AN ADAPTIVE DIFFUSION QUATERNION LMS ALGORITHM FOR DISTRIBUTED NETWORKS OF 3D AND 4D VECTOR SENSORS

Cyrus Jahanchahi and Danilo P. Mandic

Communications and Signal Processing Research Group
Imperial College London, London
Email: {cyrus.jahanchahi, d.mandic}@imperial.ac.uk

ABSTRACT

A diffusion widely linear quaternion least mean square (D-WLIQLMS) algorithm for the collaborative processing of quaternion signals over distributed networks is proposed. We show that the underlying quaternion division algebra and the widely linear model allow for a unified processing of 3D and 4D data, which can exhibit both circular and noncircular distributions. The analysis shows that the D-WLIQLMS provides a solution that is robust to link and node failures in sensor networks. Simulations on benchmark 4D signals illustrate the advantages offered by the D-WLIQLMS.

***Index Terms*—** Quaternions, widely linear models, distributed networks, diffusion algorithms.

1. INTRODUCTION

Advances in sensor technology and wireless communication have led to a growing use of sensor networks in practical applications, ranging from environmental monitoring to target localization. These networks typically comprise a number of interconnected nodes which are able to communicate with one another in order to estimate some parameter of interest from noisy measurements. Within distributed solutions to estimation problems, every node in the network communicates only with a subset of the nodes, while processing is distributed among all nodes in the network. This is opposed to a global solution where every node in the network transmits data to a central fusion center, where the processing is performed using all available information. The distributed solution, though usually marginally underperforming compared to the global solution, requires less power for communication between the sensors and is more robust to faults in the networks or fusion center [1] [2] [3] [4] . These properties are very desirable given the growing use of low power wireless sensors.

Distributed least mean square algorithms (also called diffusion LMS) have recently been proposed [1] for real-valued data. This was soon followed by the development of diffusion RLS [5] and diffusion Kalman algorithms [6] [2]. With the development in augmented complex statistics, distributed algorithms have recently been extended to the complex domain [7] allowing cooperative processing of two-dimensional data using all available second order statistics.

Albeit better equipped to deal with multidimensional data than real valued models, complex-valued signal processing techniques are not adequate where the data is three- or four-dimensional. Furthermore, a growing proportion of sensors used in applications give three-dimensional outputs, for example three-dimensional wind anemometers or inertial motion sensors. For such signals, quaternions (ordered pair of complex numbers) can be used to obtain elegant and compact solutions. Owing to their division algebra, quaternions too have the advantage of providing accurate and efficient models for three-dimensional rotation and orientation tracking. In these cases, real vectors are inadequate and may result in a loss of one degree of freedom (gimbal lock).

In the complex domain, augmented statistics is needed to capture all second order statistics and develop algorithms that are second order optimal. In the same vain, the development of the quaternion widely linear model, augmented statistics [8] [9] [10] and $\mathbb{HR}$-calculus [11] has recently received plenty of attention. Numerous quaternion adaptive filtering algorithms have been developed in the quaternion domain, including the widely linear quaternion least mean square (WLQLMS) [12] and widely linear quaternion affine projection [13]. More recently, using involution gradients, an efficient implementation of the WLQLMS (WLIQLMS) has been introduced [14], requiring half the operations of the QLMS. In this paper, the IQLMS will be used as a platform for deriving the diffusion IQLMS (D-IQLMS) and diffusion WLIQLMS (D-WLIQLMS), allowing cooperative adaptive estimation of both circular and noncircular quaternion valued signals. The advantage of the D-WLIQLMS over the strictly linear D-IQLMS is illustrated over simulations on benchmark signals.

We next review the quaternion algebra (Section 2) and quaternion widely linear model (Section 3). In Section 4 we derive the D-IQLMS, while in Section 5 we employ the widely linear model to obtain the D-WLIQLMS. Section 6 presents stability bounds, thus ensuring convergence in the mean. We conclude the work with representative simulations.

