Distributed and Collaborative Estimation over Wireless Sensor Networks

Alberto Speranzon and Carlo Fischione and Karl Henrik Johansson

Abstract— A new distributed algorithm for cooperative estimation of a slowly time-varying signal using a wireless sensor network is presented. The estimate in each node is based on a so called consensus algorithm, which weights measurements and estimates of neighboring nodes. The algorithm is therefore scalable with the number of network nodes. It requires only limited information exchange between nodes and computations in each node. The weights are locally optimized based on a minimum variance criterion. Numerical results show that the proposed algorithm exhibits good performance compared to other distributed algorithms proposed in the literature.

I. INTRODUCTION

A challenge in control of autonomous agents is to use spatially distributed information in an efficient way. It is often important to be able to track a common variable without employing a centralized strategy, since such a strategy is vulnerable to node failures. A prominent example is distributed tracking of a moving target using a wireless sensor network (WSN). In this case, sensors have to cooperate in order to accomplish accurate information of, e.g., target position, velocity, clock, etc.

In recent years, several techniques for estimation using WSNs have been proposed. These techniques rely often on some physical characteristics of the wireless propagation, on the communication protocols, or any physical reading related to the signal to estimate. For example, the distance between pairs of nodes can be estimated measuring the network connectivity [1]. A taxonomy of methods for the location estimation in WSNs from a signal processing perspective is provided in [2]. In traditional synchronization algorithms, nodes exchange data packets containing the current clock value and synchronize using received data and knowledge of the communication delay [3]. The propagation delay has been also used for the estimation of the node positions [4]. It is clear that these techniques are effective only if there are not packet losses or time-varying communication delays, if the nodes are highly connected, and if there are not clock drifts.

Collaboration can be suitable to overcome intrinsic limitations in processing only local measurements, since measurements are usually affected by noise, e.g., [5]. In fact, exploiting samples taken from different nodes and explicitly

A. Speranzon, C. Fischione and K. H. Johansson are with the Automatic Control Lab, School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden, {albspe|carlofi|kallej}@ee.kth.se.

taking into account the communication, it is possible to design distributed algorithms for which nodes cooperate to achieve better estimates. A large body of literature is available on distributed sensor fusion. Here we limit the discussion to some recent and relevant contributions on consensus filters. In [6], the problem of distributed estimation of an average by a wireless sensor network is presented. It is assumed that nodes take a set of initial samples, and then iteratively exchange the averages of the samples locally collected. Each node reaches asymptotically the global average. The approach is based on a local weighted least squares method, where the weights are derived from a fast mixing Markov chain on a graph. In [7], a more general approach is investigated. The authors propose the consensus of the average of a common time-varying signal measured by each sensor, when the signal is affected by a zero-mean noise. A convex optimization problem is posed for which the authors find a set of symmetric edge weights that minimize the least mean square deviation. The same linear filter was also considered in [8], where a faster coefficient computation is investigated. In [9], a related consensus filter for distributed low-pass sensor fusion has been proposed.

In this paper, we study cooperative and distributed estimation algorithms using sensors nodes communicating through wireless transmission. Specifically, the system model takes into account time-varying signals, and we investigate how to ensure the consensus of the estimates while minimizing the variance of the error. We propose a distributed filter where the nodes compute the estimates without central coordination. The filter design includes local to guarantee the global asymptotic stability of the estimation error. Moreover, the distributed filter is scalable with respect to the network size. The algorithm can be applied e.g. for the position estimation, temporal synchronization, as well as tracking of signals.

Compared to recent relevant work [6]-[9], our approach is original because we adopt a more general model of the filter structure, without resorting to the common heuristic of the Laplacian associated to the communication graph. Moreover, our approach differ from [7] and [8], since we are interested to investigate distributed solutions, whereas in such papers centralized algorithms are used to computed the filter coefficients.

The paper is organized as follows. In Section II we formalize the problem. We also define a model of the communication network based on undirected graphs. A centralized solution based on the minimization of the overall estimate variance is discussed in Section III. Since a centralized solution requires a large amount of data flowing from the nodes

to a central station and back, we propose a decentralized solution in Section IV. The detailed algorithm, corresponding to the decentralized solution, and some implementation issues are discussed in Section V. Numerical results, where the proposed algorithm is compared with other solutions, are reported in Section VI.

