Signal Processing 104 (2014) 33–42

Contents lists available at ScienceDirect

Signal Processing

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

Complex dual channel estimation: Cost effective widely linear adaptive filtering

Cyrus Jahanchahi, Sithan Kanna*, Danilo Mandic

Department of Electrical and Electronic Engineering, Imperial College London, UK

ARTICLE INFO

Article history:
Received 19 November 2013
Received in revised form
29 March 2014
Accepted 31 March 2014
Available online 8 April 2014

Keywords:
CLMS
ACLMS
Widely linear estimation
Complex noncircularity
Augmented statistics

ABSTRACT

Widely linear estimation for complex-valued data allows for a unified treatment of both second order circular (proper) and non-circular (improper) signals. We propose the complex dual channel (CDC) estimation technique as an alternative to widely linear estimation to both gain further insights into complex valued minimum mean square error (MMSE) estimation and to design computationally efficient adaptive filtering algorithms. This is achieved by finding two sets of optimal weights that minimize the mean square error (MSE) in estimating the real and imaginary parts of the signal independently. The concept is used in a stochastic gradient setting to design the dual channel complex least mean square (DC-CLMS). The analysis shows that any one of the sub-filters within the DC-CLMS can be used to estimate strictly linear models while the DC-CLMS is equivalent to widely linear estimation. This results in a reduction of computational complexity of complex-valued adaptive filters by a half, while providing enhanced physical insight and control over complex-valued estimation algorithms.

Notation: Lowercase letters are used to denote scalars, boldface letters for vectors and boldface uppercase letters for matrices. The symbol $(\cdot)^*$ denotes complex conjugation, $(\cdot)^T$ and $(\cdot)^H$ – transposition and conjugate transposition, $(\cdot)^{-1}$ – matrix inversions and $\mathrm{Tr}[\cdot]$ is the trace of a matrix. The operators $\Re[\cdot]$ and $\Im[\cdot]$ are used to extract respectively the real and imaginary parts of a complex variable and $j=\sqrt{-1}$. The subscript k is used as a time index and $E[\cdot]$ represents the statistical expectation operator.

1. Introduction

Complex-valued linear minimum mean square error (MMSE) estimation is an important statistical technique in communications and signal processing. It has now been accepted that the standard strictly linear model for complex data is not guaranteed to capture the complete second-order statistical relationship between the input (regressor) and the output (observations) as generic strictly linear extensions of real-valued estimators cater only for data with rotation invariant probability distributions [1–3]. To account for general complex data, we have to employ the widely linear framework [4,5], whereby by using an augmented input vector, $\mathbf{x}^a=[\mathbf{x},\ \mathbf{x}^*]^T$, that contains the input vector and its conjugate, the widely linear model is equipped with the sufficient degrees of freedom to fully exploit the second order statistics in the data [6].

Consequently, widely linear estimators are able to achieve a lower mean square error (MSE) for estimating second order circular and non-circular signals. This has inspired the development of widely linear adaptive filtering algorithms such as the augmented complex least mean square (ACLMS) [7], widely linear recursive least squares (WL-RLS) [8] and augmented affine projection algorithm (AAPA) [9].

However, owing to the use of augmented variables, widely linear algorithms require twice the number of

* Corresponding author.
E-mail address: ssk08@ic.ac.uk (S. Kanna).

http://dx.doi.org/10.1016/j.sigpro.2014.03.050
0165-1684/

coefficients to update compared to their standard strictly linear counterparts. This adds to both higher computational complexity and a larger excess mean square error (EMSE). To deal with these issues, efficient formulations of widely linear adaptive filters have been proposed – these formulations typically make use of the duality between the complex and real domain to cast the computations into the real domain [10,11]. An efficient implementation of the ACLMS in [12] is one such approach. However, in this way, the input vector no longer resides in the original complex domain and any physical meaning or performance advantage inherent in working in the complex domain might disappear.

To help circumvent these problems and at the same time provide greater insight into complex-valued MMSE estimation, we propose an estimation technique that (i) has the two degrees of freedom necessary to capture the complete second order statistics and (ii) provides physically meaningful estimates at a low computational complexity. The analysis shows that the operation of widely linear models is over-parameterized and that the proposed combination of two complex strictly linear models is sufficient for second order optimality. We refer to the proposed method as the complex dual channel (CDC) estimator and show that it can not only be used as an alternative to the strictly and widely linear frameworks for processing proper and improper data but also that it provides additional insights into the physical parameters of such data. Finally, stochastic gradient adaptive filters based on the CDC framework, referred to as the CLMSr, CLMSi (for proper signals) and the dual channel-CLMS (DC-CLMS) (for improper signals), are introduced and their convergence and stability properties are analysed. The CDC framework is shown to give physically meaningful parameters which is also confirmed through simulations.

2. Complex dual channel estimation

2.1. Problem with strictly linear estimation in $\mathbb{C}$

Minimum mean square error (MMSE) estimation aims to find the optimal second order estimate of a desired signal $d_k \in \mathbb{C}$ given the regressor vector $\mathbf{x}_k \in \mathbb{C}^{N\times 1}$. For strictly linear MMSE estimation, the data model is constrained to be strictly linear, $y_k = \mathbf{h}^H \mathbf{x}_k$, where y_k is the estimate of the desired signal, $\mathbf{h} \in \mathbb{C}^{N\times 1}$ the coefficient vector, and the estimation error is given by $e_k = d_k - y_k$. The optimal second order mean square error (MSE) fit for the data is obtained by using the cost function

$$J_{\text{MSE}} = E[|e_k|^2] = E[|d_k - y_k|^2] \tag{1}$$

and the optimal weight vector, denoted by $\mathbf{h}_o$, that minimizes the cost function in (1) is given by the Wiener solution

$$\mathbf{h}_o = \mathbf{R}^{-1}\mathbf{r} \tag{2}$$

where $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H]$ is the covariance matrix of the input and $\mathbf{r} = E[\mathbf{x}d^*]$ is the cross-covariance vector of the desired signal and the input.

Of particular importance to this work is the (often overlooked) observation that the MSE for estimating complex-valued signals combines two separate components: (i) the proportion of MSE in estimating the real part of the signal, and (ii) the proportion of MSE corresponding to estimating the imaginary part of the signal. This becomes immediately clear when the error, e_k, is rewritten in terms of its real and imaginary components as $e_k = e_{r,k} + je_{i,k}$, so that the cost function in (1) can be written as the sum of two real-valued cost functions as

$$J_{\text{MSE}} = E[e_{r,k}^2] + E[e_{i,k}^2] = J_r + J_i \tag{3}$$

Remark #1. The linear MMSE estimator aims to minimize two real-valued cost functions, J_r and J_i, using a single complex-valued weight vector $\mathbf{h}_o$. However, one weight vector in general does not have enough degrees of freedom to minimize both J_r and J_i, and the class of signals that admit such an MMSE estimator is very restrictive. The analysis in Section 2.3 shows that the standard strictly linear estimation problem is well posed only when the input vector $\mathbf{x}_k$ is jointly circular with the desired signal d_k.

To cater for a general class of complex-valued signals, a widely linear model was proposed in [4], and is given by

$$y_k = (\mathbf{w}_k^a)^H \mathbf{x}_k^a \quad \mathbf{w}_k^a = \begin{bmatrix} \mathbf{h}_k \\ \mathbf{g}_k \end{bmatrix} \quad \mathbf{x}_k^a = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{x}_k^* \end{bmatrix} \tag{4}$$

where the estimator uses an augmented input vector, $\mathbf{x}_k^a$, formed by concatenating the input, $\mathbf{x}_k$, and its conjugate, $\mathbf{x}_k^*$, resulting in the augmented weight vector, $\mathbf{w}_k^a$, that has two degrees of freedom. Based upon the augmented vector $\mathbf{x}_k^a$, the sufficient second order statistics is obtained through the augmented covariance matrix

$$\mathbf{R}^a = E\begin{bmatrix} \mathbf{x}\mathbf{x}^H & \mathbf{x}\mathbf{x}^T \\ \mathbf{x}^*\mathbf{x}^H & \mathbf{x}^*\mathbf{x}^T \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{P} \\ \mathbf{P}^* & \mathbf{R}^* \end{bmatrix} \tag{5}$$

and the augmented cross-covariance vector

$$\mathbf{r}^a = E\begin{bmatrix} \mathbf{x}d^* \\ \mathbf{x}^*d^* \end{bmatrix} = \begin{bmatrix} \mathbf{r} \\ \mathbf{u}^* \end{bmatrix}. \tag{6}$$

Observe that the standard covariance matrix $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H]$, and cross-covariance vector $\mathbf{r} = E[\mathbf{x}d^*]$ do not have the sufficient degrees of freedom to explain the complete second order information, and to capture the second order statistics we also need the pseudocovariance matrix, $\mathbf{P} = E[\mathbf{x}\mathbf{x}^T]$, and the pseudocross-covariance vector, $\mathbf{u} = E[\mathbf{x}d]$. The input and desired signal pairs which exhibit a vanishing pseudocovariance, $\mathbf{P} = \mathbf{0}$ and pseudocross-covariance, $\mathbf{u} = \mathbf{0}$, are referred to as jointly circular (in terms of second order statistics) [4,13]. In this work, we show that the advantage offered by widely linear estimators can also be obtained by using two strictly linear estimators that minimize J_r and J_i independently, thus reducing the computational cost of widely linear algorithms.