2. QUATERNION ALGEBRA

Quaternions are an extension of complex numbers (forming an ordered pair), comprising of a real part (denoted by a subscript a) and three imaginary parts (denoted by subscripts b, c and d). A quaternion variable $q \in \mathbb{H}$ can be described as

$$q = q_r + iq_\imath + jq_\jmath + kq_\kappa \qquad (1)$$

The unit axis vectors i, j and k in the quaternion domain $\mathbb{H}$ are also imaginary units, and obey the following rules

$$ij = k \quad jk = i \quad ki = j$$

$$i^2 = j^2 = k^2 = ijk = -1$$

Note that quaternion multiplication is not commutative, that is, $ij = k \neq ji = -k$.

A quaternion variable q can be conveniently written in a real-vector form as [15]

$$q = Sq + Vq$$

where, $Sq = q_r$ (denotes the scalar part of q) and $Vq = iq_\imath + jq_\jmath + kq_\kappa$ (denotes the vector part of q). Then, the quaternion product can be expressed as:

$$\begin{aligned} q_1q_2 &= (Sq_1 + Vq_1)(Sq_2 + Vq_2) \\ &= Sq_1Sq_2 - Vq_1 \bullet Vq_2 + Sq_2Vq_1 + Sq_1Vq_2 \\ &\quad + Vq_1 \times Vq_2 \end{aligned}$$

where the symbol '$\bullet$' denotes the dot-product and '$\times$' the usual cross-product in vector analysis. The quaternion conjugate, denoted by q^* is given by

$$q^* = Sq - Vq$$

The norm $\|q\|$ of a quaternion variable q, is defined as

$$\|q\| = \sqrt{qq^*} = \sqrt{q_r^2 + q_\imath^2 + q_\jmath^2 + q_\kappa^2}$$

The three-dimensional vector part Vq is also called a pure quaternion, whereas the inclusion of the real part Sq gives a full quaternion. The special algebraic structure of quaternions enables unified processing of three- and four-dimensional multivariate processes. Therefore the distributed algorithm developed here applies to both 3D and 4D data.

2.1. Quaternion Involutions

Involutions are self-inverse mappings and are defined as[1]

$$q^i = -iqi = q_r + \imath q_\imath - \jmath q_c \jmath - \kappa q_\kappa \qquad (2)$$

$$q^j = -jqj = q_r - \imath q_\imath + \jmath q_c \jmath - \kappa q_\kappa \qquad (3)$$

$$q^k = -kqk = q_r - \imath q_\imath - \jmath q_c \jmath + \kappa q_\kappa \qquad (4)$$

To verify that involutions represent self-inverse mappings, consider for instance $(q^i)^i = q$. The involution of a product is also a product of the individual involutions (i.e. $(q_1q_2)^i = q_1^iq_2^i$). It is important to realize that involutions can be seen as a quaternion counterpart of the complex conjugate, as they allow the components of a quaternion variable to be expressed in terms of the actual variable and its involutions, that is

$$q_r = \frac{1}{4}[q + q^i + q^j + q^k] \qquad q_\imath = \frac{1}{4i}[q + q^i - q^j - q^k]$$

$$q_\jmath = \frac{1}{4j}[q - q^i + q^j - q^k] \qquad q_\kappa = \frac{1}{4k}[q - q^i - q^j + q^k]$$

The above representation is a basis for the derivation of quaternion valued widely-linear adaptive filtering models.

3. QUATERNION WIDELY LINEAR MODEL

The existing (strictly linear) estimation model in the quaternion domain is given by

$$\hat{y} = \mathbf{w}^T\mathbf{x} \qquad (5)$$

Observe from (2)-(4), and the above expressions for $q_a, \ldots, q_d$ that for all the quaternion components[2]

$$\hat{y}_\eta = E[y_\eta \,|\, \mathbf{x}_r, \mathbf{x}_\imath, \mathbf{x}_\jmath, \mathbf{x}_\kappa] \qquad \eta \in \{r, \imath, \jmath, \kappa\}$$

We can therefore express the components y_r, $y_\imath$, $y_\jmath$, y_κ of a quaternion via its involutions, to yield

$$\hat{y_\eta} = E[y_\eta \,|\, \mathbf{x}, \mathbf{x}^\imath, \mathbf{x}^\jmath, \mathbf{x}^\kappa] \text{ and } \hat{y} = E[y \,|\, \mathbf{x}, \mathbf{x}^\imath, \mathbf{x}^\jmath, \mathbf{x}^\kappa]$$

