

# DISTRIBUTED MAP AND LMMSE ESTIMATION OF RANDOM SIGNALS USING AD HOC WIRELESS SENSOR NETWORKS WITH NOISY LINKS

*Ioannis D. Schizas, Georgios B. Giannakis and Alejandro Ribeiro*

Dept. of ECE, University of Minnesota, USA

## ABSTRACT

Distributed estimation of random parameter vector is dealt with using ad hoc wireless sensor networks (WSNs). The decentralized estimation problem is cast as the solution of multiple convex optimization subproblems and the alternating direction method of multipliers is employed to derive algorithms which can be decomposed into a set of simpler tasks suitable for distributed implementation. Different from existing alternatives, the novel approach does not require knowing the desired estimator in closed-form as is generally the case with the maximum a posteriori estimator (MAP). In addition, a priori information is accounted for and sensor observations are allowed to be correlated. The resulting algorithms converge to the centralized estimators under ideal channel links, while they exhibit noise robustness provably established for the distributed linear minimum mean-square error estimator (LMMSE).

## 1. INTRODUCTION

A popular application of WSNs is decentralized estimation of deterministic parameters or random signals using observation data collected across sensors. The estimation task can be performed iteratively in a distributed fashion based on successive refinements of local estimates maintained at individual sensors. Each iteration of the estimation algorithm comprises a communication step where the sensors interchange information with their neighbors, and an update step where each sensor uses this information to update its local estimate. In this context, estimation of deterministic parameters in linear Gaussian models was considered in [9] using the notion of consensus averaging. Decentralized estimation of Gaussian random parameters in a scalar linear model was also reported in [3]. Distributed estimation of Markov random fields was pursued in [4], where each sensor estimates locally a random parameter.

However, consensus averaging schemes are challenged by the presence of noise (non-ideal sensor links), exhibiting a statistical behavior similar to that of a random walk, and eventually diverging [9]. Furthermore, [3, 4] do not account for communication noise.

Here we focus on decentralized estimation of random signals in general (possibly nonlinear and/or non-Gaussian) data models. Novelty of our approach include: i) formulation of the desired estimator as the solution of judiciously designed convex minimization subproblems that have a separable structure and are thus amenable to distributed processing; ii) decentralized MAP algorithms even when the desired estimator is not available in closed form, a case where [1

and [9] do not apply; iii) provable robustness under communication noise, and iv) ability to incorporate a priori information and account for correlated sensor data (not available with [1, 6, 9]). Different from [4] our distributed MAP and LMMSE estimators account for communication noise, while LMMSE also allows for arbitrary sensor data correlations.

Specifically, we view MAP estimation as the optimal solution of a separable constrained convex minimization problem in Section 3, and utilize the alternating direction method of multipliers to construct the corresponding distributed algorithm. In Section 4 we consider distributed LMMSE estimation, for which we develop noise-robust distributed algorithms. Numerical results in Section 5 corroborate our theoretical findings.

## 2. PROBLEM FORMULATION AND PRELIMINARIES

Consider an ad hoc WSN comprising  $J$  sensors, where only single-hop communications are allowed; i.e., the  $j$ -th sensor communicates solely with nodes  $j'$  in its neighborhood  $\mathcal{N}_j \subseteq [1, J]$ . Communication links are symmetric and the WSN is modelled as an undirected graph whose vertices are the sensors and its edges represent the available links; see Fig. 1. The connectivity information is summarized in the so called adjacency matrix  $\mathbf{E} \in \mathbb{R}^{J \times J}$  with entries  $\mathbf{E}_{jj'} = \mathbf{E}_{j'j} = 1$  if  $j' \in \mathcal{N}_j$ , while  $\mathbf{E}_{jj'} = \mathbf{E}_{j'j} = 0$  if  $j' \notin \mathcal{N}_j$ .

The WSN is deployed to estimate a  $p \times 1$  random parameter vector  $\mathbf{s}$  based on distributed random observations  $\{\mathbf{x}_j \in \mathbb{R}^{L_j \times 1}\}_{j=1}^J$ . Without loss of generality both  $\mathbf{s}$  and  $\mathbf{x}_j$  for  $j = 1, \dots, J$  are assumed zero mean. Depending on the available a priori statistical information that the sensors have about  $\mathbf{s}$  we consider two estimation scenarios. In the first one, the a priori information about  $\mathbf{s}$  is summarized in the probability density function (pdf)  $p(\mathbf{s})$ , which is assumed known to all sensors. Conditioned on  $\mathbf{s}$ ,  $\{\mathbf{x}_j\}_{j=1}^J$  are assumed independent across sensors, and the associated pdfs  $p(\mathbf{x}_j|\mathbf{s})$  are known  $\forall j = 1, \dots, J$ . Under these assumptions, the MAP estimator can be written as

$$\hat{\mathbf{s}}_{\text{map}} := \arg \min_{\mathbf{s} \in \mathbb{R}^{p \times 1}} - \sum_{j=1}^J (\ln[p_j(\mathbf{x}_j|\mathbf{s})] + J^{-1} \ln[p(\mathbf{s})]). \quad (1)$$

The second scenario arises when only the first and second order statistics associated with  $\mathbf{s}$  and  $\mathbf{x}_j$  are known across the network; i.e., when sensors know only the (cross-) covariance matrices  $\mathbf{C}_{xx} := E[\mathbf{x}\mathbf{x}^T]$ ,  $\mathbf{C}_{sx} := E[\mathbf{s}\mathbf{x}^T]$  and  $\mathbf{C}_{ss} := E[\mathbf{s}\mathbf{s}^T]$ , where  $\mathbf{x} := [\mathbf{x}_1^T \dots \mathbf{x}_J^T]^T$  contains all the  $L := \sum_{j=1}^J L_j$  sensor observations. Unlike [1, 6, 9] where the sensor data are assumed to be uncorrelated,  $\mathbf{C}_{xx}$  here allows for arbitrary correlation among sensor data. Specifically, sensor  $j$  has available  $\mathbf{C}_{x_j} := [\mathbf{C}_{x_j x_1} \dots \mathbf{C}_{x_j x_J}]$  and  $\mathbf{C}_{s x_j}$ . These matrices can be acquired either from the physics of the problem, or, during a training phase. Notwithstanding, each sensor does not have to know the entire matrices  $\mathbf{C}_{xx}$  and  $\mathbf{C}_{sx}$  but only part of

Prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

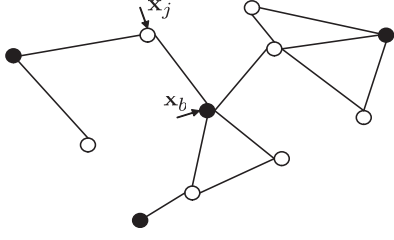


Fig. 1. An ad-hoc wireless sensor network.

them containing  $1/J$  of the total covariance information. A pertinent approach in this setup is to form the LMMSE, see e.g., [5]

$$\hat{\mathbf{s}}_{\text{lmmse}} := \mathbf{C}_{sx} \mathbf{C}_{xx}^{-1} \mathbf{x}. \quad (2)$$

We will develop iterative algorithms based on communication with one-hop neighbors that generate (local) time iterates  $\mathbf{s}_j(k)$  so that:

(s1) If the  $j$ -th sensor knows only  $p_j(\mathbf{x}_j|\mathbf{s})$  and  $p(\mathbf{s})$ , the local iterates converge, i.e.,  $\lim_{k \rightarrow \infty} \mathbf{s}_j(k) = \hat{\mathbf{s}}_{\text{map}}$ .

(s2) If only  $\mathbf{C}_{x_j}$  and  $\mathbf{C}_{sx_j}$  are known at the  $j$ -th sensor and the covariance matrix  $\mathbf{C}_{xx}$  is full rank, then  $\lim_{k \rightarrow \infty} \mathbf{s}_j(k) = \hat{\mathbf{s}}_{\text{lmmse}}$ .

The decentralized MAP estimator in (s1) is of particular importance for estimation in nonlinear data models since it is optimal in the sense of minimizing the ‘hit-or-miss’ Bayes risk (see e.g., [5, pg. 372]). The LMMSE estimator is MSE optimal when considered within the class of linear estimators; however, it remains applicable even when only  $\{\mathbf{C}_{sx_j}\}_{j=1}^J$  and  $\{\mathbf{C}_{x_j x_j}\}_{j=1}^J$  are known. Clearly, if  $\{\mathbf{x}_j\}_{j=1}^J$  and  $\mathbf{s}$  are jointly Gaussian, then  $\hat{\mathbf{s}}_{\text{map}} \equiv \hat{\mathbf{s}}_{\text{lmmse}}$ .

The local iterates  $\mathbf{s}_j(k)$  will turn out to exhibit robustness to communication noise. Specifically, if  $\mathbf{t}_{j'}^j(k) \in \mathbb{R}^{p \times 1}$  represents<sup>1</sup> a vector transmitted from the  $j$ -th to the  $j'$ -th sensor at time slot  $k$ , the corresponding vector  $\mathbf{r}_{j'}^j(k) \in \mathbb{R}^{p \times 1}$  received by the  $j'$ -th sensor is

$$\mathbf{r}_{j'}^j(k) = \mathbf{t}_{j'}^j(k) + \boldsymbol{\eta}_{j'}^j(k) \quad (3)$$

where  $\boldsymbol{\eta}_{j'}^j(k) \in \mathbb{R}^{p \times 1}$  denotes zero-mean additive white noise at sensor  $j'$ . Vector  $\boldsymbol{\eta}_{j'}^j(k)$  is assumed uncorrelated across sensors and time with covariance matrix  $\mathbf{C}_{\boldsymbol{\eta}_{j'}^j}$ . Additionally, we assume that:

(a1) The communication graph is connected.

(a2) The pdfs  $\{p(\mathbf{x}_j|\mathbf{s})\}_{j=1}^J$  and  $p(\mathbf{s})$  are log-concave in  $\mathbf{s}$ .

Similar to [1, 9], (a1) ensures utilization of all observation vectors by the decentralized scheme, while (a2) is to guarantee uniqueness even for the centralized MAP estimator and is satisfied by a number of unimodal pdfs encountered in practice; see e.g., [8].

### 3. DISTRIBUTED MAP

In this section we consider distributed computation of  $\hat{\mathbf{s}}_{\text{map}}$  in (s1), under (a1) and (a2). Since summands in (1) are coupled through  $\mathbf{s}$ , it is not straightforward to decompose the unconstrained optimization problem in (1). We overcome this coupling by introducing the auxiliary variable  $\mathbf{s}_j$  to represent the local estimate of  $\mathbf{s}$  at sensor  $j$  and consider the constrained optimization problem

$$\{\hat{\mathbf{s}}_j\}_{j=1}^J := \arg \min_{\mathbf{s}_j} - \sum_{j=1}^J (\ln[p_j(\mathbf{x}_j|\mathbf{s}_j)] + J^{-1} \ln[p(\mathbf{s}_j)]) \quad (4)$$

s. to  $\mathbf{s}_j = \bar{\mathbf{s}}_b$ ,  $b \in \mathcal{B}$ ,  $j \in \mathcal{N}_b$

<sup>1</sup>Subscripts specify the sensor at which variables are computed at, while superscripts denote the sensor at which the variable is communicated to.

where  $\mathcal{B} \subseteq [1, J]$  is a subset of ‘‘bridge’’ sensors maintaining local vectors  $\bar{\mathbf{s}}_b$  that are utilized to impose consensus among local variables  $\mathbf{s}_j$  across all sensors. If e.g.,  $\mathcal{B} = [1, J]$ , then (a1) and the constraint  $\mathbf{s}_j = \bar{\mathbf{s}}_b$ ,  $b \in \mathcal{B}$ ,  $j \in \mathcal{N}_b$  will render  $\mathbf{s}_j = \mathbf{s}_i \forall i, j$ . In such a case (1) and (4) are equivalent in the sense that  $\hat{\mathbf{s}}_j = \hat{\mathbf{s}}_{\text{map}} \forall j = 1, \dots, J$ . Interestingly, a milder requirement on  $\mathcal{B}$  is sufficient to ensure equivalence of (1) and (4), as described next:

**Definition 1** Set  $\mathcal{B}$  is a subset of bridge sensors if and only if

- (a)  $\forall j \in [1, J]$  there exists at least one  $b \in \mathcal{B}$  so that  $b \in \mathcal{N}_j$ ; and
- (b)  $\forall b_1 \in \mathcal{B}$  there exists a sensor  $b_2 \in \mathcal{B}$  such that the shortest path between  $b_1$  and  $b_2$  has at most two edges.

