Cooperative Adaptive Estimation of Distributed Noncircular Complex Signals

Dahir H. Dini and Danilo P. Mandic

Department of Electrical and Electronic Engineering
Imperial College London, Exhibition Road, London, SW7 2BT
Email:{dahir.dini, d.mandic}@imperial.ac.uk

***Abstract*—The problem of distributed (cooperative) adaptive estimation of complex signals is addressed using augmented statistics and widely linear modelling, which enables optimal second order estimation of complex signals with both circular (rotation invariant) and noncircular (rotation dependent) distributions. The widely linear distributed augmented complex Kalman filter (D-ACKF) and recursive least squares (D-ACRLS) algorithms are introduced, and shown to allow for a unified treatment of the generality of complex valued signals. Further, the D-ACKF proposed here avoids the typical assumption that the observation noises at different nodes in the network are uncorrelated; thus providing enhanced performance in real-world scenarios.**

***Index Terms*—Widely linear model, complex circularity, Kalman filter, distributed recursive least squares (RLS), distributed diffusion estimation, sensor networks**

I. Introduction

Collaborative distributed networks consisting of multiple nodes, equipped with communication capabilities that share information with their neighbouring nodes, are a standard in real-world applications, including wireless sensor networks and battlefield surveillance [1] [2]. Cooperation between the nodes leads to a more robust estimation framework than using independent uncooperative nodes. Moreover, distributed systems offer numerous advantages compared with centralised systems, such as robustness to link and node failures, and lower communication overheads. In this work, we address the problem of adaptive estimation of noncircular complex signals within a framework of cooperative distributed networks; this has recently attracted plenty of interest, as complex signals are the backbone in distributed applications such as wireless communication networks and seismic sensing [3].

In second order estimation, zero-mean complex signals $\mathbf{x}$, similar to real signals, have conventionally been characterised by their covariances $\mathbf{R}_\mathbf{x} = E\{\mathbf{x}\mathbf{x}^H\}$, where $(\cdot)^H = ((\cdot)^*)^T$ is the complex conjugate-transpose operator. This characterisation yields in algorithms suited only to complex signals with rotation invariant (circular) distributions. Complex signals are, however, typically noncircular, and their rigorous treatment requires the covariance together with a second moment known as the pseudocovariance, $\mathbf{R}_\mathbf{x} = E\{\mathbf{x}\mathbf{x}^T\}$, to fully capture the second order statistics.

To introduce an optimal second order estimator for the generality of complex signals, consider first the mean square error (MSE) estimator of a real valued random vector $\mathbf{y}$ in terms of an observed real vector $\mathbf{x}$, that is, $\hat{\mathbf{y}} = E\{\mathbf{y}|\mathbf{x}\}$. For zero-mean, jointly normal $\mathbf{y}$ and $\mathbf{x}$, the optimal estimator is linear, that is

$$\hat{\mathbf{y}} = \mathbf{A}\mathbf{x} \tag{1}$$

where $\mathbf{A} = \mathbf{R}_{\mathbf{yx}}\mathbf{R}_\mathbf{x}^{-1}$ is a coefficient matrix, and $\mathbf{R}_{\mathbf{yx}} = E\{\mathbf{y}\mathbf{x}^H\}$. Standard, 'strictly linear' estimation in $\mathbb{C}$ assumes the same model but with complex valued $\mathbf{y}, \mathbf{x}$, and $\mathbf{A}$. However, when $\mathbf{y}$ and $\mathbf{x}$ are jointly noncircular $\mathbf{P}_{\mathbf{yx}} = E\{\mathbf{y}\mathbf{x}^T\} \neq \mathbf{0}$, and $\mathbf{x}$ is noncircular $\mathbf{P}_\mathbf{x} \neq \mathbf{0}$, then the optimal estimator becomes [4] [5]

$$\hat{\mathbf{y}} = \mathbf{B}\mathbf{x} + \mathbf{C}\mathbf{x}^* = \mathbf{W}\mathbf{x}^a \tag{2}$$

where $\mathbf{B} = \mathbf{R}_{\mathbf{yx}}\mathbf{D} + \mathbf{P}_{\mathbf{yx}}\mathbf{E}^*$ and $\mathbf{C} = \mathbf{R}_{\mathbf{yx}}\mathbf{E} + \mathbf{P}_{\mathbf{yx}}\mathbf{D}^*$ are coefficient matrices, with $\mathbf{D} = (\mathbf{R}_\mathbf{x} - \mathbf{P}_\mathbf{x}\mathbf{R}_\mathbf{x}^{*-1}\mathbf{P}_\mathbf{x}^*)^{-1}$ and $\mathbf{E} = -(\mathbf{R}_\mathbf{x} - \mathbf{P}_\mathbf{x}\mathbf{R}_\mathbf{x}^{*-1}\mathbf{P}_\mathbf{x}^*)^{-1}\mathbf{P}_\mathbf{x}\mathbf{R}_\mathbf{x}^{*-1}$, while $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^H]^T$ is the augmented input vector, and $\mathbf{W} = [\mathbf{B}, \mathbf{C}]$ the optimal coefficient matrix. The estimator in (2) is optimal for the generality of complex signals, both circular and noncircular. Further, the full second order information is contained in the augmented covariance matrix

