW}}^T(k) \\ & -\mu\left(\rho_x\mathbf{g}(k)\mathbf{h}^H(k) + \rho_x^*\mathbf{h}(k)\mathbf{g}^H(k)\right) \\ & +2\mu^2\sigma_x^2\Re[\rho_x\mathbf{g}(k)\mathbf{h}^H(k) + \rho_x^*\mathbf{h}(k)\mathbf{g}^H(k)] \\ & +\mu^2\Big(\sigma_x^4\left(\|\tilde{\mathbf{h}}_{\mathrm{DW}}(k)\|_2^2 + \|\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\|_2^2\right) \\ & +2\sigma_x^2\Re\left[\rho_x\tilde{\mathbf{h}}_{\mathrm{DW}}^H(k)\tilde{\mathbf{g}}_{\mathrm{DW}}(k)\right]\Big)\mathbf{I}_M \\ & +\mu^2\Big(\sigma_x^4\left(\|\bar{\mathbf{h}}^o\|_2^2 + \|\bar{\mathbf{g}}^o\|_2^2\right) \\ & +2\sigma_x^2\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o]\Big)\mathbf{I}_M + \mu^2\sigma_q^2\sigma_x^2\mathbf{I}_M\end{aligned} \tag{43}$$

and

$$\begin{aligned}\mathbf{G}_{\mathrm{DW}}(k+1) = {} & (1 - 2\mu\sigma_x^2 + \mu^2\sigma_x^4)\mathbf{G}_{\mathrm{DW}}(k) \\ & +\mu^2|\rho_x|^2\left(\mathbf{G}_{\mathrm{DW}}^T(k) + \mathbf{H}_{\mathrm{DW}}(k)\right) + \mu^2\sigma_x^4\mathbf{H}_{\mathrm{DW}}^T(k) \\ & -\mu\left(\rho_x\mathbf{g}(k)\mathbf{h}^H(k) + \rho_x^*\mathbf{h}(k)\mathbf{g}^H(k)\right) \\ & +2\mu^2\sigma_x^2\Re[\rho_x\mathbf{g}(k)\mathbf{h}^H(k) + \rho_x^*\mathbf{h}(k)\mathbf{g}^H(k)]\end{aligned}$$

$$+\mu^2\big(\sigma_x^4\big(\|\widetilde{\mathbf{h}}_{\mathrm{DW}}(k)\|_2^2+\|\widetilde{\mathbf{g}}_{\mathrm{DW}}(k)\|_2^2\big)$$
$$+2\sigma_x^2\Re\big[\rho_x\widetilde{\mathbf{h}}_{\mathrm{DW}}^H(k)\widetilde{\mathbf{g}}_{\mathrm{DW}}(k)\big]\big)\mathbf{I}_M$$
$$+\mu^2\big(\sigma_x^4\big(\|\bar{\mathbf{h}}^o\|_2^2+\|\bar{\mathbf{g}}^o\|_2^2\big)$$
$$+2\sigma_x^2\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o]\big)\mathbf{I}_M+\mu^2\sigma_q^2\sigma_x^2\mathbf{I}_M \tag{44}$$

Remark 5. The mean square recursions for $\mathbf{H}_{\mathrm{DW}}(k)$ in (43) and $\mathbf{G}_{\mathrm{DW}}(k)$ in (44) are not statistically independent. This is due to the fact that the weight updates for the standard part and the conjugate part in ACLMS, given in (3) and (4) respectively, involve the same error $e(k)$ which comprises both the outputs of the standard and conjugate parts, as shown in (5).

2.3. Mean square stability

We shall now consider the sufficient conditions for MSE convergence of the deficient length ACLMS. For mathematical tractability, we first need to diagonalise the augmented covariance matrix $\mathbf{R_z}$ as

$$\mathbf{R_z}=\mathbf{Q}\boldsymbol{\Lambda}_{\mathbf{R}}\mathbf{Q}^H \tag{45}$$

where $\boldsymbol{\Lambda}_{\mathbf{R}}=\mathrm{diag}\{\lambda_1,\lambda_2,\dots,\lambda_{2M}\}$, $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_{2M}$, is the $2M\times 2M$ diagonal matrix of the real valued eigenvalues of $\mathbf{R_z}$, and $\mathbf{Q}$ is a $2M\times 2M$ unitary matrix composed of the eigenvectors of $\mathbf{R_z}$, so that $\mathbf{Q}\mathbf{Q}^H=\mathbf{Q}^H\mathbf{Q}=\mathbf{I}_{2M}$.

For a clearer physical insight into the eigenstructure of the augmented covariance matrix $\mathbf{R_z}$, and to reduce the computational complexity, we here employ the recently introduced approximate uncorrelating (AUT) [29], which allows us to jointly diagonalise both the covariance matrix $\mathbf{R_x}$ and the pseudo-covariance matrix $\mathbf{P_x}$ within the augmented covariance matrix $\mathbf{R_z}$ with a single singular value decomposition (SVD). The Takagi factorization states that any complex symmetric matrix, such as the pseudo-covariance matrix $\mathbf{P_x}^T=\mathbf{P_x}$, can be diagonalised as [41]

$$\mathbf{P_x}=\mathbf{Q_P}\mathbf{D_P}\mathbf{Q_P}^T \tag{46}$$

where $\mathbf{Q_P}$ is an $M\times M$ unitary matrix, $\mathbf{Q_P}\mathbf{Q_P}^H=\mathbf{Q_P}^H\mathbf{Q_P}=\mathbf{I}_M$, and $\mathbf{D_P}=\mathrm{diag}\{p_1,p_2,\dots,p_M\}$ is an $M\times M$ diagonal matrix of real-valued entries, $p_1\geq p_2\geq\cdots\geq p_M$. The approximate uncorrelating transform (AUT) establishes that the same matrix $\mathbf{Q_P}$ can be used to approximately diagonalise the covariance matrix $\mathbf{R_x}$, so that [29]

$$\mathbf{R_x}\simeq\mathbf{Q_P}\mathbf{D_R}\mathbf{Q_P}^H \tag{47}$$

where $\mathbf{D_R}=\mathrm{diag}\{r_1,r_2,\dots,r_M\}$ is a diagonal matrix of real-valued entries, and $r_1\geq r_2\geq\cdots\geq r_M$. The approximation in (47) is valid for univariate data, and the equality is achieved when $\mathbf{x}(k)$ is real-valued, that is, maximum noncircular, or doubly white [21,29]. Therefore, the matrices $\mathbf{Q}$ and $\boldsymbol{\Lambda}_{\mathbf{R}}$ in (45) can be expressed as [21]

$$\mathbf{Q}=\frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{Q_P} & -\mathbf{Q_P}\\ \mathbf{Q_P}^* & \mathbf{Q_P}^*\end{bmatrix} \tag{48}$$

and

$$\boldsymbol{\Lambda}_{\mathbf{R}}=\begin{bmatrix}\mathbf{D_R}+\mathbf{D_P} & \mathbf{0}\\ \mathbf{0} & \mathbf{D_R}-\mathbf{D_P}\end{bmatrix} \tag{49}$$

and hence

$$\mathbf{R_z}^{-1}=\mathbf{Q}\boldsymbol{\Lambda}_{\mathbf{R}}^{-1}\mathbf{Q}^H \tag{50}$$

In a similar way, the augmented pseudo-covariance matrix $\mathbf{P_z}$ in (29) can be diagonalised as [21]

$$\mathbf{P_z}=\mathbf{Q}\boldsymbol{\Lambda}_{\mathbf{P}}\mathbf{Q}^T \tag{51}$$

where

$$\boldsymbol{\Lambda}_{\mathbf{P}}=\begin{bmatrix}\mathbf{D_R}+\mathbf{D_P} & \mathbf{0}\\ \mathbf{0} & -(\mathbf{D_R}-\mathbf{D_P})\end{bmatrix} \tag{52}$$

We can now rotate the augmented weight error vector and the augmented input vector as

$$\hat{\mathbf{w}}(k)=\mathbf{Q}^H\widetilde{\mathbf{w}}(k)\quad\text{and}\quad\hat{\mathbf{z}}(k)=\mathbf{Q}^H\widetilde{\mathbf{z}}(k) \tag{53}$$

In a similar way, we can define

$$\hat{\mathbf{T}}(k)=\mathbf{Q}^H\mathbf{T}(k)\mathbf{Q}\quad\text{and}\quad\hat{\mathbf{D}}=\mathbf{Q}^H\mathbf{D}\mathbf{Q} \tag{54}$$