In other words, since every quaternion component is a function of its involutions, to capture the full second order information available we can employ the *widely linear* model

$$\hat{y} = \mathbf{u}^T\mathbf{x} + \mathbf{v}^T\mathbf{x}^\imath + \mathbf{g}^T\mathbf{x}^\jmath + \mathbf{h}^T\mathbf{x}^\kappa = \mathbf{w}^{aT}\mathbf{x}^a \qquad (6)$$

where the augmented coefficient vector $\mathbf{w}^a = [\mathbf{u}^T, \mathbf{v}^T, \mathbf{g}^T, \mathbf{h}^T]^T$ and the augmented regressor vector $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^{\imath T}, \mathbf{x}^{\jmath T}, \mathbf{x}^{\kappa T}]^T$, for more detail see [8].

Current statistical signal processing in $\mathbb{H}$ is largely based on strictly linear models, drawing upon the covariance matrix $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^H]$. However, to model both the second order circular (proper) and second order noncircular (improper) signals, based on the widely linear model in (6) we need to employ the augmented covariance matrix, given by [8]

$$\mathbf{R}_x^a = E[\mathbf{x}^a\mathbf{x}^{aH}] = \begin{bmatrix} \mathbf{R}_x & \mathbf{P}_x & \mathbf{S}_x & \mathbf{T}_x \\ \mathbf{P}_x^\imath & \mathbf{R}_x^\imath & \mathbf{T}_x^\imath & \mathbf{S}_x^\imath \\ \mathbf{S}_x^\jmath & \mathbf{T}_x^\jmath & \mathbf{R}_x^\jmath & \mathbf{P}_x^\jmath \\ \mathbf{T}_x^\kappa & \mathbf{S}_x^\kappa & \mathbf{P}_x^\kappa & \mathbf{R}_x^\kappa \end{bmatrix} \qquad (7)$$

where $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^H]$, $\mathbf{P}_x = E[\mathbf{x}\mathbf{x}^{\imath H}]$, $\mathbf{S}_x = E[\mathbf{x}\mathbf{x}^{\jmath H}]$ and $\mathbf{T}_x = E[\mathbf{x}\mathbf{x}^{\kappa H}]$.

For proper (second order circular) signals, all the pseudo-covariance matrices $\mathbf{P}_x$, $\mathbf{S}_x$ and $\mathbf{T}_x$ vanish, and such signals have probability distributions that are rotation invariant with respect to all the six possible pairs of axes (combinations of $\imath$, $\jmath$ and κ) [8], and thus equal powers in all the components.

Remark#1: The processing in $\mathbb{R}^4$ requires ten covariance matrices, as opposed to four in the quaternion domain (since only $\mathbf{R}_x$, $\mathbf{P}_x$, $\mathbf{S}_x$ and $\mathbf{T}_x$ are needed to fully describe $\mathbf{R}_x^a$).

[1] Note that the quaternion conjugate is also an involution.

[2] Throughout this paper, a vector $\mathbf{x}$ and its involutions are treated formally as independent variables. This is a usual formalism inherited from the complex domain, and in the $\mathbb{CR}$-calculus.

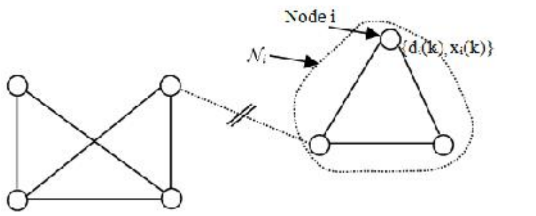

Fig. 1. Topology of a distributed network, highlighting the neighbourhood of node i.