$$\mathbf{R}_\mathbf{x}^a = E\{\mathbf{x}^a\mathbf{x}^{aH}\} = \begin{bmatrix} \mathbf{R}_\mathbf{x} & \mathbf{P}_\mathbf{x} \\ \mathbf{P}_\mathbf{x}^* & \mathbf{R}_\mathbf{x}^* \end{bmatrix} \tag{3}$$

and as such, estimation based on $\mathbf{R}_\mathbf{x}^a$ incorporates both the covariance and pseudocovariance.

Adaptive filters suited to the generality complex signals have been introduced recently, and include the augmented complex Kalman filter (ACKF) and recursive least squares (ACRLS) algorithms [5] [6] [7]. We here extend these algorithms to the case of distributed collaborative estimation of noncircular complex data, and proposed the widely linear distributed ACKF (D-ACKF) and distributed ACRLS (D-ACRLS) algorithms, inspired partly by the diffusion strategies of [1] [2] [8]. The D-ACKF proposed here accounts for possible cross-correlations between the observation noises at different nodes, a real-world scenario encountered when the nodes are operating in a common environment, such as in the presence of jamming-noise, and extends earlier solutions which assume uncorrelated nodal observation noises [1] [2] [8].

II. Distributed Recursive Least Squares

We next introduce a diffusion based, distributed augment complex (widely linear) recursive least squares algorithm (ACRLS) suited to the generality of complex signals. Consi a network consisting of N nodes distributed over some ar where at every time instant n, each node $i \in [1, N]$ has acc to a desired signal $d_{i,n} \in \mathbb{C}$ and a regressor vector $\mathbf{x}_{i,n}$ $\mathbb{C}^{p\times 1}$ which are related as

$$d_{i,n} = \mathbf{g}^{oT}\mathbf{x}_{i,n} + \mathbf{h}^{oT}\mathbf{x}^*_{i,n} + v_{i,n}$$

where $v_{i,n} \in \mathbb{C}$ is a zero-mean white noise process. The a is to estimate the unknown widely linear weight vectors $\mathbf{g}^o$ $\mathbb{C}^{p\times 1}$ and $\mathbf{h}^o \in \mathbb{C}^{p\times 1}$ that minimise the weighted least squa cost function

$$\epsilon_{i,n} = \sum_{k=0}^{n} \lambda^{n-k}|e_{i,k}|^2$$

where $0 < \lambda \leq 1$ ia an exponential forgetting (weightir factor, and the error

$$e_{i,k} = d_{i,k} - y_{i,k} \tag{6}$$

is the difference between the desired signal $d_{i,k}$ and the widely linear filter output

$$y_{i,k} = \mathbf{g}^T_{i,n}\mathbf{x}_{i,k} + \mathbf{h}^T_{i,n}\mathbf{x}^*_{i,k} \tag{7}$$

The vectors $\mathbf{h}_{i,k}$ and $\mathbf{g}_{i,k}$ are the filter coefficients (weights) at node i. The filter output (7) can be expressed in a compact form using complex augmented vectors, that is

$$y_{i,k} = \underline{\mathbf{w}}^T_{i,n}\mathbf{x}^a_{i,k} \tag{8}$$

where $\underline{\mathbf{w}}_{i,n} = [\mathbf{g}^T_{i,n}, \mathbf{h}^T_{i,n}]^T$ and $\mathbf{x}_{i,k} = [\mathbf{x}^T_{i,k}, \mathbf{x}^H_{i,k}]^T$ are the augmented weight and regressor vectors respectively.

Observe that (8) has a similar form to the strictly linear model (1), and as such, the derivation of the D-ACRLS follows in the same manner as that used to derive the standard complex RLS (CRLS) algorithm, that is, the derivative of the cost function $\epsilon_{i,n}$ with respect to $\underline{\mathbf{w}}^*_{i,n}$ is set to zero, and after some matrix manipulations the resulting expression is the augmented least squares Weiner solution

$$\underline{\mathbf{w}}_{i,n} = \left(\hat{\mathbf{R}}^a_{i,n}\right)^{-1}\hat{\mathbf{r}}^a_{i,n} \tag{9}$$

where

$$\hat{\mathbf{R}}^a_{i,n} = \sum_{k=0}^{n}\lambda^{n-k}\mathbf{x}^a_{i,k}\mathbf{x}^{aH}_{i,k} = \sum_{k=0}^{n}\lambda^{n-k}\begin{bmatrix}\mathbf{x}_{i,k}\mathbf{x}^H_{i,k} & \mathbf{x}_{i,k}\mathbf{x}^T_{i,k}\\ \mathbf{x}^*_{i,k}\mathbf{x}^H_{i,k} & \mathbf{x}^*_{i,k}\mathbf{x}^T_{i,k}\end{bmatrix}$$

$$\hat{\mathbf{r}}^a_{i,n} = \sum_{k=0}^{n}\lambda^{n-k}d_{i,k}\mathbf{x}^{a*}_{i,k} = \sum_{k=0}^{n}\lambda^{n-k}\begin{bmatrix}d_{i,k}\mathbf{x}_{i,k}\\ d_{i,k}\mathbf{x}^*_{i,k}\end{bmatrix}$$