Therefore, the term $\mathrm{tr}\big(\mathbf{R_z}\mathbf{K}(k)\big)$ on the RHS of (28) can be decomposed as

$$\mathrm{tr}\big(\mathbf{R_z}\mathbf{K}(k)\big)=E[\widetilde{\mathbf{w}}^H(k)\mathbf{R_z}\widetilde{\mathbf{w}}(k)]=E[\hat{\mathbf{w}}^H(k)\boldsymbol{\Lambda}\hat{\mathbf{w}}(k)]=\boldsymbol{\lambda}^T\mathbf{v}(k) \tag{55}$$

where the $2M\times 1$ vector $\boldsymbol{\lambda}=[\lambda_1,\lambda_2,\dots,\lambda_{2M}]^T$, and $\mathbf{v}(k)$ is the $2M\times 1$ second order moment vector, the components of which are the diagonal elements of $E[\hat{\mathbf{w}}(k)\hat{\mathbf{w}}^H(k)]$. Then, based on (28), the evolution of $\mathbf{v}(k)$ can be expressed as

$$\mathbf{v}(k+1)=\mathbf{A}\mathbf{v}(k)+\mathbf{c}(k)+\mu^2\mathbf{k}+\mu^2\sigma_q^2\boldsymbol{\lambda} \tag{56}$$

with

$$\begin{aligned}\mathbf{A}&=\mathbf{I}_{2M}-2\mu\boldsymbol{\Lambda}_{\mathbf{R}}+\mu^2\boldsymbol{\Lambda}_{\mathbf{R}}^2+\mu^2\boldsymbol{\Lambda}_{\mathbf{P}}^2+\mu^2\boldsymbol{\lambda}\boldsymbol{\lambda}^T\\&=\mathbf{I}_{2M}-2\mu\boldsymbol{\Lambda}_{\mathbf{R}}+2\mu^2\boldsymbol{\Lambda}_{\mathbf{R}}^2+\mu^2\boldsymbol{\lambda}\boldsymbol{\lambda}^T\end{aligned} \tag{57}$$

$$\mathbf{c}(k)=2\mu\Re\big[\mathbf{Y}E[\hat{\mathbf{w}}^*(k)]\big]-2\mu^2\Re[\mathbf{p}(k)] \tag{58}$$

$$\mathbf{Y}=\mathrm{diag}\{\hat{b}_1,\hat{b}_2,\dots,\hat{b}_{2M}\} \tag{59}$$

where $\hat{b}_i$ is the ith component of $\hat{\mathbf{b}}=\mathbf{Q}^H\mathbf{b}$. The $2M\times 1$ vector $\mathbf{p}(k)=[\hat{T}_{1,1}(k),\hat{T}_{2,2}(k),\dots,\hat{T}_{2M,2M}(k)]^T$, in which $\hat{T}_{i,i}(k)$ is the ith diagonal element of $\hat{\mathbf{T}}(k)=\mathbf{Q}^H\mathbf{T}(k)\mathbf{Q}$, and the $2M\times 1$ vector $\mathbf{k}=[\hat{D}_{1,1},\hat{D}_{2,2},\dots,\hat{D}_{2M,2M}]^T$, in which $\hat{D}_{i,i}$ is the ith diagonal element of $\hat{\mathbf{D}}=\mathbf{Q}^H\mathbf{D}\mathbf{Q}$.

Remark 6. The convergence of the recursion in (56) is subject to two conditions: 1) all eigenvalues of the $2M\times 2M$ transition matrix $\mathbf{A}$ are less than unity; 2) the $2M\times 1$ vector $\mathbf{c}(k)$ is bounded.

By considering (14), (30) and (58), we find that the boundedness of $\mathbf{c}(k)$ is guaranteed if $E[\hat{\mathbf{w}}(k)]$ is bounded. Since $E[\hat{\mathbf{w}}(k)]=\mathbf{Q}^H E[\widetilde{\mathbf{w}}(k)]$, by considering (20), condition 2) is satisfied if $0<\mu<(2/\mathrm{tr}(\mathbf{R_z}))$. It is well known in the literature that condition 1) results in the following bound [42]:

$$0<\mu<\frac{2}{3\mathrm{tr}(\mathbf{R_z})} \tag{60}$$

The upper bound in (60) is tighter than that in condition 2), and therefore, mean square convergence of the deficient length ACLMS is guaranteed if the step-size μ satisfies (60). As expected, this bound has an identical form to that obtained for mean-square stability of the deficient length LMS due to the fact that the transition matrix $\mathbf{A}$ in (56) is identical in both cases [25].

2.4. Steady state analysis

To investigate the steady state performance of the deficient length ACLMS, based on (55), we shall first rewrite the MSE in (26) as

$$J(k)=\sigma_q^2+u-2\Re\big[\mathbf{b}^H E[\widetilde{\mathbf{w}}(k)]\big]+\boldsymbol{\lambda}^T\mathbf{v}(k) \tag{61}$$

where

$$u=\bar{\mathbf{w}}^{oH}\mathbf{R_{\bar{z}}}\bar{\mathbf{w}}^o \tag{62}$$

Suppose that μ is chosen such that the mean square stability is guaranteed; then using (22), (50) and $\mathbf{b}=\mathbf{Q}\hat{\mathbf{b}}$, the steady state MSE, that is, $J(\infty)$, can be derived as

$$J(\infty) = \sigma_q^2 + u - 2\mathbf{b}^H\mathbf{R}_\mathbf{z}^{-1}\mathbf{b} + \boldsymbol{\lambda}^T\mathbf{v}(\infty)$$
$$= \sigma_q^2 + u - 2\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j} + \boldsymbol{\lambda}^T\mathbf{v}(\infty) \tag{63}$$

On the other hand, the minimum MSE, denoted by $J_{\min}$, of the deficient length ACLMS can be obtained by substituting $E[\widetilde{\mathbf{w}}(\infty)]$ for $\widetilde{\mathbf{w}}(k)$ in (61), to give

$$J_{\min} = \sigma_q^2 + u - \sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j} \tag{64}$$

which provides an exact measure of the effect of the level of system undermodelling on the minimum MSE. For comparison, the minimum MSE attained by the sufficient length ACLMS is σ_q^2 [18,19].

By definition, the steady state MSE of the LMS-type adaptive algorithm is described by [43]

$$J(\infty) = J_{\min} + J_{\text{ex}}(\infty) \tag{65}$$

where $J_{\text{ex}}(\infty)$ is the steady state excess MSE. From (63) and (64), $J_{\text{ex}}(\infty)$ of the deficient length ACLMS assumes the form

$$J_{\text{ex}}(\infty) = \boldsymbol{\lambda}^T\mathbf{v}(\infty) - \sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j} \tag{66}$$

where using (64) in (56) and considering that in the steady state $n\to\infty$, we arrive at

$$\mathbf{v}(\infty) = (\mathbf{I}_{2M}-\mathbf{A})^{-1}\mathbf{d}(\infty)$$
$$= (2\mu\boldsymbol{\Lambda}_\mathbf{R} - 2\mu^2\boldsymbol{\Lambda}_\mathbf{R}^2 - \mu^2\boldsymbol{\lambda}\boldsymbol{\lambda}^T)^{-1}\mathbf{d}(\infty) \tag{67}$$

where

$$\mathbf{d}(\infty) = \mathbf{c}(\infty) + \mu^2\mathbf{k} + \mu^2\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j}\boldsymbol{\lambda} - \mu^2 u\boldsymbol{\lambda} + \mu^2 J_{\min}\boldsymbol{\lambda} \tag{68}$$

Upon defining

$$\boldsymbol{\Lambda} = 2\boldsymbol{\Lambda}_\mathbf{R} - 2\mu\boldsymbol{\Lambda}_\mathbf{R}^2 \tag{69}$$

and employing the matrix inversion lemma, after a few manipulations, from (67) we arrive at

$$\mathbf{v}(\infty) = \left(\mathbf{I}_{2M} + \frac{\mu\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}\boldsymbol{\lambda}^T}{1-\boldsymbol{\lambda}^T\mu\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}}\right)\mu^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{d}(\infty) \tag{70}$$

and hence

$$\boldsymbol{\lambda}^T\mathbf{v}(\infty) = \frac{\mu^{-1}\boldsymbol{\lambda}^T\boldsymbol{\Lambda}^{-1}\mathbf{d}(\infty)}{1-\mu\boldsymbol{\lambda}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{\lambda}} \tag{71}$$

Substituting (71) into (66) and following the analysis given in Appendix C, we obtain

$$J_{\text{ex}}(\infty) = \frac{\sum_{i=1}^{2M}\frac{\mu f_i}{2-2\mu\lambda_i} + \sum_{i=1}^{2M}\frac{\mu\lambda_i J_{\min}}{2-2\mu\lambda_i}}{1-\sum_{i=1}^{2M}\frac{\mu\lambda_i}{2-2\mu\lambda_i}} \tag{72}$$

where

$$f_i = \hat{D}_{i,i} - 2\Re[\hat{T}_{i,i}(\infty)] + 2\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j}\lambda_i + 2|\hat{b}_i|^2 - u\lambda_i \tag{73}$$

Eq. (72) is an expression for the steady state excess MSE of the deficient length ACLMS for correlated and pseudo-correlated Gaussian input data.

