DISTRIBUTED SIGNAL ESTIMATION IN A WIRELESS SENSOR NETWORK WITH PARTIALLY-OVERLAPPING NODE-SPECIFIC INTERESTS OR SOURCE OBSERVABILITY

Jorge Plata-Chaves, Alexander Bertrand, Marc Moonen

KU Leuven, Dept. Electrical Engineering ESAT
Stadius Center for Dynamical Systems, Signal Processing and Data Analytics
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
E-mails: {jplata, alexander.bertrand, marc.moonen}@esat.kuleuven.be

ABSTRACT

We study a distributed node-specific signal estimation problem where the node-specific desired signals and/or the sensor observations can have partially-overlapping latent signal subspaces. First, we provide the minimum number of linear combinations of observed sensor signals that each node can broadcast to still let all other nodes achieve the network-wide Linear Minimum Mean-Square Error (LMMSE) estimate of their node-specific desired signals. Later, for a fully-connected wireless sensor network, we derive a distributed algorithm that, under some settings, allows each node to achieve the LMMSE estimate of its node-specific desired signals by broadcasting the smallest number of signals. Unlike the existing algorithms, the proposed algorithm deals with the problem of partially-overlapping node-specific interests and incomplete observability of all latent sources at the nodes. Finally, the effectiveness of the proposed technique is shown through numerical simulations.

Index Terms— Distributed signal estimation, wireless sensor networks, distributed compression

1. INTRODUCTION

In a wireless sensor network (WSN), the estimation of a set of parameters or signals is traditionally performed in a central unit which collects all the sensor signal observations of all the nodes in the network. To reduce energy consumption, and to improve robustness and scalability, more recent approaches (e.g. [1]-[2]) rely on distributed algorithms based on in-network processing of the sensor signals.

In most distributed estimation problems, it is generally assumed that the nodes in a WSN have the same interest, i.e. the estimation of a global vector of parameters or a network-wide signal (e.g. [1]-[5]). However, motivated by applications such as speech enhancement in acoustic sensor networks [6] or cooperative spectrum sensing in cognitive radio networks [7]-[8], special attention is being paid to more general distributed estimation problems where the nodes have different but overlapping estimation interests.

In the growing literature on node-specific estimation problems in adaptive networks, works such as [9] apply consensus strategies to solve node-specific parameter estimation (NSPE) problems where there are parameters of common interest to a subset of nodes in the network. For scenarios where there are parameters of local interest to a node in addition to parameters of common and/or network-wide interest, there are also several NSPE algorithms based on adaptive filtering techniques under an incremental [10] or a diffusion [11] mode of cooperation. Other recent works solving different NSPE problems based on adaptive filtering techniques can be found in [12]-[13].

Rather than NSPE problems, we consider node-specific signal estimation (NSSE) problems, which are fundamentally different and require different techniques to solve them (see [6] for a detailed comparison). In particular, we focus on linear NSSE techniques that estimate the samples of a node-specific desired signal by performing a filter-and-sum operation on all the sensor signals in the WSN. For a fully-connected and a tree network, the authors in [14] and [15] propose a distributed adaptive node-specific signal estimation (DANSE) algorithm that significantly reduces the communication bandwidth, while still letting each node achieve the network-wide LMMSE estimate of its node-specific desired signals. To do so, these distributed algorithms consider a fully-overlapping NSSE (FO-NSSE) problem where (a) all node-specific desired signals fully span the same latent low-dimensional signal subspace and where (b) all nodes observe *all* the latent sources in their sensor signals. However, when one of these assumptions does not hold, convergence of the DANSE algorithm to the network-wide optimal solution is not ensured [16]. Furthermore, for such a scenario, it is rather unclear *how many* compressed signals have to be broadcast and *how* the optimal compression rules can be found to let all of them achieve the node-specific LMMSE estimate.