For the WSN in Fig. 1 a possible selection of sensors to form a bridge sensor subset  $\mathcal{B}$  is represented by the black nodes. Henceforth, the set of bridge neighbors of the  $j$ -th sensor will be denoted as  $\mathcal{B}_j := \mathcal{N}_j \cap \mathcal{B}$ , and its cardinality by  $|\mathcal{B}_j|$  for  $j = 1, \dots, J$ . In words, condition (a) in Definition 1 ensures that every node has at least one bridge sensor neighbor; while condition (b) ensures that all the bridge variables  $\{\bar{\mathbf{s}}_b\}_{b \in \mathcal{B}}$  are equal. Together, they provide a necessary and sufficient condition for the equivalence between (1) and (4) as asserted by the following proposition [7].

**Proposition 1** The optimal solutions of (1) and (4) coincide; i.e.,

$$\hat{\mathbf{s}}_{\text{map}} = \hat{\mathbf{s}}_j, \quad \forall j = 1, \dots, J, \quad (5)$$

if and only if  $\mathcal{B}$  is a subset of bridge sensors.

#### 3.1. Alternating Direction Method of Multipliers

Here we solve (1) using the alternating direction method of multipliers, through which we obtain a distributed algorithm for computing  $\hat{\mathbf{s}}_{\text{map}}$ . The method of multipliers exploits the decomposable structure of the augmented Lagrangian (see e.g., [2, pg. 253]). To this end, let  $\mathbf{v}_j^b$  denote the Lagrange multiplier associated with the constraint  $\mathbf{s}_j = \bar{\mathbf{s}}_b$ , where the multipliers  $\{\mathbf{v}_j^b\}_{b \in \mathcal{B}_j}$  are kept at the  $j$ -th sensor. The augmented Lagrangian for (4) is

$$\mathcal{L}_a[\mathbf{s}, \bar{\mathbf{s}}, \mathbf{v}] = - \sum_{j=1}^J (\ln[p_j(\mathbf{x}_j|\mathbf{s}_j)] + J^{-1} \ln[p(\mathbf{s}_j)]) \quad (6)$$

$$+ \sum_{b \in \mathcal{B}} \sum_{j \in \mathcal{N}_b} (\mathbf{v}_j^b)^T (\mathbf{s}_j - \bar{\mathbf{s}}_b) + \sum_{b \in \mathcal{B}} \sum_{j \in \mathcal{N}_b} \frac{c_j}{2} \|\mathbf{s}_j - \bar{\mathbf{s}}_b\|_2^2$$

where  $\mathbf{s} := \{\mathbf{s}_j\}_{j=1}^J$ ,  $\bar{\mathbf{s}} := \{\bar{\mathbf{s}}_b\}_{b \in \mathcal{B}}$ ,  $\mathbf{v} := \{\mathbf{v}_j^b\}_{j \in [1, J], b \in \mathcal{B}_j}$  and the constants  $\{c_j > 0\}_{j=1}^J$  are penalty coefficients corresponding to the constraints  $\mathbf{s}_j = \bar{\mathbf{s}}_b$ ,  $\forall b \in \mathcal{B}_j$ . Combining the method of multipliers with a block coordinate descent iteration [2, Chpt.3], we have established the following result [7].

**Proposition 2** For a time index  $k$ , let iterates  $\mathbf{v}_j^b(k)$ ,  $\mathbf{s}_j(k)$  and  $\bar{\mathbf{s}}_b(k)$  be defined by the recursions

$$\mathbf{v}_j^b(k) = \mathbf{v}_j^b(k-1) + c_j [\mathbf{s}_j(k) - \bar{\mathbf{s}}_b(k)], \quad b \in \mathcal{B}_j \quad (7)$$

$$\mathbf{s}_j(k+1) = \arg \min_{\mathbf{s}_j} [-\ln[p_j(\mathbf{x}_j|\mathbf{s}_j)] - J^{-1} \ln[p(\mathbf{s}_j)]$$

$$+ \sum_{b \in \mathcal{B}_j} \frac{c_j}{2} \|\mathbf{s}_j - \bar{\mathbf{s}}_b(k)\|_2^2 + \sum_{b \in \mathcal{B}_j} (\mathbf{v}_j^b(k))^T [\mathbf{s}_j - \bar{\mathbf{s}}_b(k)] \quad (8)$$

$$\bar{\mathbf{s}}_b(k+1) = \sum_{j \in \mathcal{N}_b} \frac{1}{\sum_{\beta \in \mathcal{N}_b} c_\beta} \left[ \mathbf{v}_j^b(k) + c_j \mathbf{s}_j(k+1) \right], \quad b \in \mathcal{B} \quad (9)$$

$\forall j = 1, \dots, J$ ; and let the initial values of the Lagrange multipliers  $\{\mathbf{v}_j^b(-1)\}_{b \in \mathcal{B}_j}$ , the local estimates  $\{\mathbf{s}_j(0)\}_{j=1}^J$  and the consensus variables  $\{\bar{\mathbf{s}}_b(0)\}_{b \in \mathcal{B}}$  be arbitrary. Under (a1), (a2) and ideal communication links, the iterates  $\mathbf{s}_j(k)$  converge to the centralized  $\hat{\mathbf{s}}_{\text{map}}$  as  $k \rightarrow \infty$  and the WSN reaches consensus; i.e.,

$$\lim_{k \rightarrow \infty} \mathbf{s}_j(k) = \lim_{k \rightarrow \infty} \bar{\mathbf{s}}_b(k) = \hat{\mathbf{s}}_{\text{map}}, \quad \forall j \in [1, J], b \in \mathcal{B}. \quad (10)$$

Local recursions (7)-(9) form the distributed (D-) MAP algorithm whereby all sensors  $j \in [1, J]$  keep track of the local estimate  $\mathbf{s}_j(k)$  along with the Lagrange multipliers  $\{\mathbf{v}_j^b(k)\}_{b \in \mathcal{B}_j}$ . The sensors belonging to  $\mathcal{B}$  also update the consensus-enforcing variables  $\bar{\mathbf{s}}_b(k)$ . During iteration  $k$ , the  $j$ -th sensor receives the consensus variables  $\bar{\mathbf{s}}_b(k)$  from all its bridge neighbors in  $b \in \mathcal{B}_j$ . Based on these consensus variables, it updates the Lagrange multipliers  $\{\mathbf{v}_j^b(k)\}_{b \in \mathcal{B}_j}$  using (7), which are then utilized to compute  $\mathbf{s}_j(k+1)$  via (8). Then sensor  $j$  transmits to each of its neighbors  $b \in \mathcal{B}_j$  the vector  $\mathbf{v}_j^b(k) + c_j \mathbf{s}_j(k+1)$ . Each sensor  $b \in \mathcal{B}$  receives  $\mathbf{v}_j^b(k) + c_j \mathbf{s}_j(k+1)$  from all its neighbors  $j \in \mathcal{N}_b$ , and proceeds to compute  $\bar{\mathbf{s}}_b(k+1)$  using (9). This completes the  $k$ -th iteration, after which all sensors in  $\mathcal{B}$  transmit  $\bar{\mathbf{s}}_b(k+1)$  to all their neighbors  $j \in \mathcal{N}_b$ , which can proceed to the  $(k+1)$ -st iteration. Notice, that the optimization problem (8) is strictly convex, implying that the optimal solution  $\mathbf{s}_j(k+1)$  is unique and can be obtained using e.g., Newton's method.