Remark 7. A complete view on the steady state MSE performance of deficient length ACLMS can only be obtained by considering (64), (65) and (72) together. Observe that the steady state MSE is influenced by the filter length $2M$, values of the system impulse response w_j^o, $j=2M+1, 2M+2, \ldots, 2N$, which are not considered by ACLMS due to the system undermodelling problem, full second order statistics of input data (both covariance and pseudo-covariance coefficients), step-size and noise variance.

For doubly white Gaussian input data, it can be verified from (73) that $f_i=0$, since $\hat{T}_{i,i}(\infty)=0$, $\hat{b}_i=0$, $\hat{D}_{i,i}=u\lambda_i$, and according to the eigenstructure of the augmented covariance matrix in (45), we have[2]

$$\lambda_i = \begin{cases} \sigma_x^2 + |\rho_x|, & i=1,2,\ldots,M \\ \sigma_x^2 - |\rho_x|, & i=M+1,M+2,\ldots,2M \end{cases} \tag{74}$$

Consider a measure, κ, of the degree of noncircularity within the complex-valued doubly white random variable x, defined as the ratio of the absolute value of its pseudo-covariance $|\rho_x|$ to its covariance σ_x^2, giving $\kappa=|\rho_x|/\sigma_x^2$, which is bounded within [0, 1) [44,45]. Then

$$\lambda_i = \begin{cases} \sigma_x^2(1+\kappa), & i=1,2,\ldots,M \\ \sigma_x^2(1-\kappa), & i=M+1,M+2,\ldots,2M \end{cases} \tag{75}$$

Therefore, the first term in (72) vanishes, leading to the following expression for the steady state excess MSE

$$J_{\text{ex}_{\text{DW}}}(\infty) = \frac{\sum_{i=1}^{2M}\frac{\mu\lambda_i J_{\min_{\text{DW}}}}{2-2\mu\lambda_i}}{1-\sum_{i=1}^{2M}\frac{\mu\lambda_i}{2-2\mu\lambda_i}}$$
$$= \frac{\left(\frac{M\mu\sigma_x^2(1+\kappa)}{2-2\mu\sigma_x^2(1+\kappa)} + \frac{M\mu\sigma_x^2(1-\kappa)}{2-2\mu\sigma_x^2(1-\kappa)}\right)J_{\min_{\text{DW}}}}{1-\frac{M\mu\sigma_x^2(1+\kappa)}{2-2\mu\sigma_x^2(1+\kappa)} - \frac{M\mu\sigma_x^2(1-\kappa)}{2-2\mu\sigma_x^2(1-\kappa)}} \tag{76}$$

so that from (64), (62) and (35), we finally have

$$J_{\min_{\text{DW}}} = \sigma_q^2 + u_{\text{DW}}$$
$$= \sigma_q^2 + \sigma_x^2\left(\|\bar{\mathbf{h}}^o\|_2^2 + \|\bar{\mathbf{g}}^o\|_2^2\right) + 2\Re[\rho_x\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o] \tag{77}$$

Therefore, the steady state MSE of the deficient length ACLMS for a doubly white second order noncircular Gaussian input can be obtained as

$$J_{\text{DW}}(\infty) = J_{\min_{\text{DW}}} + J_{\text{ex}_{\text{DW}}}(\infty) \tag{78}$$

For comparison, the steady state performance of the 'sufficient length' ACLMS is given by [18,21]

$$J_{\text{ex}_{\text{DW}}}(\infty) = \frac{\sum_{i=1}^{2N}\frac{\mu\lambda_i J_{\min_{\text{DW}}}}{2-2\mu\lambda_i}}{1-\sum_{i=1}^{2N}\frac{\mu\lambda_i}{2-2\mu\lambda_i}}$$
$$= \frac{\left(\frac{N\mu\sigma_x^2(1+\kappa)}{2-2\mu\sigma_x^2(1+\kappa)} + \frac{N\mu\sigma_x^2(1-\kappa)}{2-2\mu\sigma_x^2(1-\kappa)}\right)J_{\min_{\text{DW}}}}{1-\frac{N\mu\sigma_x^2(1+\kappa)}{2-2\mu\sigma_x^2(1+\kappa)} - \frac{N\mu\sigma_x^2(1-\kappa)}{2-2\mu\sigma_x^2(1-\kappa)}}$$

[2] Note that the diagonal elements $\{p_i\}, i=1,2,\ldots,M$ in $\boldsymbol{\Lambda}_\mathbf{P}$ obtained by the Takagi factorization are nonnegative square roots of $\mathbf{P}_\mathbf{x}\mathbf{P}_\mathbf{x}^H$ [41]. Therefore, for doubly white second order noncircular data, we have $\mathbf{P}_\mathbf{x}\mathbf{P}_\mathbf{x}^H = |\rho_x|^2\mathbf{I}_M$, and hence $p_i=|\rho_x|$ for $\forall i$.

$$\begin{aligned} J_{\min_{\mathrm{DW}}} &= \sigma_q^2 \\ J_{\mathrm{DW}}(\infty) &= J_{\min_{\mathrm{DW}}} + J_{\mathrm{ex}_{\mathrm{DW}}}(\infty) \end{aligned} \tag{79}$$

Remark 8. The steady state MSE of the sufficient length ACLMS increases with an increase in noncircularity of the doubly white Gaussian input signal. This can be verified through the first derivative of $J_{\mathrm{DW}}(\infty)$ in (79) with respect to the degree of noncircularity κ, which always gives $\partial J_{\mathrm{DW}}(\infty)/\partial\kappa > 0$ for κ bounded within [0, 1) [21].

Remark 9. The relationship between the steady state MSE $J_{\mathrm{DW}}(\infty)$ of the deficient length ACLMS and the degree of input noncircularity κ is more complicated. This is because κ is implicitly, and to some extent nonlinearly, involved in the bias term u_{DW}, which arises within the minimum MSE $J_{\min_{\mathrm{DW}}}$ in (77) due to system undermodelling. The steady state MSE performance may therefore decrease or increase as the input noncircularity κ increases.

To clarify this observation, consider a doubly white Gaussian input, $x(k)$, whose pseudo-covariance ρ_x is real-valued, i.e., $E[x_r x_i] = 0$. Without loss generality, we further assume that ρ_x is positive, i.e. $\sigma_{x_r}^2 > \sigma_{x_i}^2$ [46]. In this case, the minimum MSE $J_{\min_{\mathrm{DW}}}$ in (77) can be rewritten as

$$\begin{aligned} J_{\min_{\mathrm{DW}}} &= \sigma_q^2 + u_{\mathrm{DW}} \\ &= \sigma_q^2 + \sigma_x^2\Big(\big(\|\bar{\mathbf{h}}^{\mathrm{o}}\|_2^2 + \|\bar{\mathbf{g}}^{\mathrm{o}}\|_2^2\big) + 2\kappa\Re[\bar{\mathbf{h}}^{\mathrm{o}H}\bar{\mathbf{g}}^{\mathrm{o}}]\Big) \end{aligned} \tag{80}$$

If the value of $\Re[\bar{\mathbf{h}}^{\mathrm{o}H}\bar{\mathbf{g}}^{\mathrm{o}}]$ in (80) is positive, it is easily verified that $\partial J_{\min_{\mathrm{DW}}}/\partial\kappa > 0$, $\partial J_{\mathrm{ex}_{\mathrm{DW}}}(\infty)/\partial\kappa > 0$ and hence $\partial J_{\mathrm{DW}}(\infty)/\partial\kappa > 0$, which indicates that the steady state MSE $J_{\mathrm{DW}}(\infty)$ of the deficient length ACLMS increases with the increase in the input noncircularity. However, if $\Re[\bar{\mathbf{h}}^{\mathrm{o}H}\bar{\mathbf{g}}^{\mathrm{o}}]$ is negative, we have $\partial J_{\min_{\mathrm{DW}}}/\partial\kappa < 0$, but the sign of $\partial J_{\mathrm{ex}_{\mathrm{DW}}}(\infty)/\partial\kappa$ can be either positive or negative, and hence the steady state MSE $J_{\mathrm{DW}}(\infty)$ may decrease as the input noncircularity κ increases. Note that this term, u_{DW}, vanishes in sufficient length ACLMS, due to the model matching.