Here, we consider a partially-overlapping NSSE (PO-NSSE) problem where the latent signal subspaces of the node-specific desired signals are only partially overlapping and/or where the nodes do not observe all latent sources. To do so, we first show the number of linearly independent signals that every node should at least broadcast to let all nodes achieve the LMMSE estimates of their node-specific desired signals as if they had access to all sensor signal observations of the network. This number can be viewed as the linear compression bound to still obtain network-wide LMMSE estimates. However, the proof does not describe how the aforementioned bound can be achieved in practice. Nevertheless, we provide a distributed algorithm where, under some settings, all nodes achieve

This work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council CoE PFV/10/002 (OPTEC), Concerted Research Action GOA-MaNet, the Interuniversity Attractive Poles Programme initiated by the Belgian Science Policy Office IUAP P7/23 'Belgian network on stochastic modeling analysis design and optimization of communication systems' (BESTCOM) 2012-2017, Research Project FWO nr. G.0763.12 'Wireless Acoustic Sensor Networks for Extended Auditory Communication', Project FWO nr. G.0931.14 'Design of distributed signal processing algorithms and scalable hardware platforms for energy-vs-performance adaptive wireless acoustic sensor networks' and EU/FP7 project HANDiCAMS. The project HANDiCAMS acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 323944. The scientific responsibility is assumed by its authors.

the network-wide LMMSE estimates of their node-specific desired signals by broadcasting the number of signals established by this bound. Unlike the DANSE algorithm, the proposed algorithm has a guaranteed convergence in PO-NSSE problems. Finally, some numerical simulations illustrate the effectiveness of the algorithm.

2. PROBLEM FORMULATION

We consider a WSN with N nodes that are randomly deployed over some region in which a set of Q complex-valued and mutually uncorrelated latent source signals, $\{s_j\}_{j=1}^{Q}$, are generated, in addition to background noise. For instance, in Fig. 1 we plot a network with $N = 5$ nodes and $Q = 2$ latent source signals generating s_1 and s_2. At each time t, each node k collects an observation $\mathbf{y}_k[t]$ of an M_k-channel signal $\mathbf{y}_k$. Each $y_{k,m}[t]$, with $m \in \{1, 2, \ldots, M_k\}$, of $\mathbf{y}_k[t]$ corresponds to the observation collected by the m-th sensor of node k at time t. We assume that all sensor signals are ergodic and stationary in short term at least, in which case the theory should be applied to finite signal segments. Moreover, we will omit the time index when referring to a signal, and we will only write it when referring to a specific observation of the signal.

From the set of sensor signals $\{\mathbf{y}_k\}_{k=1}^{N}$, each node k aims at estimating a node-specific desired signal d_k, which consists of a linear mixture of the latent source signals. To make this more concrete, we define the Q-channel signal $\mathbf{s}$ in which all Q signals s_j are stacked, i.e., $\mathbf{s} = \mathrm{col}\{\{s_j\}_{j=1}^{Q}\}$. Modelling measurement noise at node k, we also define $\mathbf{n}_k$ as a zero mean noise component that is statistically independent of $\mathbf{s}$ and is possibly correlated to $\mathbf{n}_\ell$ with $k \neq \ell$. Then, the signals observed by node k are described by

$$\mathbf{y}_k = \mathbf{B}_k \mathbf{s} + \mathbf{n}_k \tag{1}$$

with $\mathbf{B}_k$ an unknown $M_k \times Q$ steering matrix to the M_k sensors of node k. Due to the attenuation properties, in practice a node k only observes a latent source s_j if it is located within its area of influence, which is denoted by the ordered set of node indices $\mathcal{B}_j$. Thus, it may happen that a node k only observes Q_k out of Q latent sources, in which case the matrix $\mathbf{B}_k$ contains zero-columns. Additionally, the node-specific desired signal d_k is given by

$$d_k = \mathbf{a}_k^H \mathbf{s} \tag{2}$$

where the superscript H denotes the conjugate transpose operator and where $\mathbf{a}_k$ is a mixing vector that specifies the interests of node k. Although a multi-channel desired signal could be considered for each node k, for the sake of an easy exposition, we will assume that d_k is a single-channel signal. In a practical scenario, note that a node k may not be interested in estimating a filtered version of all latent source signals, i.e., the vector $\mathbf{a}_k$ may contain zeros at the entries corresponding to sources in which node k is not interested. As a result, a latent source is not necessarily within the interest of all nodes of the network. We will use the set $\mathcal{A}_j$ to denote the set of nodes interested in estimating a linear mixture including the latent source signal s_j. Note that the sets $\{\mathcal{A}_j\}_{j=1}^{Q}$ are not necessarily related to the sets $\{\mathcal{B}_j\}_{j=1}^{Q}$. Nonetheless, there might exist scenarios where d_k consists of a linear combination of a subset of the latent source signals observed by node k as they impinge on one of its sensors, called the reference sensor. In that case, $\mathbf{a}_k$ would be composed of entries of a column of $\mathbf{B}_k^H$ and $\mathcal{A}_j \subseteq \mathcal{B}_j$. For instance, for the network in Fig. 1 all nodes that observe latent source signal s_1, i.e., $\mathcal{B}_1 = \{1, 2, 3, 4\}$, are interested in estimating s_1. As a result, $\mathcal{A}_1 = \mathcal{B}_1$. On the contrary, within the set of nodes that observe s_2, i.e., $\mathcal{B}_2 = \{2, 3, 4, 5\}$, only the nodes in $\mathcal{A}_2 = \{3, 4, 5\} \subset \mathcal{B}_2$ are interested in s_2.