2.2. Proposed solution: complex dual channel estimation

Based on the MSE cost function in (3), we propose to use two strictly linear estimators that minimise the cost functions $J_r = E[e_{r,k}^2]$ and $J_i = E[e_{i,k}^2]$ independently. This makes it possible to have individual Wiener solutions, with the corresponding optimal weights $\mathbf{w}_{cr}^o$ and $\mathbf{w}_{ci}^o$ that minimize the cost functions J_r and J_i independently, and therefore J_{MSE} in (3).

Remark #2. By minimising J_r and J_i independently, the estimator is no longer required to compromise between minimising J_r and J_i. Such an estimator has the desired two degrees of freedom, and performs optimal estimation for both second order circular and non-circular complex-valued signals.

The proposed CDC estimator is formed by combining the real part of $y_{cr,k}=\mathbf{w}_{cr}^H\mathbf{x}_k$ and imaginary part of $y_{ci,k}=\mathbf{w}_{ci}^H\mathbf{x}_k$. Notice that since $\mathbf{w}_{cr}^o$ minimizes the error in the real part of an estimator, it is natural to use only the real part of $y_{cr,k}$. The same reasoning also applies for the imaginary part of $y_{ci,k}$.

The optimal weights $\mathbf{w}_{cr}^o$ and $\mathbf{w}_{ci}^o$ are next derived using a standard gradient methodology. The mean square error in estimating the real part of the signal, J_r, can be written as

$$J_r=E[e_{r,k}^2]=E[(\Re\{d_k\}-\Re\{\mathbf{w}_{cr}^H\mathbf{x}_k\})^2] \tag{7}$$

Upon finding the minimum of (7) with respect to $\mathbf{w}_{cr}$, a closed form solution for the optimal weights is obtained in the form

$$\mathbf{w}_{cr}^o=2[\mathbf{R}-\mathbf{P}\mathbf{R}^{*-1}\mathbf{P}^*]^{-1}[\mathbf{p}_{rx}-\mathbf{P}\mathbf{R}^{*-1}\mathbf{p}_{rx}^*] \tag{8}$$

where $\mathbf{P}=E[\mathbf{x}\mathbf{x}^T]$ is the pseudocovariance matrix of the input, $\mathbf{R}=E[\mathbf{x}\mathbf{x}^H]$ is the covariance matrix, $\mathbf{p}_{rx}=E[\mathbf{x}d_r]$ is the cross-covariance between the real part of the desired signal and the complex input vector $\mathbf{x}_k$ and $\mathbf{p}_{ix}=E[\mathbf{x}d_i]$ is the cross-covariance between the imaginary part of the desired signal and the input.

Similarly, the mean square error in estimating the imaginary part of the signal, J_i, is given by

$$J_i=E[e_{i,k}^2]=E[(\Im\{d_k\}-\Im\{\mathbf{w}_{ci}^H\mathbf{x}_k\})^2] \tag{9}$$

and the corresponding optimal estimator weights are

$$\mathbf{w}_{ci}^o=-2j[\mathbf{R}-\mathbf{P}\mathbf{R}^{*-1}\mathbf{P}^*]^{-1}[\mathbf{p}_{ix}-\mathbf{P}\mathbf{R}^{*-1}\mathbf{p}_{ix}^*] \tag{10}$$

Remark #3. Observe from (8) and (10) that, in general, the weights that minimize J_r in (7) are not those that minimize J_i in (9). This justifies the proposed individual minimisation of the contributing terms in the total MSE cost function, J_{MSE} in (3), and also plays a major role in the analysis of convergence of complex-valued adaptive filters.

2.3. Degrees of freedom in complex MMSE

To further illustrate the limitations of using a strictly linear model to minimize $J_{\text{MSE}}=E[|e_k|^2]$, we shall now set $\mathbf{w}_{cr}=\mathbf{w}_{ci}$ in order to find the condition for one set of weights to be able to minimize both J_r, J_i (and therefore J_{MSE}). Upon equating (8) and (10), we obtain

$$[\mathbf{p}_{rx}-\mathbf{P}\mathbf{R}^{*-1}\mathbf{p}_{rx}^*]=-j[\mathbf{p}_{ix}-\mathbf{P}\mathbf{R}^{*-1}\mathbf{p}_{ix}^*]$$
$$\mathbf{p}_{rx}+j\mathbf{p}_{ix}=\mathbf{P}\mathbf{R}^{*-1}(\mathbf{p}_{rx}^*+j\mathbf{p}_{ix}^*).$$

Recognising that $\mathbf{p}_{rx}^*+j\mathbf{p}_{ix}^*=E[\mathbf{x}^*d]=\mathbf{r}^*$ is the complex valued cross-covariance and $\mathbf{p}_{rx}+j\mathbf{p}_{ix}=E[\mathbf{x}d]=\mathbf{u}$ is the pseudocross-covariance, we arrive at the condition

$$\mathbf{u}=\mathbf{P}\mathbf{R}^{*-1}\mathbf{r}^*. \tag{11}$$

Remark #4. When estimating the real and imaginary parts of a signal with one set of complex weights, optimal performance is only possible when the desired signal d_k, and the regressor vector, $\mathbf{x}_k$ are jointly circular, that is, the pseudocovariance $\mathbf{P}=\mathbf{0}$, and the pseudocross-covariance, $\mathbf{u}=\mathbf{0}$. This is a well known result, but is addressed here from a different viewpoint.

2.4. Equivalence of complex dual channel and widely linear estimators

The equivalence of the proposed complex dual channel (CDC) model and the widely linear model can be established by factorising the output of the CDC estimator as

$$\hat{y}=\Re\{\mathbf{w}_{cr}^H\mathbf{x}\}+j\Im\{\mathbf{w}_{ci}^H\mathbf{x}\} \tag{12}$$

$$\hat{y}=\left(\frac{\mathbf{w}_{cr}+\mathbf{w}_{ci}}{2}\right)^H\mathbf{x}+\left(\frac{\mathbf{w}_{cr}-\mathbf{w}_{ci}}{2}\right)^T\mathbf{x}^*=\mathbf{h}^H\mathbf{x}+\mathbf{g}^H\mathbf{x}^* \tag{13}$$

where $\mathbf{h}\triangleq(\mathbf{w}_{cr}+\mathbf{w}_{ci})/2$ and $\mathbf{g}^*\triangleq(\mathbf{w}_{cr}-\mathbf{w}_{ci})/2$. This form is equivalent to the widely linear model in (4), for more detail see also [11].