3. Simulations

Numerical examples were conduced in the MATLAB programming environment in order to evaluate the theoretical findings on both the mean and mean square convergence performance of the deficient length ACLMS algorithm. Experiments were performed in a nonlinear system identification setting [4], where the system to be identified was a strictly linear FIR channel of length N, for which the weight coefficients $\mathbf{h}_N^{\mathrm{o}}$ were drawn from a uniformly distributed complex-valued vector random variable, the desired signal $d(k) = 2\Re[\mathbf{h}_N^{\mathrm{o}H}\mathbf{x}_N(k)] + q(k)$, where the system input $\mathbf{x}_N(k)$ was assumed to be a zero-mean complex-valued second order noncircular Gaussian process and $q(k)$ was another zero-mean complex-valued doubly white circular Gaussian noise with $\sigma_q^2 = E[|q(k)|^2]$. Note that the advantage of the widely linear model based ACLMS over the strictly linear CLMS for this task was straightforward to justify, since the desired signal $d(k)$ can be described in the widely linear sense as $d(k) = \mathbf{h}_N^{\mathrm{o}H}\mathbf{x}_N(k) + \mathbf{g}_N^{\mathrm{o}H}\mathbf{x}_N^*(k) + q(k) = \mathbf{w}_{2N}^{\mathrm{o}H}\mathbf{z}_{2N}(k) + q(k)$, with $\mathbf{g}_N^{\mathrm{o}} = \mathbf{h}_N^{\mathrm{o}*}$, $\mathbf{w}(k) = [\mathbf{h}^T(k), \mathbf{g}^T(k)]^T$ and $\mathbf{z}(k) = [\mathbf{x}^T(k), \mathbf{x}^H(k)]^T$. The weights within the ACLMS were initialised with zeros, and simulation results were obtained by averaging over 1000 independent trials.

3.1. Correlated and pseudo-correlated Gaussian input

In the first experiment, both the unknown system and a widely linear FIR filter trained by the ACLMS algorithm were fed with a complex-valued correlated and pseudo-correlated second order noncircular Gaussian signal, given by

$$x(k) = 0.8x(k-1) + 0.1x^*(k-1) + \eta(k)$$

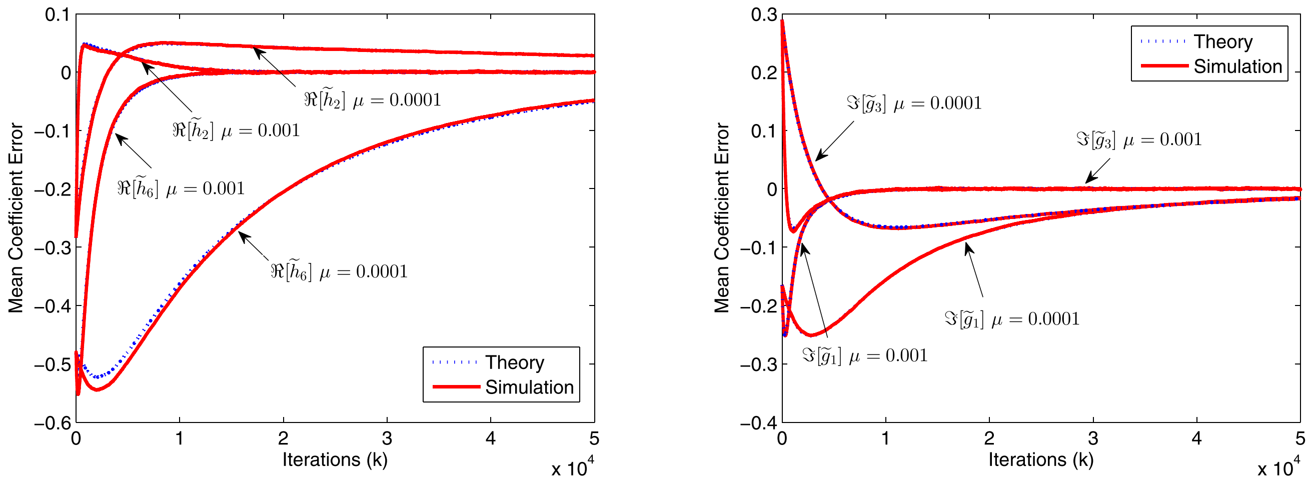
where $\eta(k)$ is a zero-mean doubly white circular Gaussian noise process with unit variance. The real and the imaginary parts of $x(k)$ therefore obey different first order autoregressive processes.

We first chose $N = 10$, $M = 8$, $\sigma_q^2 = 0.0001$, and considered a small step-size $\mu = 0.0001$ and a larger step-size $\mu = 0.001$. Fig. 1 (a) and (b) illustrate the mean convergence behaviour, in terms of real and imaginary parts of several weight error coefficients, of the deficient length ACLMS, obtained from the simulations and theoretical analysis from (14). Fig. 2 (a) and (b) show respectively the theoretical and simulated convergence behaviours of the weight error coefficient power, $\mathrm{tr}\big(\mathbf{K}(k)\big)$, and the MSE $J(k)$, whose theoretical evolutions are respectively given in (28) and (61). It can be observed from Fig. 1 and Fig. 2 that our theoretical analysis accurately describes both the mean and mean square evolutions of the weight error coefficients, as well as the MSE performance, of the deficient length ACLMS for correlated and pseudo-correlated second order noncircular Gaussian inputs. In the next stage, we investigated the validity of the proposed steady state mean square performance analysis of the deficient length ACLMS. In order to achieve the closed-form expression of the steady state MSE $J(\infty)$, given in (65), the approximate uncorrelating transform (AUT) [29] was employed to approximately diagonalise the input covariance matrix $\mathbf{R_x}$ by using the same orthogonal matrix $\mathbf{Q_p}$ from the Takagi factorisation of its pseudo-covariance $\mathbf{P_x}$. Within $J(\infty)$ in (65), its corresponding minimum MSE, $J_{\min}$, and excess MSE, $J_{\mathrm{ex}}(\infty)$, were respectively evaluated by using (64) and (72), as discussed in **Remark 7**. Fig. 3 illustrates both the theoretically evaluated $J(\infty)$ and the simulated one of the deficient length ACLMS for different values of the filter length M, where $\mu = 0.002$, $\sigma_q^2 = 0.001$. In all the cases, a good agreement between the analytical and empirical results can be observed, which also implies a high accuracy of AUT for the joint diagonalisation task. As expected, the steady state MSE performance of the deficient length ACLMS improved as the filter length increased, which is mainly due to a decrease in the bias term $u = \bar{\mathbf{w}}^{\mathrm{o}H}\mathbf{R}_{\bar{\mathbf{z}}}\bar{\mathbf{w}}^{\mathrm{o}}$ within the minimum MSE $J_{\min}$ in (64), which results from the system undermodelling.

