An Adaptive Diffusion Augmented CLMS Algorithm for Distributed Filtering of Noncircular Complex Signals

Yili Xia, *Student Member, IEEE*, Danilo P. Mandic, *Senior Member, IEEE*, and Ali H. Sayed, *Fellow, IEEE*

Abstract—An adaptive diffusion augmented complex least mean square (D-ACLMS) algorithm for collaborative processing of the generality of complex signals over distributed networks is proposed. The algorithm enables the estimation of both second order circular (proper) and noncircular (improper) signals within a unified framework of augmented complex statistics. The analysis shows that the performance advantage of the widely linear D-ACLMS over the strictly linear D-CLMS increases with the degree of noncircularity while maintaining similar performance for proper data. Simulations on both synthetic benchmark and real world noncircular data support the approach.

Index Terms—Adaptive diffusion, distributed estimation, noncircular complex signals, widely linear modelling.

I. INTRODUCTION

A Number of modern applications, such as those arising in environmental monitoring, target localization, and sensor networks, require new classes of estimation techniques, based on distributed networks of nodes embedded with cooperative data processing. In such cases, we can employ limited cooperation, incremental like techniques, or adaptive diffusion strategies to implement cooperation among individual adaptive nodes [1], [2]. These nodes are equipped with local learning capabilities: they produce local estimates for the parameter of interest and share information with only their neighbors. When larger scale communication resources are available, distributed adaptive algorithms can be derived that exploit more fully the network connectivity and increase the degree of cooperation among nodes. Owing to their localized and real-time mode of operation, adaptive diffusion techniques offer enhanced robustness to link and node failures and are scalable.

Standard complex-valued adaptive filtering algorithms are normally considered generic extensions of the corresponding algorithms in $\mathbb{R}$. However, consider a real-valued conditional mean squared error (MSE) estimator $\hat{y} = E[y|\mathbf{x}]$, which estimates the signal y in terms of another observation $\mathbf{x}$. For zero mean, jointly normal y and $\mathbf{x}$, the optimal solution is the linear

model given by $\hat{y} = \mathbf{x}^T \mathbf{h}$, where $\mathbf{h} = [h_1, \ldots, h_L]^T$ is a vector of fixed filter coefficients, $\mathbf{x} = [x_1, \ldots, x_L]^T$ the regressor vector, and $(\cdot)^T$ the vector transpose operator.

In the complex domain, since both the real and imaginary parts of complex variables are real, we have $\Re(\hat{y}) = E[\Re(y)|\Re(\mathbf{x}), \Im(\mathbf{x})]$, $\Im(\hat{y}) = E[\Im(y)|\Re(\mathbf{x}), \Im(\mathbf{x})]$, and thus $\hat{y} = E[\Re(\hat{y})|\Re(\mathbf{x}), \Im(\mathbf{x})] + \jmath E[\Im(\hat{y})|\Re(\mathbf{x}), \Im(\mathbf{x})]$, where the operators $\Re(\cdot)$ and $\Im(\cdot)$ extract respectively the real and imaginary parts of a complex variable. Substitute $\Re(\mathbf{x}) = (\mathbf{x} + \mathbf{x}^*)/2$ and $\Im(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^*)/2\jmath$ to arrive at

$$\hat{y} = E[\Re(y)|\mathbf{x}, \mathbf{x}^*] + \jmath E[\Im(y)|\mathbf{x}, \mathbf{x}^*] = E[y|\mathbf{x}, \mathbf{x}^*] \quad (1)$$

leading to the *widely linear* estimator for complex-valued data [3]–[5]

$$\hat{y} = \mathbf{h}^T \mathbf{x} + \mathbf{g}^T \mathbf{x}^* = \mathbf{x}^T \mathbf{h} + \mathbf{x}^H \mathbf{g} \quad (2)$$

where $\mathbf{h}$ and $\mathbf{g}$ are complex-valued coefficient vectors. Such a widely linear estimator is optimal for the generality of complex signals (both proper and improper), and it simplifies into the strictly linear model ($\mathbf{g} = \mathbf{0}$) for proper data. In practice, the widely linear estimate in (2) is based on a regressor vector produced by concatenating the input vector $\mathbf{x}$ with its conjugate $\mathbf{x}^*$, to give an augmented $2L \times 1$ input vector $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^H]^T$, together with the corresponding augmented coefficient vector $\mathbf{w}^a = [\mathbf{h}^T, \mathbf{g}^T]^T$. The corresponding $2L \times 2L$ augmented covariance matrix then becomes [6]