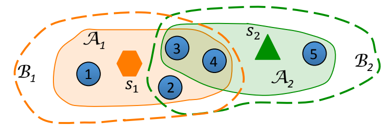

Fig. 1. WSN where nodes have partially-overlapping interests and where the nodes observe different latent source signals.

Without making any assumption on the probability distributions of the involved signals, we consider the following node-specific LMMSE estimator[1] to estimate d_k

$$\widehat{\mathbf{w}}_k = \underset{\mathbf{w}_k}{\mathrm{argmin}} \left\{ J_k(\mathbf{w}_k) \right\} = \underset{\mathbf{w}_k}{\mathrm{argmin}} \left\{ E \parallel d_k - \mathbf{w}_k^H \mathbf{y} \parallel^2 \right\} \tag{3}$$

where $\mathbf{w}_k$ is an unknown complex $M \times 1$ vector and where $\mathbf{y}$ is the M-channel signal in which all $\mathbf{y}_k$ are stacked with $M = \sum_{k=1}^{N} M_k$, i.e., $\mathbf{y} = \mathrm{col}\{\{\mathbf{y}_k\}_{k=1}^{N}\}$. It is assumed that the node-specific desired signals are unknown and, possibly, different for any two nodes k and ℓ with $k \neq \ell$, i.e. $d_k \neq d_\ell$. In the case where d_k is a linear mixture of the latent source signals as they impinge on the reference sensor of node k, the idea of (3) is to perform a denoising of the sensor signals, while preserving the spatial information (the local mixture of the latent source signals as observed in the reference sensor) at each node. This could be important, e.g. for directional hearing in hearing aid applications [17], or when the denoising step is followed by a localization procedure [18].

Assuming that the correlation matrix $\mathbf{R}_{yy} = E\{\mathbf{y}\mathbf{y}^H\}$ is full rank, which is generally satisfied due to sensor noise, the unique solution of (3) is [19]:

$$\hat{\mathbf{w}}_k = \mathbf{R}_{\mathbf{yy}}^{-1} \, \mathbf{r}_{\mathbf{y}d_k} \tag{4}$$

where $\mathbf{r}_{\mathbf{y}d_k} = E\{\mathbf{y} \, d_k^H\}$. Since the signals are assumed to be ergodic, $\mathbf{R}_{\mathbf{yy}}$ can be estimated directly from the sensor signal observations by time averaging. Since we assume that the signal d_k is unknown, we will assume that $\mathbf{r}_{\mathbf{y}d_k}$ is estimated indirectly based on the sensor observations of $\mathbf{y}$. For instance, if the signals $\{d_k\}_{k=1}^{N}$ have an ON-OFF behavior (as it is the case for, e.g. speech signals), then the nodes are able to observe noise-only segments in their sensor signals. This allows to compute the noise covariance, from which $\mathbf{r}_{\mathbf{y}d_k}$ can be estimated as long as $\mathbf{a}_k$ corresponds to a column of $\mathbf{B}_\ell^H$ for some $\ell \in \{1, 2, \ldots, N\}$ (see [14] and [17] for further details).

To find the solution given in (4) at each node k, in principle, each node k needs to have access to all the M sensor signals in $\mathbf{y}$. Therefore, each node would need to broadcast all observations of its M_k-channel signal $\mathbf{y}_k$ to the other nodes of the network. Alternatively, to increase the energy efficiency, the nodes can broadcast linearly compressed versions of their sensor signal observations. When determining the linear compressors that let each node k find (4) in a distributed fashion, the existing works (e.g. [14]-[15]) consider a FO-NSSE problem where each node k can observe *all* latent source signals and has a desired signal d_k that consists of a linear mixture of *all* the latent source signals as they impinge on its reference sensor. Instead, this paper studies how many and which signals should be broadcast by each node in a PO-NSSE problem where some nodes may not observe *all* latent source signals (i.e., $\mathbf{B}_k$ may contain zero columns) and where each d_k may only consist of a mixture of a *subset* of the latent source signals (i.e., the vector $\mathbf{a}_k$ may contain zeros). It is noted that the DANSE algorithm in [14]-[15] cannot deal with this kind of scenarios [16].