3.2. Doubly white Gaussian input

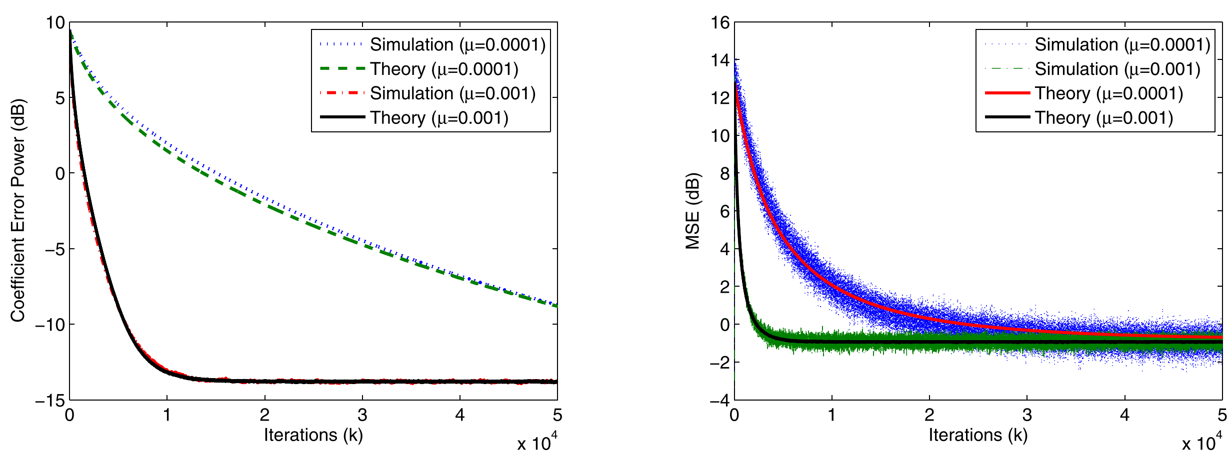
We next considered the case where both the unknown system to be identified and the adaptive filter were excited by a zero-mean second order noncircular doubly white Gaussian signal $x(k)$ with unit variance, for which both the covariance coefficient $r(i)$ and the pseudo-covariance coefficient $p(i)$ vanish for lags $i \neq 0$. The length of the unknown channel impulse response N was fixed to $N = 40$. We first designed $x(k)$ to have pseudo-covariance $\rho = 0.2$, to give a degree of noncircularity $\kappa = \rho/\sigma^2 = 0.2$, and chose the length of adaptive filter $M = 25$, step-size $\mu = 0.001$, and system noise variance $\sigma_q^2 = 0.0001$. For illustration, Fig. 4 shows the mean behaviour of the fifth and twenty-fifth weight error coefficient in terms of their real part and that of the tenth and twentieth weight error coefficient in terms of their imaginary part, obtained from the simulations and theoretical analysis given in (24). Note that in this case, the vector $\mathbf{b}$ in (24) vanishes since the Gaussian input is doubly white, enabling the deficient length ACLMS adaptive filter to converge in the mean sense to the first $2M$ weight coefficients within the unknown system impulse response vector $\mathbf{w}_{2N}^{\mathrm{o}}$, that is, $\mathbf{w}^{\mathrm{o}}$ in (9), in an unbiased manner, as discussed in Remark 3 and illustrated in Fig. 4. We next compared the steady state MSE given by (78) and its simulated counterpart for different values of M, μ and σ_q^2. As illustrated in Table 1, a good agreement between the theoretical and simulated steady state MSE was achieved, with the maximum performance difference of only 0.006 dB.

As discussed in Remark 8 and Remark 9, by virtue of AUT, we were able to diagonalise both the covariance matrix and the pseudo-covariance matrix within the augmented covariance matrix



(a) Real part of the weight error coefficients (b) Imaginary part of the weight error coefficients

Fig. 1. Comparison of the theoretical and simulated curves of the mean evolution of some weight error coefficients of the deficient length ACLMS. Both correlated and pseudo-correlated Gaussian input data were considered, with $N = 10$, $M = 8$, and $\sigma_q^2 = 0.0001$. (a) Real part, and (b) Imaginary part.



(a) Weight error power (b) MSE

Fig. 2. Comparison of the theoretical and simulated curves for the mean square behaviours of the deficient length ACLMS for correlated and pseudo-correlated Gaussian input data, with $N = 10$, $M = 8$, and $\sigma_q^2 = 0.0001$. (a) Weight error power, and (b) MSE.

Table 1
Comparison of theoretical and simulated steady state MSE of the deficient length ACLMS for doubly white second order noncircular Gaussian input data, and for different values of M, μ and σ_q^2.

μ	σ_q^2	Theoretical $J(\infty)$ (dB) in (78)			Simulated $J(\infty)$ (dB)		
		$M = 10$	$M = 20$	$M = 30$	$M = 10$	$M = 20$	$M = 30$
0.0001	0.0001	14.317	12.520	10.090	14.321	12.526	10.101
0.0005	0.001	14.334	12.561	10.143	14.337	12.566	10.145
0.001	0.01	14.358	12.607	10.214	14.360	12.613	10.218

with a single SVD, so as to build up an intuitive and explicit link between the theoretical steady state MSE of the deficient length ACLMS algorithm and the degree of noncircularity κ of the doubly white Gaussian input data. Unlike the sufficient length ACLMS, for which the steady state MSE always increases with an increase in the noncircularity κ of the doubly white input, the steady state MSE of its deficient length counterpart may even decrease as the input noncircularity κ increases, if the value of $\Re[\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^{o}]$ in (80) is negative. To illustrate this phenomenon, in the final experiment, we compared the steady state MSE performances of both the deficient length ACLMS and its sufficient length counterpart for varying degrees of input noncircularity. Since the doubly white Gaussian input with unit variance was used in the simulations, for a fair comparison, the different degrees of input noncircularity κ were directly achieved by varying its pseudo-covariance ρ_x. The upper and the lower panel of Fig. 5 give respectively both the simulated and theoretical steady state MSEs of the deficient length and sufficient length ACLMS for varying degrees of input noncircularity κ, which conforms with the corresponding analysis in Remark 9 and Remark 8. The agreement between the simulated and theoretical MSE performances can also be observed in Fig. 5 for both the deficient length ACLMS and its sufficient length counterpart.

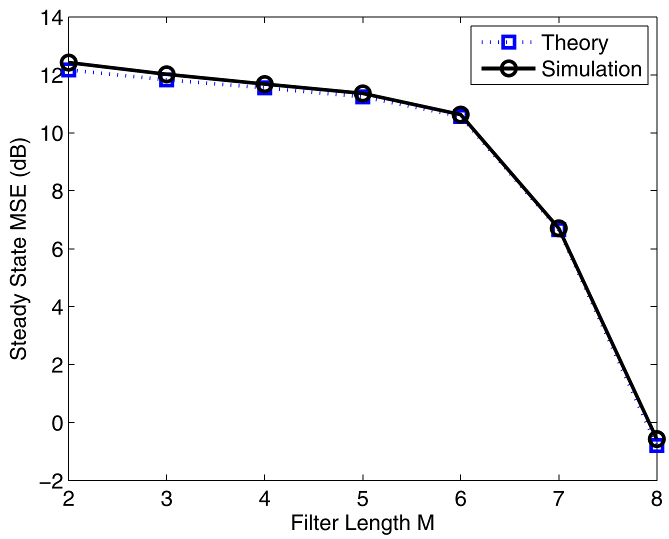


Fig. 3. Comparison of the theoretical and simulated steady state MSE (dB) of the deficient length ACLMS for both correlated and pseudo-correlated Gaussian input data, with $N = 10$, $\mu = 0.002$, $\sigma_q^2 = 0.001$, and for different filter lengths M.

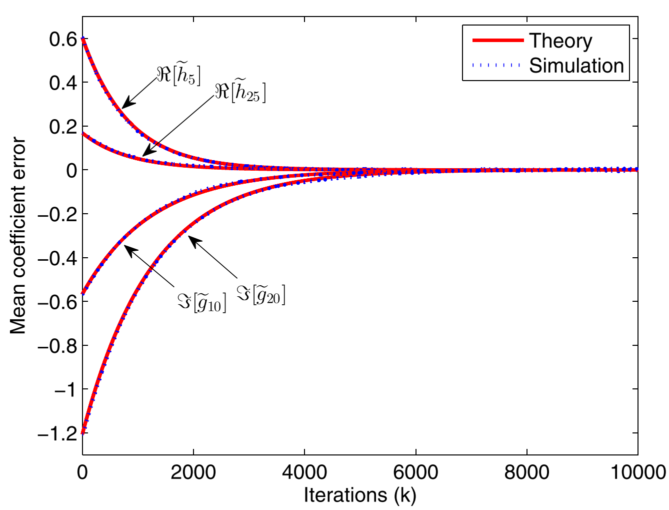


Fig. 4. Comparison of the theoretical and simulated curves for the mean behaviour of several weight error coefficients of the deficient length ACLMS for doubly white Gaussian input data, with $N = 40$, $M = 25$, $\mu = 0.001$, and $\sigma_q^2 = 0.0001$.