In the presence of communication errors,  $\bar{\mathbf{s}}_b(k)$  and  $\mathbf{s}_j(k)$  in (7)-(9) are corrupted as described in (3). Then, these recursions can be thought of as a stochastic gradient algorithm; see e.g., [2, Sec. 7.8]. This implies that noise causes  $\mathbf{s}_j(k)$  to fluctuate around  $\hat{\mathbf{s}}_{\text{map}}$  with the magnitude of fluctuations being proportional to the noise variance. However,  $\mathbf{s}_j(k)$  is guaranteed to remain within a ball around  $\hat{\mathbf{s}}_{\text{map}}$  with high probability [7]. This should be contrasted with the consensus averaging scheme in [9] which suffers from catastrophic noise propagation.

#### 4. NOISE-ROBUST DISTRIBUTED LMMSE

In this section we consider distributed computation of  $\hat{\mathbf{s}}_{\text{lmmse}}$  in (s2) under (a1), following an approach similar to the one in Section 3. To this end, we have shown in [7] that  $\hat{\mathbf{s}}_{\text{lmmse}}$  can be obtained as the optimal solution of

$$\begin{aligned} \{\hat{\mathbf{s}}_j, \hat{\mathbf{y}}_j\}_{j=1}^J &:= \arg \min \sum_{j=1}^J \|\mathbf{s}_j - J\mathbf{C}_{s_{x_j}} \mathbf{y}_j^j\|_2^2 \quad (11) \\ \text{s. to } \mathbf{s}_j &= \bar{\mathbf{s}}_b, \quad \mathbf{y}_j = \bar{\mathbf{y}}_b, \quad b \in \mathcal{B}, \quad j \in \mathcal{N}_b \\ \mathbf{C}_{x_j} \mathbf{y}_j &= \mathbf{x}_j, \quad j = 1, \dots, J \end{aligned}$$

where  $\hat{\mathbf{s}}_j = \hat{\mathbf{s}}_{\text{lmmse}}$  and  $\hat{\mathbf{y}}_j = \mathbf{C}_{x_j}^{-1} \mathbf{x}_j$ , while  $\mathbf{y}_j := [(\mathbf{y}_j^1)^T \dots (\mathbf{y}_j^J)^T]^T$  with  $\mathbf{y}_j^i \in \mathbb{R}^{L_i \times 1}$  for  $i = 1, \dots, J$ . Note that the optimization formulation in (11) involves the additional local variables  $\mathbf{y}_j$  at sensor  $j$ , as well as their consensus counterpart  $\bar{\mathbf{y}}_b$  located at the  $b$ -th bridge sensor. Intuitively, the role of  $\{\mathbf{y}_j\}_{j=1}^J$  is to encapsulate the covariance information embedded in  $\mathbf{C}_{x_j}^{-1} \mathbf{x}_j$ , which is vital for computing  $\hat{\mathbf{s}}_{\text{lmmse}}$  in (2). Introducing also the Lagrange multiplier  $\mathbf{w}_j^b$  associated with the constraint  $\mathbf{y}_j = \bar{\mathbf{y}}_b$  that is maintained at the  $j$ -th sensor with  $b \in \mathcal{B}_j$ , and letting  $d_j$  denote the corresponding penalty coefficient, we can form the augmented Lagrangian function for (11) as in (6). Utilizing once more, the alternating direction method of multipliers, we have proved the following result [7].

**Proposition 3** For each sensor  $j \in [1, J]$ , and iteration  $k$  let iterates  $\mathbf{v}_j^b(k)$ ,  $\mathbf{w}_j^b(k)$ ,  $\mathbf{s}_{y,j}(k)$  and  $\bar{\mathbf{s}}_{y,b}(k)$  be defined by the recursions

$$[\mathbf{v}_j^b(k)^T, \mathbf{w}_j^b(k)^T]^T = [(\mathbf{v}_j^b(k-1))^T, (\mathbf{w}_j^b(k-1))^T]^T \quad (12)$$

$$+ [c_j(\mathbf{s}_j(k) - \bar{\mathbf{s}}_b(k))^T, d_j(\mathbf{y}_j(k) - \bar{\mathbf{y}}_b(k))^T]^T, \quad b \in \mathcal{B}_j$$

$$\begin{aligned} \mathbf{s}_{y,j}(k+1) &:= [(\mathbf{s}_j(k+1))^T, (\mathbf{y}_j(k+1))^T]^T \\ &= \hat{\mathbf{x}}_j - \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\zeta_j(k) \end{aligned} \quad (13)$$

$$\bar{\mathbf{s}}_{y,b}(k+1) := [(\bar{\mathbf{s}}_b(k+1))^T, (\bar{\mathbf{y}}_b(k+1))^T]^T \quad (14)$$

$$\begin{aligned} &= \sum_{j \in \mathcal{N}_b} \text{diag}((\sum_{\beta \in \mathcal{N}_b} c_\beta)^{-1} \mathbf{I}_p, (\sum_{\beta \in \mathcal{N}_b} d_\beta)^{-1} \mathbf{I}_L) \\ &\times \left( [(\mathbf{v}_j^b(k))^T, (\mathbf{w}_j^b(k))^T]^T + \text{diag}(c_j \mathbf{I}_p, d_j \mathbf{I}_L) \mathbf{s}_{y,j}(k+1) \right) \end{aligned}$$

where  $\text{diag}(\cdot)$  denotes a diagonal matrix and

$$\mathbf{F}_j := 2 \begin{bmatrix} (1 + 0.5c_j |\mathcal{B}_j|) \mathbf{I}_p & -J\bar{\mathbf{C}}_{s_{x_j}}^T \\ -J\bar{\mathbf{C}}_{s_{x_j}} & J^2 \bar{\mathbf{C}}_{s_{x_j}} \bar{\mathbf{C}}_{s_{x_j}}^T + 0.5d_j |\mathcal{B}_j| \mathbf{I}_L \end{bmatrix}$$