$$\mathbf{C}_{\mathbf{xx}}^a = E\left[\mathbf{x}^a \mathbf{x}^{aH}\right] = E\begin{bmatrix} \mathbf{x} \\ \mathbf{x}^* \end{bmatrix} \begin{bmatrix} \mathbf{x}^H \mathbf{x}^T \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{\mathbf{xx}} & \mathbf{P}_{\mathbf{xx}} \\ \mathbf{P}_{\mathbf{xx}}^* & \mathbf{C}_{\mathbf{xx}}^* \end{bmatrix} \quad (3)$$

and contains the full second order statistical information. Hence, in addition to the standard covariance matrix $\mathbf{C}_{\mathbf{xx}} = E[\mathbf{xx}^H]$, we also need to consider the pseudo-covariance matrix, $\mathbf{P}_{\mathbf{xx}} = E[\mathbf{xx}^T]$; processes with vanishing pseudo-covariance, $\mathbf{P}_{\mathbf{xx}} = \mathbf{0}$, are termed second order circular (proper). The augmented covariance matrix in (3) considers the improperness of the observation process $\mathbf{x}$ in the widely linear estimation problem in (2). A more general solution where the estimate and observation are not jointly proper can be found in [7].

The augmented complex statistics has opened the possibility to design adaptive filtering algorithms based on widely linear models, suitable for the processing of both circular and noncircular signals. These algorithms are usually termed "widely linear" or "augmented" algorithms, such as the WL-LMS algorithm in communications [8], and the augmented CLMS (ACLMS) algorithm in adaptive filtering area [9], [10].

In this letter, the Diffusion ACLMS (D-ACLMS) algorithm is introduced in order to provide cooperative adaptive estimation of noncircular complex-valued signals, where the distributed nodes cooperate for the estimation of the same global task. The advantage of the widely linear D-ACLMS over the strictly linear
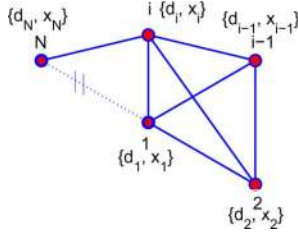
Fig. 1.　Topology of a distributed network with N nodes.

D-CLMS is illustrated over simulations on both benchmark and real-world noncircular wind signals.

II. DERIVATION OF D-ACLMS ALGORITHM

Consider a network with N distributed nodes, where each node has access to the local information $\{d_i(k), \mathbf{x}_i(k)\}$ at time instant k, where $d_i(k)$ is the desired signal, and $\mathbf{x}_i(k)$ is a regressor input vector of length L, defined as $\mathbf{x}_i(k) = [x_i(k-1), \ldots, x_i(k-L)]^T$, both associated with node i. As illustrated in Fig. 1, each node in the network has access only to their peer neighbors, and the diffusion protocols allow every node i to combine information collected from its neighborhood $\mathcal{N}_i$, which is the set of all the nodes linking to node i, including itself. By linearly combining the weight estimates in the neighborhood of node i, and following the derivation in [2], we can motivate replacing the local weights for standard and conjugate parts, that is $\mathbf{h}_i(k)$ and $\mathbf{g}_i(k)$, with the weighted estimates

$$\underline{\mathbf{h}}_i(k) = \sum_{l \in \mathcal{N}_i} c(i,l)\mathbf{h}_l(k) \text{ and } \underline{\mathbf{g}}_i(k) = \sum_{l \in \mathcal{N}_i} c(i,l)\mathbf{g}_l(k) \quad (4)$$

where the italic symbols with underscores denote the combined weights vectors $\underline{\mathbf{h}}_i(k)$ and $\underline{\mathbf{g}}_i(k)$ at node i for some combining coefficients $\{c(i,l) \geq 0\}$ and satisfying $\sum_{l \in \mathcal{N}_i} c(i,l) = 1$. This aggregate estimation at node i can be interpreted as a weighted linear combination which fuses information from nodes across the network into node i. The resulting aggregated weight vectors $\underline{\mathbf{h}}_i(k)$ and $\underline{\mathbf{g}}_i(k)$ at node i, can then be fed into its local adaptive filter in order to respond to local information and update the local weights $\mathbf{h}_i(k)$ and $\mathbf{g}_i(k)$. This way, the resulting cooperative adaptive network provides a peer-to-peer estimation framework that is robust to node and link failures and exploits network connectivity.