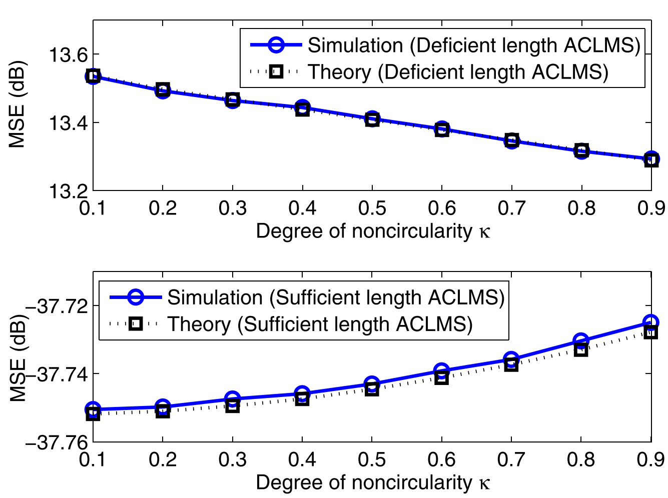


Fig. 5. Steady state MSE performances of both the deficient length ACLMS and its sufficient length counterpart for varying degrees of input noncircularity κ with $N = 40$, $M = 20$, $\mu = 0.01$, $\sigma_q^2 = 0.0001$, and $\Re[\bar{\mathbf{h}}^{oH}\bar{\mathbf{g}}^o] = -0.622$.

4. Conclusions

The mean and mean square convergence behaviours of the deficient length ACLMS have been analysed for second order noncircular Gaussian input data. The analysis has provided theoretical mean and mean square evolutions for the weight error coefficients, as well as the MSE performance, and has shown that the deficient length ACLMS exhibits a different convergence behaviour as compared with the sufficient length ACLMS for both correlated and pseudo-correlated Gausian input data. For doubly white Gaussian input data, the deficient length ACLMS has been shown to exhibit similar mean convergence properties to its sufficient length counterpart, in the sense that its weight coefficients unbiasedly converge to the corresponding coefficients of an unknown system. Mean square stability conditions and closed-form expressions for the steady state MSE and excess MSE of the deficient length ACLMS have been derived using the recently introduced approximate uncorrelating transform (AUT), which diagonalises the covariance and pseudo-covariance matrices with a single singular value decomposition (SVD). We have also shown that unlike the sufficient length ACLMS, for which the steady state MSE always increases with an increase in the degree of noncircularity of doubly white Gaussian inputs, the steady state MSE of its deficient length counterpart may even decrease as the input noncircularity increases. Simulations in system identification tasks support the analysis.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grants 61401094 and 61771124, the Natural Science Foundation of Jiangsu Province under Grant BK20140645, and the Fundamental Research Funds for the Central Universities (Grant No. 2242016K41050).

Appendix A. The Detailed Form of T(k) in (30)

Since $\bar{\mathbf{w}}^o = [\bar{\mathbf{h}}^{oT}, \bar{\mathbf{g}}^{oT}]^T$ and $\bar{\mathbf{z}}(k) = [\bar{\mathbf{x}}^T(k), \bar{\mathbf{x}}^H(k)]^T$, observe that on the RHS of (30), the term $\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)$ can be decomposed as

$$\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k) = \sum_{j=M+1}^{N} \left(h_j^{o*}x(k-j+1) + g_j^{o*}x^*(k-j+1)\right) \tag{81}$$

Similarly, since $\mathbf{z}(k) = [\mathbf{x}^T(k), \mathbf{x}^H(k)]^T$ and $\tilde{\mathbf{w}}(k) = [\tilde{\mathbf{h}}^T(k), \tilde{\mathbf{g}}^T(k)]^T$, the term $\mathbf{z}^H(k)\tilde{\mathbf{w}}(k)$ can be expressed as

$$\mathbf{z}^H(k)\tilde{\mathbf{w}}(k) = \sum_{i=1}^{M} \left(\tilde{h}_i(k)x^*(k-i+1) + \tilde{g}_i(k)x(k-i+1)\right) \tag{82}$$

Using the independence assumptions, $\mathbf{T}(k)$ in (30) can be written as

$$\begin{aligned}\mathbf{T}(k) = & \sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\tilde{h}_i(k)]E\Big[\big(h_j^{o*}x^*(k-i+1)x(k-j+1)\\ & +g_j^{o*}x^*(k-i+1)x^*(k-j+1)\big)\mathbf{z}(k)\mathbf{z}^H(k)\Big]\\ & +\sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\tilde{g}_i(k)]E\Big[\big(h_j^{o*}x(k-i+1)x(k-j+1)\\ & +g_j^{o*}x(k-i+1)x^*(k-j+1)\big)\mathbf{z}(k)\mathbf{z}^H(k)\Big]\end{aligned} \tag{83}$$

Now, using the fourth order moment factoring theorem for zero mean Gaussian input data, $\mathbf{T}(k)$ can be expressed as

$$\mathbf{T}(k) = \begin{bmatrix} \mathbf{E}(k) & \mathbf{F}(k) \\ \mathbf{G}(k) & \mathbf{H}(k) \end{bmatrix} \tag{84}$$

in which, the (m, l)th elements of the $M\times M$ sub-matrices $\mathbf{E}(k)$, $\mathbf{F}(k)$, $\mathbf{G}(k)$ and $\mathbf{H}(k)$ are respectively given by

$$\begin{aligned} E_{m,l}(k) &= \sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{h}_i(k)]\Big(h_j^{o*}\big(r(i-j)r(l-m) \\ &\quad +r(i-m)r(l-j)+p^*(i-l)p(j-m)\big)+g_j^{o*}\big(p^*(j-i)r(l-m) \\ &\quad +r(i-m)p^*(l-j)+p^*(i-l)r(j-m)\big)\Big) \\ &\quad +\sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{g}_i(k)]\Big(h_j^{o*}\big(p(i-j)r(l-m)+p(i-m)r(l-j) \\ &\quad +r(l-i)p(j-m)\big)+g_j^{o*}\big(r(j-i)r(l-m)+p(i-m)p^*(l-j) \\ &\quad +r(l-i)r(j-m)\big)\Big) \end{aligned} \tag{85}$$

$$\begin{aligned} F_{m,l}(k) &= \sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{h}_i(k)]\Big(h_j^{o*}\big(r(i-j)p(m-l) \\ &\quad +r(i-m)p(j-l)+r(i-l)p(j-m)\big) \\ &\quad +g_j^{o*}\big(p^*(i-j)p(m-l) \\ &\quad +r(i-m)r(j-l)+r(i-l)r(j-m)\big)\Big) \\ &\quad +\sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{g}_i(k)]\Big(h_j^{o*}\big(p(i-j)p(m-l) \\ &\quad +p(i-m)p(j-l)+p(i-l)p(j-m)\big) \\ &\quad +g_j^{o*}\big(r(j-i)p(m-l) \\ &\quad +p(i-m)r(j-l)+p(i-l)r(j-m)\big)\Big) \end{aligned} \tag{86}$$

$$\begin{aligned} G_{m,l}(k) &= \sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{h}_i(k)]\Big(h_j^{o*}\big(r(i-j)p^*(m-l) \\ &\quad +p^*(i-m)r(l-j)+p^*(i-l)r(m-j)\big) \\ &\quad +g_j^{o*}\big(p^*(i-j)p^*(m-l) \\ &\quad +p^*(i-m)p^*(j-l)+p^*(i-l)p^*(j-m)\big)\Big) \\ &\quad +\sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{g}_i(k)]\Big(h_j^{o*}\big(p(i-j)p^*(m-l) \\ &\quad +r(m-i)r(l-j)+r(l-i)r(m-j)\big) \\ &\quad +g_j^{o*}\big(r(j-i)p^*(m-l)+r(m-i)p^*(j-l) \\ &\quad +r(l-i)p^*(j-m)\big)\Big) \end{aligned} \tag{87}$$

$$\begin{aligned} H_{m,l}(k) &= \sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{h}_i(k)]\Big(h_j^{o*}\big(r(i-j)r(m-l) \\ &\quad +p^*(i-m)p(j-l)+r(i-l)r(m-j)\big) \\ &\quad +g_j^{o*}\big(p^*(i-j)r(m-l) \\ &\quad +p^*(i-m)r(j-l)+r(i-l)p^*(j-m)\big)\Big) \\ &\quad +\sum_{i=1}^{M}\sum_{j=M+1}^{N} E[\widetilde{g}_i(k)]\Big(h_j^{o*}\big(p(j-i)r(m-l) \\ &\quad +r(m-i)p(j-l)+p(i-l)r(m-j)\big) \\ &\quad +g_j^{o*}\big(r(j-i)r(m-l)+r(m-i)r(j-l) \\ &\quad +p(i-l)p^*(j-m)\big)\Big) \end{aligned} \tag{88}$$