3. The design of adaptive filters using the CDC framework

We now provide a new framework to derive complex-valued adaptive filtering algorithms using the CDC estimation model. Our focus is on the stochastic gradient descent type of algorithms, and on benchmarking the CDC framework against the complex least mean square (CLMS) and augmented CLMS (ACLMS), which both minimize the global mean square error cost function $J_{\text{MSE}}=E[|e_k|^2]$. The CLMS uses a strictly linear model of the data [14] and is given by

$$y_k=\mathbf{w}_k^H\mathbf{x}_k$$
$$e_k=d_k-y_k$$
$$\mathbf{w}_{k+1}=\mathbf{w}_k+\mu e_k^*\mathbf{x}_k \tag{14}$$

The ACLMS, proposed in [7], uses a widely linear model that employs both $\mathbf{x}_k$ and $\mathbf{x}_k^*$, and is given by

$$y_k=\mathbf{h}_k^H\mathbf{x}_k+\mathbf{g}_k^H\mathbf{x}_k^*$$
$$e_k=d_k-y_k$$
$$\mathbf{h}_{k+1}=\mathbf{h}_k+\mu e_k^*\mathbf{x}_k$$
$$\mathbf{g}_{k+1}=\mathbf{g}_k+\mu e_k^*\mathbf{x}_k^* \tag{15}$$

where $\mathbf{h}_k$ and $\mathbf{g}_k$ are complex-valued coefficient vectors.

3.1. Proposed algorithm: dual channel CLMS

The dual channel-CLMS (DC-CLMS) is formed by a combining two stochastic gradient descent algorithms: the CLMSr and CLMSi. The CLMSr uses a strictly linear model $y_{cr,k}=\mathbf{w}_{cr,k}^H\mathbf{x}_k$ and aims to find the minimum of the instantaneous cost function $J_{r,k}=(\Re[d_k-y_{cr,k}])^2$ that corresponds to the error in estimating the real part of the signal. The gradient is found using the conjugate derivative, $\nabla_{\mathbf{w}^*}$, of $J_{r,k}$ [15,16,10] to give

$$\mathbf{w}_{cr,k+1} = \mathbf{w}_{cr,k} - \mu_2 \nabla_{\mathbf{w}_{cr}^*} J_{r,k} \tag{16}$$

that results in the weight update

$$\mathbf{w}_{cr,k+1} = \mathbf{w}_{cr,k} + \mu_2 \Re[e_{cr,k}]\mathbf{x}_k \tag{17}$$

where $e_{cr,k} = d_k - y_{cr,k}$.

Similarly, the CLMSi uses a strictly linear model $y_{ci,k} = \mathbf{w}_{ci,k}^H \mathbf{x}_k$ and aims to find the minimum of the instantaneous cost function $J_{i,k} = (\Im[d_k - y_{ci,k}])^2$ that corresponds to the error in estimating the imaginary part of the signal. The CLMSi weight update is therefore given by

$$\mathbf{w}_{ci,k+1} = \mathbf{w}_{ci,k} - \mu_2 j \Im[e_{ci,k}]\mathbf{x}_k \tag{18}$$

where $e_{ci,k} = d_k - y_{ci,k}$.

Remark #5. Notice that both the weight vectors $\mathbf{w}_{cr}$ and $\mathbf{w}_{ci}$ are complex-valued, and do not represent the real or imaginary parts of the CLMS or ACLMS weights.

The DC-CLMS is summarized in Algorithm 1 and depicted in Fig. 1, and operates in a collaborative fashion by combining the real output of CLMSr with the imaginary output of CLMSi to obtain an output identical to that of the ACLMS.

Algorithm 1. The Dual Channel CLMS (DC-CLMS).

CLMSr

$y_{cr,k} = \mathbf{w}_{cr,k}^H \mathbf{x}_k$

$e_{cr,k} = d_k - y_{cr,k}$

$\mathbf{w}_{cr,k+1} = \mathbf{w}_{cr,k} + \mu_2 \Re[e_{cr,k}]\mathbf{x}_k$

CLMSi

$y_{ci,k} = \mathbf{w}_{ci,k}^H \mathbf{x}_k$

$e_{ci,k} = d_k - y_{ci,k}$

$\mathbf{w}_{ci,k+1} = \mathbf{w}_{ci,k} - j\mu_2 \Im[e_{ci,k}]\mathbf{x}_k$

Output

$y_k = \Re[y_{cr,k}] + j\Im[y_{ci,k}]$

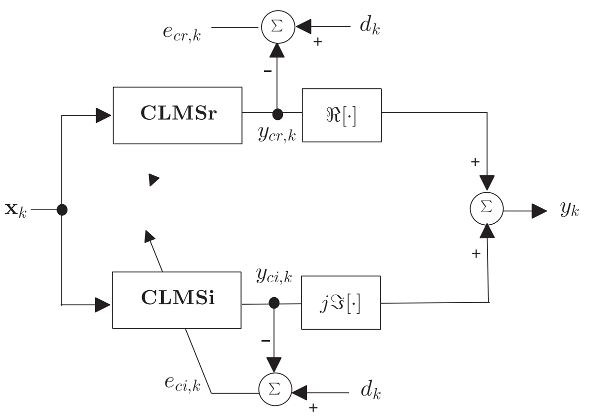

Fig. 1. Architecture of the DC-CLMS.

3.2. Equivalence of the CLMSr/i and CLMS

Following on Remark 5, we next establish the relationship between the optimal CLMSr and CLMSi weight estimates and the optimal CLMS weights, in order to show that either the CLMSr or CLMSi sub-filter can replace the traditional CLMS filter when estimating strictly linear models. This results in a computationally efficient implementation of CLMS. To show this, observe that in steady state, the CLMSr and CLMSi sub-filters satisfy

$$\text{CLMSr:} \ \lim_{k\to\infty} E[\mathbf{w}_{cr,k+1}] = E[\mathbf{w}_{cr,k}] = \mathbf{w}_{cr,\text{opt}}$$
$$\text{CLMSi:} \ \lim_{k\to\infty} E[\mathbf{w}_{ci,k+1}] = E[\mathbf{w}_{ci,k}] = \mathbf{w}_{ci,\text{opt}}. \tag{19}$$

The steady state condition for the CLMSr weight update is satisfied when

$$E[\Re[e_{cr,k}]\mathbf{x}_k] = \mathbf{0}. \tag{20}$$

From $\Re[e_{cr,k}] = \frac{1}{2}\left(e_{cr,k} + e_{cr,k}^*\right)$ and substituting for $e_{cr,k} = d_k - \mathbf{w}_{cr,k}^H \mathbf{x}_k$, we arrive at

$$E[\mathbf{x}_k d_k] - E[\mathbf{x}_k \mathbf{x}_k^T]E[\mathbf{w}_{cr,k}^*] = E[\mathbf{x}_k \mathbf{x}_k^H]E[\mathbf{w}_{cr,k}] - E[\mathbf{x}_k d_k^*]$$

which allows us to express the condition for an unbiased estimate for the CLMSr in the form

$$\mathbf{u} - \mathbf{P}E[\mathbf{w}_{cr,k}^*] = \mathbf{R}E[\mathbf{w}_{cr,k}] - \mathbf{r}. \tag{21}$$

Next, upon substituting $E[\mathbf{w}_{cr,k}] = \mathbf{R}^{-1}\mathbf{r}$ into (21), the condition for the mean weight vector of the CLMS in steady state to be also a valid CLMSr weight vector becomes

$$\mathbf{u} = \mathbf{P}\mathbf{R}^{*-1}\mathbf{r}^*. \tag{22}$$

Notice that this condition is identical to (11) and establishes the equivalence of the CLMSr and CLMSi weights in the steady state. The condition for steady state equivalence between the CLMSr, CLMSi and CLMS is identical to that introduced by Picinbono in [4] which addresses the suitability of strictly linear models to estimate general complex signals.

Remark #6. The above result shows that for jointly circular signals ($\mathbf{u} = 0$ and $\mathbf{P} = 0$), the traditional CLMS algorithm can be replaced with the more efficient CLMSr or CLMSi filters. In other words, when identifying strictly linear models, the CLMSr or CLMSi algorithms will achieve the same steady state mean square error while requiring only half the operations.

3.3. Equivalence of the ACLMS and DC-CLMS

By finding a recursive expression for $\Re[y_k]$ and $\Im[y_k]$ in ACLMS and comparing it with the recursive expression for the outputs $\Re[y_{cr,k}]$ and $\Im[y_{ci,k}]$ of the DC-CLMS, we now show that the ACLMS and DC-CLMS perform the same operation.

