EUSIPCO 2013 1569746681

DISTRIBUTED INCREMENTAL-BASED RLS FOR NODE-SPECIFIC PARAMETER ESTIMATION OVER ADAPTIVE NETWORKS

Jorge Plata-Chaves, Nikola Bogdanović, Kostas Berberidis

Department of Computer Engineering and Informatics
University of Patras & C.T.I RU-8, 26500, Rio - Patra, Greece
E-mails: {jorge, bogdanovic, berberid}@ceid.upatras.gr

ABSTRACT

In this paper, we introduce an adaptive distributed technique that attains the exact (Recursive Least Squares) RLS solution of a node-specific parameter estimation problem where each node is interested in a set of parameters of local interest and a set of global parameters. To do so, each node of the network relies on its own local data as well as the communication with its intermediate neighbor under an incremental mode of cooperation. Since the required inter-node communication of the new scheme may be prohibitive for networks with scarce energy resources, an alternative low-cost scheme is derived to reduce the communication burden. It is shown that this approximate strategy may attain the exact RLS solution in the steady state. To illustrate the effectiveness of the proposed techniques we provide some indicative simulation results.

Index Terms— Adaptive distributed networks, incremental algorithm, node-specific parameter estimation.

1. INTRODUCTION

To enable low-complex distributed implementations of the estimation task, there have been considerable research efforts on optimization techniques that generally rely on some separable structure of a cost function defining a linear estimator. Based on them, eliminating the need to embed powerful processors at the nodes, different algorithms based on Least Mean Squares (LMS) have been developed to estimate a set of parameters of global interest to all nodes of the network [1]-[3]. However, available processors continuously decrease in cost and have more powerful batteries as well as computational capabilities. To provide better estimation performance at the expense of increased computational complexity and energy consumption, RLS-type algorithms have also been derived under the approaches considered in the development of the aforementioned LMS-type strategies [4]-[6].

Up to now, in many of the distributed estimation problems, it is considered that the nodes are interested in estimating the same vector of global parameters. This scenario can be viewed as a special case of a more general problem where the nodes of the network have overlapped but different estimation interests. To motivate this node-specific setting, let us consider an example related to environmental monitoring. Specifically, consider a set of applications related to localization of diffusive sources which may include the temperature field, toxic chemicals in air and water, pollutants etc. Now, let us assume multiple diffusive sources affecting an area covered by N nodes (see Fig.1). Each node k is located at a known position and can measure the substance concentration:

$$y_k(r_{k_g}, r_{k_l}, t) = c_{k_g}(r_{k_g}, t) + c_{k_l}(r_{k_l}, t) + e(t)$$

where $e(t)$ denotes measurement noise, c_{k_g} represents the concentration of the global source at node k, while c_{k_l} denotes the substance concentration that originates from a local source affecting only the area of node k. By leveraging the dependence of c_{k_g} and c_{k_l} on the distances between a node and the aforementioned sources, denoted with r_{k_g} and r_{k_l}, the estimation task of each node would be the localization of its neighboring (local) source and of the global source.

Despite the fact that Node-Specific Parameter Estimation (NSPE) problems appear in several applications, only few relevant papers may be found in the literature. In one of these few works [7], under the consensus approach, the authors use the alternating direction method of multipliers in order to enforce the nodes to reach an agreement when estimating parameters of common interest. In [8], the authors apply diffusion adaptation and scalarization techniques to a specific multi-objective optimization problem that appears in a NSPE setting where there is no prior knowledge regarding the overlapped estimation interests. More related to present paper, the authors in [9] consider a NSPE problem where, as it happens in many applications, there is available some prior knowledge about the estimation interests associated with each node. To solve this problem, they rely on adaptive filtering techniques to derive a distributed incremental-based LMS that converges to the minimum mean-square error (MMSE) solution of the corresponding centralized problem.

Here, we firstly propose an incremental-based distributed algorithm that implements the exact RLS solution of a NSPE

The work was partially supported by the European SmartEN ITN project (Grant No. 238726) under the Marie Curie ITN FP7 program and by the University of Patras.