[1] It is noted that we consider complex variables, hence this can be viewed as a frequency-domain implementation of a multi-channel linear filtering operation in the time domain.

3. MINIMUM NUMBER OF BROADCAST SIGNALS

From $\hat{\mathbf{w}}_k$ given in (4), we can check that the centralized LMMSE estimate of the node-specific desired signal d_k is expressed as

$$\hat{d}_k = \hat{\mathbf{w}}_k^H \mathbf{y} = \sum_{\ell=1}^{N} \hat{\mathbf{w}}_{k,\ell}^H \mathbf{y}_\ell \tag{5}$$

where $\hat{\mathbf{w}}_{k,\ell}$ denotes the sub-vector of $\hat{\mathbf{w}}_k$ that is applied to $\mathbf{y}_\ell$. To be able to use all the sensor signal observations and find $\hat{d}_k$ in a distributed fashion, we consider algorithms where each node broadcasts linearly compressed observations of its M_k-channel signal $\mathbf{y}_k$. In this case, each node has access to its own sensor signal observations, $\mathbf{y}_k$, as well as to a compressed version of the sensor signal observations at the other $N-1$ nodes, which are stacked in

$$\mathbf{z}_{-k} = [\mathbf{z}_1^T \cdots \mathbf{z}_{k-1}^T\, \mathbf{z}_{k+1}^T \cdots \mathbf{z}_N^T]^T \tag{6}$$

where $\mathbf{z}_\ell = \mathbf{C}_\ell^H \mathbf{y}_\ell$ and where $\mathbf{C}_\ell$ is a $M_\ell \times L_\ell$ compression matrix with $L_\ell \leq M_\ell$ and $\ell \neq k$.

For a FO-NSSE problem, previous works (see e.g. [14]-[15]) have designed algorithms where the observations to be broadcast by each node k can be linearly compressed by a factor M_k/L_k where $L_k = \min(M_k, Q)$ without any loss of optimality. An intuitive reason why each node k should broadcast observations of Q signals might be because $\mathbf{z}_k$ should fully capture the Q-dimensional signal subspace spanned by the node-specific interests. However, it is not clear if this intuition is correct. This unsolved question is even more uncertain if we consider the PO-NSSE problem of Section 2, where node k observes Q_k out of Q latent source signals and/or where the node-specific interests do not share the same latent signal subspace. For instance, since the complete Q-dimensional latent signal subspace is not captured anyway by the sensor signals of node k when $\mathbf{B}_k$ is rank deficient, one might be tempted to think that node k only needs to broadcast observations of a Q_k-channel compressed signal $\mathbf{z}_k$. However, this is not generally true as it can be deduced from the following theorem (the proof is omitted due to space constraints).

Theorem 1. *Define $\mathcal{N}^M$ as the space of the M-channel noise signal $\mathbf{n}$ in which all $\mathbf{n}_k$ are stacked. Also assume that the $M \times Q$ stacked matrix $\mathbf{B} = \left[\mathbf{B}_1^T\, \mathbf{B}_2^T \cdots \mathbf{B}_N^T\right]^T$ is full rank with $M \geq Q$. Then, to be able to achieve the optimal LMMSE estimate* (5) *at each node $k \in \{1, 2, \ldots, N\}$ for any possible $\mathbf{n} \in \mathcal{N}^M$, each node k has to broadcast observations of at least*

$$L_k^o = \min\{M_k, P_k\} \tag{7}$$

linearly independent signals where

$$P_k = \mathrm{rank}(\mathbf{A}_{-k}) \leq Q \tag{8}$$

with $\mathbf{A}_{-k} = \left[\mathbf{a}_1\, \mathbf{a}_2 \cdots \mathbf{a}_{k-1}\, \mathbf{a}_{k+1} \cdots \mathbf{a}_N\right]$. Furthermore, L_k^o is a tight bound, i.e., if node k broadcasts observations of less than L_k^o signals, then the LMMSE estimate (5) *cannot be achieved at all other nodes for any possible $\mathbf{n} \in \mathcal{N}^M$.*

A surprising result of Theorem 1 is the independence between the number of latent source signals observed by node k, i.e., Q_k, and the number of signals L_k^o of which observations have to be broadcast by node k to let all other nodes achieve the LMMSE estimate of their node-specific desired signals. According to Theorem 1, note that even if node k observes, e.g., only one of the latent signals, it should still broadcast observations of at least an L_k^o-channel signal to ensure optimality in all the NSSE problems. This can be explained by the fact that the noise may be correlated across different nodes. In this case, node k may help other nodes to achieve better estimates by providing good noise references, even if node k does not observe all the desired latent source signals that are within the interest of the rest of the nodes. For example, although node 5 observes one of the two latent source signals present in the network of Fig. 1, if there is no prior knowledge about noise covariance matrix, it needs to broadcast observations of at least $P_5 = 2$ signals to ensure optimality in all other NSSE problems.