II. PROBLEM FORMULATION

Let us consider $N > 1$ nodes randomly distributed in the plane. We assume that each node can measure a common scalar signal $d(t)$ corrupted by additive noise:

$$u_i(t) = d(t) + v_i(t), \qquad i = 1, \ldots, N,$$

with $t \in \mathbb{N} \cup \{0\}$ and where v_i is a zero-mean Gaussian random variable. This is a common assumption to characterize the noise fluctuations, e.g., [10], [9], and can be motivated by the central limit theorem. Defining the vectors $u(t) = (u_1(t), \ldots, u_N(t))^T$ and $v(t) = (v_1(t), \ldots, v_N(t))^T$, we can rewrite the previous equation as

$$u(t) = d(t)\mathbb{1} + v(t),$$

where $\mathbb{1} = (1, \ldots, 1)^T$. We assume that the covariance matrix of the random vector $v(t)$ is $\Sigma = \sigma^2 I$, so that v_i and v_j, for $i \neq j$, are uncorrelated.

Since the nodes are connected through a communication network, each node has available extra data, transmitted by the neighbors, in order to reconstruct the signal $d(t)$. We thus assume that a node i builds an estimate, $x_i(t)$, of the signal $d(t)$ as

$$x_i(t) = \sum_{j=1}^{N} k_{ij}(t)x_j(t-1) + \sum_{j=1}^{N} h_{ij}(t)u_j(t). \quad \text{(II.1)}$$

Thus each node computes an new estimate by linearly combining its previous estimate and current measurement with previous estimates and current measurements received from neighbors nodes. If node i is not connected with node j, then $k_{ij}(t) = k_{ji}(t) = 0$ and $h_{ij}(t) = h_{ji}(t) = 0$, for all $t \geq 0$. This estimation model is similar to the one proposed in [9].

Remark 2.1: From the model (II.1) that we propose, it is clear that one could try to design $k_{ii}(t)$ and $h_{ii}(t)$ so that a single node, without exchanging data with neighbors, is able to estimate $d(t)$. This would have the advantage of saving power for communication. However for a single node it would require a longer time before achieving a good estimate of $d(t)$. Moreover measurements taken too close in time, by the same node, are generally corrupted by correlated noise

$$E\{v_i(t)v_i(t-\tau)\} = r(\tau),$$

where $r(\tau)$ is the autocorrelation function of the noise. Measurements taken by different nodes are instead corrupted by uncorrelated noise.

We rewrite the estimator (II.1) in a more compact way as

$$x(t) = K(t)x(t-1) + H(t)u(t) \quad \text{(II.2)}$$

where $x(t) = (x_1(t), \ldots, x_N(t))^T$, $[K(t)]_{ij} = k_{ij}(t)$, and $[H(t)]_{ij} = h_{ij}(t)$.

It is convenient to model the communication network as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, N\}$ is the vertex set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ the edge set. Note that $(i, j) \in \mathcal{E}$ implies that $(j, i) \in \mathcal{E}$ since the graph is undirected. The graph $\mathcal{G}$ is said to be connected if there is a sequence of edges in $\mathcal{E}$ that can be traversed to go from any vertex to any other vertex. We associate to each edge $(i, j) \in \mathcal{E}$ a time-varying weight $w_{ji}(t)$. In general, it may hold that the weights $w_{ij}(t)$ and $w_{ji}(t)$ are different. We introduce the adjacency matrix $W(t)$ as

$$[W(t)]_{ij} = \begin{cases} w_{ij}(t), & \text{if } (j, i) \in \mathcal{E} \\ 0, & \text{otherwise}. \end{cases}$$

We say that a matrix $W(t)$ is compatible with $\mathcal{G}$, if $W(t)$ defines an adjacency matrix for $\mathcal{G}$. We denote this by $W(t) \simeq \mathcal{G}$. We interpret the matrices $K(t)$ and $H(t)$ of equation (II.2) as the adjacency matrices of two weighted graphs, one associated to the communication of estimates $x(t)$ and the other associated to the communication of measurements $u(t)$. It is convenient to introduce the neighbors of a node i as the set $\mathcal{N}_i$ of all nodes that can communicate with i, namely

$$\mathcal{N}_i = \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}.$$

We can now state the main problem of the paper. Given a wireless sensor network modelled by a connected graph $\mathcal{G}$, find time-varying matrices $K(t)$ and $H(t)$, compatible with $\mathcal{G}$, such that the signal $d(t)$ is consistently estimated and the variance of the estimate is minimized. Moreover, the solution should be distributed in the sense that the computation of $k_{ij}(t)$ and $h_{ij}(t)$ should be performed by node i.