4. DERIVATION OF D-IQLMS

Consider a network of N nodes distributed over some area, as shown in Figure 1. At time instant k, each node i has access to the local teaching signal $d_i(k)$ and a regressor input vector, defined as $\mathbf{x}_i(k) = [x_i(k-1), ..., x_i(k-L)]^T$. In addition to each node i having access to their own data ($d_i(k)$ and $\mathbf{x}_i(k)$ and $w(k)$), nodes can also access to data from their direct neighbours, where the neighbourhood $\mathcal{N}_i$ is defined as the set of all nodes linking to node i, including itself. A diffusion protocol describes how data from the neighbourhood $\mathcal{N}_i$ is combined. For a global (where all data is available to all nodes) least squares solution to the network in Figure 1, the centralized IQLMS (c-IQLMS) weight update can be written as [1]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \sum_{i=1}^{N} e_i(k)\mathbf{x}_i^*(k) \tag{8}$$

where the $e_i(k) = d_i(k) - y_i(k)$ is the error between the teaching signal $d_i(k)$ and output signal $y_i(k)$ at node i. In a distributed system, only data from the neighbourhood of a node is available and so the global weight update in (8) is replaced by local weight updates

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu \sum_{l\in\mathcal{N}_i} c_{l,i} e_l(k)\mathbf{x}_l^*(k) \tag{9}$$

where the weighting coefficient $c_{l,i}$ is always greater than 0, being equal to 0 when nodes l and k are not connected, and satisfies $\sum_{l\in\mathcal{N}_i} c_{l,i} = 1$. The weighting coefficient $c_{l,i}$ can be conveniently described as a matrix $\mathbf{C}$ where $\mathbb{1}^T\mathbf{C} = \mathbb{1}^T$ and $\mathbf{C}\mathbb{1} = \mathbb{1}$. The weight update at node i can be interpreted as a weighted linear combination of neighbouring estimates of the local cost function gradient. The advantage of this local weight update over the global weight update is that it is robust to node and link failures. The local weight estimates $w_i(k)$ in (9) can be further improved by combining the local weight updates iteratively as follows

$$\begin{aligned}\boldsymbol{\phi}_i(k+1) &= \mathbf{w}_i(k-1) + \mu \sum_{l\in\mathcal{N}_i} c_{l,i} e_l(k)\mathbf{x}_l^*(k) \\ \mathbf{w}_i(k) &= \sum_{l\in\mathcal{N}_i} a_l(k)\boldsymbol{\phi}_l(k) \end{aligned}\tag{10}$$

where $\sum_{l\in\mathcal{N}_i} a_l = 1$ (more conveniently written as$\mathbf{A}\mathbb{1} = \mathbb{1}$).

5. THE D-WLIQLMS

The D-IQLMS in (10) is only adequate when the underlying model is strictly linear. To process a signal generated from a widely linear system, or from noncircular statistics, we employ the widely linear model in (6) to augment the input vector. The D-WLIQLMS can then be compactly written as

$$\boldsymbol{\phi}_i^a(k+1) = \mathbf{w}_i^a(k-1) + \mu \sum_{l\in\mathcal{N}_i} c_{l,i}\mathbf{x}_l^{a*}(k)e_l(k) \tag{11}$$

$$\mathbf{w}_i^a(k) = \sum_{l\in\mathcal{N}_i} a_l(k)\boldsymbol{\phi}_l^a(k) \tag{12}$$

where $\mathbf{x}_l^a = [\mathbf{x}_l^{\imath T}, \mathbf{x}_l^{\imath T}, \mathbf{x}_l^{\jmath T}, \mathbf{x}_l^{\kappa T}]^T$.

6. STABILITY ANALYSIS OF D-WLIQLMS

To investigate the convergence in the mean of the D-WLIQLMS algorithm, we rewrite the weight update in (12) as

$$\boldsymbol{\phi}^{aT}(k) = \mathbf{w}^{aT}(k-1) + \left(\mathbf{d}^T - \mathbf{w}^{aT}(k-1)\mathbf{X}^a(k)\right)\mathbf{X}^{aH}\mathbf{MD} \tag{13}$$

where

$$\begin{aligned}\boldsymbol{\phi}^a(k) &= [\boldsymbol{\phi}_1^{aT}(k), \ldots, \boldsymbol{\phi}_N^{aT}(k)] \\ \mathbf{w}^a(k) &= [\mathbf{w}_1^{aT}(k), \ldots, \mathbf{w}_N^{aT}(k)]^T \\ \mathbf{X}^a(k) &= diag\{[\mathbf{x}_1^{aT}(k), \ldots, \mathbf{x}_N^{aT}(k)]^T\} \\ \mathbf{d}(k) &= [d_1(k), \ldots, d_N(k)]^T \\ \mathbf{D} &= diag\{[\mu_1\mathbf{I}_{4L}, \ldots, \mu_N\mathbf{I}_{4L}]\} \\ \mathbf{M} &= \mathbf{C}\otimes\mathbf{I}_{4L}\end{aligned}$$