In the context of widely linear adaptive filtering, the output signal $y_i(k)$ at node i of a cooperative filter is given by

$$y_i(k) = \mathbf{x}_i^T(k)\underline{\mathbf{h}}_i(k) + \mathbf{x}_i^H(k)\underline{\mathbf{g}}_i(k) \quad (5)$$

whereas the corresponding output error at the node i is defined as $e_i(k) = d_i(k) - y_i(k)$. Based on the global cost function we wish to optimize $\mathcal{J}_{\text{glob}}(\underline{\mathbf{h}}, \underline{\mathbf{g}}) = \sum_{i=1}^N E|e_i(k)|^2 = \sum_{i=1}^N E|d_i(k) - \mathbf{x}_i^T(k)\underline{\mathbf{h}}_i(k) - \mathbf{x}_i^H(k)\underline{\mathbf{g}}_i(k)|^2$, using the $\mathbb{CR}$ calculus [11], the local weight updates for the standard part and conjugate part at each node i can be derived as[1]

$$\mathbf{h}_i(k+1) = \underline{\mathbf{h}}_i(k) + \mu_i e_i(k)\mathbf{x}_i^*(k) \quad (6)$$
$$\mathbf{g}_i(k+1) = \underline{\mathbf{g}}_i(k) + \mu_i e_i(k)\mathbf{x}_i(k) \quad (7)$$

[1]Notice that the standard Cauchy–Riemann derivative of a real function of complex variables, such as the error power in the cost function does not exist, however, the $\mathbb{CR}$ calculus shows that the steepest descent is in the direction of the conjugate weight vector, i.e., $\nabla_{\mathbf{h}^*}\mathcal{J}$[11].

TABLE I
COMPARISON OF THE COMPUTATIONAL COMPLEXITIES OF THE PROPOSED
D-ACLMS AND THE STRICTLY LINEAR D-CLMS (PER ITERATION)

Algorithms	No. of real multiplications	No. of real additions
D-CLMS	$(2N_i+8)L+2$	$(2N_i+6)L$
D-ACLMS	$(4N_i+16)L+4$	$(4N_i+14)L-2$

The computational complexities, in terms of real multiplications and additions[2], of the proposed D-ACLMS and the strictly linear D-CLMS [1] are summarised in Table I, where N_i is the number of the neighboring nodes of node i.

III. STABILITY ANALYSIS OF D-ACLMS ALGORITHM

To provide insight to the convergence of the D-ACLMS algorithm, we shall next investigate the role of the degree of cooperation and network topology on the system performance. To that end, rewrite the local and global quantities as

$$\begin{aligned}\mathbf{h}(k) &= \text{col}\{\mathbf{h}_1(k), \ldots, \mathbf{h}_N(k)\};\\ \mathbf{g}(k) &= \text{col}\{\mathbf{g}_1(k), \ldots, \mathbf{g}_N(k)\}\\ \underline{\mathbf{h}}(k) &= \text{col}\{\underline{\mathbf{h}}_1(k), \ldots, \underline{\mathbf{h}}_N(k)\};\\ \underline{\mathbf{g}}(k) &= \text{col}\{\underline{\mathbf{g}}_1(k), \ldots, \underline{\mathbf{g}}_N(k)\}\\ \mathbf{X}(k) &= \text{diag}\{\mathbf{x}_1(k), \ldots, \mathbf{x}_N(k)\};\\ \mathbf{d}(k) &= \text{col}\{d_1(k), \ldots, d_N(k)\}\end{aligned}$$

in terms of the quantities that appear in (4), (6), and (7), and denote by $\mathbf{D} = \text{diag}\{\mu_1\mathbf{I}_L, \mu_2\mathbf{I}_L, \ldots, \mu_N\mathbf{I}_L\}$, a diagonal matrix containing the local step-sizes.