III. CENTRALIZED ESTIMATION

Let us assume that $x(0)$ and $u(0)$ are independent and identically distributed random variables. Let us consider the estimation error $e(t) = x(t) - d(t)\mathbb{1}$. We have

$$e(t) = K(t)e(t) + K(t)d(t-1)\mathbb{1} + (H(t) - I)d(t)\mathbb{1} + H(t)v(t).$$

Let us assume that $d(t) = d(t-1) + \delta(t)$ where $|\delta(t)| < \bar{\delta} \ll 1$, taking the expected value with respect to the noise $v(t)$, we obtain

$$\mathbb{E}\, e(t) = K(t)\,\mathbb{E}\, e(t-1) + (K(t) + H(t) - I)\mathbb{1}\, d(t-1) + (H(t) - I)\delta(t)\mathbb{1}. \quad \text{(III.1)}$$

Under the assumption that $\delta(t)$ is not known, the convergence of $\mathbb{E}\, e(t)$ to zero is guaranteed if $\gamma_{\max}(K(t)) < 1$ for[1] all t, and if the matrices $K(t)$ and $H(t)$ fulfill

$$(K(t) + H(t))\mathbb{1} = \mathbb{1} \quad \text{(III.2)}$$

$$H(t)\mathbb{1} = \mathbb{1}. \quad \text{(III.3)}$$

Under this conditions, it is possible to show that the filter is unbiased and the minimum variance is achieved with $K(t) =$

---

[1]With $\gamma_{\max}(A)$ we denote the largest singular value of the matrix A.

0 and with $H(t)$ such that

$$h_{ij}(t) = h_{ji}(t) = \begin{cases} \dfrac{1}{|\mathcal{N}_i|} & \text{if } j \in \mathcal{N}_i \\ 0 & \text{otherwise} . \end{cases}$$

In order to take advantage of previous estimates we need to relax the condition on the bias. Consequently, we have to tradeoff the bias for the variance reduction, as for example discussed [11]. We then remove the condition on the matrix $H(t)$ in (III.3). It can be shown that if the signal is slowing varying, as assumed, then the bias is negligible.

The degree of freedom in the choice of $K(t)$ and $H(t)$, can be used to minimize the variance of the estimate. For this purpose we study how the covariance matrix changes with time. Let us assume that $x(t)$ and $u(t)$ are independent random vectors. Introduce the matrix

$$P(t) = \mathbb{E}\, (e(t) - \mathbb{E}\, e(t))(e(t) - \mathbb{E}\, e(t))^T .$$

Then,

$$P(t) = K(t)P(t-1)K(t)^T + \sigma^2 H(t)H(t)^T . \quad \text{(III.4)}$$

One option now is to choose $K(t)$ and $H(t)$ such that $P(t)$ is minimized at each time instance. Hence, we have the following optimization problem

$$\min_{K(t), H(t)} \quad \text{tr}\,(K(t)P(t-1)K(t)^T) + \sigma^2 \text{tr}\,(H(t)H(t)^T)$$
$$\text{(III.5)}$$

$$\text{s.t.} \quad (K(t) + H(t))\mathbb{1} = \mathbb{1} ,$$
$$\gamma_{\max}(K(t)) < 1 ,$$
$$K(t) \simeq \mathcal{G} ,$$
$$H(t) \simeq \mathcal{G} .$$

This optimization problem is solved iteratively, starting with some initial guess $P(0)$. Notice that the objective function is quadratic in $K(t)$ and $H(t)$ for a given $P(t-1)$. The first constraint is the linear matrix equality (III.2). The second constraint, which ensures that the estimation error converges to zero, can be written as a linear matrix inequality using Schur complement [12]. The last constraints, impose structure on the matrices $K(t)$ and $H(t)$. They are represented by equalities to zero of some elements of these matrices.

Although the optimization problem (III.5) could be solved using standard numerical optimization tools, it clearly requires a powerful central node collecting data, computing new weights and dispatching them to the sensors.

Beside the typical disadvantage of a centralized solution, which is not robust to failures, we would also have large delays due to the propagation of the data from the farthest nodes to the central node. Although this could be overcome having directed paths from every node to the central node, this would require, in general, a total power consumption which is prohibitive for small nodes. We propose in the following a decentralized solution where each node computes its weights minimizing the variance of its estimate.