Appendix B. The Detailed Form of D in (31)

Observe first that on the RHS of (31), the term $|\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)|^2$ can be decomposed as

$$\begin{aligned} |\bar{\mathbf{w}}^{oH}\bar{\mathbf{z}}(k)|^2 = \sum_{i=M+1}^{N}\sum_{j=M+1}^{N} &\big(h_i^{o}x^*(k-i+1)+g_i^{o}x(k-i+1)\big) \\ &\cdot\big(h_j^{o*}x^*(k-j+1)+g_j^{o*}x(k-j+1)\big) \end{aligned} \tag{89}$$

Again, by using the Gaussian fourth order moment factoring theorem, $\mathbf{D}$ in (31) can be expressed as

$$\mathbf{D} = \begin{bmatrix} \mathbf{O} & \mathbf{P} \\ \mathbf{U} & \mathbf{V} \end{bmatrix} \tag{90}$$

in which, the (m, l)th elements of the $M\times M$ sub-matrices $\mathbf{O}$, $\mathbf{P}$, $\mathbf{U}$ and $\mathbf{V}$ are respectively given by

$$\begin{aligned} O_{m,l} &= \sum_{i=M+1}^{N}\sum_{j=M+1}^{N} \Big(h_i^{o}h_j^{o*}\big(r(i-j)r(l-m) \\ &\quad +r(i-m)r(l-j)+p^*(i-l)p(j-m)\big) \\ &\quad +h_i^{o}g_j^{o*}\big(p^*(i-j)r(l-m) \\ &\quad +r(i-m)p^*(j-l)+p^*(i-l)r(j-m)\big) \\ &\quad +g_i^{o}h_j^{o*}\big(p(i-j)r(l-m) \\ &\quad +p(i-m)r(l-j)+r(l-i)p(j-m)\big) \\ &\quad +g_i^{o}g_j^{o*}\big(r(j-i)r(l-m) \\ &\quad +p(i-m)p^*(j-l)+r(l-i)r(j-m)\big)\Big) \end{aligned} \tag{91}$$

$$\begin{aligned} P_{m,l} &= \sum_{i=M+1}^{N}\sum_{j=M+1}^{N} \Big(h_i^{o}h_j^{o*}\big(r(i-j)p(m-l) \\ &\quad +r(i-m)p(j-l)+r(i-l)p(j-m)\big) \\ &\quad +h_i^{o}g_j^{o*}\big(p^*(i-j)p(m-l) \\ &\quad +r(i-m)r(j-l)+r(i-l)r(j-m)\big) \\ &\quad +g_i^{o}h_j^{o*}\big(p(i-j)p(m-l) \\ &\quad +p(i-m)p(j-l)+p(i-l)p(j-m)\big) \\ &\quad +g_i^{o}g_j^{o*}\big(r(j-i)p(m-l) \\ &\quad +p(i-m)r(j-l)+p(i-l)r(j-m)\big)\Big) \end{aligned} \tag{92}$$

$$\begin{aligned} U_{m,l} &= \sum_{i=M+1}^{N}\sum_{j=M+1}^{N} \Big(h_i^{o}h_j^{o*}\big(r(i-j)p^*(m-l) \\ &\quad +p^*(i-m)r(l-j)+p^*(i-l)r(m-j)\big) \end{aligned}$$

$$
\begin{aligned}
&+h_i^o g_j^{o*}\Big(p^*(i-j)p^*(m-l)\\
&+p^*(i-m)p^*(j-l)+p^*(i-l)p^*(j-m)\Big)\\
&+g_i^o h_j^{o*}\Big(p(i-j)p^*(m-l)\\
&+r(m-i)r(l-j)+r(l-i)r(m-j)\Big)\\
&+g_i^o g_j^{o*}\Big(r(j-i)p^*(m-l)\\
&+r(m-i)p^*(j-l)+r(l-i)p^*(j-m)\Big)\Big)
\end{aligned}
\tag{93}
$$

$$
\begin{aligned}
&V_{m,l}\\
&=\sum_{i=M+1}^{N}\sum_{j=M+1}^{N}\Big(h_i^o h_j^{o*}\Big(r(i-j)r(m-l)\\
&+p^*(i-m)p(j-l)+r(i-l)r(m-j)\Big)\\
&+h_i^o g_j^{o*}\Big(p^*(i-j)r(m-l)\\
&+p^*(i-m)r(j-l)+r(i-l)p^*(j-m)\Big)\\
&+g_i^o h_j^{o*}\Big(p(i-j)r(m-l)\\
&+r(m-i)p(j-l)+p(i-l)r(m-j)\Big)\\
&+g_i^o g_j^{o*}\Big(r(j-i)r(m-l)\\
&+r(m-i)r(j-l)+p(i-l)p^*(j-m)\Big)\Big)
\end{aligned}
\tag{94}
$$

Appendix C. The Derivation of $J_{\mathrm{ex}}(\infty)$ in (72)

Following the approach in [26], substituting (68) into (71) yields

$$
\boldsymbol{\lambda}^T\mathbf{v}(\infty)=\frac{\displaystyle\sum_{i=1}^{2M}\frac{\frac{2|\hat{b}_i|^2}{\lambda_i}+\mu\hat{D}_{i,i}-2\mu\Re[\hat{T}_{i,i}(\infty)]+\mu\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j}\lambda_i-\mu u\lambda_i}{2-2\mu\lambda_i}+\sum_{i=1}^{2M}\frac{\mu\lambda_i J_{\min}}{2-2\mu\lambda_i}}{\displaystyle 1-\sum_{i=1}^{2M}\frac{\mu\lambda_i}{2-2\mu\lambda_i}}
\tag{95}
$$

Since

$$
\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j}=\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2(2-2\mu\lambda_j)}{\lambda_j(2-2\mu\lambda_j)}
\tag{96}
$$

We have

$$
\begin{aligned}
&J_{\mathrm{ex}}(\infty)=\boldsymbol{\lambda}^T\mathbf{v}(\infty)-\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j}\\
&=\frac{\displaystyle\sum_{i=1}^{2M}\frac{\frac{2|\hat{b}_i|^2}{\lambda_i}+\mu\hat{D}_{i,i}-2\mu\Re[\hat{T}_{i,i}(\infty)]+2\mu\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2}{\lambda_j}\lambda_i-\mu u\lambda_i}{2-2\mu\lambda_i}-\sum_{j=1}^{2M}\frac{|\hat{b}_j|^2(2-2\mu\lambda_j)}{\lambda_j(2-2\mu\lambda_j)}}{\displaystyle 1-\sum_{i=1}^{2M}\frac{\mu\lambda_i}{2-2\mu\lambda_i}}\\
&+\frac{\displaystyle\sum_{i=1}^{2M}\frac{\mu\lambda_i J_{\min}}{2-2\mu\lambda_i}}{\displaystyle 1-\sum_{i=1}^{2M}\frac{\mu\lambda_i}{2-2\mu\lambda_i}}\\
&=\frac{\displaystyle\sum_{i=1}^{2M}\frac{\mu f_i}{2-2\mu\lambda_i}+\sum_{i=1}^{2M}\frac{\mu\lambda_i J_{\min}}{2-2\mu\lambda_i}}{\displaystyle 1-\sum_{i=1}^{2M}\frac{\mu\lambda_i}{2-2\mu\lambda_i}}
\end{aligned}
\tag{97}
$$

where f_i is defined in (73).