The output of the ACLMS is first written in terms of its real and imaginary parts where the operator $\mathcal{O}[\cdot]$ takes the form of either the real operator $\Re[\cdot]$ or imaginary operator $\Im[\cdot]$, so that

$$\begin{aligned}\mathcal{O}[y_k] &= \mathcal{O}[\mathbf{h}_k^H \mathbf{x}_k] + \mathcal{O}[\mathbf{g}_k^H \mathbf{x}_k^*] \\ &= \mathcal{O}[(\mathbf{h}_0^H + \mu_1 e_0 \mathbf{x}_0^H + \cdots + \mu_1 e_{k-1}\mathbf{x}_{k-1}^H)\mathbf{x}_k] \\ &\quad + \mathcal{O}[(\mathbf{g}_0^H + \mu_1 e_0 \mathbf{x}_0^T + \cdots + \mu_1 e_{k-1}\mathbf{x}_{k-1}^T)\mathbf{x}_k^*].\end{aligned} \tag{23}$$

Table 1
Computational requirements for the complex LMS algorithms considered, where N is the length of the complex-valued input vector $\mathbf{x}_k$.

Algorithm	Multiplications	Additions
Identification of strictly linear models		
CLMS	$8N+2$	$8N$
Proposed: CLMSr	$4N+1$	$4N$
Proposed: CLMSi	$4N+1$	$4N$
Identification of widely linear models		
ACLMS	$16N+2$	$16N$
DCRLMS [11]	$8N+4$	$8N$
RC-WL-LMS [12]	$8N+2$	$8N$
Proposed: DC-CLMS	$8N+2$	$8N$

Similarly, for the DC-CLMS, we have

$$\begin{aligned}\mathcal{O}[y_O] &= \mathcal{O}[\mathbf{w}_{O,k}^H\mathbf{x}_k]\\ &= \mathcal{O}[(\mathbf{w}_{O,0}^H+\mu_2\eta\mathcal{O}[e_0]\mathbf{x}_0^H+\cdots+\mu_2\eta\mathcal{O}[e_{k-1}]\mathbf{x}_{k-1}^H)\mathbf{x}_k]\end{aligned} \quad (24)$$

where $y_{O,k}=y_{cr,k}$, $\mathbf{w}_{O,k}=\mathbf{w}_{cr,k}$ for $\mathcal{O}[\cdot]=\Re[\cdot]$ and $y_{O,k}=y_{ci,k}$, $\mathbf{w}_{O,k}=\mathbf{w}_{ci,k}$ for $\mathcal{O}[\cdot]=\Im[\cdot]$, and $\eta=\{1,j\}$ for $\mathcal{O}[\cdot]=\{\Re[\cdot],\Im[\cdot]\}$

From (23) and (24), the output of ACLMS is equivalent to that of DC-CLMS, subject to $\mu_2=2\mu_1$ and $\mathbf{h}_0=\mathbf{g}_0=\mathbf{w}_{cr,0}=\mathbf{w}_{ci,0}$.

3.4. Comparison with the existing reduced complexity algorithms

For rigour, we now compare the CLMSr/i and DC-CLMS algorithms with the two existing reduced complexity augmented complex least mean square (ACLMS) algorithms:

(a) Dual-channel real-valued LMS (DCRLMS) algorithm, a real-valued algorithm that exploits the duality between $\mathbb{R}^2$ and $\mathbb{C}$ [11].
(b) Reduced complexity widely linear LMS (RC-WL-LMS), a complex-valued algorithm which employs a complex-valued weight vector and a real-valued input vector (constructed by augmenting the real and imaginary parts of the original complex-valued input vector) [12].

Although DCRLMS [11] has sufficient degrees of freedom to model complex-valued signals, its parameters reside in $\mathbb{R}$, and any physical insight that can be gained from the signal model is obscured. On the other hand, the RC-WL-LMS [12] does operate in the complex domain but has a limitation since its aim is only to reduce the computational complexity. In contrast to the DCRLMS and RC-WL-LMS, the framework presented in this paper allows for a unified and efficient formulation of both the strictly linear CLMS algorithm in the form of the CLMSr/i sub-filters and the ACLMS algorithm by combining the CLMSr/i sub-filters. Table 1 compares the number of operations required per iteration for all the complex LMS algorithms considered in this paper.

4. Transient performance of the DC-CLMS

The weight error vectors, $\tilde{\mathbf{w}}_{cr,k}\triangleq\mathbf{w}_o-\mathbf{w}_{cr,k}$ and $\tilde{\mathbf{w}}_{ci,k}\triangleq\mathbf{w}_o-\mathbf{w}_{ci,k}$ for the CLMSr and CLMSi respectively are given by

$$\text{CLMSr:}\ \ \tilde{\mathbf{w}}_{cr,k+1}=\tilde{\mathbf{w}}_{cr,k}-\mu_2\Re[e_{cr,k}]\mathbf{x}_k \quad (25)$$

$$\text{CLMSi:}\ \ \tilde{\mathbf{w}}_{ci,k+1}=\tilde{\mathbf{w}}_{ci,k}+j\mu_2\Im[e_{ci,k}]\mathbf{x}_k. \quad (26)$$

These recursions can be simplified by expanding the output errors in terms of the weight error. The output error of CLMSr can be expressed as $e_{cr,k}=\tilde{\mathbf{w}}_{cr,k}^H\mathbf{x}_k+\eta_k$, where η_k is complex-valued white Gaussian noise, $\eta\sim\mathcal{N}(0,\sigma_\eta^2)$, so that

$$\begin{aligned}&\text{CLMSr:}\ \ \tilde{\mathbf{w}}_{cr,k+1}=\tilde{\mathbf{w}}_{cr,k}-\mu_2\mathbf{x}_k\Re[\mathbf{x}_k^H\tilde{\mathbf{w}}_{cr,k}+\eta_k^*]\\ &\text{CLMSi:}\ \ \tilde{\mathbf{w}}_{ci,k+1}=\tilde{\mathbf{w}}_{ci,k}+j\mu_2\mathbf{x}_k\Im[\mathbf{x}_k^T\tilde{\mathbf{w}}_{ci,k}^*+\eta_k].\end{aligned} \quad (27)$$

Upon applying the statistical expectation operator and using the standard independence assumptions, we have

$$\begin{aligned}&\text{CLMSr:}\ \ E[\tilde{\mathbf{w}}_{cr,k+1}]=(\mathbf{I}-\mu_2\mathbf{R})E[\tilde{\mathbf{w}}_{cr,k}]-\mu_2\mathbf{P}E[\tilde{\mathbf{w}}_{cr,k}^*]\\ &\text{CLMSi:}\ \ E[\tilde{\mathbf{w}}_{ci,k+1}]=(\mathbf{I}-\mu_2\mathbf{R})E[\tilde{\mathbf{w}}_{ci,k}]+\mu_2\mathbf{P}E[\tilde{\mathbf{w}}_{ci,k}^*]\end{aligned} \quad (28)$$

where $\mathbf{R}=E[\mathbf{x}\mathbf{x}^H]$ is the covariance matrix and $\mathbf{P}=E[\mathbf{x}\mathbf{x}^T]$ the pseudocovariance matrix of the input data. Recall that the weight error recursion for the CLMS is given by [1]

$$\text{CLMS:}\ \ E[\tilde{\mathbf{w}}_{k+1}]=(\mathbf{I}-\mu\mathbf{R})E[\tilde{\mathbf{w}}_k]. \quad (29)$$

Remark #7. By comparing the weight error recursions for DC-CLMS in (28) and the CLMS from (29), observe that the evolution of the weight vectors will be identical if the data is proper (i.e. when $\mathbf{P}=\mathbf{0}$). This is supported by the simulation result in Fig. 2, which shows the average weight trajectories along the error surfaces of the CLMSr, CLMSi and CLMS.

An important property of the DC-CLMS algorithm, which will be instrumental in the transient analysis, is that when the data is improper, the CLMSr and CLMSi weights follow opposite paths, forming a mirror image (see (28) and Fig. 3(a)). Fig. 3(b) illustrates that although the cost functions of the CLMSr and CLMSi have the same minimum, the corresponding error surfaces are affected by the impropriety of the data.