$$\zeta_j(k) := \left[ \sum_{b \in \mathcal{B}_j} (\mathbf{v}_j^b(k) - c_j \bar{\mathbf{s}}_b(k))^T \sum_{b \in \mathcal{B}_j} (\mathbf{w}_j^b(k) - d_j \bar{\mathbf{y}}_b(k))^T \right]^T$$

$$\hat{\mathbf{x}}_j := \mathbf{F}_j^{-1} \bar{\mathbf{C}}_{x_j}^T (\bar{\mathbf{C}}_{x_j} \mathbf{F}_j^{-1} \bar{\mathbf{C}}_{x_j}^T)^{-1} \mathbf{x}_j$$

$$\mathbf{G}_j := \bar{\mathbf{C}}_{x_j}^T (\bar{\mathbf{C}}_{x_j} \mathbf{F}_j^{-1} \bar{\mathbf{C}}_{x_j}^T)^{-1} \bar{\mathbf{C}}_{x_j} \mathbf{F}_j^{-1}$$

with  $\bar{\mathbf{C}}_{s_{x_j}} := [\mathbf{0}_{p \times L_1} \dots \mathbf{C}_{s_{x_j}} \dots \mathbf{0}_{p \times L_J}]^T$  and  $\bar{\mathbf{C}}_{x_j} := [\mathbf{0}_{L_j \times p} \mathbf{C}_{x_j}]$ , while  $\{c_j, d_j > 0\}_{j=1}^J$ . Assuming ideal links and the validity of (a1), the iterates  $\mathbf{s}_j(k)$  converge to the LMMSE as  $k \rightarrow \infty$ ; i.e., for all  $j \in [1, J]$  and  $b \in \mathcal{B}$

$$\begin{aligned} \lim_{k \rightarrow \infty} [\mathbf{s}_j(k)^T, \mathbf{y}_j(k)^T]^T &= \lim_{k \rightarrow \infty} [\bar{\mathbf{s}}_b(k)^T, \bar{\mathbf{y}}_b(k)^T]^T \\ &= [\hat{\mathbf{s}}_{\text{lmmse}}^T (\mathbf{C}_{x_j}^{-1} \mathbf{x}_j)^T]^T \end{aligned}$$

Following steps analogous to those described for D-MAP, the algorithm described by the local recursions (12)-(14) performs distributed (D-) LMMSE estimation. Indeed, recursions (12)-(14) exhibit similar structure with those in (7)-(9). Thus, the message exchanging protocol that takes place among neighboring sensors during iteration  $k$  of the D-MAP, can also be applied here after replacing  $\mathbf{s}_j(k)$  with  $\mathbf{s}_{y,j}(k)$ ,  $\bar{\mathbf{s}}_b(k)$  with  $\bar{\mathbf{s}}_{y,b}(k)$ , and  $(\mathbf{v}_j^b(k))^T$  with  $[(\mathbf{v}_j^b(k))^T, (\mathbf{w}_j^b(j))^T]^T$ .

Relative to [3], the D-LMMSE algorithm is more general since it neither requires linearity nor Gaussianity in the data model. Furthermore, [3] requires each sensor, say the  $j$ -th, to communicate with all these sensors whose observation data are correlated with  $\mathbf{x}_j$ . For arbitrary  $\mathbf{C}_{ss}$  and  $\mathbf{C}_{xx}$  this may require multi-hop communication. However, D-LMMSE guarantees convergence to  $\hat{\mathbf{s}}_{\text{lmmse}}$  across all sensors through single-hop communication. Notice also that D-LMMSE incorporates distributed inversion of  $\mathbf{C}_{xx}$ , through the local variables  $\mathbf{y}_j(k)$  which is essential since every sensor has only a portion of  $\mathbf{C}_{xx}$ . A fair comparison between [3] and the present D-LMMSE algorithm does not appear possible, since the former utilizes information from the inverse covariance matrices which facilitates estimation but leaves open the question of how and how costly acquiring this information is.

#### 4.1. Noise-Resilient D-LMMSE

Starting from the D-LMMSE algorithm in Proposition 3, we will derive a provably noise resilient distributed algorithm that is amenable

to convergence analysis. Toward this end, let us initialize the recursions (12)-(14) with  $\{\mathbf{v}_j^b(-1) = \mathbf{0}, \mathbf{w}_j^b(-1) = \mathbf{0}\}_{j \in [1, J], b \in \mathcal{B}_j}^b$ ,  $\{\bar{\mathbf{s}}_{y,b}(-1) = \mathbf{0}\}_{b \in \mathcal{B}}$  and  $\{\mathbf{s}_{y,j}(0) = \hat{\mathbf{x}}_j\}_{j=1}^J$ . Upon substituting (12) in (13) and (14), we find that the D-LMMSE local recursions can be rewritten in a simpler form as follows:

$$\mathbf{s}_{y,j}(k+1) = [\mathbf{I}_{L+p} - \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^1]\mathbf{s}_{y,j}(k) + \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^2 \sum_{b \in \mathcal{B}_j} [2\bar{\mathbf{s}}_{y,b}(k) - \bar{\mathbf{s}}_{y,b}(k-1)], j \in [1, J] \quad (15)$$

$$\bar{\mathbf{s}}_{y,b}(k+1) = \sum_{j \in \mathcal{N}_b} \mathbf{D}_j^2 \mathbf{D}_b \mathbf{s}_{y,j}(k+1), b \in \mathcal{B} \quad (16)$$

where  $\mathbf{D}_j^1 := \text{diag}(c_j |\mathcal{B}_j| \mathbf{I}_p, d_j |\mathcal{B}_j| \mathbf{I}_L)$ ,  $\mathbf{D}_j^2 := \text{diag}(c_j \mathbf{I}_p, d_j \mathbf{I}_L)$  and  $\mathbf{D}_b := \text{diag}((\sum_{\beta \in \mathcal{N}_b} c_\beta)^{-1} \mathbf{I}_p, (\sum_{\beta \in \mathcal{N}_b} d_\beta)^{-1} \mathbf{I}_L)$ . Observe that (15)-(16) are simpler than (12)-(14) in the sense that they avoid the multipliers  $\mathbf{v}_j^b(k)$  and  $\mathbf{w}_j^b(k)$ .