To evaluate performance, without loss in generality, consider an improper teaching signal $d_i(k)$, given by [5]

$$d_i(k) = \mathbf{x}_i^T(k)\mathbf{h}_o + \mathbf{x}_i^H(k)\mathbf{g}_o + \nu_i(k) \quad (8)$$

where $\mathbf{h}_o$ and $\mathbf{g}_o$ are respectively the unknown optimal local standard and conjugate weight vectors, and the symbol $\nu_i(k)$ denotes samples of doubly white[3] noise with variance σ_ν^2, statistically independent of the input sequence $\mathbf{x}_i(k)$. The D-ACLMS algorithm can now be written in a compact form

$$\begin{aligned}\underline{\mathbf{h}}(k) &= \mathbf{G}\mathbf{h}(k), \qquad \underline{\mathbf{g}}(k) = \mathbf{G}\mathbf{g}(k)\\ \mathbf{h}(k+1) &= \underline{\mathbf{h}}(k)\\ &\quad + \mathbf{D}(\mathbf{d}(k) - \mathbf{X}^T(k)\underline{\mathbf{h}}(k) - \mathbf{X}^H(k)\underline{\mathbf{g}}(k))\mathbf{X}^*(k)\\ \mathbf{g}(k+1) &= \underline{\mathbf{g}}(k)\\ &\quad + \mathbf{D}(\mathbf{d}(k) - \mathbf{X}^T(k)\underline{\mathbf{h}}(k) - \mathbf{X}^H(k)\underline{\mathbf{g}}(k))\mathbf{X}(k)\end{aligned}$$

where $\mathbf{G} = \mathbf{C} \otimes \mathbf{I}_L$ is the transition matrix and $\mathbf{C} = [c(i,l)]$ is the $N \times N$ network connection matrix representing the network topology: a nonzero entry $c(i,l)$ means that node i is connected to node l. Moreover, $\mathbf{C}\mathbf{q}_N = \mathbf{q}_N$, where $\mathbf{q}_N = \text{col}\{1, \ldots, 1\}_N$ is an $N \times 1$ column vector.

The standard and conjugate weight error vectors for all the nodes are given by

$$\widetilde{\mathbf{h}}(k) = \mathbf{h}_{(o)} - \mathbf{h}(k) \qquad \text{and} \qquad \widetilde{\mathbf{g}}(k) = \mathbf{g}_{(o)} - \mathbf{g}(k) \quad (9)$$

[2]Note that one complex multiplication requires four real multiplications and two real additions, while one complex additions requires two real additions.

[3]The term "doubly white" refers to the white and mutually statistically independent real and imaginary parts, with variance $\sigma_\nu^2 = \sigma_{\Re(\nu)}^2 + \sigma_{\Im(\nu)}^2$.

where $\mathbf{h}_{(o)} = \mathbf{q}_N \otimes \mathbf{h}_o$, $\mathbf{g}_{(o)} = \mathbf{q}_N \otimes \mathbf{g}_o$, and their evolution is governed by

$$\widetilde{\mathbf{h}}(k+1) = \mathbf{G}\widetilde{\mathbf{h}}(k) - \mathbf{D}\mathbf{X}^*(k)\left(\mathbf{X}^T(k)\mathbf{G}\widetilde{\mathbf{h}}(k)\right.$$
$$\left. + \mathbf{X}^H(k)\mathbf{G}\widetilde{\mathbf{g}}(k) + \boldsymbol{\nu}(k)\right)$$

$$\widetilde{\mathbf{g}}(k+1) = \mathbf{G}\widetilde{\mathbf{g}}(k) - \mathbf{D}\mathbf{X}(k)\left(\mathbf{X}^T(k)\mathbf{G}\widetilde{\mathbf{h}}(k)\right.$$
$$\left. + \mathbf{X}^H(k)\mathbf{G}\widetilde{\mathbf{g}}(k) + \boldsymbol{\nu}(k)\right).$$

Therefore, the evolution of the augmented weight error vector $\widetilde{\mathbf{w}}(k+1) = [\widetilde{\mathbf{h}}^T(k+1), \widetilde{\mathbf{g}}^T(k+1)]^T$ is described by

$$\begin{bmatrix} \widetilde{\mathbf{h}}(k+1) \\ \widetilde{\mathbf{g}}(k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{I} - \mathbf{D}\mathbf{X}^*(k)\mathbf{X}^T(k) & -\mathbf{X}^*(k)\mathbf{X}^H(k) \\ -\mathbf{X}(k)\mathbf{X}^T(k) & \mathbf{I} - \mathbf{D}\mathbf{X}(k)\mathbf{X}^H(k) \end{bmatrix}$$
$$\times \mathbf{G}\begin{bmatrix} \widetilde{\mathbf{h}}(k) \\ \widetilde{\mathbf{g}}(k) \end{bmatrix} - \mathbf{D}\begin{bmatrix} \mathbf{X}^*(k) \\ \mathbf{X}(k) \end{bmatrix}\boldsymbol{\nu}(k) \quad (10)$$

and can be written more compactly as

$$\widetilde{\mathbf{w}}(k+1) = \left(\mathbf{I} - \mathbf{D}\mathbf{X}^{a*}(k)\mathbf{X}^{aT}(k)\right)\mathbf{G}\widetilde{\mathbf{w}}(k) - \mathbf{D}\mathbf{X}^{a*}(k)\boldsymbol{\nu}(k)$$

where $\mathbf{X}^a(k) = [\mathbf{X}^T(k), \mathbf{X}^H(k)]^T$.