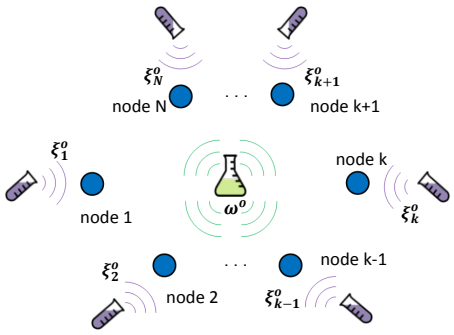

Fig. 1. Network with node-specific parameter estimation interests.

problem where each node of the network is interested in estimating some parameters of local interest and some parameters of global interest. The resulting algorithm outperforms the LMS solution [9] at the expense of increased transmission requirements. Later, since the involved communications cost may be prohibitive in some applications [10], we modify the initial scheme to reduce the communication burden. Finally, the performance of the proposed schemes is illustrated via indicative computer simulations.

We use boldface letters for random variables and normal fonts for deterministic quantities. Capital letters refer to matrices and small letters refer to both vectors and scalars. Moreover, the Hermitian transposition is denoted by $(\cdot)^H$. We use the weighted norm notation $\|x\|^2_\Sigma \triangleq x^H \Sigma x$ where x a vector and $\Sigma > 0$. Finally, $0_{L\times M}$ denotes a $L \times M$ zero matrix.

2. PROBLEM FORMULATION

To begin with, let us consider that a network of N nodes is randomly deployed over some region in order to estimate some unknown vectors of parameters (see Fig.1). At every time instant i, each node k has access to data $\{d_{k,i}, U_{k,i}\}$. These data are assumed to be related to the unknown vectors of parameters by the following model

$$\mathbf{d}_{k,i} = \mathbf{U}_{k,i} w_k^o + \mathbf{v}_{k,i} = \begin{bmatrix}\mathbf{U}_{k_g,i} & \mathbf{U}_{k_l,i}\end{bmatrix} \begin{bmatrix} w^o \\ \xi_k^o \end{bmatrix} + \mathbf{v}_{k,i} \quad (1)$$

where, for each time instant i,

- w_k^o equals the vector of dimension M_k that contains all the parameters of interest for node k. This vector is formed by w^o, which is a sub-vector of dimension $M_g \times 1$ consisting of all the parameters of global interest, and by ξ_k^o, which is a sub-vector of dimension $M_{k_l} \times 1$ that gathers all the parameters of local interest,
- $\mathbf{v}_{k,i}$ is measurement and/or model noise with zero mean and covariance matrix $R_{v_k,i}$ of dimensions $L_k \times L_k$,
- $\mathbf{d}_{k,i}$ and $\mathbf{U}_{k,i}$ are zero-mean random variables with dimensions $L_k \times 1$ and $L_k \times M_k$, respectively. Forming the matrix $\mathbf{U}_{k,i}$, the matrices $\mathbf{U}_{k_g,i}$ and $\mathbf{U}_{k_l,i}$, of dimensions $L_k \times M_g$ and $L_k \times M_{k_l}$, might be correlated, and consist of the columns $\mathbf{U}_{k,i}$ associated with w^o and ξ_k^o, respectively.

In most existing works, e.g., [5] and [6], the derived adaptation schemes assume $w_k^o = w^o$, for all $k \in \{1, 2, \ldots, N\}$. As in our previous work [9], we consider a scenario where the node-specific parameters of interest, i.e., $\{w_k^o\}_{k=1}^N$, may, in general, be different. Hence, the objective for each node k is to estimate its specific unknown vector w_k^o from the data $\{d_{k,i}, U_{k,i}\}$, $k \in \{1, 2, \ldots, N\}$. In particular, as shown in Fig. 1, each vector $\{w_k^o\}_{k=1}^N$ consists of globally common components as well as components of local interest for sensor k. The parameters of global interest in the network may account for an event common to all nodes. In contrast, the parameters of local interest for each node k may represent an influence of some local phenomena that is different for each node.