and $\otimes$ denotes the Kronecker product operator. To evaluate the performance of the D-WLIQLMS, without loss in generality, consider an improper teaching signal $d(k)$, given by

$$d(k) = \mathbf{w}_o^{aT}(k)\mathbf{X}^a(k) + \mathbf{n}(k) \tag{14}$$

where $\mathbf{w}_o^{aT}(k)$ is the optimal (but unknown) local augmented weight vector, and $\mathbf{n}(k)$ is quadruply white Gaussian noise with zero mean and variance σ_n^2, uncorrelated with $\mathbf{X}(k)$. Substituting the teaching (14) into the weight update (13) and multiplying through we obtain

$$\begin{aligned}\boldsymbol{\phi}^{aT}(k) = \mathbf{w}^{aT}(k-1) &+ \big(\mathbf{w}_o^{aT}(k) \\ &- \mathbf{w}^{aT}(k-1)\big)\mathbf{X}^a(k)\mathbf{X}^{aH}\mathbf{MD} + \mathbf{n}(k)\mathbf{X}^{aH}\mathbf{MD}\end{aligned}$$

while a further multiplication of both sides by -1 and addition of $\mathbf{w}_o^{aT}(k)$ yields

$$\begin{aligned}\widetilde{\boldsymbol{\phi}}^{aT}(k) &= \widetilde{\mathbf{w}}^{aT}(k-1) - \widetilde{\mathbf{w}}^{aT}(k-1)\left(\mathbf{X}^a(k)\mathbf{X}^{aH}\mathbf{MD}\right. \\ &+ \left.\mathbf{n}(k)\mathbf{X}^{aH}\mathbf{MD}\right)\end{aligned}$$

where the weight error vectors $\widetilde{\boldsymbol{\phi}} = \mathbf{w}_o(k) - \boldsymbol{\phi}(k)$ and $\widetilde{\mathbf{w}} = \mathbf{w}_o(k) - \mathbf{w}(k)$. Substituting for $\widetilde{\boldsymbol{\phi}}$ using (12) we have

$$\begin{aligned}\widetilde{\mathbf{w}}^{aT}(k-1) = \widetilde{\mathbf{w}}^{aT}(k-1)\mathbf{P} - \widetilde{\mathbf{w}}^{aT}(k-1)&\left(\mathbf{X}^a(k)\mathbf{X}^{aH}\mathbf{MD}\right. \\ &+ \left.\mathbf{n}(k)\mathbf{X}^{aH}\mathbf{MD}\right)\mathbf{P}\end{aligned}$$

where $\mathbf{P} = \mathbf{A} \otimes \mathbf{I}_{4L}$. Upon applying the statistical expectation operator we have

$$E[\widetilde{\mathbf{w}}^{aT}(k-1)] \quad = \quad E[\widetilde{\mathbf{w}}^{aT}(k-1)]\,(\mathbf{I} - \mathbf{R}^a_{xx}\mathbf{MD})\,\mathbf{P}$$

where

$$\mathbf{R}^a_{xx} = E[\mathbf{X}^a(k)\mathbf{X}^{aH}] = diag\{E[\mathbf{x}^a_1\mathbf{x}^{aH}_1], \ldots, E[\mathbf{x}^a_N\mathbf{x}^{aH}_N]\}$$