Under the assumption that the noise is statistically independent of the regressor vectors and the independence among the regressor vectors over space and time, upon applying the statistical expectation operator on both sides of the above equation, the evolution of the mean weight error vector is governed by

$$E[\widetilde{\mathbf{w}}(k+1)] = (\mathbf{I} - \mathbf{D}\mathcal{C}^{a*}_{\mathbf{XX}})\mathbf{G}E[\widetilde{\mathbf{w}}(k)] \quad (11)$$

where $\mathcal{C}^a_{\mathbf{XX}} = \mathrm{diag}\{\mathcal{C}^a_{\mathbf{x}_1\mathbf{x}_1}, \ldots, \mathcal{C}^a_{\mathbf{x}_N\mathbf{x}_N}\}$ is block diagonal and $\mathcal{C}^a_{\mathbf{x}_i\mathbf{x}_i} = E[\mathbf{x}^a_i \mathbf{x}^{aH}_i]$ is the augmented covariance matrix as defined in (3). Hence, for the expression (11) to converge, that is, $E[\widetilde{\mathbf{w}}(k+1)] \to E[\widetilde{\mathbf{w}}(k)]$, $k \to \infty$, for stability in the mean $|\lambda_i(\mathbf{BG})| < 1$, where λ_i are the eigenvalues for $i = 1, \ldots, N$ and $\mathbf{B} = \mathbf{I} - \mathbf{D}\mathcal{C}^{a*}_{\mathbf{XX}}$. Following the analysis in [1], it can be shown that $|\lambda_{\max}(\mathbf{BG})| \le |\lambda_{\max}(\mathbf{B})|$. In other words, the spectral radius of $\mathbf{BG}$ is generally smaller than that of $\mathbf{B}$, and therefore cooperation under the diffusion protocol has a stabilizing effect on the network.

A. Performance Dependence on the Degree of Noncircularity

The strictly linear D-CLMS can be summarised as [1]

$$y_i(k) = \mathbf{x}^T_i(k)\underline{\mathbf{h}}_i(k), \quad e_i(k) = d_i(k) - y_i(k)$$
$$\underline{\mathbf{h}}_i(k) = \sum_{l \in \mathcal{N}_i} c(i,l)\mathbf{h}_l(k)$$
$$\mathbf{h}_i(k+1) = \underline{\mathbf{h}}_i(k) + \mu_i e_i(k)\mathbf{x}^*_i(k) \quad (12)$$

and the evolution of its weight error vector follows the same generic form as that in (11), with the standard covariance matrix replacing the augmented covariance matrix. Since it does not account for the pseudocovariance, it cannot explain the noncircularity (rotation dependent distribution) of input signals, and performs suboptimally for improper data. Locally, for every node in the network, the analysis in [5], [10] shows that the advantage of the widely linear ACLMS over standard CLMS is more pronounced with higher degrees of noncircularity. Since the diffusion topology facilitates mutual communication between nodes,
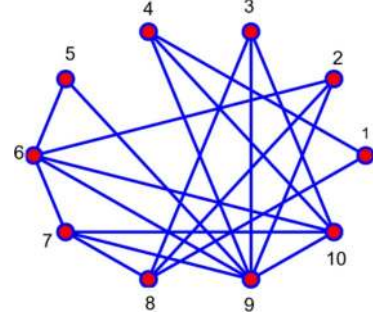

Fig. 2. Network topology used in the simulations.

and each distributed node performs local adaptive estimation on the same global task, the performance advantage of D-ACLMS over D-CLMS is also more pronounced when processing improper signals, whereas for proper signals their performances are identical, as shown in the simulations.