4.1. Convergence in the mean of the DC-CLMS

For the DC-CLMS to be stable, both its sub-filters, CLMSr and CLMSi, must also be stable. To analyse the stability of CLMSr and CLMSi, we shall first split the recursions given in (28) into their real and imaginary parts. Assuming that the real and imaginary parts of the input data are uncorrelated, the covariance and pseudocovariance matrices are expressed in terms of the covariances of the real and imaginary parts of the input vector as

$$\begin{aligned}&\mathbf{R}=E[\mathbf{x}_k\mathbf{x}_k^H]=\mathbf{R}_{rr}+\mathbf{R}_{ii}\\ &\mathbf{P}=E[\mathbf{x}_k\mathbf{x}_k^T]=\mathbf{R}_{rr}-\mathbf{R}_{ii}.\end{aligned} \quad (30)$$

Upon expressing the weight error recursion for the CLMSr in terms of its real and imaginary parts, we have

$$E[\Re[\tilde{\mathbf{w}}_{cr,k+1}]]=[\mathbf{I}-\mu\mathbf{R}_{rr}]E[\Re[\tilde{\mathbf{w}}_{cr,k}]] \quad (31)$$

$$E[\Im[\tilde{\mathbf{w}}_{cr,k+1}]]=[\mathbf{I}-\mu\mathbf{R}_{ii}]E[\Im[\tilde{\mathbf{w}}_{cr,k}]]. \quad (32)$$

For the CLMSr to be stable in the mean, both the real and imaginary parts of its weights must converge. These recursions are stable if the maximum eigenvalues of the

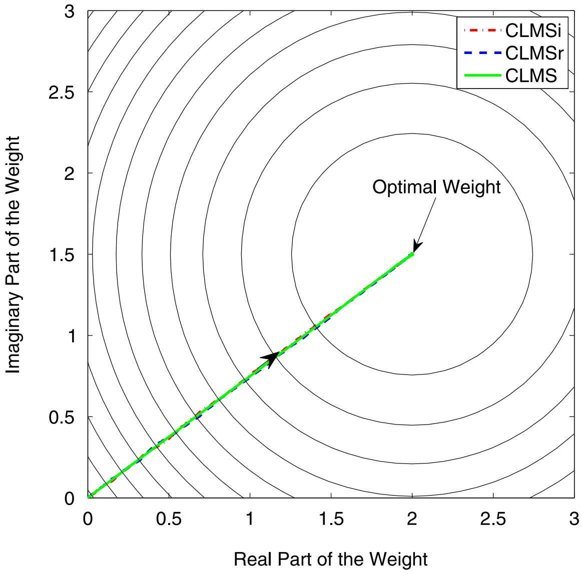

Fig. 2. Averaged weight trajectories along the error performance surface for the estimation of a strictly linear MA(1) system driven by circular white Gaussian noise.

matrices $[\mathbf{I}-\mu\mathbf{R}_{rr}]$ and $[\mathbf{I}-\mu\mathbf{R}_{rr}]$ are less than unity, which gives the condition on the learning rate

$$\text{CLMSr:}\ \ 0<\mu<\frac{2}{\max(\lambda_i,\lambda_r)} \tag{33}$$

where λ_i and λ_r are respectively the maximum eigenvalues for the covariance matrices $\mathbf{R}_{rr}$ and $\mathbf{R}_{ii}$.

Remark #8. Unlike the CLMS, for improper data where the statistics of $\Re[\mathbf{x}_k]$ are different from the statistics of $\Im[\mathbf{x}_k]$ $(\lambda_i\neq\lambda_r)$, the CLMSr has different convergence rates for the real and imaginary components of the weights. As a consequence, the learning curve of the CLMSr will exhibit two different regions converging at different rates.

Similarly, for the CLMSi sub-filter, the recursion

$$E[\Re[\mathbf{w}_{ci,k+1}]]=[\mathbf{I}-\mu\mathbf{R}_{ii}]E[\Re[\mathbf{w}_{ci,k}]] \tag{34}$$

$$E[\Im[\mathbf{w}_{ci,k+1}]]=[\mathbf{I}-\mu\mathbf{R}_{rr}]E[\Im[\mathbf{w}_{ci,k}]] \tag{35}$$

gives the same stability bound as in (33). Observe that the real component of the weight vector of CLMSr has the same convergence rate as the imaginary component of the weight vector of CLMSi.

It was previously stated that the ACLMS is equivalent to the DC-CLMS when the step size of the DC-CLMS is twice that of the ACLMS (see Section 3.3). Therefore for the ACLMS to converge, both the CLMSr and CLMSi must converge, and thus for the convergence of ACLMS in the mean, the step-size must be bounded by

$$0<\mu_{\text{ACLMS}}<\frac{1}{\max(\lambda_i,\lambda_r)} \tag{36}$$

Remark #9. By analysing the stability range for the ACLMS from the transient characteristics of the DC-CLMS, we no longer require the strong uncorrelating transform (SUT) [17,18]. The simplicity of the DC-CLMS approach also proves to be more insightful in evaluating the mean behaviour ACLMS.

5. Steady state performance of the DC-CLMS

Following the standard analysis, the steady state mean square error is given by

$$\text{MSE}=\lim_{k\to\infty}E[|e_k|^2]=\text{EMSE}+J_{\min}$$

where the EMSE (excess mean square error) results from the mismatch between the filter weights and the true system model while $J_{\min}=\sigma_\eta^2$ is the power of the measurement noise, η_k. Since the measurement noise is not influenced by the adaptive filter, we will only analyse the excess mean square error (EMSE) to quantify the performance of the DC-CLMS in steady state.

5.1. Excess mean square error of CLMSr

Since the CLMSr only lends the real part of its output to the DC-CLMS, we shall estimate the EMSE with this in mind. The EMSE is defined as the power of the *a priori* error, $e_{a,k}$, and its steady state value is [19]

$$\text{EMSE:}\ \ \lim_{k\to\infty}E[e_{a,k}^2]. \tag{37}$$

Since EMSE is caused by a mismatch between the filter weights $\mathbf{w}_{cr,k}$ and the optimal system coefficients, $\mathbf{w}_o$, this error is given by

$$e_{a,k}\triangleq\Re[\mathbf{x}_k^T(\mathbf{w}_o-\mathbf{w}_{cr,k})^*]=\Re[\mathbf{x}_k^T\tilde{\mathbf{w}}_{cr,k}^*]. \tag{38}$$

Upon conjugating the weight error recursion in (25), we have

$$\tilde{\mathbf{w}}_{cr,k+1}^*=\tilde{\mathbf{w}}_{cr,k}^*-\mu\Re[e_{cr,k}]\mathbf{x}_k^* \tag{39}$$

and pre-multiplying both sides by $\mathbf{x}_k^T$ gives $\mathbf{x}_k^T\tilde{\mathbf{w}}_{cr,k+1}^*=\mathbf{x}_k^T\tilde{\mathbf{w}}_{cr,k}^*-\mu\Re[e_{cr,k}]\mathbf{x}_k^T\mathbf{x}_k^*$. Next, we define $e_{p,k}\triangleq\Re[\mathbf{x}_k^T\tilde{\mathbf{w}}_{cr,k+1}^*]$ as the *a posteriori* error which gives us

$$e_{p,k}=e_{a,k}-\mu\Re[e_{cr,k}]\|\mathbf{x}_k\|^2 \tag{40}$$

where $e_{a,k}=\Re[\mathbf{x}_k^T\tilde{\mathbf{w}}_{cr,k}^*]$. Thus, the term $\Re[e_{cr,k}]$ becomes

$$\Re\left[e_{cr,k}\right]=\frac{e_{a,k}-e_{p,k}}{\mu\|\mathbf{x}_k\|^2} \tag{41}$$

and by substituting for $\Re[e_{cr,k}]$ in (39), we obtain

$$\tilde{\mathbf{w}}_{cr,k+1}+\frac{e_{a,k}}{\|\mathbf{x}_k\|^2}\mathbf{x}_k=\tilde{\mathbf{w}}_{cr,k}+\frac{e_{p,k}}{\|\mathbf{x}_k\|^2}\mathbf{x}_k \tag{42}$$

Squaring both sides and applying the statistical expectation operator gives

$$E\left[\|\tilde{\mathbf{w}}_{cr,k+1}\|^2\right]+E\left[\frac{e_{a,k}^2}{\|\mathbf{x}_k\|^2}\right]=E\left[\|\tilde{\mathbf{w}}_{cr,k}\|^2\right]+E\left[\frac{e_{p,k}^2}{\|\mathbf{x}_k\|^2}\right]. \tag{43}$$