Given the previous observation model, the goal for each node k is to collect the measurements and regressors from time 0 up to time i, i.e.,

$$\mathcal{Y}_{k,i} = \text{col}\{d_{k,0}, d_{k,1}, \ldots, d_{k,i}\}$$

and

$$\mathcal{H}_{k,i} = \text{col}\{U_{k,0}, U_{k,1}, \ldots, U_{k,i}\},$$

respectively, in order to obtain the node-specific estimators $\{w_k\}_{k=1}^N$ that minimize the associated weighted, regularized, least-squares cost

$$\sum_{k=1}^{N} \left(\lambda^{i+1} \|w_k\|^2_{\Pi_k} + \|\mathcal{Y}_{k,i} - \mathcal{H}_{k,i} w_k\|^2_{\mathcal{W}_{k,i}} \right), \quad (2)$$

where $\Pi_k = \delta^{-1} I_{M_k}$ and

$$\mathcal{W}_{k,i} = \text{diag}\{\lambda^i \Gamma_k, \lambda^{i-1}\Gamma_k, \ldots, \lambda\Gamma_k, \Gamma_k\}$$

with $\delta > 0$ equal to a large constant, $\Gamma_k = R^{-1}_{v_k,i}$ and forgetting factor $0 \ll \lambda \leq 1$. Thus, after particularizing for global and local vector of parameters, our NSPE problem can be casted as

$$\begin{aligned} &\{\hat{w}, \{\hat{\xi}_k\}_{k=1}^N\} \\ &= \underset{w, \{\xi_k\}_{k=1}^N}{\text{argmin}} \left\{ \sum_{k=1}^{N} \lambda^{i+1} \begin{bmatrix} w^H & \xi_k^H \end{bmatrix} \begin{bmatrix} \Pi_{k,g} & 0 \\ 0 & \Pi_{k,l} \end{bmatrix} \begin{bmatrix} w \\ \xi_k \end{bmatrix} \right. \\ &\qquad \left. + \sum_{k=1}^{N} \|\mathcal{Y}_{k,i} - \mathcal{H}^g_{k,i} w - \mathcal{H}^l_{k,i} \xi_k\|^2_{\mathcal{W}_{k,i}} \right\} \end{aligned} \quad (3)$$

where

$$\mathcal{H}^g_{k,i} = \text{col}\{U_{k_g,0}, U_{k_g,1}, \ldots, U_{k_g,i}\}$$

and

$$\mathcal{H}^l_{k,i} = \text{col}\{U_{k_l,0}, U_{k_l,1}, \ldots, U_{k_l,i}\}.$$

3. A SOLUTION OF THE NSPE PROBLEM

In this section, we firstly derive a centralized solution of the optimization problem (3), and then we develop a distributed strategy that converges to this centralized solution. For the sake of simplicity and without losing generality, we assume that $M_k = M$, $M_g = M_g$, $M_{k_l} = M_l$ and $L_k = L$ for all $k \in \{1, 2, \ldots, N\}$.

3.1. Centralized solution

To solve the considered NSPE problem in (3), we have to optimize a scalar real-valued cost function with respect to (w.r.t.) multiple vector variables, i.e., $\{w, \{\xi_k\}_{k=1}^N\}$. After defining the following augmented vector

$$\tilde{w} = \left[w^T \ \xi_1^T \ \xi_2^T \cdots \xi_N^T\right]^T \quad (\widetilde{M} \times 1)$$

and gathering all the data up to time i into

$$\mathcal{Y}_i = \text{col}\{y_0, y_1, \ldots, y_i\} \quad (N \cdot L \cdot (i+1) \times 1)$$

and

$$\widetilde{\mathcal{H}}_i = \text{col}\{\widetilde{H}_0, \widetilde{H}_1, \ldots, \widetilde{H}_i\} \quad (N \cdot L \cdot (i+1) \times \widetilde{M}) \quad (4)$$