IV. SIMULATIONS

The proposed D-ACLMS was compared with the standard D-CLMS in a one step ahead prediction setting of both proper and improper signals. Out of the several methods to select the combination coefficients $c(i,l)$, such as the Metropolis [1], Laplacian [12] and nearest neighbor rules [1], we used the Metropolis network topology described by (see Fig. 2)

$$c(i,l) = \begin{cases} \frac{1}{\max(n_i, n_l)}, & \text{if } l \ne i \text{ are connected} \\ 0, & \text{if } l \ne i \text{ are not connected} \\ 1 - \sum_{l \in \mathcal{N}_{i/i}} c(i,l), & \text{for } l = i \end{cases} \quad (13)$$

where n_i and n_l are respectively the numbers of nodes in the neighborhood of nodes i and l. A network with $N = 10$ nodes was considered, with filter length $L = 4$, and the global step size $\mu = 0.01$. Learning curves comparing D-CLMS and D-ACLMS were produced by averaging 100 iterations of independent trials, performed on benchmark complex-valued stable circular and noncircular signals. Single trial simulations were performed on real-world noncircular wind signals[4].

The benchmark circular complex signal used in simulations was a stable autoregressive AR(4) process, with coefficient vector $\mathbf{a} = [1.79, -1.85, 1.27, -0.41]$ and complex-valued doubly white Gaussian driving noise $q(k)$, with zero mean and unit variance [5]. The benchmark improper complex signal was an autoregressive moving average ARMA(4,1) complex process, which combined the improper MA(1) model in [13] with the above stable AR(4) circular process, to give

$$r(k) = \mathbf{a}^T\mathbf{r}(k) + 2q(k) + 0.5q^*(k) + q(k-1) + 0.9q^*(k-1) \quad (14)$$

where $\mathbf{r}(k) = [r(k-1), \ldots, r(k-4)]^T$, and the driving noise covariance $E\{q(k-i)q^*(k-j)\} = \delta(i-j)$ and pseudocovariance $E\{q(k-i)q(k-j)\} = C\delta(i-j)$, with the degree of improperness $C = 0.95$.

Another benchmark noncircular signal considered was the Ikeda signal (nonlinear and with coupled states), given by

$$x(k+1) = 1 + u(x(k)\cos[t(k)] - y(k)\sin[t(k)])$$
$$y(k+1) = u(x(k)\sin[t(k)] + y(k)\cos[t(k)]) \quad (15)$$

[4]In $\mathbb{C}$, the wind vector $w(k)$ can be represented as $w(k) = |v(k)|e^{j\Phi(k)}$, where $v(k)$ denotes the speed and $\Phi(k)$ the direction.

TABLE II
COMPARISON OF DEGREES OF NONCIRCULARITY s
FOR THE VARIOUS CLASSES OF SIGNALS

	Circular AR(4)	Noncircular ARMA	Ikeda map
s	0.0016	0.9450	0.8936

	Wind (*low*)	Wind (*medium*)	Wind (*high*)
s	0.1583	0.4305	0.8117

TABLE III
COMPARISON OF PREDICTION GAINS R_p FOR THE VARIOUS
CLASSES OF CIRCULAR AND NONCIRCULAR SIGNALS

R_p [dB]	Circular AR(4)	Noncircular ARMA	Ikeda map	Wind (*low*)	Wind (*medium*)	Wind (*high*)
CLMS	3.5605	3.3118	2.9585	5.1487	7.3286	13.3386
ACLMS	3.5516	4.1351	6.3959	5.3557	7.5811	13.6833
D-CLMS	4.0577	3.8433	3.3482	5.6685	7.9849	14.2435
D-ACLMS	4.0496	4.6681	7.1925	5.9347	8.3352	14.6105

where $u = 0.8$ and $t(k) = 0.4 - 6/1 + x^2(k) + y^2(k)$.

The real world noncircular wind signals were with different dynamics, identified as regions *high*, *medium* or *low* based on the changes in the wind intensity, for more detail see [5].

For a quantitative measurement of the degree of noncircularity of a complex vector $\mathbf{x}$, we used the index $s \triangleq 1 - \det(\mathbf{C}_{\mathbf{x}^a\mathbf{x}^a})\det^{-2}(\mathbf{C}_{\mathbf{x}\mathbf{x}})$[14], where $\det(\cdot)$ denotes the matrix determinant operator; the degree of noncircularity s is normalised to within $[0, 1]$ with the value of zero indicating perfect circularity. Table II illustrates the degrees of noncircularity s for the signals considered.