At the steady state we can use the energy conservation principle [19], which states that $E[\|\tilde{\mathbf{w}}_{cr,k+1}\|^2]=E[\|\tilde{\mathbf{w}}_{cr,k}\|^2]$, to obtain

$$E\left[\frac{e_{a,k}^2}{\|\mathbf{x}_k\|^2}\right]=E\left[\frac{e_{p,k}^2}{\|\mathbf{x}_k\|^2}\right] \tag{44}$$

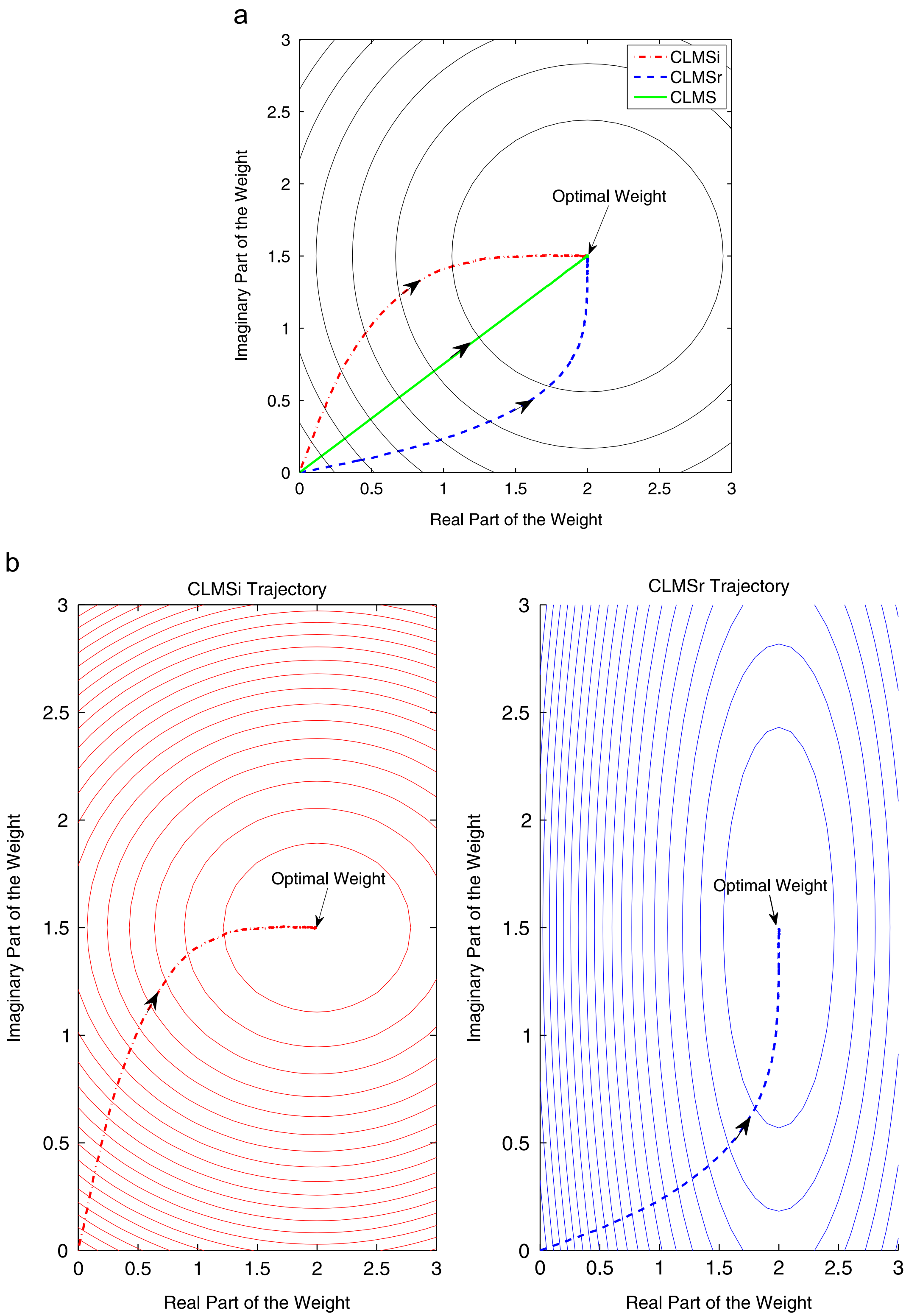

Fig. 3. Averaged weight trajectories for the estimation of a strictly linear MA(1) process driven by non-circular white Gaussian noise. (a) Trajectories for the CLMS, CLMSr and CLMSi for improper input data. (b) Individual error surfaces for the CLMSr and CLMSi for improper input data.

Inserting the relationship between the *a posteriori* error and the *a priori* error, $e_{p,k}=e_{a,k}-\mu\Re[e_{cr,k}]\|\mathbf{x}_k\|^2$, from (40) gives

$$E\left[\frac{e_{a,k}^2}{\|\mathbf{x}_k\|^2}\right]=E\left[\frac{(e_{a,k}-\mu\Re[e_{cr,k}]\|\mathbf{x}_k\|^2)^2}{\|\mathbf{x}_k\|^2}\right]. \quad (45)$$

To find an expression that includes only $e_{a,k}$ as the error term, we expand the real part of the error from the CLMSr, $\Re[e_{cr,k}]=\Re[d_k-y_{cr,k}]$, where the desired signal is, $d_k=\mathbf{w}_o^H\mathbf{x}_k+\eta_k$. The error term $\Re[e_{cr,k}]$ and the *apriori* error, $e_{a,k}$, are then related as

$$\Re[e_{cr,k}]=\Re[(\mathbf{w}_o-\mathbf{w}_{cr,k})^H\mathbf{x}_k+\eta_k]=e_{a,k}+\eta_{r,k} \quad (46)$$

where $\eta_{r,k}$ is the real part of the system noise. Replacing (46) into (45) gives

$$E\left[\frac{e_{a,k}^2}{\|\mathbf{x}_k\|^2}\right]=E\left[\frac{(e_{a,k}-\mu e_{a,k}\|\mathbf{x}_k\|^2-\mu\eta_{r,k}\|\mathbf{x}_k\|^2)^2}{\|\mathbf{x}_k\|^2}\right]. \quad (47)$$

Assuming that the noise term η_k is statistically independent from the *a priori* error, $e_{a,k}$, expression (47) simplifies into

$$E\left[e_{a,k}^2\right]=\frac{\mu}{2}E\left[e_{a,k}^2\|\mathbf{x}_k\|^2\right]+\frac{\mu}{2}\,\mathrm{Tr}[\mathbf{R}]\sigma_{\eta_r}^2 \quad (48)$$

where $\sigma_{\eta_r}^2$ is the power of the real part of the system noise. For a small step size μ we can assume that the term $(\mu/2)\|\mathbf{x}_k\|^2E[e_{a,k}^2]$ is negligible compared to $(\mu/2)\,\mathrm{Tr}[\mathbf{R}]\sigma_{\eta_r}^2$ and hence

$$\mathrm{EMSE}_{\mathrm{CLMSr}}^{\mathrm{small}\ \mu}:\ \lim_{k\to\infty}E\left[e_{a,k}^2\right]=\frac{\mu}{2}\,\mathrm{Tr}[\mathbf{R}]\sigma_{\eta_r}^2. \quad (49)$$

For a large step size, the term $\mu/2\|\mathbf{x}_k\|^2E[e_{a,k}^2]$ is not negligible. Instead, we make a further assumption that at steady state $\|\mathbf{x}_k\|^2$ is statistically independent from $e_{a,k}^2$ [19], to obtain

$$\mathrm{EMSE}_{\mathrm{CLMSr}}^{\mathrm{large}\ \mu}:\ \lim_{k\to\infty}E\left[e_{a,k}^2\right]=\frac{\mu\,\mathrm{Tr}[\mathbf{R}]\sigma_{\eta_r}^2}{2-\mu\,\mathrm{Tr}[\mathbf{R}]}. \quad (50)$$

5.2. Excess mean square error of CLMSi

The analysis for the EMSE of the CLMSi is similar to that for the CLMSr. The difference is in the definitions for the *a priori* and *a posteriori* errors, which for the CLMSi are respectively $e_{a,k}=\Im[\mathbf{x}_k^T\tilde{\mathbf{w}}_{ci,k}^*]$ and $e_{p,k}=\Im[\mathbf{x}_k^T\tilde{\mathbf{w}}_{ci,k+1}^*]$. Then, the EMSE of the CLMSi takes the form

$$\mathrm{EMSE}_{\mathrm{CLMSi}}^{\mathrm{small}\ \mu}:\ \lim_{k\to\infty}E\left[e_{a,k}^2\right]=\frac{\mu}{2}\,\mathrm{Tr}[\mathbf{R}]\sigma_{\eta_i}^2$$

for small step sizes, and

$$\mathrm{EMSE}_{\mathrm{CLMSi}}^{\mathrm{large}\ \mu}:\ \lim_{k\to\infty}E\left[e_{a,k}^2\right]=\frac{\mu\,\mathrm{Tr}[\mathbf{R}]\sigma_{\eta_i}^2}{2-\mu\,\mathrm{Tr}[\mathbf{R}]}$$

for large step sizes, where $\sigma_{\eta_i}^2$ is the power of the imaginary part of the system noise.