It then follows from (15) that to ensure that the power of the weight error vector $E[\widetilde{\mathbf{w}}^{aT}(k-1)]$ converges to 0 as $k \to \infty$, the following condition must hold

$$||\,(\mathbf{I} - \mathbf{R}^a_{xx}\mathbf{MD})\,\mathbf{P}||_2 < 1$$

Since the 2-norm is sub-multiplicative and $||P|| = 1$ (this results from the fact that the rows and columns add to 1) we can write

$$||\,(\mathbf{I} - \mathbf{R}^a_{xx}\mathbf{MD})\,\mathbf{P}||_2 < ||\,(\mathbf{I} - \mathbf{R}^a_{xx}\mathbf{MD})\,||_2$$

and obtain a more conservative condition for convergence in the form

$$||\,(\mathbf{I} - \mathbf{R}^a_{xx}\mathbf{MD})\,||_2 < 1$$

For the case where all nodes have the same step size, i.e. matrix $\mathbf{D}$ is a scaled identity matrix, the above convergence condition becomes $||\{(\mathbf{I} - \mu\mathbf{R}^a_{xx}\mathbf{M})\}||_2 < 1$, yielding the following bound for the step size

$$\mu < \frac{2}{\lambda_{max}(\mathbf{R}^a_{xx}\mathbf{M})} \qquad (15)$$

7. SIMULATIONS

The one step ahead prediction performances of the proposed D-WLIQLMS and D-IQLMS were compared to the global IQLMS (where a centralized processing unit has access to all nodes) and to the non-cooperative IQLMS, where nodes do not share information with neighbouring nodes (i.e. where $\mathbf{M} = \mathbf{I}$). The network topology consisted of 6 nodes and is shown in Figure. Numerous methods exist to select the combination coefficients $c_{l,i}$, such as the Metropolis [1], Laplacian and nearest neighbour rule. For our simulations the Metropolis rule was used, described by

$$c_{l,i} = \begin{cases} \frac{1}{max(n_i,n_l)} & \text{if } l \neq i \text{ are connected} \\ 0 & \text{if } l \neq i \text{ are not connected} \\ 1 - \sum_{l \in N_i} c_{l,i} & \text{if } l = i \end{cases} \qquad (16)$$

where n_i and n_l are respectively the numbers of nodes in the neighbourhood of nodes i and l. Following the Metropolis rule in (16), the combination matrix $\mathbf{C}$ for the network topology in Fig. 2 becomes

$$\mathbf{C} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

The simulations were carried out for two signals, one generated by a strictly linear AR(4) model and the other generated by the three dimensional Lorenz system.

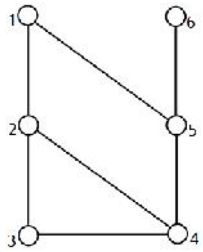

Fig. 2. Network topology.

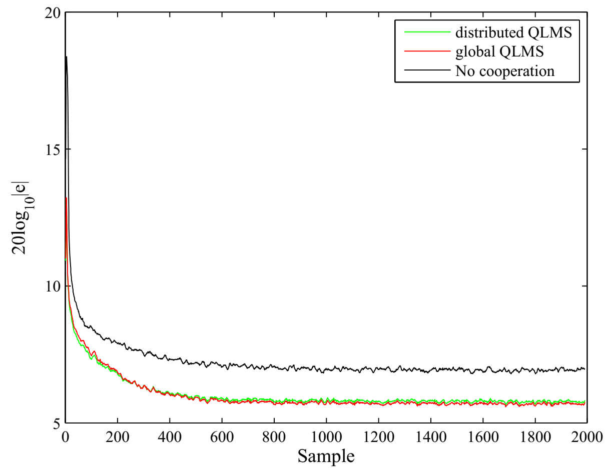

Fig. 3. Comparison of MSEs of the non-cooperative IQLMS, global IQLMS and D-IQLMS on the circular AR(4) process.

7.1. A circular AR(4) process

The benchmark circular quaternion signal used in simulations was generated by an AR(4) model described by

$$\begin{aligned} y(k) \quad = \quad & 1.79y(k-1) - 1.85y(k-2) + 1.27y(k-3) \\ & -0.41y(k-4) + n(k) \end{aligned}$$

where the driving noise $n(k)$ was quaternion valued circular white Gaussian noise. Figure 3 shows the evolution of the mean square error (MSE) for the one step ahead prediction of the AR(4) process when the filter length is 4. Observe that as expected the distributed QLMS achieves significantly better steady state performance compared to the case where there is no cooperation between the nodes. It is also worth noting that the steady state performance of the distributed QLMS is only marginally worse than that of the global QLMS, while being more robust to failures in the network and requiring fewer mathematical operations.