IV. DECENTRALIZED ESTIMATION

In order to state a decentralized optimization problem, we need to introduce some notation.

Let M_i denote the number of neighbors of node i, i.e., M_i is the cardinality of $\mathcal{N}_i = \{i_1, \ldots, i_{M_i}\}$. Collect the estimation errors available at node i in the vector $\epsilon_i(t) \in \mathbb{R}^{M_i}$. The elements of ϵ_i are ordered according to the node indices:

$$\epsilon_i(t) = (e_{i_1}(t), \ldots, e_{i_{M_i}}(t))^T, \qquad i_1 < \cdots < i_{M_i} .$$

Similarly, we introduce vectors $\kappa_i(t), \eta_i(t) \in \mathbb{R}^{M_i}$ corresponding to the non-zero elements of row i of the matrices $K(t)$ and $H(t)$, respectively, and ordered according to node indices. It follows from (III.4) that the variance of $e_i(t)$ can be evaluated as

$$\mathbb{E}\, e_i(t)^2 = \kappa_i^T(t)\Gamma_i(t-1)\kappa_i(t) + \sigma^2 \eta_i^T(t)\eta_i(t) ,$$

where $\Gamma_i(t) = \mathbb{E}\, \epsilon_i(t)\epsilon_i(t)^T$. To minimize the variance of the estimation error in each node, we propose that $\kappa_i(t)$ and $\eta_i(t)$ are chosen to minimize this expression. To obtain consensus and convergence as well, the following optimization problem should be solved at each time t and in each node i:

$$\min_{\kappa_i(t), \eta_i(t)} \quad \kappa(t)_i^T \Gamma_i(t)\kappa_i(t) + \sigma^2 \eta_i(t)^T \eta_i(t) \quad \text{(IV.1)}$$

$$\text{s.t.} \quad (\kappa_i(t) + \eta_i(t))^T \mathbb{1} = 1 \quad \text{(IV.2)}$$

$$\gamma_{\max}(K(t)) < 1 . \quad \text{(IV.3)}$$

Note, however, that the inequality constraint is global, since $K(t)$ depends on all $\kappa_i(t)$, $i = 1, \ldots, N$.

In order to make the optimization problem (IV.1) distributed we need to find conditions on the $\kappa_i(t)$ and $\eta_i(t)$ that would guarantee that $\gamma_{\max}(K(t)) < 1$. In particular we use the following result.

Proposition 4.1: If it holds

$$\sum_{j=1}^N k_{ij}^2(t) < 1/2 , \quad \text{and} \quad \sum_{j=1}^N |k_{ij}(t)| < 1 \quad \text{(IV.4)}$$

then $\gamma_{\max}(K(t)) < 1$, for any $t \geq 0$.

Proof: Since the result does not depend on t, in order to have lighter notation, in the following we drop the time dependence. We need to prove that if the given inequalities hold, then $\gamma_{\max}(K) < 1$. This means that it must hold $\lambda_{\max}(KK^T) < 1$, where with $\lambda_{\max}(.)$ we denote the maximum eigenvalue. We have that

$$[KK^T]_{ij} = \sum_r k_{ir}k_{jr}$$

and in particular

$$[KK^T]_{ii} = \sum_s k_{is}^2 .$$

Let λ be a generic eigenvalues of KK^T. Then we know,

from the Gerschgorin's circle theorem, that

$$\lambda \in \left\{ z \in \mathbb{C} \ : \ \bigcup_{i=1}^{N} |z - \sum_s k_{is}^2| \leq \sum_{j \neq i} |\sum_s k_{is}k_{js}| \right. $$
$$\left. \leq \sum_{j \neq i} \sum_s |k_{is}| |k_{js}| \right\}.$$

If we now consider the following sum we have that

$$\left(\sum_s |k_{is}| \right)^2 = \sum_s k_{is}^2 + \sum_s \sum_{s \neq \ell} |k_{is}| |k_{i\ell}|.$$

Since $(\sum_s |k_{is}|)^2 < 1$ by hypothesis then we have that

$$\sum_s \sum_{s \neq \ell} |k_{is}| |k_{i\ell}| < 1 - \sum_s k_{is}^2,$$

and that

$$\lambda \in \left\{ z \in \mathbb{C} \ : \ \bigcup_{i=1}^{N} |z - \sum_s k_{is}^2| < 1 - \sum_s k_{is}^2 \right\}.$$

Since, by hypothesis, $\sum_s k_{is}^2 < 1/2$, we easily conclude that $|\lambda| < 1$, and thus we have the sought results. ∎

Remark 4.2: Notice that the bounds previous are quite conservative since Gerschgorin's bounds are in general not tight.