where $\widetilde{M} = M_g + N \cdot M_l$. In the definition of $\mathcal{Y}_i$ and $\widetilde{\mathcal{H}}_i$

$$y_i = \text{col}\{d_{1,i}, d_{2,i}, \ldots, d_{N,i}\} \quad (N \cdot L \times 1)$$

$$\widetilde{H}_i = \text{col}\{\widetilde{U}_{1,i}, \widetilde{U}_{2,i}, \ldots, \widetilde{U}_{N,i}\} \quad (N \cdot L \times \widetilde{M})$$

and the augmented regressor is expressed as

$$\widetilde{U}_{k,i} = \begin{bmatrix} U_{k_g,i} & 0_{L \times M_a} & U_{k_l,i} & 0_{L \times M_b} \end{bmatrix} \quad (5)$$

with $M_a = (k-1)M_l$ and $M_b = (N-k)M_l$. Now, we can easily verify that our optimization problem is equivalent to

$$\widehat{\tilde{w}} = \underset{\tilde{w}}{\text{argmin}} \left\{ \lambda^{i+1} \|\tilde{w}\|_{\widetilde{\Pi}}^2 + \|\mathcal{Y}_i - \widetilde{\mathcal{H}}_i \tilde{w}\|_{\mathcal{W}_i}^2 \right\} \quad (6)$$

where

$$\mathcal{W}_i = \text{diag}\{\lambda^i D, \lambda^{i-1} D, \ldots, \lambda D, D\}$$

and

$$\widetilde{\Pi} = \text{diag}\left\{ \sum_{k=1}^{N} \Pi_{k,g}, \Pi_{1,l}, \Pi_{2,l}, \ldots, \Pi_{N,l} \right\},$$

with

$$D = \text{diag}\{\Gamma_1, \Gamma_2, \ldots, \Gamma_N\}.$$

It is well-known that the solution $\widehat{\tilde{w}}_i$ is given by [11]:

$$\widehat{\tilde{w}}_i = \widetilde{P}_i \widetilde{\mathcal{H}}_i^H \mathcal{W}_i \mathcal{Y}_i, \quad (7)$$

where

$$\widetilde{P}_i = \left(\lambda^{i+1} \widetilde{\Pi} + \widetilde{\mathcal{H}}_i^H \mathcal{W}_i \widetilde{\mathcal{H}}_i \right)^{-1}. \quad (8)$$

However, this centralized batch solution requires the inversion of a square matrix whose dimension is actually proportional to the number of nodes N. In addition, it requires that we store in memory all data available until time i. Hence, a prohibitively high computational and memory-wise cost is needed.

3.2. Distributed solution

With the aim of increasing energy efficiency and improving robustness and scalability it is highly desirable to design a distributed and adaptive scheme in order to update $\widehat{\tilde{w}}_{i-1}$ to $\widehat{\tilde{w}}_i$. Toward this goal, we firstly develop a distributed recursion for $\widetilde{P}_i$.

By following the approach described in [5], we firstly express the relation (8) as follows

$$\begin{aligned} \widetilde{P}_i^{-1} &= \lambda^{i+1} \widetilde{\Pi} + \widetilde{\mathcal{H}}_i^H \mathcal{W}_i \widetilde{\mathcal{H}}_i \\ &= \lambda \left(\lambda^i \widetilde{\Pi} + \widetilde{\mathcal{H}}_{i-1}^H \mathcal{W}_{i-1} \widetilde{\mathcal{H}}_{i-1} \right) + \widetilde{H}_i^H D \widetilde{H}_i \\ &= \lambda \widetilde{P}_{i-1}^{-1} + \widetilde{H}_i^H D \widetilde{H}_i. \end{aligned} \quad (9)$$