The standard prediction gain $R_p \triangleq 10 \log_{10} \left(\sigma_x^2 / \hat{\sigma}_e^2 \right)$ [dB] was employed to assess the performance, where σ_x^2 and and $\hat{\sigma}_e^2$ denote respectively the variance of the input signal $x(k)$ and the forward prediction error $e(k)$ which is the averaged error over all the nodes. Table III compares averaged prediction gains Rp (dB) over 100 independent trials for the D-CLMS and D-ACLMS trained cooperative networks. Observe that for the circular AR(4) signal, the performances of the standard algorithms (noncooperative CLMS [15] and cooperative D-CLMS) were similar to those of their widely linear counterparts (ACLMS and D-ACLMS), whereas for the noncircular signals, there was an improvement in the prediction gain when the widely linear algorithms were employed. In all the cases, due to the cooperative mode of operation, the diffusion algorithms outperformed their single node counterparts.

Fig. 3 shows the evolution of mean-square errors (MSEs) for the noncircular ARMA process and Ikeda map for all the algorithms considered—observe that the diffusion algorithms achieved both faster convergence and smaller MSEs than their noncooperative counterparts, and that the widely linear algorithms outperformed the strictly linear ones.

V. CONCLUSION

We have introduced a diffusion ACLMS (D-ACLMS) algorithm for distributed adaptive estimation of general complex-valued signals (both circular and noncircular) in a cooperative fashion. The advantage of the diffusion topology over noncooperative structures in terms of convergence speed and smaller MSEs has been illustrated via analysis and simulations on both synthetic and real world noncircular complex signals.
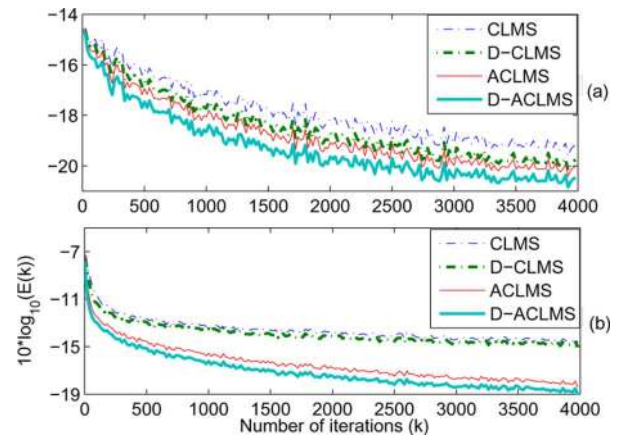

Fig. 3. Comparison of MSEs of the noncooperative CLMS and ACLMS and cooperative D-CLMS and D-ACLMS algorithms on noncircular signals (a) Noncircular ARMA and (b) Ikeda map.

REFERENCES

[1] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3122–3136, 2008.

[2] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1035–1048, 2010.

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

[4] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[5] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ: Wiley, 2009.

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

[7] J. Navarro-Moreno, M. D. Estudillo-Martínez, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Estimation of improper complex-valued random signals in colored noise by using the Hilbert space theory," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2859–2867, 2009.

[8] R. Schober, W. H. Gerstacker, and L. H. J. Lampe, "Date-aided and blind stochastic gradients for widely linear MMSE MAI supression for DS-CDMA," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 746–756, 2004.

[9] D. P. Mandic, S. Javidi, S. L. Goh, A. Kuh, and K. Aihara, "Complex-valued prediction of wind profile using augmented complex statistics," *Renewable Energy*, vol. 34, no. 1, pp. 196–210, 2009.

[10] D. P. Mandic, Y. Xia, and S. C. Douglas, "Steady state analysis of the CLMS and augmented CLMS algorithms for noncircular complex signals," in *Proc. 44th Asilomar Conf. Signals, Sys., Comp.*, 2010, pp. 1635–1639.

[11] K. Kreutz-Delgado, "The complex gradient operator and the $\mathbb{CR}$ calculus," *Lecture Suppl. ECE275A*, pp. 1–74, 2006.

[12] R. Bru, L. Elsner, and M. Neumann, "Convergence of infinite products of matrices and inner-outer interation schemes," *Elec. Trans. Numer. Anal.*, vol. 2, pp. 183–193, 1994.

[13] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3061–3068, 2008.

[14] P. J. Schreier, "Bounds on the degree of impropriety of complex random vectors," *IEEE Signal Process. Lett.*, vol. 15, pp. 190–193, 2008.

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