From the definition of $\kappa_i(t)$ and $\eta_i(t)$ it follows that the two inequalities (IV.4) are equivalent to

$$\sum_{j=1}^{M_i} \kappa_{i_j}^2(t) < 1/2, \quad \text{and} \quad \sum_{j=1}^{M_i} |\kappa_{i_j}(t)| < 1$$

since the other $N - M_i$ coefficient are zero. Let us define the following set

$$\Theta_i = \left\{ (\theta_1, \dots, \theta_{M_i}) : \sum_{j=1}^{M_i} \theta_j^2 < \frac{1}{2} \wedge \sum_{j=1}^{M_i} |\theta_j| < 1 \right\}. \tag{IV.5}$$

It is not difficult to see that the set Θ is convex. We can then rewrite (IV.1) as the following problem, where all the variable to be optimized are local for node i

$$\min_{\kappa_i(t),\eta_i(t)} \quad \kappa_i(t)^T \Gamma_i(t-1)\kappa_i(t) + \sigma^2 \eta_i(t)^T \eta_i(t) \tag{IV.6}$$
$$\text{s.t.} \quad \left(\kappa_i(t)^T + \eta_i(t)^T \right) \mathbb{1} = 1$$
$$\kappa_i(t) \in \Theta$$

Notice that the problem is quadratic in $\kappa_i(t)$ and $\eta_i(t)$ for a given $\Gamma_i(t-1)$, with one linear constraint and a nonlinear constraint. Since the cost function of the primal optimization problem (IV.6) is convex in Θ, the constraint $(\kappa_i(t)^T + \eta_i(t)^T) \mathbb{1} = 1$ is linear in $\kappa_i(t)$ and $\eta_i(t)$, and if we assume the solution is in the set Θ, then strong duality follows [13]. The problem can then be transform in its dual.

Let us consider the Lagrangian

$$L(\eta, \kappa_i(t), \eta_i(t)) = \kappa_i(t)^T \Gamma_i \kappa_i(t) + \sigma^2 \eta_i(t)^T \eta_i(t)(t) $$
$$+ \zeta((\kappa_i(t)^T + \eta_i(t)^T) \mathbb{1} - 1)$$

where $\zeta \in \mathbb{R}$ is a Lagrangian multiplier. We introduce the dual function

$$g(\eta) = \inf_{\kappa_i(t),\eta_i(t)} L(\zeta, \kappa_i(t), \eta_i(t)),$$

and the dual optimization problem becomes

$$\max_\zeta g(\zeta).$$

In particular we obtain that

$$\kappa_i(t) = -\frac{\zeta \Gamma_i(t-1)^{-1} \mathbb{1}}{2} \tag{IV.7}$$
$$\eta_i(t) = -\frac{\zeta \mathbb{1}}{2\sigma^2}. \tag{IV.8}$$

We then minimize the dual optimization function $g(\zeta)$ with respect to ζ. Simple algebraic calculations yield the optimal values

$$\kappa_i(t) = \frac{\Gamma_i(t-1)^{-1} \mathbb{1}}{\sigma^{-2} M_i + \mathbb{1}^T \Gamma_i(t-1)^{-1} \mathbb{1}}$$
$$\eta_i(t) = \frac{\mathbb{1}}{M_i + \sigma^2 \mathbb{1}^T \Gamma_i(t-1)^{-1} \mathbb{1}}.$$

which are the optimal weights for each sensor, for a given $\Gamma_i(t-1)$. We should now show that these optimal weights are feasible for the problem (IV.6), which corresponds to prove that the solution belong to the set Θ defined in (IV.5). This so far has not been shown formally, however extensive simulations show that the obtained solution is feasible.

As we have pointed out before the optimal weights $\kappa_i(t)$ and $\eta_i(t)$ depend on the covariance matrix $\Gamma_i(t-1)$. Since each node receives measurements and estimates from the neighbors, it is possible to compute, or estimate, the covariance matrix $\Gamma_i(t-1)$ at each time step.