Within the RLS framework, the inverse of the input covariance matrix is required, and for the augmented complex case, $\hat{\mathbf{R}}^a_{i,n}$, it can be shown that

$$\left(\hat{\mathbf{R}}^a_{i,n}\right)^{-1} = \begin{bmatrix}\hat{\mathbf{R}}_{i,n} & \hat{\mathbf{P}}_{i,n}\\ \hat{\mathbf{P}}^*_{i,n} & \hat{\mathbf{R}}^*_{i,n}\end{bmatrix}^{-1} = \begin{bmatrix}\hat{\mathbf{C}}_{i,n} & \hat{\mathbf{D}}_{i,n}\\ \hat{\mathbf{C}}^*_{i,n} & \hat{\mathbf{D}}^*_{i,n}\end{bmatrix}$$

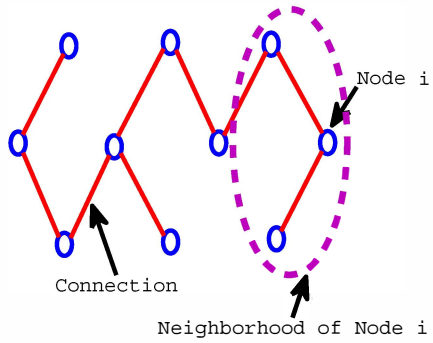


Fig. 1. A distributed network with $N = 10$ nodes.

Note that the inverse matrix $\left(\hat{\mathbf{R}}^a_{i,n}\right)^{-1}$ maintains the same block-conjugate structure as the augmented covariance matrix $\hat{\mathbf{R}}^a_{i,n}$. Continuing with the derivation, the ACRLS recursions for the node i take on the following form:

$$\mathbf{z}^a_{i,n} = \hat{\mathbf{R}}^a_{i,n-1}\mathbf{x}^{a*}_{i,n} \tag{10}$$

$$\mathbf{k}_{i,n} = \frac{\mathbf{z}^a_{i,n}}{\lambda + \mathbf{x}^{aT}_{i,n}\mathbf{z}^a_{i,n}} \tag{11}$$

$$\underline{\mathbf{w}}_{i,n} = \mathbf{w}_{i,n-1} + \mathbf{k}_{i,n}\left(d_{i,n} - \mathbf{w}^T_{i,n-1}\mathbf{x}^a_{i,n}\right) \tag{12}$$

$$\hat{\mathbf{R}}^a_{i,n} = \frac{1}{\lambda}\left(\hat{\mathbf{R}}^a_{i,n-1} - \mathbf{k}_{i,n}\mathbf{z}^{aH}_{i,n}\right) \tag{13}$$

After each node has estimated the weights for time n, the next step is the diffusion scheme used to enable collaboration between the nodes in the distributed network. Let $\mathcal{N}_i$ denote the neighbourhood of a node i, that is, the set of nodes connected with the node i (including itself) - see Figure 1. At time instant n, each node i forms an estimate $\underline{\mathbf{w}}_{i,n}$, using its desired signal $d_{i,n}$ and input vector $\mathbf{x}_{i,n}$, of the optimal weight vector $\mathbf{w}^o_{i,n} = [\mathbf{g}^{oT}_{i,k}, \mathbf{h}^{oT}_{i,k}]^T$, and utilises the following fusion (diffusion) scheme

$$\mathbf{w}_{i,n} = \sum_{k\in\mathcal{N}_i} c_{k,i}\underline{\mathbf{w}}_{k,n} \tag{14}$$

where $\mathbf{w}_{i,n}$ is the diffused estimate at node i computed as weighted average of the estimates from its neighbourhood $\mathcal{N}_i$, and $c_{k,i} \geq 0$ are the weighting coefficients satisfying $\sum_{k\in\mathcal{N}_i} c_{k,i} = 1$. Hence, the D-ACRLS algorithm requires each node to communicate their local estimates $\underline{\mathbf{w}}_{i,n}$ to their neighbours, which are then used to form the diffusion estimate $\mathbf{w}_{i,n}$. A number of fusion schemes exist, including the Laplacian and nearest neighbour method [2], however, these schemes are generally not optimal, as the determination of the optimal weighting coefficients for an arbitrary network is a difficult problem [9].

Observe that the D-ACRLS in (10)-(13), is based on the use of augmented vectors and covariance matrices which conform to block-conjugate structures[1], and as such, D-ACRLS is over-parameterised and has unnecessary excess computational

[1]The lower halves of augmented vectors and covariance matrices can be fully determined from their upper halves, or vice-versa.