Additionally, Theorem 1 shows the optimality of the compression factor $M_k/\min\{M_k, Q\}$ applied by each node k when implementing the algorithms derived in [14]-[15] for a FO-NSSE problem where $P_k = Q$ holds. For the more general PO-NSSE problem where a node k may only be interested in some of the latent source signals and/or may only observe $Q_k \leq Q$ latent source signals, it may still occur that $P_k = Q$, in which case node k should still broadcast observations of Q signals to ensure optimality in all other NSSE problems. For instance, this is the case of the network in Fig. 1. However, none of the previous settings may not occur. As an example, consider a network where $Q = 3$ latent source signals are observed by 5 nodes equipped with 4 sensors. If nodes {2,3,4,5} are only interested in s_3 and node 1 is interested in estimating a linear mixture of all the 3 latent sources, we can check that $P_1 = 1$ and $P_k = 2$ for $k \neq 1$. In this case, according to Theorem 1, the compression factor $M_k/\min\{M_k, Q\}$ is not necessarily optimal.

4. DANSE ALGORITHM FOR PO-NSSE

Here, we briefly describe a distributed algorithm that solves a PO-NSEE problem of Section 2. Without losing optimality in any of the NSSE problems, the algorithm compresses the signals $\mathbf{y}_k$ into Q-channel signals, which means that the algorithm broadcasts the minimum number of signals when $P_k = Q$ for all nodes (see Theorem 1). However, constructing an algorithm that achieves the bound L_k^o when $P_k < Q$ for some k, remains an open problem. From now on, to avoid straightforward solutions, we will assume that $M_k > Q$.

Our starting point is the DANSE algorithm that solves a FO-NSSE problem in a fully connected network [14]. In summary, this algorithm allows each node to obtain the LMMSE estimate of a node-specific Q-channel signal $\mathbf{d}_k$ from linearly compressed observations of other nodes (note that this is a generalization of (2) for multi-channel desired signals). To do so, the following optimization problem is solved at each iteration $i \geq 1$

$$\begin{bmatrix} \mathbf{W}_{k,k}^{i+1} \\ \mathbf{G}_{k,-k}^{i+1} \end{bmatrix} = \underset{\mathbf{W}_{k,k}, \mathbf{G}_{k,-k}}{\operatorname{argmin}}\; E \left\| \mathbf{d}_k - \left[\ \mathbf{W}_{k,k}^H \mid \mathbf{G}_{k,-k}^H\ \right] \tilde{\mathbf{y}}_k^i \right\|^2 \tag{9}$$

where $k = \mathrm{mod}(i-1, N) + 1$, $\mathbf{G}_{k,-k} = \mathrm{col}\{\{\mathbf{G}_{k,\ell}\}_{\ell=1;\ell\neq k}^N\}$ and $\tilde{\mathbf{y}}_k^i = \mathrm{col}\{\mathbf{y}_k, \mathbf{z}_{-k}^i\}$ with $\mathbf{z}_k^i = [\mathbf{W}_{k,k}^i]^H \mathbf{y}_k$. It is noted that $\mathbf{W}_{k,k}^i$ acts both as the compressor matrix $\mathbf{C}_k$ at iteration i and as a part of the estimator of $\mathbf{d}_k$, i.e., the observations of the compressed signal $\mathbf{z}_k^i$ that is broadcast by node k is also used in the estimation of $\mathbf{d}_k$ at node k itself. Similar to (4), the solution of (9) is

$$\begin{bmatrix} \mathbf{W}_{k,k}^{i+1} \\ \mathbf{G}_{k,-k}^{i+1} \end{bmatrix} = \mathbf{R}_{\tilde{\mathbf{y}}_k^i, \tilde{\mathbf{y}}_k^i}^{-1} \mathbf{R}_{\tilde{\mathbf{y}}_k^i, \mathbf{d}_k} \tag{10}$$