After noting that (9) can be rewritten as a sequence of rank-L updates, we can apply the matrix inversion lemma so that the following distributed recursion for $\widetilde{P}_i$ is obtained

$$\left\{ \begin{array}{l} \widetilde{P}_{0,i} \leftarrow \lambda^{-1} \widetilde{P}_{N,i-1} \\ \text{for } k = 1 : N \\ \quad \widetilde{G}_{k,i} = \left(\Gamma_k^{-1} + \widetilde{U}_{k,i} \widetilde{P}_{k-1,i} \widetilde{U}_{k,i}^H \right)^{-1} \\ \quad \widetilde{P}_{k,i} = \widetilde{P}_{k-1,i} \\ \qquad + \widetilde{P}_{k-1,i} \widetilde{U}_{k,i}^H \widetilde{G}_{k,i} \widetilde{U}_{k,i} \widetilde{P}_{k-1,i} \\ \text{end} \end{array} \right. \quad (10)$$

where $\widetilde{P}_{k,i}$ denotes the local estimate of $\widetilde{P}_i$ at node k at some time instant i.

Now, let us focus on the distributed update of $\widehat{\tilde{w}}_{i-1}$ to $\widehat{\tilde{w}}_i$. Toward this goal, let us define the intermediate global matrices $\mathcal{Y}_i^k$ and $\widetilde{\mathcal{H}}_i^k$ that stack $\mathcal{Y}_{i-1}$ and $\widetilde{\mathcal{H}}_{i-1}$ in addition to the measurements and regressors collected across the network at time i up to node k, respectively,

$$\mathcal{Y}_i^k = \left[\begin{array}{c} \mathcal{Y}_{i-1} \\ \hline d_{1,i} \\ d_{2,i} \\ \vdots \\ d_{k,i} \end{array} \right] \quad \text{and} \quad \widetilde{\mathcal{H}}_i^k = \left[\begin{array}{c} \widetilde{\mathcal{H}}_{i-1} \\ \hline \widetilde{U}_{1,i} \\ \widetilde{U}_{2,i} \\ \vdots \\ \widetilde{U}_{k,i} \end{array} \right].$$

Given the previous definitions, we can easily check that the local estimate $\tilde{w}_i$ at node k, i.e., $\tilde{\psi}_k^{(i)}$, at any time instant i is equal to the solution of the following LS problem

$$\tilde{\psi}_k^{(i)} = \underset{\tilde{w}}{\text{argmin}} \left\{ \lambda^{i+1} \|\tilde{w}\|_{\widetilde{\Pi}}^2 + \|\mathcal{Y}_i^k - \widetilde{\mathcal{H}}_i^k \tilde{w}\|_{\mathcal{W}_i^k}^2 \right\} \quad (11)$$

where

$$\mathcal{W}_i^k = \begin{bmatrix} \lambda \mathcal{W}_{i-1} & 0 \\ 0 & D_k \end{bmatrix}$$

and $D_k = \text{diag}\{\Gamma_1, \ldots, \Gamma_k\}$. Next, if we take into account that the solution of (11) is equal to

$$\tilde{\psi}_k^{(i)} = \widetilde{P}_{k,i} [\widetilde{\mathcal{H}}_i^k]^H \mathcal{W}_i^k \mathcal{Y}_i^k, \tag{12}$$

by using (10) and noting that

$$\mathcal{W}_i^k = \begin{bmatrix} \mathcal{W}_i^{k-1} & 0 \\ 0 & \Gamma_k \end{bmatrix}$$

and

$$[\widetilde{\mathcal{H}}_i^k]^H \mathcal{W}_i^k \mathcal{Y}_i^k = [\widetilde{\mathcal{H}}_i^{k-1}]^H \mathcal{W}_i^{k-1} \mathcal{Y}_i^{k-1} + \widetilde{U}_{k,i}^H \Gamma_k d_{k,i},$$

several algebraic manipulations similar to the ones undertaken in [5] yield the following distribute recursion for $\tilde{\psi}_k^{(i)}$

$$\tilde{\psi}_k^{(i)} = \tilde{\psi}_{k-1}^{(i)} + \widetilde{P}_{k-1,i} \widetilde{U}_{k,i}^H \widetilde{G}_{k,i} \left(d_{k,i} - \widetilde{U}_{k,i} \tilde{\psi}_{k-1}^{(i)} \right). \tag{13}$$