Algorithm 1 D-ACRLS

Initialisation: For each node $i = 1, 2, \ldots, N$

$$\text{Forgetting factor: } 0 < \lambda \leq 1$$
$$\mathbf{g}_{i,0} = \mathbf{0}; \quad \mathbf{h}_{i,0} = \mathbf{0}$$
$$\hat{\mathbf{C}}_{i,0} = \delta \mathbf{I}; \quad \hat{\mathbf{D}}_{i,0} = \mathbf{0}; \quad \delta > 0$$

For every time instant $n = 1, 2, \ldots$

– Evaluate at each node $i = 1, 2, \ldots, N$

$$\mathbf{z}_{i,n} = \hat{\mathbf{C}}_{i,n-1}\mathbf{x}_{i,n}^* + \hat{\mathbf{D}}_{i,n-1}\mathbf{x}_{i,n} \tag{15}$$
$$\mathbf{k}_{i,n} = \frac{\mathbf{z}_{i,n}}{\lambda + 2\text{Re}\{\mathbf{x}_{i,n}^T \mathbf{z}_{i,n}\}} \tag{16}$$
$$e_{i,n} = d_{i,n} - \mathbf{g}_{i,n-1}^T \mathbf{x}_{i,n} - \mathbf{h}_{i,n-1}^T \mathbf{x}_{i,n}^* \tag{17}$$
$$\underline{\mathbf{g}}_{i,n} = \mathbf{g}_{i,n-1} + \mathbf{k}_{i,n} e_{i,n} \tag{18}$$
$$\underline{\mathbf{h}}_{i,n} = \mathbf{h}_{i,n-1} + \mathbf{k}_{i,n}^* e_{i,n} \tag{19}$$
$$\hat{\mathbf{C}}_{i,n} = \frac{1}{\lambda}\big(\hat{\mathbf{C}}_{i,n-1} - \mathbf{k}_{i,n}\mathbf{z}_{i,n}^H\big) \tag{20}$$
$$\hat{\mathbf{D}}_{i,n} = \frac{1}{\lambda}\big(\hat{\mathbf{D}}_{i,n-1} - \mathbf{k}_{i,n}\mathbf{z}_{i,n}^T\big) \tag{21}$$

– For every node i, compute the diffusion update

$$\mathbf{g}_{i,n} = \sum_{k \in \mathcal{N}_i} c_{k,i} \underline{\mathbf{g}}_{k,n}; \quad \mathbf{h}_{i,n} = \sum_{k \in \mathcal{N}_i} c_{k,i} \underline{\mathbf{h}}_{k,n} \tag{22}$$

complexity. However, after some algebraic manipulations, an alternative equivalent version of D-ACRLS with reduced complexity can be achieved, based on the work in [6], and is summarised in Algorithm 1. The D-ACRLS presented in Algorithm 1 requires approximately half the number multiplications as that required by (10)-(13).

III. Distributed Kalman Filter

Consider the standard linear state space corresponding to a node i in a distributed system [10],

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{w}_n \tag{23a}$$
$$\mathbf{y}_{i,n} = \mathbf{H}_{i,n}\mathbf{x}_n + \mathbf{v}_{i,n} \tag{23b}$$

where $\mathbf{x}_n \in \mathbb{C}^L$ and $\mathbf{y}_{i,n} \in \mathbb{C}^K$ are the state vector at time instant n and observation (measurement) vector at node i, respectively, while $\mathbf{F}_n$ and $\mathbf{H}_{i,n}$ are the state transition and observation matrices, whereas $\mathbf{w}_n \in \mathbb{C}^L$ and $\mathbf{v}_{i,n} \in \mathbb{C}^K$ are the white state and measurement noises at node i, respectively, and are assumed to be uncorrelated and zero-mean, with covariances and pseudocovariances defined as

$$E\begin{bmatrix}\mathbf{w}_n \\ \mathbf{v}_{i,n}\end{bmatrix}\begin{bmatrix}\mathbf{w}_k \\ \mathbf{v}_{i,k}\end{bmatrix}^H = \begin{bmatrix}\mathbf{Q}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{i,n}\end{bmatrix}\delta_{nk} \tag{24}$$
$$E\begin{bmatrix}\mathbf{w}_n \\ \mathbf{v}_{i,n}\end{bmatrix}\begin{bmatrix}\mathbf{w}_k \\ \mathbf{v}_{i,k}\end{bmatrix}^T = \begin{bmatrix}\mathbf{P}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{i,n}\end{bmatrix}\delta_{nk} \tag{25}$$

where δ_{nk} is the Kronecker delta function. Note that the nodal observation noises are not assumed to be uncorrelated, that is, $E\{\mathbf{v}_{i,n}\mathbf{v}_{k,n}^H\} = \mathbf{0}$ for $i \neq k$, which was used to derive the earlier distributed Kalman filtering algorithms in [2] [1] [8].