where $\mathbf{R}_{\tilde{\mathbf{y}}_k^i, \tilde{\mathbf{y}}_k^i} = E\{\tilde{\mathbf{y}}_k^i [\tilde{\mathbf{y}}_k^i]^H\}$ and $\mathbf{R}_{\tilde{\mathbf{y}}_k^i, \mathbf{d}_k} = E\{\tilde{\mathbf{y}}_k^i \mathbf{d}_k^H\}$. For further details concerning the estimation of these second order statistics, we refer to [14] and [17]. In the particular case of speech enhancement, note that the estimation of $\mathbf{R}_{\tilde{\mathbf{y}}_k^i, \mathbf{d}_k}$ may require a multi-speaker voice activity detection, e.g., using [21]-[22].

To ensure that the estimates provided by the DANSE algorithm,

$$\hat{\mathbf{d}}_k^i = \left[\mathbf{W}_{k,k}^i\right]^H \mathbf{y}_k + \left[\mathbf{G}_{k,-k}^i\right]^H \mathbf{z}_{-k}^i \tag{11}$$

converge to the LMMSE estimate $\hat{\mathbf{d}}_k = \widehat{\mathbf{W}}_k \mathbf{y}$, where $\widehat{\mathbf{W}}_k = \mathbf{R}_{\mathbf{yy}}^{-1}\mathbf{R}_{\mathbf{y}\mathbf{d}_k}$, it is assumed that each node-specific signals $\mathbf{d}_k$ fully captures the Q-dimensional latent signal subspace, i.e.,

$$\mathbf{d}_k = \mathbf{A}_k \mathbf{s} \tag{12}$$

with $\mathbf{A}_k$ a full rank $Q \times Q$ matrix. As stated in Section 2, only one signal in $\mathbf{d}_k$ might be of actual interest for node k, while the other signals in $\mathbf{d}_k$ can be seen as auxiliary signals that allow to capture the entire Q-dimensional latent signal subspace.

In many practical situations, node k may observe Q_k out of Q latent source signals with $Q_k < Q$. For instance, in Fig. 1, nodes $\{1, 5\}$ only observe 1 out of 2 latent sources, respectively. In these settings, if the node-specific signal $\mathbf{d}_k$ corresponds to a filtered version of the latent source signals $\mathbf{s}$ as they impinge on the sensors of node k, (i.e., $\mathbf{A}_k$ consists of Q rows of $\mathbf{B}_k$), (12) only holds for a rank-deficient matrix $\mathbf{A}_k$. In this case, the convergence of the DANSE algorithm to the LMMSE estimates $\{\hat{\mathbf{d}}_k\}_{k=1}^N$ cannot be ensured and any possible convergence point can be shown to be suboptimal [16]. Next, we will show how to overcome this difficulty.

Assuming that the non-zero columns in $\mathbf{A}_k$ are drawn from a continuous distribution, the rank of $\mathbf{A}_k$ equals Q_k almost surely. Hence, from a signal $\mathbf{x}_{k,\mathrm{dir}}$ that consists of the desired component of Q_k channels of the $\mathbf{y}_k$, node k can almost surely capture Q_k out of the total Q dimensions of the latent signal subspace. Fortunately, by using a technique similar to the one employed in [17], the remaining $Q - Q_k$ dimensions of the latent signal subspace can be captured by a $(Q - Q_k)$-channel signal $\mathbf{x}_{k,\mathrm{ind}}$ received from the other nodes in the network. In particular, the entries of $\mathbf{x}_{k,\mathrm{ind}}$ correspond to the desired components in the signal(s) $\mathbf{z}_\ell^i$ with $\ell \neq k$ and ℓ belonging to a set $\mathcal{B}_j$ where node k is not included. In this way, although node k may not observe all latent source signals, it can still verify (12) if it re-defines its node-specific Q-channel signal as follows

$$\mathbf{d}_k^i = \begin{bmatrix} \mathbf{x}_{k,\mathrm{dir}} \\ \mathbf{x}_{k,\mathrm{ind}}^i \end{bmatrix} = \mathbf{A}_k^i \mathbf{s} \tag{13}$$

where $\mathbf{A}_k^i$ equals a full rank $Q \times Q$ matrix, where one of the entries in $\mathbf{x}_{k,\mathrm{dir}}$ equals the node-specific desired signal d_k, and where, for $p \in \{1, 2, \ldots, Q - Q_k\}$, $l \in \{1, 2, \ldots, Q_\ell\}$,