Recursions (15) and (16) constitute an intermediate step based on which we build next a distributed noise-robust algorithm for D-LMMSE. The idea is to introduce an auxiliary local variable  $\varphi_j(k)$  which replaces  $\mathbf{s}_{y,j}(k)$ , and is kept at the  $j$ -th sensor. We will show that the mean of successive differences of  $\varphi_j(k)$  converge to the LMMSE; i.e.,  $\lim_{k \rightarrow \infty} E[\varphi_j(k+1) - \varphi_j(k)] = [\hat{\mathbf{s}}_{\text{lmmse}}^T (\mathbf{C}_{xx}^{-1} \mathbf{x})^T]^T$ , while the covariance matrix of this difference remains bounded. Intuitively, noise terms that propagate from  $\varphi_j(k)$  to  $\varphi_j(k+1)$  cancel when considering the difference  $\varphi_j(k+1) - \varphi_j(k)$ , thus achieving the desired robustness to noise. This is akin to the principle for noise suppression utilized in the approach of coupled oscillators in [1], where a *continuous-time* differential (state) equation is involved per sensor, and the information is encoded in the derivative of the state. The desired *discrete-time* recursion for  $\varphi_j(k)$  is introduced in the following lemma [7].

**Lemma 1** *If  $\varphi_j(0) = \hat{\mathbf{x}}_j$  and  $\varphi_j(-1) = \varphi_j(-2) = \mathbf{0}$ , the second-order recursions*

$$\varphi_j(k+1) = \hat{\mathbf{x}}_j + (\mathbf{I}_{L+p} - \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^1) \varphi_j(k) + \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^2 \sum_{b \in \mathcal{B}_j} [2\bar{\varphi}_b(k) - \bar{\varphi}_b(k-1)] \quad (17)$$

$$\bar{\varphi}_b(k+1) = \sum_{j \in \mathcal{N}_b} \mathbf{D}_j^2 \mathbf{D}_b \varphi_j(k+1), b \in \mathcal{B}, k \geq -1 \quad (18)$$

yield iterates  $\varphi_j(k)$  and  $\bar{\varphi}_b(k)$  whose differences  $\delta\varphi_j(k) := \varphi_j(k) - \varphi_j(k-1)$  and  $\delta\bar{\varphi}_b(k) := \bar{\varphi}_b(k) - \bar{\varphi}_b(k-1)$  equal the iterates  $\mathbf{s}_{y,j}(k)$  and  $\bar{\mathbf{s}}_{y,b}(k)$  in (15) and (16), respectively.

We already know from Proposition 3 that under (a1) and ideal links  $\lim_{k \rightarrow \infty} \mathbf{s}_{y,j}(k) = [\hat{\mathbf{s}}_{\text{lmmse}}^T, (\mathbf{C}_{xx}^{-1} \mathbf{x})^T]^T$ ; thus, using Lemma 1 we deduce that as  $k \rightarrow \infty$  the differences  $\delta\varphi_j(k)$  and  $\delta\bar{\varphi}_b(k)$  converge to the vector  $[\hat{\mathbf{s}}_{\text{lmmse}}^T, (\mathbf{C}_{xx}^{-1} \mathbf{x})^T]^T$ , i.e.,  $\lim_{k \rightarrow \infty} \delta\varphi_j(k) = \lim_{k \rightarrow \infty} \delta\bar{\varphi}_b(k) = [\hat{\mathbf{s}}_{\text{lmmse}}^T, (\mathbf{C}_{xx}^{-1} \mathbf{x})^T]^T$ .

Next, we deal with noisy communication links. Define the quantities  $\bar{\psi}_b(k) := 2\bar{\varphi}_b(k) - \bar{\varphi}_b(k-1)$  and  $\psi_j(k+1) := 2\varphi_j(k+1) - \varphi_j(k)$ . Note that  $\bar{\psi}_b(k)$  is equal to the  $b$ -th term inside the sum of (17), and recursion (18) can be replaced by

$$\bar{\psi}_b(k+1) = \sum_{j \in \mathcal{N}_b} \mathbf{D}_j^2 \mathbf{D}_b \psi_j(k+1), b \in \mathcal{B}. \quad (19)$$

Notice, that (17)-(18), and the pair (17) and (19) produce the same sequence of  $\varphi_j(k)$ 's. Actually, when sensors communicate to their neighbors the vectors  $\psi_j(k+1)$  and  $\bar{\psi}_b(k)$  the communication noise incorporated in the recursions on a per step basis is lower than when

exchanging  $\varphi_j(k)$ ,  $\bar{\varphi}_b(k)$  and  $\varphi_j(k-1)$ ,  $\bar{\varphi}_b(k-1)$  separately [7]. In the presence of noise, (17) and (19) become

$$\begin{aligned} \varphi_j(k+1) &= \hat{\mathbf{x}}_j + (\mathbf{I}_{L+p} - \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^1) \varphi_j(k) \\ &+ \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^2 \sum_{b \in \mathcal{B}_j} \bar{\psi}_b(k) \\ &+ \mathbf{F}_j^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_j)\mathbf{D}_j^2 \sum_{b \in \mathcal{B}_j \setminus \{j\}} \eta_j^b(k) \end{aligned} \quad (20)$$

$$\bar{\psi}_b(k+1) = \sum_{j \in \mathcal{N}_b} \mathbf{D}_j^2 \mathbf{D}_b \psi_j(k+1) + \sum_{j \in \mathcal{N}_b, j \neq b} \mathbf{D}_j^2 \mathbf{D}_b \bar{\eta}_b^j(k+1). \quad (21)$$

The steps involved in implementing locally (17) and (19) are: (i) all sensors  $j$  receive the vectors  $\bar{\psi}_b(k) + \eta_j^b(k)$  from  $b \in \mathcal{B}_j$  to form a (noisy) iterate  $\varphi_j(k+1)$ ; and (ii) bridge sensors receive the vector  $\psi_j(k+1) + \bar{\eta}_b^j(k+1)$  from  $j \in \mathcal{N}_b$  to form the (noisy) iterate  $\bar{\psi}_b(k+1)$ . The distributed algorithm resulting from the local recursions (20)-(21) is abbreviated as resilient distributed (RD) LMMSE.