Following on from the result in [5], the widely linear counterpart to (23) is given by [5]

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{A}_{n-1}\mathbf{x}_{n-1}^* + \mathbf{w}_n \tag{26a}$$
$$\mathbf{y}_{i,n} = \mathbf{H}_{i,n}\mathbf{x}_n + \mathbf{B}_{i,n}\mathbf{x}_n^* + \mathbf{v}_{i,n} \tag{26b}$$

which leads to an augmented vector representation

$$\mathbf{x}_n^a = \mathbf{F}_{n-1}^a \mathbf{x}_{n-1}^a + \mathbf{w}_n^a \tag{27a}$$
$$\mathbf{y}_{i,n}^a = \mathbf{H}_{i,n}^a \mathbf{x}_n^a + \mathbf{v}_{i,n}^a \tag{27b}$$

where $\mathbf{x}_n^a = [\mathbf{x}_n^T, \mathbf{x}_n^H]^T$ and $\mathbf{y}_n^a = [\mathbf{y}_n^T, \mathbf{y}_n^H]^T$, while,

$$\mathbf{F}_n^a = \begin{bmatrix}\mathbf{F}_n & \mathbf{A}_n \\ \mathbf{A}_n^* & \mathbf{F}_n^*\end{bmatrix} \text{ and } \mathbf{H}_{i,n}^a = \begin{bmatrix}\mathbf{H}_{i,n} & \mathbf{B}_{i,n} \\ \mathbf{B}_{i,n}^* & \mathbf{H}_{i,n}^*\end{bmatrix}$$

Observe that for $\mathbf{A}_n = \mathbf{0}$ and $\mathbf{B}_{i,n} = \mathbf{0}$, the widely linear (augmented) state space model become strictly linear, however, the augmented state space representation should still be preferred in order to account for the pseudocovariance of the signals. The covariance matrices of $\mathbf{w}_n^a = [\mathbf{x}_n^T, \mathbf{w}_n^H]^T$ and $\mathbf{v}_{i,n}^a = [\mathbf{v}_{i,n}^T, \mathbf{v}_{i,n}^H]^T$ are given by

$$\mathbf{Q}_n^a = E\{\mathbf{w}_n^a \mathbf{w}_n^{aH}\} = \begin{bmatrix}\mathbf{Q}_n & \mathbf{P}_n \\ \mathbf{P}_n^* & \mathbf{Q}_n^*\end{bmatrix} \tag{28}$$
$$\mathbf{R}_{i,n}^a = E\{\mathbf{v}_{i,n}^a \mathbf{v}_{i,n}^{aH}\} = \begin{bmatrix}\mathbf{R}_{i,n} & \mathbf{U}_{i,n} \\ \mathbf{U}_{i,n}^* & \mathbf{R}_{i,n}^*\end{bmatrix} \tag{29}$$

To enable collaborative estimation of the state within distributed networks, we propose the use of neighbourhood observation equations, comprising of all the observation data from the neighbourhood of each node i, that is

$$\underline{\mathbf{y}}_{i,n} = \underline{\mathbf{H}}_{i,n}\mathbf{x}_n + \underline{\mathbf{B}}_{i,n}\mathbf{x}_n^* + \underline{\mathbf{v}}_{i,n} \tag{30}$$

with the collective (neighbourhood) variables defined as

$$\underline{\mathbf{y}}_{i,n} = \Big[\mathbf{y}_{i_1,n}^T, \mathbf{y}_{i_2,n}^T, \vdots, \mathbf{y}_{i_M,n}^T\Big]^T$$
$$\underline{\mathbf{H}}_{i,n} = \Big[\mathbf{H}_{i_1,n}^T, \mathbf{H}_{i_2,n}^T, \vdots, \mathbf{H}_{i_M,n}^T\Big]^T$$
$$\underline{\mathbf{B}}_{i,n} = \Big[\mathbf{B}_{i_1,n}^T, \mathbf{B}_{i_2,n}^T, \vdots, \mathbf{B}_{i_M,n}^T\Big]^T$$
$$\underline{\mathbf{v}}_{i,n} = \Big[\mathbf{v}_{i_1,n}^T, \mathbf{v}_{i_2,n}^T, \vdots, \mathbf{v}_{i_M,n}^T\Big]^T$$

where $\{i_1, i_2, \ldots, i_M\}$ are all the nodes in the neighbourhood $\mathcal{N}_i$. Compared with a centralised Kalman filter observation equation, which consists of all the observation data in the network, the neighbourhood observation equation in (30) essentially mimics it at a local level, using only the observation data within the neighbourhoods. The second order statistics of