$$\mathbf{x}_{k,\mathrm{ind}}^i(p) = \left[\mathbf{W}_{\ell,\ell}^i(l)\right]^H \mathbf{B}_\ell \mathbf{s} \tag{14}$$

with $\mathbf{x}_{k,\mathrm{ind}}^i(p)$ equal to the p-th entry of $\mathbf{x}_{k,\mathrm{ind}}^i$ and $\mathbf{W}_{\ell,\ell}^i(l)$ denoting the l-th column of $\mathbf{W}_{\ell,\ell}^i$. In order to have a full-rank matrix $\mathbf{A}_k^i$, note that the indices ℓ and l in (14) need to be suitably chosen.

Since each entry of $\mathbf{x}_{k,\mathrm{ind}}^i$ is an output of the adaptive filter $\mathbf{W}_{\ell,\ell}^i(l)$ in another node ℓ, the full-rank matrix $\mathbf{A}_k^i$ varies at each iteration i instead of being fixed, as it is considered for the convergence of the DANSE algorithm derived in [14]. Despite this fact, from the results in [17], it can be easily shown that the re-definition of the node-specific signals as in (13) ensures the convergence and optimality of the DANSE algorithm to the optimal LMMSE estimates $\{\hat{d}_k\}_{k=1}^N$ in the PO-NSSE problem of Section 2 (details omitted). Additionally, from Theorem 1 we can prove that the proposed strategy achieves the optimal compression rate M_k/L_k^o if $P_k = Q$ for all k. Although the proposed algorithm still converges to the optimal solution in a setting where $P_k < Q$, note that its factor of compression is not optimal anymore. In this case, future research is needed to design algorithms that achieve better rates of compression.

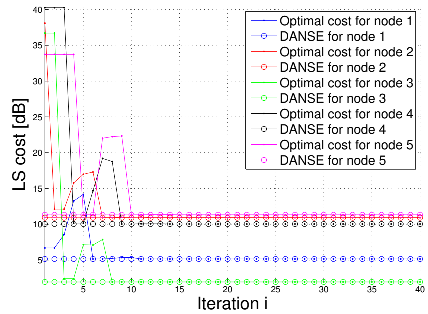

Fig. 2. LS error for each node in the network of Fig. 1.

5. SIMULATIONS

In this section, we illustrate the effectiveness of the proposed algorithm based on the network of Fig. 1, which consists of 5 nodes and $Q = 2$ latent sources. We have implemented a batch-mode version of the algorithm, meaning that $\mathbf{R}_{\tilde{\mathbf{y}}_k^i, \tilde{\mathbf{y}}_k^i}$ and $\mathbf{R}_{\tilde{\mathbf{y}}_k^i, \mathbf{d}_k}$, are computed over the full signal length, T, in all iterations. Assuming that d_k consists of the mixture of the latent source signals that are observed in a reference sensor of node k, the Q-dimensional latent signal subspace can be captured by each node if the node-specific desired signals in (13) are defined as follows. For each node $k \in \{1, 2, \ldots, 5\}$, the first entry of $\mathbf{x}_{k,\mathrm{dir}}$ corresponds with d_k, while the rest of the entries equal the desired component in the signal observed by $Q_k - 1$ auxiliary sensors of node k. Moreover, $\mathbf{x}_{1,\mathrm{ind}}^i = \left[\mathbf{W}_{2,2}^i(1)\right]^H \mathbf{B}_2\mathbf{s}$ and $\mathbf{x}_{5,\mathrm{ind}}^i = \left[\mathbf{W}_{3,3}^i(1)\right]^H \mathbf{B}_3\mathbf{s}$. Since nodes 2, 3, 4 can observe all latent sources, notice that $\mathbf{d}_k^i = \mathbf{x}_{k,\mathrm{dir}}$ for $k \in \{2, 3, 4\}$.