Pertinent to the ensuing RD-LMMSE convergence analysis is the global RD-LMMSE recursion formed by stacking  $\varphi_j(k+1)$  for  $j = 1, \dots, J$ . For future use, define also the matrices  $\mathbf{A}_1 := \mathbf{F}_G \text{diag}(\mathbf{D}_1^1 \dots \mathbf{D}_J^1) - 2\mathbf{F}_G \mathbf{W}_E$ ,  $\mathbf{A}_2 := \mathbf{F}_G \mathbf{W}_E$  with  $\mathbf{F}_G := \text{diag}(\mathbf{F}_1^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_1) \dots \mathbf{F}_J^{-1}(\mathbf{I}_{L+p} - \mathbf{G}_J))$  and  $\mathbf{W}_E$  as follows

$$\mathbf{W}_E = \mathbf{D}^2 \sum_{b \in \mathcal{B}} (\mathbf{I}_J \otimes \mathbf{D}_b) (\mathbf{e}_b \otimes \mathbf{I}_{L+p}) (\mathbf{e}_b \otimes \mathbf{I}_{L+p})^T \mathbf{D}^2 \quad (22)$$

where  $\mathbf{e}_b$  denotes the  $b$ -th column of the adjacency matrix  $\mathbf{E}$ ,  $\mathbf{D}^2 := \text{diag}(\mathbf{D}_1^2, \dots, \mathbf{D}_J^2)$ , and  $\otimes$  is the Kronecker product. Upon substituting (21) in (20), and stacking the local variables  $\varphi_j(k)$  for  $j = 1, \dots, J$  in the vector  $\varphi(k) := [(\varphi_1(k))^T \dots (\varphi_J(k))^T]^T$  the RD-LMMSE global recursion is written as [7]

$$\varphi(k+1) = \hat{\mathbf{x}} + \varphi(k) - \mathbf{A}_1 \varphi(k) - \mathbf{A}_2 \varphi(k-1) + \bar{\eta}(k) + \bar{\eta}_b(k) \quad (23)$$

where  $\hat{\mathbf{x}} := [\hat{\mathbf{x}}_1^T, \dots, \hat{\mathbf{x}}_J^T]^T$ , and the noise vectors  $\bar{\eta}(k) := [\bar{\eta}_1^T(k) \dots \bar{\eta}_J^T(k)]^T$  and  $\bar{\eta}_b(k) := [\bar{\eta}_{b,1}^T(k) \dots \bar{\eta}_{b,J}^T(k)]^T$  have entries

$$\bar{\eta}_j(k) = \tilde{\mathbf{F}}_j \sum_{b \in \mathcal{B}_j \setminus \{j\}} \eta_j^b(k), \quad \bar{\eta}_{b,j}(k) := \tilde{\mathbf{F}}_j \sum_{\substack{b \in \mathcal{B}_j, \\ j' \in \mathcal{N}_b, j' \neq b}} \mathbf{D}_{j'}^2 \mathbf{D}_b \bar{\eta}_b^{j'}(k)$$

where  $\tilde{\mathbf{F}}_j := \mathbf{F}_j^{-1}(\mathbf{I} - \mathbf{G}_j)\mathbf{D}_j^2$ .

Note that (23) is a recursion with memory 2. Stacking  $\varphi(k+1)$  and  $\varphi(k)$ , we obtain a first-order recursion that has a  $2J(p+L) \times 2J(p+L)$  transition matrix  $\mathbf{A}$  formed by the  $J(L+p) \times J(L+p)$  submatrices  $[\mathbf{A}]_{11} = \mathbf{I}_{Jp} - \mathbf{A}_1$ ,  $[\mathbf{A}]_{12} = -\mathbf{A}_2$ ,  $[\mathbf{A}]_{21} = \mathbf{I}_{Jp}$  and  $[\mathbf{A}]_{22} = \mathbf{0}_{Jp}$ . This recursion as well as the associated transition matrix  $\mathbf{A}$  are particularly important for the convergence analysis of RD-LMMSE. Due to space limitations, we omit the details and provide directly the main result establishing the noise-resilience of RD-LMMSE. Let  $\lambda_{A,i}$ ,  $\mathbf{u}_{A,i}$  and  $\mathbf{v}_{A,i}$  denote the  $i$ th largest in magnitude eigenvalue of  $\mathbf{A}$  and the corresponding right and left eigenvectors, respectively. Define also  $\mathbf{C}_{\bar{\eta}\bar{\eta}_b} = \text{diag}(\mathbf{C}_{\bar{\eta}\bar{\eta}}, \mathbf{C}_{\bar{\eta}_b\bar{\eta}_b})$  and  $\bar{\mathbf{C}}_{\bar{\eta}\bar{\eta}_b} = \text{diag}(\mathbf{C}_{\bar{\eta}\bar{\eta}} + \mathbf{C}_{\bar{\eta}\bar{\eta}_b}, \mathbf{0}_{J(L+p)})$ , whose entries have finite magnitude (since  $\{\mathbf{C}_{\eta_j \eta_j}\}_{j=1}^J$  are finite). Moreover, let  $\delta\bar{\varphi}(k) := [(\delta\varphi(k))^T (\delta\varphi(k-1))^T]^T$  with  $\delta\varphi(k) := \varphi(k) - \varphi(k-1)$ . Based on these definitions, we have shown that [7]:

**Proposition 4** *The RD-LMMSE algorithm summarized in (23) reaches consensus in the mean i.e.,*

$$\lim_{k \rightarrow \infty} E[\delta\varphi_j(k)] := [\hat{\mathbf{s}}_{\text{lmmse}}^T (\mathbf{C}_{xx}^{-1} \mathbf{x})^T]^T, j = 1, \dots, J.$$

Matrix  $\mathbf{C}_\eta(k) := E[(\delta\bar{\varphi}(k) - E[\delta\bar{\varphi}(k)])(\delta\bar{\varphi}(k) - E[\delta\bar{\varphi}(k)])^T]$  converges to

$$\lim_{k \rightarrow \infty} \mathbf{C}_\eta(k) = \bar{\mathbf{C}}_{\bar{\eta}\bar{\eta}_b} + \sum_{i=L+p+1}^{2J(L+p)} \sum_{i'=L+p+1}^{2J(L+p)} \frac{\mathbf{u}_{A,i} \mathbf{u}_{A,i'}^T}{1 - \lambda_{A,i} \lambda_{A,i'}} \mathbf{v}_{A,i}^T \cdot \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_1 \\ -\mathbf{I} & -\mathbf{I} \end{bmatrix} \mathbf{C}_{\bar{\eta}\bar{\eta}_b} \begin{bmatrix} \mathbf{A}_1^T & -\mathbf{I} \\ \mathbf{A}_1^T & -\mathbf{I} \end{bmatrix} \mathbf{v}_{A,i'}. \quad (24)$$

Furthermore, the entries of  $\mathbf{C}_\eta(k)$  are bounded.