5.3. Excess mean square error of the DC-CLMS

Since the DC-CLMS is formed from the real part of CLMSr and imaginary part of CLMSi, the EMSE of the DC-CLMS is $\mathrm{EMSE}_{\mathrm{DC-CLMS}}=\mathrm{EMSE}_{\mathrm{CLMSr}}+\mathrm{EMSE}_{\mathrm{CLMSi}}$ and $\sigma_\eta^2=\sigma_{\eta_r}^2+\sigma_{\eta_i}^2$, and consequently

$$\mathrm{EMSE}_{\mathrm{DC-CLMS}}^{\mathrm{small}\ \mu}=\frac{\mu}{2}\,\mathrm{Tr}[\mathbf{R}]\sigma_\eta^2=\mathrm{EMSE}_{\mathrm{CLMS}}^{\mathrm{small}\ \mu} \quad (51)$$

for small step-sizes. For large step-sizes, this becomes

$$\mathrm{EMSE}_{\mathrm{DC-CLMS}}^{\mathrm{large}\ \mu}=\frac{\mu\,\mathrm{Tr}[\mathbf{R}]\sigma_\eta^2}{2-\mu\,\mathrm{Tr}[\mathbf{R}]}=\mathrm{EMSE}_{\mathrm{CLMS}}^{\mathrm{large}\ \mu}. \quad (52)$$

When the step-size of the DC-CLMS is equal to that of the CLMS, they achieve the same steady state excess mean square error.

5.4. Excess mean square error of the ACLMS

We have shown that the ACLMS behaves like the DC-CLMS if the corresponding learning rates satisfy the condition $\mu_{\mathrm{DC-CLMS}}=2\mu_{\mathrm{ACLMS}}$. This makes it possible to write the EMSE for ACLMS in the form

$$\mathrm{EMSE}_{\mathrm{ACLMS}}^{\mathrm{small}\ \mu}=\mu\,\mathrm{Tr}[\mathbf{R}]\sigma_\eta^2=2\times\mathrm{EMSE}_{\mathrm{CLMS}}^{\mathrm{small}\ \mu} \quad (53)$$

for small step sizes, and

$$\mathrm{EMSE}_{\mathrm{ACLMS}}^{\mathrm{large}\ \mu}=\frac{\mu\,\mathrm{Tr}[\mathbf{R}]\sigma_\eta^2}{1-\mu\,\mathrm{Tr}[\mathbf{R}]}>\mathrm{EMSE}_{\mathrm{CLMS}}^{\mathrm{large}\ \mu}$$

for large step sizes. This result reveals that when modelling a strictly linear system, the ACLMS has a larger steady state error compared to the CLMS. This is attributed to the gradient noise introduced by the additional filter coefficients needed by the ACLMS.

6. Simulations

To verify the analyses, all the filters considered were evaluated in the system identification setting with the step-size $\mu=0.02$. The mean square error (MSE) of the algorithms was calculated at each time instant, k, by averaging the error power from 100 independent trials to give

$$\mathrm{MSE}_k=\frac{1}{100}\sum_{\ell=1}^{100}\left|d_k^{(\ell)}-y_k^{(\ell)}\right|^2 \quad (54)$$

where $d_k^{(\ell)}$ is the desired signal and $y_k^{(\ell)}$ is the estimate given by the algorithms considered at trial (realisation) ℓ. The performances of CLMS, CLMSr and CLMSi were assessed for identifying a strictly linear MA(4) model described by

$$y_k=b_0x_k+b_1x_{k-1}+b_2x_{k-2}+b_3x_{k-3}+\eta_k \quad (55)$$

for which the coefficients were

$$b_0=6-6j,\quad b_1=0.5+j,\quad b_2=-2+j,\quad b_3=2+3j \quad (56)$$

and the statistics of the data and noise

$$x\sim\mathcal{N}(0,1)+j\mathcal{N}(0,1),\quad \eta\sim\mathcal{N}(0,0.1)+j\mathcal{N}(0,0.1). \quad (57)$$

Fig. 4 shows that CLMSr and CLMSi achieved the same steady state mean square error as the CLMS while requiring only half the operations of the CLMS. This is consistent with Remark 6 and implies that the more efficient CLMSr or CLMSi filters can substitute the standard CLMS in strictly linear estimation.

Fig. 5 shows the performances of the DC-CLMS, CLMS and ACLMS when estimating a widely linear system, in

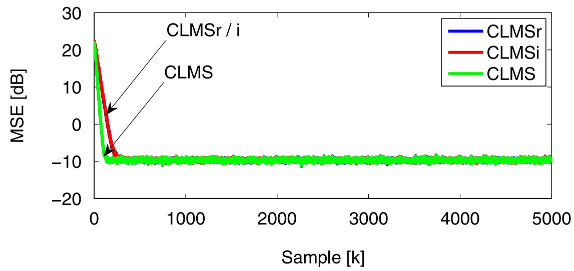

Fig. 4. The proposed CLMSr and CLMSi have the same steady state mean square error as the conventional CLMS.

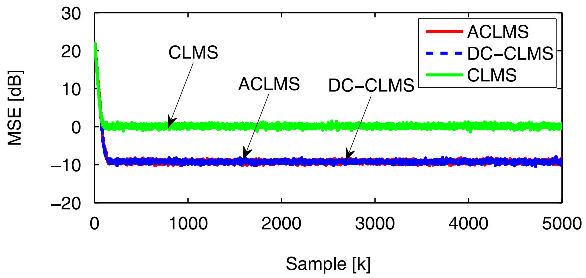

Fig. 5. Mean square error curves of the CLMS, ACLMS and the proposed DC-CLMS, show that the DC-CLMS and ACLMS have the same mean square error performance when modelling widely linear systems.

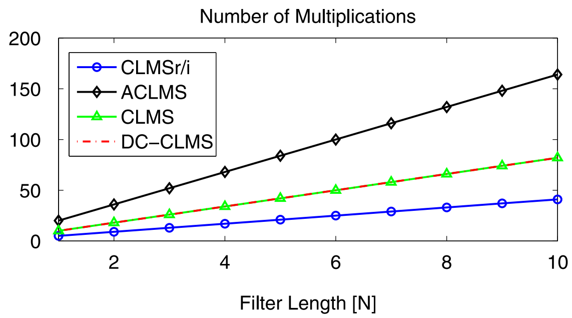

Fig. 6. Number of multiplications as a function of the filter length.

terms of the mean square error (MSE). The widely linear MA(4) system used to generate the signal is given by

$$y_k = b_0 x_k + b_1 x_{k-1} + b_2 x_{k-2} + b_3 x_{k-3} + \overline{b}_0 x_k^* + \overline{b}_1 x_{k-1}^* + \overline{b}_2 x_{k-2}^* + \overline{b}_3 x_{k-3}^* + \eta_k \quad (58)$$

where the same system coefficients b_i, $\{i = 0, \ldots, 3\}$ described in (56) and $\overline{b}_0 = 0.2 - 0.2j$, $\overline{b}_1 = 0.1 + 0.1j$, $\overline{b}_2 = 2$, $\overline{b}_3 = 0.4j$, were used. The statistics of the input data and noise are described in (57).