The results of the computer simulations are provided in Fig. 2, which shows the least-squares (LS) cost function of each node k, i.e. $\sum_{t=0}^{T} |d_k[t] - [\mathbf{W}_k(1)^i]^H \mathbf{y}[t]\,|^2$ with $\mathbf{W}_k(1)^i$ equal to the first column of $\mathbf{W}_k^i$, as a function of the iteration index i. For the simulations shown in Fig. 2, the matrices $\mathbf{W}_{k,k}^0$ and $\mathbf{G}_{k,-k}^0$ have been randomly initialized. We have also considered that $T = 1000$, that the elements in $\mathbf{B}_k$ (and hence also of $\mathbf{a}_k$) are generated by a uniform random process on the unit interval and that each latent source signal s_j is a uniformly random process on the interval [-0.5,0.5]. Moreover, the noise component in each of the sensor signals has been independently generated according to a Gaussian distribution of zero mean and variance chosen so that the Signal-to-Noise Ratio (SNR) at each node ranges from 5 to 10 dB. As expected, in Fig. 2 the proposed DANSE algorithm converges to the optimal linear LS solution of a PO-NSSE problem where the nodes have partially-overlapping interests and/or where some nodes do not observe all latent sources.

6. CONCLUSION

In this paper, we have addressed a generalization of the NSSE problem in a WSN where the nodes may have partially overlapping estimation interests and/or may not observe all latent sources. We have provided the minimum number of signals that at least have to be broadcast by node k in order to allow the other nodes to achieve the network-wide LMMSE estimates of its node-specific desired signals. Additionally, we have described a distributed algorithm that achieves the network-wide solution of the considered NSSE problem by broadcasting the minimum number of signals per node under some settings. Finally, the effectiveness of the proposed algorithm has been illustrated through computer simulations.

7. REFERENCES

[1] 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.

[2] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[3] 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.

[4] 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.

[5] S. Chouvardas, K. Slavakis, and S. Theodoridis, "Adaptive robust distributed learning in diffusion sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4692–4707, 2011.

[6] A. Bertrand, *Signal processing algorithms for Wireless Acoustic Sensor Networks*, Ph.D. thesis, KU Leuven, University of Leuven, Belgium, May, 2011.

[7] P. Di Lorenzo, S. Barbarossa, and A. H. Sayed, "Bio-inspired swarming for dynamic radio access based on diffusion adaptation," in *IEEE 19th European Signal Conference, 2011. EUSIPCO 2011*, 2012, pp. 1–6.

[8] P. Di Lorenzo, S. Barbarossa, and A. H. Sayed, "Bio-inspired decentralized radio access based on swarming mechanisms over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 12, pp. 3183–3197, 2013.

[9] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2013.

[10] N. Bogdanovic, J. Plata-Chaves, and K. Berberidis, "Distributed incremental-based LMS for node-specific adaptive parameter estimation," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5382–5397, 2014.

[11] J. Plata-Chaves, N. Bogdanovic, and K. Berberidis, "Distributed diffusion-based LMS for node-specific parameter estimation over adaptive networks," *Available: http://arxiv.org/abs/1408.3354*, 2014.

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

[13] R. Abdolee, B. Champagne, and A. H. Sayed, "Diffusion LMS for source and process estimation in sensor networks," in *IEEE/SP 17th Workshop on Statistical Signal Processing, 2012. SSP 2012*, 2012, pp. 165–168.

[14] A. Bertrand and M. Moonen, "Distributed adaptive node-specific signal estimation in fully connected sensor networks - part I: Sequential node updating," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5277–5291, 2010.

[15] A. Bertrand and M. Moonen, "Distributed adaptive estimation of node-specific signals in wireless sensor networks with a tree topology," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2196–2210, 2011.

[16] A. Bertrand and M. Moonen, "Distributed signal estimation in sensor networks where nodes have different interests," *Signal Processing*, vol. 92, no. 7, pp. 1679–1690, 2012.

[17] A. Bertrand and M. Moonen, "Robust distributed noise reduction in hearing aids with external acoustic sensor nodes," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 12, 2009.

[18] A. Hassani, A. Bertrand, and M. Moonen, "Cooperative integrated noise reduction and node-specific direction-of-arrival estimation in a fully connected wireless acoustic sensor network," *Signal Processing*, vol. 107, no. 2, pp. 68–81, 2015.

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

[20] A. Bertrand and M. Moonen, "Distributed adaptive node-specific signal estimation in fully connected sensor networks - part II: Simultaneous and asynchronous node updating," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5292–5306, 2010.

[21] A. Bertrand and M. Moonen, "Energy-based multi-speaker voice activity detection with an ad hoc microphone array," in *IEEE 35th International Conference on Acoustics, Speech and Signal Processing, 2010. ICASSP 2010*, 2010.

[22] S. Chouvardas, M. Muma, K. Hamaidi, S. Theodoridis, and A. M. Zoubir, "Distributed robust labeling of audio sources in heterogeneous wireless sensor networks," in *To appear in IEEE 40th International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015.