Finally, if we group the recursions (10) and (13), we obtain a distributed incremental-based RLS algorithm that provides the exact solution to the centralized NSPE problem (6). The new algorithm is summarized as follows

Incremental-Based NSPE RLS (I-NSPE RLS)

- Initialization: $\psi_N^{(-1)} = 0, \quad \widetilde{P}_{N,-1} = \widetilde{\Pi}^{-1}$.
- At each time $i \geq 0$, for each $k \in \{1, \ldots, N\}$ execute

$$\begin{aligned} \tilde{\psi}_0^{(i)} &\leftarrow \tilde{\psi}_N^{(i-1)}; \qquad \widetilde{P}_{0,i} \leftarrow \lambda^{-1} \widetilde{P}_{N,i-1} \\ e_{k,i} &= d_{k,i} - \widetilde{U}_{k,i} \tilde{\psi}_{k-1}^{(i)} \\ \widetilde{G}_{k,i} &= \left(\Gamma_k^{-1} + \widetilde{U}_{k,i} \widetilde{P}_{k-1,i} \widetilde{U}_{k,i}^H \right)^{-1} \\ \tilde{\psi}_k^{(i)} &= \tilde{\psi}_{k-1}^{(i)} + \widetilde{P}_{k-1,i} \widetilde{U}_{k,i}^H \widetilde{G}_{k,i} e_{k,i} \\ \widetilde{P}_{k,i} &= \widetilde{P}_{k-1,i} + \widetilde{P}_{k-1,i} \widetilde{U}_{k,i}^H \widetilde{G}_{k,i} \widetilde{U}_{k,i} \widetilde{P}_{k-1,i} \end{aligned} \tag{14}$$

3.3. Low-communication distributed RLS

Although the previously developed distributed algorithm (14) provides the exact RLS solution to the NSPE problem (6), it needs $O(\widetilde{M}^2)$ transmission complexity which can be prohibitive for applications with strict energy constraints. Motivated by this fact, at the expense of some performance degradation, we propose a simplified implementation that approximates $\widetilde{P}_{k,i}$ by a block diagonal matrix to reduce the number of transmitted parameters between two neighboring nodes.

A careful inspection of (5) together with (14) reveals that, under a block diagonal approximation for $\widetilde{P}_{k,i}$, only two subvectors of $\tilde{\psi}_k^{(i)}$ and two submatrices of $\widetilde{P}_{k,i}$ are updated at each node and at a some specific time instant i. In particular, according to the recursions (10) and (13), only the subvectors associated with the local estimates of w^o and ξ_k^o at node k and time instant i, denoted as $\psi_k^{(i)}$ and $\xi_k^{(i)}$, respectively, are updated from $\psi_{k-1}^{(i)}$, $\xi_k^{(i-1)}$,

$$\widetilde{P}_{k-1,i}(1 : M_g) = P_{(k-1)_g,i}$$

and

$$\widetilde{P}_{k,i-1}(M_g + M_a + 1 : M_g + M_l + M_a) = P_{k_l,i-1}$$

where $A(l_a : l_b)$ equals a square submatrix defined by the rows and columns $l_a, l_a + 1, l_a + 2, \ldots, l_b$ with $l_a \leq l_b$. Similarly, at each time instant i each node k updates the submatrices $P_{k_g,i}$ and $P_{k_l,i}$ based on $P_{(k-1)_g,i}$ and $P_{k_l,i-1}$. Therefore, without any loss of optimality w.r.t. the steady state performance of I-NSPE RLS when $\mathbf{U}_{k_g,i}$ and $\mathbf{U}_{k_l,i}$ are independent, the previous facts allow to properly modify (14) in order to obtain the subsequent incremental-based algorithm