$\underline{\mathbf{v}}_{i,n}$ are given as

$$\underline{\mathbf{R}}_{i,n} = E\{\underline{\mathbf{v}}_{i,n}\underline{\mathbf{v}}_{i,n}^H\} = \begin{bmatrix} \mathbf{R}_{i_1,n} & \mathbf{R}_{i_1 i_2,n} & \cdots & \mathbf{R}_{i_1 i_M,n} \\ \mathbf{R}_{i_2 i_1,n} & \mathbf{R}_{i_2,n} & \cdots & \mathbf{R}_{i_2 i_M,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{i_M i_1,n} & \mathbf{R}_{i_M i_2,n} & \cdots & \mathbf{R}_{i_M,n} \end{bmatrix}$$

$$\underline{\mathbf{U}}_{i,n} = E\{\underline{\mathbf{v}}_{i,n}\underline{\mathbf{v}}_{i,n}^T\} = \begin{bmatrix} \mathbf{U}_{i_1,n} & \mathbf{U}_{i_1 i_2,n} & \cdots & \mathbf{U}_{i_1 i_M,n} \\ \mathbf{U}_{i_2 i_1,n} & \mathbf{U}_{i_2,n} & \cdots & \mathbf{U}_{i_2 i_M,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{U}_{i_M i_1,n} & \mathbf{U}_{i_M i_2,n} & \cdots & \mathbf{U}_{i_M,n} \end{bmatrix}$$

where $\mathbf{R}_{i_a,n} = E\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_a,n}^H\}$, $\mathbf{R}_{i_a i_b,n} = E\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_b,n}^H\}$, $\mathbf{U}_{i_a,n} = E\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_a,n}^T\}$ and $\mathbf{U}_{i_a i_b,n} = E\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_b,n}^T\}$, for $a, b \in \{1, 2, \ldots, M\}$. The augmented neighbourhood observation equations can now be written as

$$\underline{\mathbf{y}}_{i,n}^a = \underline{\mathbf{H}}_{i,n}^a \mathbf{x}_n^a + \underline{\mathbf{v}}_{i,n}^a \tag{31}$$

where

$$\underline{\mathbf{y}}_{i,n}^a = \begin{bmatrix} \underline{\mathbf{y}}_{i,n} \\ \underline{\mathbf{y}}_{i,n}^* \end{bmatrix}, \quad \underline{\mathbf{H}}_{i,n}^a = \begin{bmatrix} \underline{\mathbf{H}}_{i,n} & \underline{\mathbf{B}}_{i,n} \\ \underline{\mathbf{B}}_{i,n}^* & \underline{\mathbf{H}}_{i,n}^* \end{bmatrix}, \quad \underline{\mathbf{v}}_{i,n}^a = \begin{bmatrix} \underline{\mathbf{v}}_{i,n} \\ \underline{\mathbf{v}}_{i,n}^* \end{bmatrix} \tag{32}$$

while the covariance of $\underline{\mathbf{v}}_{i,n}^a$ is defined as

$$\underline{\mathbf{R}}_{i,n}^a = E\{\underline{\mathbf{v}}_{i,n}^a \underline{\mathbf{v}}_{i,n}^{aH}\} = \begin{bmatrix} \underline{\mathbf{R}}_{i,n} & \underline{\mathbf{U}}_{i,n} \\ \underline{\mathbf{U}}_{i,n}^* & \underline{\mathbf{R}}_{i,n}^* \end{bmatrix} \tag{33}$$

and caters for the complete second order statistics of the observation noises within the neighbourhood $\mathcal{N}_i$, together with their cross-correlations.

Let $\widehat{\underline{\mathbf{x}}}_{i,n|n}^a$ be the augmented (widely linear) complex Kalman filter (ACKF) state estimate at time n, based on the neighbourhood augmented observation equation for node i, then the diffused state estimate $\widehat{\mathbf{x}}_{i,n|n}^a$ for the neighbourhood $\mathcal{N}_i$, similar to the D-ACRLS, is then given by

$$\widehat{\mathbf{x}}_{i,n|n}^a = \sum_{k \in \mathcal{N}_i} c_{k,i} \widehat{\underline{\mathbf{x}}}_{k,n|n}^a \tag{34}$$

The distributed ACKF (D-ACKF), based on widely linear neighbourhood equations, is summarised in Algorithm 2.

IV. Application Examples

We compared the performance D-ACKF with its strictly linear counterpart, the distributed conventional complex Kalman filter (D-CCKF), which caters for the cross-correlations between nodal observation noises, but does not account for the pseudocovariances of the signals. Consider the decentralised network in Figure 1 with the objective of filtering the complex 3rd order autoregressive process, AR(3), defined as

$$z_n = 0.92 z_{n-1} - 0.53 z_{n-2} + 0.24 z_{n-2} + u_n$$

where u_n is the noncircular complex white Gaussian driving noise with unit variance $E\{|u_n|^2\} = 1$. The observations corresponding to each node i was of the form

$$y_{i,n} = z_n + v_{i,n}$$

Algorithm 2 D-ACKF

Initialisation: For each node $i = 1, 2, \ldots, N$

$$\widehat{\mathbf{x}}_{i,0|0}^a = \left[E\{\mathbf{x}_0\}^T, E\{\mathbf{x}_0\}^H\right]^T$$
$$\mathbf{M}_{i,0|0}^a = E\left\{(\mathbf{x}_0^a - \widehat{\mathbf{x}}_{i,0|0}^a)(\mathbf{x}_0^a - \widehat{\mathbf{x}}_{i,0|0}^a)^{aH}\right\}$$