Proposition 4 guarantees convergence of the RD-LMMSE in the mean sense, while the noise covariance  $\mathbf{C}_\eta(k)$  remains bounded and converges to a fixed matrix as  $k \rightarrow \infty$ . Existing schemes, e.g. consensus averaging, are not able to provide distributed computation of the LMMSE without assuming knowledge of  $\mathbf{C}_{x\bar{x}}^{-1}$ . But even in that case consensus averaging suffers from catastrophic noise propagation [9]. RD-LMMSE on the other hand, offers a distributed matrix inversion procedure, through the updating of  $\mathbf{y}_j(k)$ , while guaranteeing bounded noise variance. Finally, notice that Proposition 4 holds universally for general data models, allowing for arbitrary correlation patterns among sensor data.

## 5. NUMERICAL EXAMPLES

Here we test the convergence of D-LMMSE and RD-LMMSE, along with their noise resilience properties in the presence of communication errors. We consider a WSN with  $J = 30$  sensors. Nodes in the WSN are randomly placed in the unit square  $[0, 1] \times [0, 1]$  with uniform distribution. Sensor  $j$  acquires  $L_j = 5$  observations, and  $\mathbf{s}$  has dimensionality  $p = 2$  while  $\mathbf{C}_{ss} = \text{diag}(0.2, 0.3)$ . We consider a linear model  $\mathbf{x}_j = \mathbf{H}_j \mathbf{s} + \mathbf{n}_j$ , where the entries of  $\mathbf{H}_j$  are random uniformly distributed over  $[-0.5, 0.5]$  and  $\{\mathbf{n}_j\}_{j=1}^J$  are zero-mean AWGN with  $\mathbf{C}_{n_j n_j} = 0.5 \mathbf{I}_{L_j \times L_j}$ . We further let  $\mathbf{C}_{\eta_j \eta_j} = \sigma^2 \mathbf{I}_{(L+p)}$ , where  $\sigma^2$  is adjusted so that the  $\text{SNR} := 10 \log_{10}[\hat{\mathbf{s}}_{\text{lmmse}} / (p\sigma^2)]$  assumes specific values. Fig. 2 depicts the total normalized error

$$e_{\text{norm}}(k) = \sum_{j=1}^J \|\mathbf{s}_j(k) - \hat{\mathbf{s}}_{\text{lmmse}}\|_2 / \|\hat{\mathbf{s}}_{\text{lmmse}}\|_2$$

versus iteration index  $k$  for different SNR values. The penalty coefficients of both D-LMMSE and RD-LMMSE, namely  $d_j, c_j$ , are set equal to 4. Notice that under ideal channel links both D-LMMSE and RD-LMMSE iterates coincide as suggested by Lemma 1, and  $e_{\text{norm}}(k) \rightarrow 0$  as  $k \rightarrow \infty$  corroborating Proposition 3. In the presence of reception noise, we average  $e_{\text{norm}}(k)$  over 50 independent D-LMMSE and RD-LMMSE estimates. As expected from Proposition 4,  $e_{\text{norm}}(k)$  obtained from RD-LMMSE exhibits an error floor confirming that the noise covariance converges to a matrix with bounded entries. Note that the D-LMMSE exhibits also noise resilience at the expense of higher steady-state variance than RD-LMMSE.

## 6. CONCLUSIONS

We developed distributed algorithms for estimation of random signals using ad hoc WSNs based on successive refinement of local estimates. The essence of our approach is to express the desired estimator, either MAP or LMMSE, as the solution of pertinent convex optimization problems. We then used the alternating direction multipliers method to enable decentralized implementation. Our framework does not require knowledge of the estimator in closed form, and allows for distributed computation, even of nonlinear estimators. Furthermore, our schemes exhibit resilience to communication errors. When it comes to decentralized computation of linear estimators, namely the LMMSE, we constructed noise-resilient algorithms

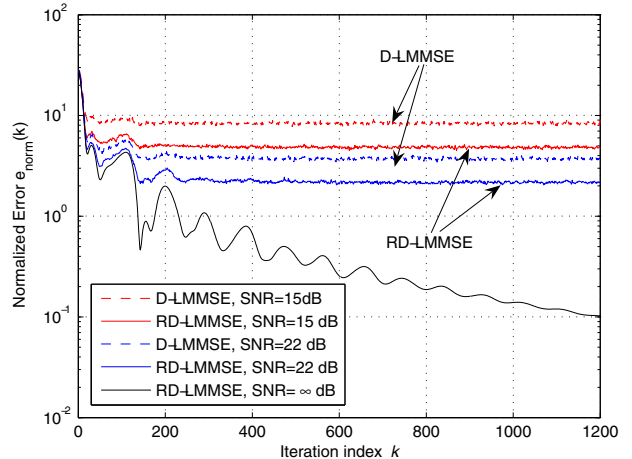


Fig. 2. Normalized error vs.  $k$  for D-LMMSE and RD-LMMSE.

that offer distributed estimation even when the observations across sensors are correlated.<sup>2</sup>

## 7. REFERENCES

- [1] S. Barbarossa and G. Scutari, "Decentralized Maximum Likelihood Estimation for Sensor Networks Composed of Nonlinearly Coupled Dynamical Systems," *IEEE-TSP*, 2007.
- [2] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, 1999.
- [3] V. Delouille, R. Neelamani, and R. Baraniuk, "Robust Distributed Estimation in Sensor Networks Using the Embedded Polygons Algorithm," *IEEE-TSP*, pp. 2998–3010, Aug. 2006.
- [4] A. Dogandžić and B. Zhang, "Distributed Estimation and Detection for Sensor Networks Using Hidden Markov Random Field Models," *IEEE-TSP*, pp. 3200–3215, Aug. 2006.
- [5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.
- [6] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus Based Distributed Parameter Estimation Using Ad Hoc Wireless Sensor Networks with Noisy Links," in *Proc. of ICASSP*, Honolulu, HA, 2007.
- [7] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, and A. Ribeiro, "Consensus in Ad Hoc Wsn with Noisy Links - Part II: Distributed Estimation and Smoothing of Random Signals," *IEEE-TSP*, submitted January 2007, <http://spincom.ece.umn.edu/>.
- [8] S. F. Shah, A. Ribeiro, and G. B. Giannakis, "Bandwidth-Constrained MAP Estimation for Wireless Sensor Networks," in *Proc. of 39th Asilomar Conf. On Signals, Systems and Computers*, Monterey, CA, Oct. 2005, pp. 215–219.
- [9] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed Average Consensus with Least-Mean-Square Deviation," *Journal of Parallel and Distributed Computing*, vol. 67, pp. 33–46, Jan. 2007.

<sup>2</sup>The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies of the Army Research Laboratory or the U. S. Government.