References

[1] B. Widrow, J. McCool, M. Ball, The complex LMS algorithm, Proc. IEEE 63 (4) (1975) 719–720.
[2] F.D. Neeser, J.L. Massey, Proper complex random processes with applications to information theory, IEEE Trans. Inf. Theory 39 (4) (1993) 1293–1302.
[3] B. Picinbono, Second-order complex random vectors and normal distributions, IEEE Trans. Signal Process. 44 (10) (1996) 2637–2640.
[4] B. Picinbono, P. Chevalier, Widely linear estimation with complex data, IEEE Trans. Signal Process. 43 (8) (1995) 2030–2033.
[5] P.J. Schreier, L.L. Scharf, Second-order analysis of improper complex random vectors and processes, IEEE Trans. Signal Process. 51 (3) (2003) 714–725.
[6] D.P. Mandic, S.L. Goh, Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models, Wiley, New York, 2009.
[7] P.J. Schreier, L.L. Scharf, Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals, Cambridge University Press, 2010.
[8] S.C. Douglas, Widely-linear recursive least-squares algorithm for adaptive beamforming, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010, pp. 2041–2044.
[9] R.M. Fernandez-Alcala, J. Navarro-Moreno, J.C. Ruiz-Molina, J.A. Espinosa-Pulido, Linear and nonlinear smoothing algorithms for widely factorizable signals, Signal Process. 93 (4) (2013) 897–903.
[10] J. Navarro-Moreno, R.M. Fernandez-Alcala, J.C. Ruiz-Molina, J.M. Quesada-Rubio, Linear and nonlinear filters based on the improper karhunen-loève expansion, Signal Process. 92 (10) (2012) 2422–2431.
[11] S. Javidi, M. Pedzisz, S.L. Goh, D.P. Mandic, The augmented complex least mean square algorithm with application to adaptive prediction problems, in: 1st IARP Workshop on Cognitive Information Processing, 2008, pp. 54–57.
[12] D.P. Mandic, S. Javidi, S.L. Goh, A. Kuh, K. Aihara, Complex-valued prediction of wind profile using augmented complex statistics, Renewable Energy 34 (1) (2009) 44–54.
[13] S. Javidi, D.P. Mandic, A. Cichocki, Complex blind source extraction from noisy mixtures using second-order statistics, IEEE Trans. Circuits Systems I, Regular Papers 57 (7) (2010) 1404–1416.
[14] Y. Xia, D.P. Mandic, A full mean square analysis of CLMS for second order noncircular inputs, IEEE Trans. Signal Process. 65 (21) (2017) 5578–5590.
[15] R. Schober, W.H. Gerstacher, L.H.-J. Lampe, Data-aided and blind stochastic gradient algorithm for widely linear MMSE MAI suppersion for DS-CDMA, IEEE Trans. Signal Process. 52 (3) (2004) 746–756.
[16] Y. Xia, S.C. Douglas, D.P. Mandic, Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators, IEEE Signal Processing Magazine 29 (5) (2012) 44–54.
[17] Y. Xia, D.P. Mandic, Widely linear adaptive frequency estimation of unbalanced three-phase power systems, IEEE Trans. Instrum. Measurement 61 (1) (2012) 74–83.
[18] S.C. Douglas, D.P. Mandic, Performance analysis of the conventional complex LMS and augmented complex LMS algorithms, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010, pp. 3494–3497.
[19] D.P. Mandic, Y. Xia, S.C. Douglas, Steady state analysis of the CLMS and augmented CLMS algorithms for noncircular complex signals, in: Proceedings of 44th Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2010, pp. 1635–1639.
[20] Y. Xia, C.C. Took, D.P. Mandic, An augmented affine projection algorithm for the filtering of noncircular complex signals, Signal Processing 90 (6) (2010) 1788–1799.
[21] D.P. Mandic, S. Kanna, S.C. Douglas, Mean square analysis of the CLMS and ACLMS for non-circular signals: The approximate uncorrelating transform approach, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 3531–3535.
[22] Y.-M. Shi, L. Huang, C. Qian, H.C. So, Shrinkage linear and widely linear complex-valued least mean squares algorithms for adaptive beamforming, IEEE Trans. Signal Process. 63 (1) (2015) 119–131.
[23] A. Khalili, A. Rastegarnia, S. Sanei, Quantized augmented complex least-mean square algorithm: Derivation and performance analysis, Signal Process. 121 (2016) 54–59.
[24] Y. Xia, D.P. Mandic, Complementary mean square analysis of augmented CLMS for second order noncircular gaussian signals, IEEE Signal Process. Lett. 24 (9) (2017) 1413–1417.
[25] K. Mayyas, Performance analysis of the deficient length LMS adaptive algorithm, IEEE Trans. Signal Process. 53 (8) (2005) 2727–2740.
[26] K. Mayyas, Analysis of the transform domain LMS algorithm with insufficient length adaptive filter, J. Circuits, Syst. Comput. 14 (3) (2005) 469–481.
[27] M. Wu, J. Yang, Y. Xu, X. Qiu, Steady-state solution of the deficient length constrained FBLMS algorithm, IEEE Trans. Signal Process. 60 (12) (2012) 6681–6687.
[28] S. Wei, D.L. Goechel, P.A. Kelly, Convergence of the complex envelope of bandlimited OFDM signals, IEEE Transactions on Information Theory 56 (10) (2010) 4893–4904.
[29] C. Cheong-Took, S.C. Douglas, D.P. Mandic, On approximate diagonalization of correlation matrices in widely linear signal procesing, IEEE Transactions on Signal Processing 60 (3) (2012) 1469–1473.
[30] J. Navarro-Moreno, J. Moreno-Kayser, R. Fernandez-Alcala, J. Ruiz-Molina, Widely linear estimation algorithms for second-order stationary signals, IEEE Trans. Signal Process. 57 (12) (2009) 4930–4935.

[31] P. Chevalier, A. Blin, Widely linear MVDR beamformers for the reception of an unknown signal corrupted by noncircular interferences, IEEE Trans. Signal Process. 55 (11) (2007) 5323–5336.
[32] S.C. Douglas, D.P. Mandic, Mean and mean-square analysis of the complex LMS algorithm for non-circular gaussian signals, in: Proceedings of Digital Signal Processing Workshop and 5th IEEE Signal Processing Educducation Workshop (DSP/SPE), 2009, pp. 101–106.
[33] M. Valkama, M. Renfors, V. Koivunen, Blind signal estimation in conjugate signal models with application to i/q imbalance compensation, IEEE Signal Process. Lett. 12 (11) (2005) 733–736.
[34] L. Anttila, M. Valkama, M. Renfors, Circularity-based i/q imbalance compensation in wideband direct-conversion receivers, IEEE Trans. Vehicular Technol. 57 (4) (2008) 2099–2113.
[35] Z. Li, Y. Xia, W. Pei, K. Wang, Y. Huang, D.P. Mandic, Noncircular measurement and mitigation of i/q imbalance for OFDM-based WLAN transmitters, IEEE Trans. Instrum. Meas. 66 (3) (2017) 383–393.
[36] E. Ewada, Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels, IEEE Trans. Signal Process. 42 (11) (1994) 2937–2944.
[37] J.C.M. Bermudez, N.J. Bershad, Transient and tracking performance analysis of the quantized LMS algorithm for time-varying system identification, IEEE Trans. Signal Process. 44 (8) (1996) 1990–1997.
[38] N.R. Yousef, A.H. Sayed, A unified approach to the steady-state and tracking analyses of adaptive filters, IEEE Trans. Signal Process. 49 (2) (2001) 314–324.
[39] A.H. Sayed, Fundamentals of Adaptive Filtering, Wiley, New York, 2003.
[40] T.Y. Al-Naffouri, A.H. Sayed, Transient analysis of data-normalized adaptive filters, IEEE Trans. Signal Process. 51 (3) (2003) 639–652.
[41] R.A. Horn, Topics in Matrix Analysis, Cambridge University Press, 1991.
[42] A. Feuer, E. Weinstein, Convergence analysis of LMS filters with uncorrelated gaussian data, IEEE Trans. Acoustics, Speech Signal Process. 33 (1) (1985) 222–230.
[43] B. Widrow, S.D. Stearns, Adaptive Signal Processing, prentice-hall, new jersey, 1985.
[44] E. Ollila, On the circularity of a complex random variable, IEEE Signal Process. Lett. 15 (2008) 841–844.
[45] E. Ollila, J. Eriksson, V. Koivunen, Complex elliptically symmetric random variables - generation, characterization, and circularity tests, IEEE Trans. Signal Process. 59 (1) (2011) 58–69.
[46] S.C. Douglas, Fixed-point algorithms for the blind separation of arbitrary complex-valued non-Gaussian signal mixtures, EURASIP J. Adv. Signal Process. 2007 (2007) Article ID 036525, 15 Pages.