$$\left\{ \begin{array}{l} \psi_0^{(i)} \leftarrow \psi_N^{(i-1)}; \qquad P_{0_g,i} \leftarrow \lambda^{-1} P_{N_g,i-1} \\ \text{for } k = 1 : N \\ \quad e_{k,i} = d_{k,i} - U_{k,i} \begin{bmatrix} \psi_{k-1}^{(i)} \\ \xi_k^{(i-1)} \end{bmatrix} \\ \quad G_{k,i} = \Big(\Gamma_k^{-1} + U_{k_g,i} P_{(k-1)_g,i} U_{k_g,i}^H \\ \qquad\qquad + \lambda^{-1} U_{k_l,i} P_{k_l,i-1} U_{k_l,i}^H \Big)^{-1} \\ \quad \begin{bmatrix} \psi_k^{(i)} \\ \xi_k^{(i)} \end{bmatrix} = \begin{bmatrix} \psi_{k-1}^{(i)} \\ \xi_k^{(i-1)} \end{bmatrix} + \begin{bmatrix} P_{(k-1)_g,i} U_{k_g,i}^H \\ \lambda^{-1} P_{k_l,i-1} U_{k_l,i}^H \end{bmatrix} G_{k,i} e_{k,i} \\ \quad P_{k_g,i} = P_{k-1_g,i} - P_{(k-1)_g,i} U_{k_g,i}^H G_{k,i} U_{k_g,i} P_{(k-1)_g,i} \\ \quad P_{k_l,i} = \lambda^{-1} P_{k_l,i-1} - \lambda^{-2} P_{k_l,i-1} U_{k_l,i}^H G_{k,i} U_{k_l,i} P_{k_l,i-1} \\ \text{end} \end{array} \right. \tag{15}$$

Note that, under the above strategy, each node only needs to transmit the blocks of $\widetilde{P}_{k,i}$ and $\tilde{\psi}_k^{(i)}$ corresponding to the global vector of parameters, i.e., $P_{k_g,i}$ and $\psi_k^{(i)}$ respectively. Since the dimension of the vector of global parameters equals M_g, the scheme in (15) requires only $O(M_g^2)$ transmission complexity, which is not dependent on the network size.

4. SIMULATIONS

We assume a network with $N = 10$ nodes where the measurements follow the observation model (1) with $M_{l_k} = 8$, $M_g = 10$ and $L_k = 1$ for all $k = \{1, \ldots, N\}$. We have also considered a forgetting factor $\lambda = 1$ and that the background noise $\mathbf{v_{k,i}}$ has variance $\sigma_{v,k}^2 = \sigma_v^2 = 10^{-3}$ across the network. Additionally, we have assumed that the regressors $\mathbf{u}_{k_g,i}$ and $\mathbf{u}_{k_l,i}$ are independently generated according to a time-correlated spatially independent Gaussian distribution. In particular, both $\mathbf{u}_{k_g,i}$ and $\mathbf{u}_{k_l,i}$ follow a stationary first-order autoregressive (AR) model with correlation functions $r_{k_g}(i) =$

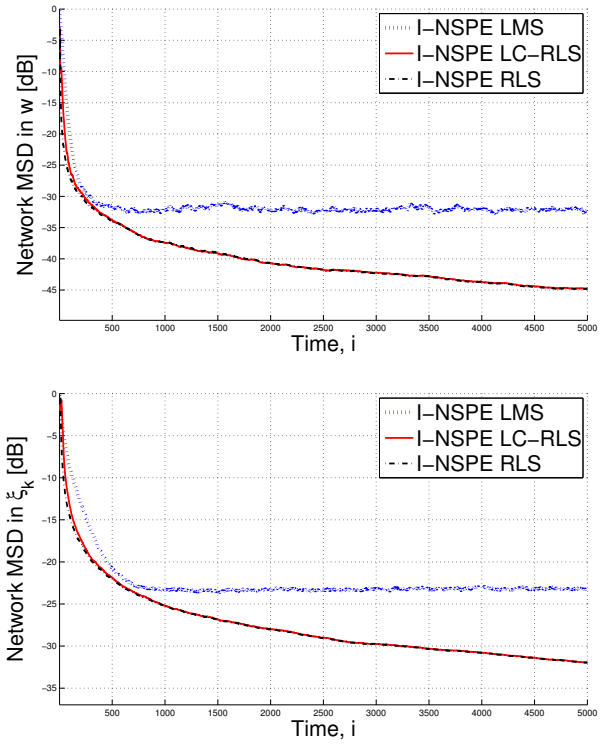

Fig. 2. Learning behavior of network MSD.