For every time instant $n = 1, 2, \ldots$

– Evaluate at each node $i = 1, 2, \ldots, N$

$$\widehat{\mathbf{x}}_{i,n|n-1}^a = \mathbf{F}_{n-1}^a \widehat{\mathbf{x}}_{i,n-1|n-1}^a \tag{35}$$
$$\mathbf{M}_{i,n|n-1}^a = \mathbf{F}_{n-1}^a \mathbf{M}_{i,n-1|n-1}^a \mathbf{F}_{n-1}^{aH} + \mathbf{Q}_n^a \tag{36}$$
$$\mathbf{G}_{i,n}^a = \mathbf{M}_{i,n|n-1}^a \underline{\mathbf{H}}_{i,n}^{aH} \left(\underline{\mathbf{H}}_{i,n}^a \mathbf{M}_{i,n|n-1}^a \underline{\mathbf{H}}_{i,n}^{aH} + \underline{\mathbf{R}}_{i,n}^a\right)^{-1} \tag{37}$$
$$\widehat{\underline{\mathbf{x}}}_{i,n|n}^a = \widehat{\mathbf{x}}_{i,n|n-1}^a + \mathbf{G}_{i,n}^a \left(\underline{\mathbf{y}}_{i,n}^a - \underline{\mathbf{H}}_{i,n}^a \widehat{\mathbf{x}}_{i,n|n-1}^a\right) \tag{38}$$
$$\mathbf{M}_{i,n|n}^a = (\mathbf{I} - \mathbf{G}_{i,n}^a \underline{\mathbf{H}}_{i,n}^a) \mathbf{M}_{i,n|n-1}^a \tag{39}$$

– For every node i, compute the diffusion update as

$$\widehat{\mathbf{x}}_{i,n|n}^a = \sum_{k \in \mathcal{N}_i} c_{k,i} \widehat{\underline{\mathbf{x}}}_{k,n|n}^a \tag{40}$$

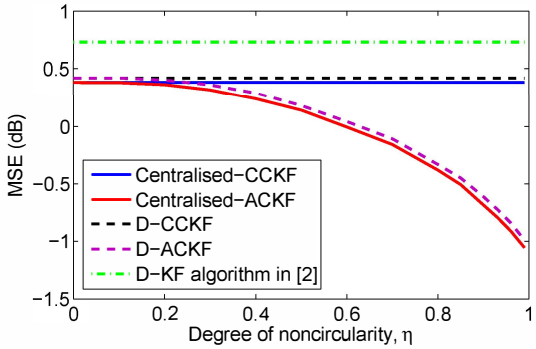
where $v_{i,n}$ is a noncircular complex white Gaussian observation noise associated with node i with the variance $R_{i,n} = E\{|v_{i,n}|^2\} = 3 + \sqrt{i}$ and cross-correlation $R_{ik,n} = E\{v_{i,n} v_{k,n}^*\} = 3$ for $i, k \in \{1, 2, \ldots, N\}$ and $i \neq k$.

We used the ratio of pseudocovariance magnitude to covariance, that is $\eta_u = |E\{u^2\}|/E\{|u|^2\}$, as a measure for the degree of noncircularity of a (zero-mean) complex signal $u = u_r + j u_i$, where a signal is circular for $\eta_u = 0$ and maximally noncircular for $\eta_u = 1$. The average mean square error (MSE) of all the nodes was used for a quantitative performances assessment of the algorithms. The nearest neighbour diffusion scheme was used, and is as follows [2]. Let $|\mathcal{N}_i|$ denote the number of a nodes in the neighbourhood $\mathcal{N}_i$, then to calculate the diffused state estimate for node i, the weight associated with a neighbour k is proportional to $|\mathcal{N}_k|$, that is

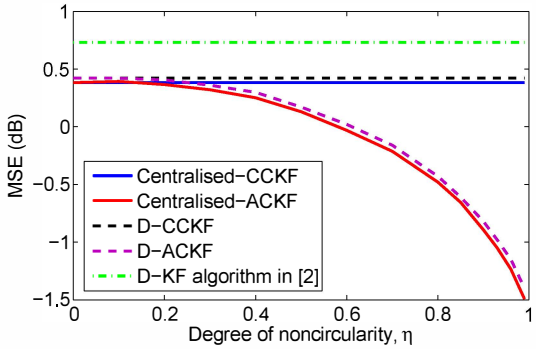
$$c_{k,i} = \begin{cases} |\mathcal{N}_k|/\alpha_i & \text{if } k \in \mathcal{N}_i \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha_i = \sum_{k \in \mathcal{N}_i} |\mathcal{N}_k|$ is a normalisation parameter which ensures that $\sum_{k \in \mathcal{N}_i} c_{k,i} = 1$.