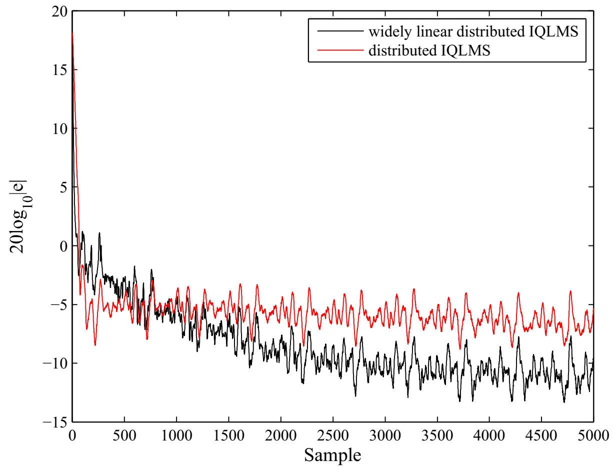

Fig. 4. Comparison of MSEs of the D-WLIQLMS and DIQLMS on the noncircular Lorenz signal.

7.2. Noncircular Lorenz attractor

The Lorenz signal, originally used to model atmospheric turbulence, is the output of the three-dimensional nonlinear system

$$\frac{\partial x}{\partial t} = \alpha(y - x) \ , \ \frac{\partial y}{\partial t} = x(\rho - z) - y \ , \ \frac{\partial z}{\partial t} = xy - \beta z$$

where $\alpha, \beta, \rho > 0$. Figure 4 shows the evolution of the MSE for the D-IQLMS and D-WLIQLMS. Observe that due to the widely linear nature of the Lorenz signal, the D-WLIQLMS has significantly lower MSE at steady state.

8. CONCLUSION

We have introduced a diffusion WLQLMS (D-WLQLMS) algorithm for the distributed processing of general quaternion valued signals (both circular and noncircular) in a cooperative fashion. The advantage of the diffusion topology over non-cooperative and fully-cooperative (global-QLMS) networks in terms of robustness, convergence speed and steady state performance has been shown in analysis and simulations on synthetic signals. The proposed strategy is particularly useful for data coming from vector sensors, such as 3D anemometers and 3D inertial body sensors.

9. REFERENCES

[1] F. Cattivelli and A. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1035–1048, 2010.

[2] F. Cattivelli and A. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, 2010.

[3] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.

[4] S. Kar, J. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3575–3605, 2012.

[5] F. Cattivelli, C. Lopes, and A. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1865–1877, 2008.

[6] B. Rao and H. Durrant-Whyte, "Fully decentralised algorithm for multisensor kalman filtering," *IEE Proceedings on Control Theory and Applications*, vol. 138, no. 5, pp. 413–420, 1991.

[7] Y. Xia, D. P. Mandic, and A. Sayed, "An adaptive diffusion augmented clms algorithm for distributed filtering of noncircular complex signals," *IEEE Signal Processing Letters*, vol. 18, no. 11, pp. 659–662, 2011.

[8] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Processing*, vol. 91, no. 2, pp. 214–224, 2011.

[9] J. Via, D. Ramirez, and I. Santamaria, "Properness and widely linear processing of quaternion random vectors," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502 –3515, 2010.

[10] J. Via, D. Palomar, L. Vielva, and I. Santamaria, "Quaternion ICA from second-order statistics," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1586 –1600, 2011.

[11] D. P. Mandic, C. Jahanchahi, and C. Cheong Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47 –50, 2011.

[12] C. C. Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4427–4431, 2010.

[13] C. Jahanchahi, C. Took, and D. Mandic, "A class of quaternion valued affine projection algorithms," *Signal Processing*, vol. 93, no. 7, pp. 1712 – 1723, 2013.

[14] C. Took, C. Jahanchahi, and D. Mandic, "A unifying framework for the analysis of quaternion valued adaptive filters," *Forty Fifth Asilomar Conference on Signals, Systems and Computers*, pp. 1771–1774, 2011.

[15] J. Ward, *Quaternions and Cayley Numbers*. Kluwer Academic Publishers, 1997.