$\sigma^2_{u,k_g}\alpha_g^{|i|}$ and $r_{k_l}(i) = \sigma^2_{u,k_l}\alpha_l^{|i|}$, respectively. In the previous AR models, the parameters $\{\alpha_g, \alpha_l\}$ and $\{\sigma^2_{u,k_g}, \sigma^2_{u,k_l}\}$ have been randomly chosen in [0,1) so that the signal-to-noise-ratio at each node is different and varies from 14 dB to 17 dB.

We compare the performance of the I-NSPE RLS strategy provided in (14), its low-communication version summarized in (15) and denoted as I-NSPE LC-RLS, as well as the I-NSPE LMS algorithm derived in [9]. Specifically, for each one of these I-NSPE algorithms, Figure 2 depicts the learning behaviour of the network mean-square deviation (MSD) associated with the estimation of w^o and ξ_k^o. The curves have been generated by averaging 50 independent experiments. We can note that, at the expense of increased computational complexity, both I-NSPE RLS and I-NSPE LC-RLS outperform I-NSPE LMS in terms of steady state floor and rate of convergence. Since the processes of estimating w^o and ξ_k^o are coupled though the observation model, the improved performance appears in both estimation tasks. Furthermore, as it was expected during the derivations of the proposed schemes, under the considered scenario, I-NSPE LC-RLS achieves identical steady state performance with I-NSPE RLS. In fact, as a result of reducing the transmission complexity of I-NSPE RLS, in the considered setting the I-NSPE LC-RLS scheme suffers only a small performance loss in the rate of convergence.

5. CONCLUSION

We addressed a novel NSPE problem where the estimation interests of the nodes comprise a set of local parameters and a set of network global parameters. Toward this goal, we initially proposed an incremental-type scheme that, in a distributed fashion, implements the exact RLS solution of a central unit processing the data of all the nodes. Next, we derived a scheme with lower transmission complexity for applications where the communications and energy resources are scarce. Additionally, it was shown that this simplified scheme may converge to the exact RLS solution. Finally, by performing computer simulations we showed the effectiveness of the proposed algorithms. Due to limited space we have not included any theoretical analysis concerning the performance of the proposed techniques. A complete analysis is currently being conducted and will appear in a forthcoming paper.

6. REFERENCES

[1] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2365–2382, 2009.

[2] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, 2007.

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

[4] G. Mateos, I. D. Schizas, and G. B. Giannakis, "Distributed recursive least-squares for consensus-based in-network adaptive estimation," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4583–4588, 2009.

[5] A. H. Sayed and C. G. Lopes, "Distributed recursive least-squares strategies over adaptive networks," in *IEEE 40th Asilomar Conference on Signals, Systems and Computers, ACSSC*, 2006, pp. 233–237.

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

[7] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *Submitted to IEEE Transactions on Power Systems*, 2012 [Online]. Available: http://arxiv.org/abs/1204.0991.

[8] J. Chen and A. H. Sayed, "Distributed Pareto-optimal solutions via diffusion adaptation," in *IEEE Statistical Signal Processing Workshop, SSP.*, 2012, pp. 648–651.

[9] N. Bogdanovic, J. Plata-Chaves, and K. Berberidis, "Distribtued incremental-based LMS for node-specific parameter estimation over adaptive networks," in *Accepted for publication in IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, 2013.

[10] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, 2001, vol. 4, pp. 2033–2036.

[11] A. H. Sayed, *Adaptive filters*, Wiley-IEEE Press, 2011.