Figure 2 compares the steady state performance of the diffusion Kalman filter in [2] (Algorithm 2), D-CCKF and D-ACKF algorithms, along with the centralised CCKF and ACKF (C-CCKF and C-ACKF), with access to all the observation data from all the nodes at each time instant. Figure 2a shows the results for circular observation noises ($U_{i,n} = E\{v_{i,n}^2\} = 0$ for $i = 1, 2, \ldots, N$) and a state (driving) noise with various degrees of noncircularity, whereas Figure 2b illustrates the performance for noncircular observation noises with a circular state noise ($P_n = 0$). For circular noises ($\eta_w = 0$ and $\eta_{v_i} = 0$), the strictly linear D-CCKF and widely linear D-ACKF algorithms have identical performances, while for noncircular noises the D-ACKF had decreasing MSE for increasing degree of noise noncircularity, whereas changes in the circularity of



(a) Noncircular state noise



(b) Noncircular observation noise

Fig. 2. Steady state performance comparison for the problem of denoising the AR(3) process in the cases of: (a) circular observation noises and a noncircular driving noise with varying degrees of noncircularity; (b) circular state noise and noncircular observation noises with varying degrees of noncircularity (all nodes have same degree of observation noise noncircularity).

the noises did not effect D-CCKF since it does not account for pseudocovariances. Similarly, C-CCKF only matched the performance of C-ACKF for circular signals.

The D-CCKF and D-ACKF outperformed the distributed Kalman filter in [2] (Algorithm 2), because they accounted for the cross-correlations between the nodal observation noises ($R_{ik,n} = E\{v_{i,n}v_{k,n}^*\}$). However, for uncorrelated nodal observation noises, and circular state and observation noises, the D-CCKF, D-ACKF and the diffusion Kalman filter in [2] will all have identical performances.

Next, we compared the strictly linear distributed complex RLS (D-CRLS) and the widely linear distributed augmented complex RLS (D-ACRLS) for the problem of predicting real-world Wind data[2] (recorded in Oregon USA, 2011), using an AR(4) model for a network consisting for two sensors.

The D-ACRLS was able to capture the underlying dynamics of the data better than D-CRLS, as indicated by its enhanced prediction performance. This can be attributed to the use of the widely linear model, which is better suited to capturing the noncircular nature of the Wind data.

[2]The Wind signal (s_n) has a magnitude (intensity) (ν_n) and direction (ϕ_n), and is naturally represented as complex signal ($s_n = \nu_n e^{\phi_n}$).

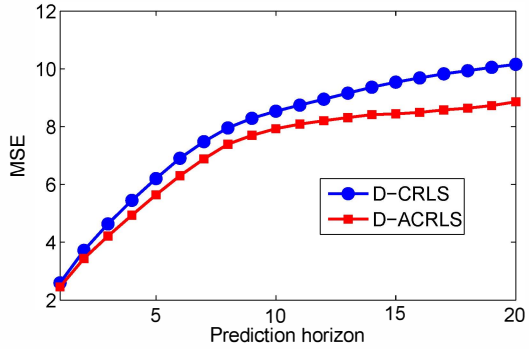


Fig. 3. Multistep ahead prediction of real-world Wind data.

V. Conclusion

We have addressed the problem of adaptive estimation of complex signals with both circular and noncircular distributions within distributed collaborative networks. The distributed (widely linear) augmented complex Kalman filter (D-ACKF) and recursive least squares (D-ACRLS) algorithms have been introduced for the sequential estimation of the generality of complex signals. Moreover, unlike existing algorithms, the D-ACKF proposed here is based on a framework which caters for correlations between the observation noises experienced by different nodes in the distributed network, thus providing enhanced performance in real-world scenarios. The analysis is supported by simulations on synthetic and real world data.

Acknowledgment

This work is part of the University Defence Research Centre (UDRC) at Imperial College London, supported by the MoD and DSTL.

References

[1] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *46th IEEE Conference on Decision and Control*, pp. 5492 –5498, Dec. 2007.

[2] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, pp. 2069 –2084, Sept. 2010.

[3] Z. Gao, H.-Q. Lai, and K. J. R. Liu, "Differential space-time network coding for multi-source cooperative communications," *IEEE Transactions on Communications*, vol. 59, pp. 3146 –3157, Nov. 2011.

[4] B. Picinbono and P. Bondon, "Second-order Statistics of Complex Signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 411–420, 1997.

[5] D. H. Dini and D. P. Mandic, "A class of widely linear complex Kalman filters," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, pp. 775 –786, May 2012.

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

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

[8] S. Kar and J. M. F. Moura, "Gossip and distributed Kalman filtering: Weak consensus under weak detectability," *IEEE Transactions on Signal Processing*, vol. 59, pp. 1766 –1784, April 2011.

[9] X.R. Li, Y. Zhu, J. Wang, and C. Han, "Optimal linear estimation fusion : Part I: Unified fusion rules," *IEEE Transactions on Information Theory*, vol. 49, pp. 2192 – 2208, Sep. 2003.

[10] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, 1996.