Conforming with the analysis, the CLMS exhibited a bias due to the inherent under-modelling (see [20]) while the DC-CLMS and ACLMS were able to model the underlying widely linear system correctly. It is important to re-emphasize that the DC-CLMS was able to achieve the same mean square performance as the ACLMS by using the half the number of operations required by the ACLMS.

The computational complexity of the algorithms considered, represented by the number of multiplications per iteration, is shown in Fig. 6. The number of multiplications of the CLMSr and CLMSi are half of that of the CLMS making the CLMSr or CLMSi more efficient than the CLMS for estimating a strictly linear system. Similarly, the DC-CLMS requires approximately half the number of operations of the ACLMS, and is more efficient than the ACLMS at modelling widely linear systems.

Table 2 compares the steady state MSE of the CLMS to that of the ACLMS for estimating the strictly linear MA(4) model in (55) when both filters have the same step-size. Notice how the CLMS achieves lower steady state MSE than the ACLMS, thus supporting the analysis in Section 5. Note the close match between the theoretical EMSE, as measured by (53), and the simulated EMSE and how, as expected, the simulated EMSE for the ACLMS is approximately twice that of the CLMS.

Finally, the behaviour DC-CLMS was analysed for modelling strictly linear systems driven by noncircular inputs. A strictly linear MA(1) signal given by

$$y_k = b_0 x_k + \eta_k, \quad b_0 = 6 + 4.8j$$
$$x \sim 0.5\mathcal{N}(0, 0.9) + j0.15\mathcal{N}(0, 0.9) \quad (59)$$

was driven by a noncircular driving noise, x_k, as shown in (59).

Fig. 7 shows that the learning curve for the DC-CLMS has two distinct regions: Region 1 in the MSE evolution, shown in the top panel, has a steeper gradient compared

Table 2
MSE and EMSE for the CLMS and ACLMS for the identification of a strictly linear MA(4) system.

	J_{min}	CLMS		ACLMS	
		MSE	EMSE	MSE	EMSE
Simulations	0.1	0.1085	0.0085	0.1195	0.0195
Theory	0.1	0.1080	0.0080	0.1160	0.0160

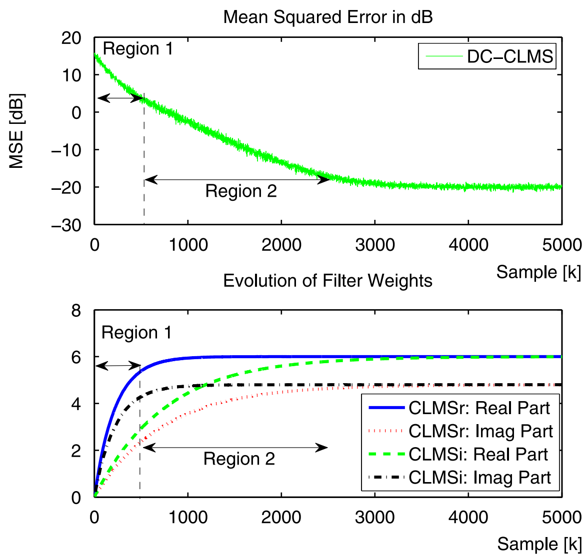

Fig. 7. Mean square error of the DC-CLMS for noncircular signals (top panel) shows two distinct regions in the transient regime due to the difference in the convergence rates of the CLMSr and CLMSi subfilters (bottom panel).

to Region 2. To explain this phenomenon, we also show the convergence rates of the real and imaginary parts of the weights of the CLMSr/i in the bottom panel of Fig. 7.

Taking the CLMSr as an example, we can observe that in Region 1, the real part of the CLMSr weight adapts more rapidly compared to its imaginary part. This causes the average adaptation rate to be higher in Region 1. Region 2 begins after the real part of the CLMSr weight has converged and the average adaptation rate is lower because of the slower adaptation of the imaginary part of the CLMSr weight. A similar behaviour is observed for the CLMSi filter, where the imaginary part of the CLMSi weight adapts more rapidly compared to its real part.

7. Conclusion

We have introduced an alternative formulation for widely linear estimation and have developed a corresponding adaptive filter referred to as the DC-CLMS. The CDC estimation framework splits the MSE cost function into the contributions from estimating the real and imaginary parts of the signal. By optimizing individually for those parts, the CDC estimator obtains the degrees of freedom necessary for widely linear estimation. The adaptive DC-CLMS has been shown to be identical to the ACLMS, while only requiring approximately half the mathematical operations. For jointly circular signals, the CLMSr and CLMSi have been shown to provide the same steady state solution as the CLMS, while requiring half the number of numerical operations. In addition, the analysis of the two sub-filters (CLMSr and CLMSi) has allowed us to derive expressions for the stability range and EMSE of the ACLMS which are simpler and physically more intuitive than the existing analyses. Simulations in the system identification setting support the approach.

References

[1] D.P. Mandic, S.L. Goh, Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models, John Wiley and Sons Ltd., 2009, http://dx.doi.org/10.1002/9780470742624.

[2] P.J. Schreier, L.L. Scharf, Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals, Cambridge University Press, 2012, http://dx.doi.org/10.1080/00107514.2011.628049.

[3] F. Neeser, J. Massey, Proper complex random processes with applications to information theory, IEEE Trans. Inf. Theory 39 (4) (1993) 1293–1302, http://dx.doi.org/10.1109/18.243446.

[4] B. Picinbono, P. Chevalier, Widely linear estimation with complex data, IEEE Trans. Signal Process. 43 (8) (1995) 2030–2033.

[5] J. Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm, IEEE Trans. Signal Process. 56 (7) (2008) 3061–3068.

[6] 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 , http://dx.doi.org/10.1109/TSP.2009.2027461.

[7] S. Javidi, M. Pedzisz, S.L. Goh, D.P. Mandic, The augmented complex least mean square algorithm with application to adaptive prediction problems, in: Proceedings of the 1st IARP Workshop on Cognitive Information Processing, 2008, pp. 54–57.

[8] S.C. Douglas, Widely-linear recursive least-squares algorithm for adaptive beamforming, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2009, pp. 2041–2044.

[9] Y.X., 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.

[10] K. Kreutz-Delgado, The Complex Gradient Operator and the $\mathbb{CR}-$Calculus, Technical Report, University of California, San Diego, 2006.

[11] D. Mandic, S. Still, S. Douglas, Duality between widely linear and dual channel adaptive filtering, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, 2009, pp. 1729–1732.

[12] G.A.N. Fernando, V.H. Nascimiento, M.T.M. Silva, Reduced-complexity widely linear adaptive estimation, Proceedings of the Seventh International Symposium on Wireless Communication Systems, 2010, pp. 399–403.

[13] J. Eriksson, E. Ollila, V. Koivunen, Essential statistics and tools for complex random variables, IEEE Trans. Signal Process. 58 (10) (2010) 5400–5408, http://dx.doi.org/10.1109/TSP.2010.2054085.

[14] B. Widrow, J. McColl, M. Ball, The complex LMS algorithm, in: Proceedings of the IEEE, vol. 63, 1975, pp. 719–720.

[15] B.D.H. Brandwood, A complex gradient operator and its applications in adaptive array theory, in: IEE Proceedings of the Communications, Radar and Signal Proceedings, vol. 130 (1), 1983, pp. 11–16.

[16] A. van den Bos, Complex gradient and Hessian, in: IEE Proceedings of the Vision, Image and Signal Processing, vol. 141 (6), 1994, pp. 380–383.

[17] J. Eriksson, V. Koivunen, Complex random vectors and ICA models: identifiability, uniqueness, and separability, IEEE Trans. Inf. Theory 52 (3) (2006) 1017–1029, http://dx.doi.org/10.1109/TIT.2005.864440.

[18] S.C. Douglas, D.P. Mandic, Performance analysis of the conventional complex LMS and augmented complex LMS algorithms, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing vol. 3, 2010, pp. 3794–3797.

[19] N. Yousef, A. Sayed, A unified approach to the steady-state and tracking analyses of adaptive filters, IEEE Trans. Signal Process. 49 (2) (2001) 314–324.

[20] D. Mandic, Y. Xia, S. Douglas, Steady state analysis of the CLMS and augmented CLMS algorithms for noncircular complex signals, in: Proceedings of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2010, pp. 1635–1639.