V. IMPLEMENTATION ISSUES

Let us discuss the estimation algorithm in detail. The implementation of the consensus algorithm is shown as Algorithm 1. First, each sensor initializes the local mean estimation error $\hat{m}_\epsilon = 0$ (see line 2) and its local covariance matrix estimate with the noise covariance, i.e. $\widehat{\Gamma}_i(0) = \sigma^2 I$ (see line 2 in the algorithm), where we remark that we are using the "hat" since this are sample estimates of the real mean and covariance of the error. The optimal weights are computed using the equations (IV.8) (lines 8 and 9). In line 10 we compute the optimal estimate. Lines from 12 to 26 implement the covariance update based on the available data.

In order to compute $\hat{m}_\epsilon$ and $\widehat{\Gamma}_i(t)$, an estimate of the estimation error needs to be computed. This is done computing an average between the current available estimates and measurements (see line 18). As an estimate of $d(t)$ we cannot use directly $x(t)$ since it is biased and we can underestimate the actual error. On the other hand, using only measurement is also not a good choice, since their noise

variance is very high. A good heuristic is the one proposed, where we combine low variance slightly biased estimates and unbiased measurements corrupted noises with high variance. In particular the mean and covariance matrix are estimated

---

Algorithm 1 Estimation algorithm for node i

---
1. $t := 0$
2. $\hat{m}_\epsilon := 0$
3. $\widehat{\Gamma}_i(0) := \sigma^2 I$
4. $x_i(0) := u_i(0)$
5. **while** forever **do**
6. $\quad M_i := |\mathcal{N}_i|$
7. $\quad t := t + 1$
8. $\quad \kappa_i(t) := \dfrac{\widehat{\Gamma}_i(t-1)^{-1}\mathbb{1}}{\sigma^{-2}M_i + \mathbb{1}^T\widehat{\Gamma}_i(t-1)^{-1}\mathbb{1}}$
9. $\quad \eta_i(t) := \dfrac{\mathbb{1}}{M_i + \sigma^2\,\mathbb{1}^T\widehat{\Gamma}_i(t)^{-1}\mathbb{1}}$
10. $\quad x_i(t) := \sum_{j\in\mathcal{N}_i} \kappa_{i_j}(t)x_j(t-1) + \eta_{i_j}(t)u_j(t)$
11. $\quad \hat{d} := 0$
12. $\quad$**if** $t \leq 2$ **then**
13. $\quad\quad$**for** $j \in \mathcal{N}_i$ **do**
14. $\quad\quad\quad \hat{d} := \hat{d} + \dfrac{u_j(t)}{M_i}$
15. $\quad\quad$**end for**
16. $\quad$**else**
17. $\quad\quad$**for** $j \in \mathcal{N}_i$ **do**
18. $\quad\quad\quad \hat{d} := \hat{d} + \dfrac{x_j(t) + u_j(t)}{2M_i}$
19. $\quad\quad$**end for**
20. $\quad$**end if**
21. $\quad$**for** $j \in \mathcal{N}_i$ **do**
22. $\quad\quad e_{i_j}(t) := x_j(t) - \hat{d}$
23. $\quad$**end for**
24. $\quad \epsilon_i := \{e_{i_1}, \ldots, e_{i_{M_i}}\}$
25. $\quad \hat{m}_{\epsilon_i}(t) := \dfrac{t-1}{t}\hat{m}_{\epsilon_i}(t-1) + \dfrac{1}{t}\epsilon_i(t)$
26. $\quad \widehat{\Gamma}_i(t) := \dfrac{t-1}{t}\widehat{\Gamma}_i(t-1) + \dfrac{1}{t}(\epsilon_i(t) - \hat{m}_{\epsilon_i}(t))(\epsilon_i(t) - \hat{m}_{\epsilon_i}(t))^T$
27. **end while**

---

from the samples as shown on line 25 and 26.

In the algorithm, the inversion of the covariance matrix should be computed. This is not a difficult operation in resource constrained sensor networks, since each node has generally a rather limited number of neighbors, and thus the size of the matrix $\widehat{\Gamma}_i$ is small.

Note that the algorithm is implemented under the assumption that each node is able to compute and communicate data within the sampling instance.

VI. NUMERICAL RESULTS

Numerical simulations have been carried out to compare the proposed algorithm with other possible approaches. Specifically, the proposed algorithm is compared with two other algorithms, one we called *arithmetic mean estimator*, namely such that

$$k_{ij}(t) = h_{ij}(t) = \begin{cases} \dfrac{1}{2\,|\mathcal{N}_i|} & \text{if } j \in \mathcal{N}_i \\ 0 & \text{otherwise}. \end{cases}$$

The other one is based on the Laplacian matrix [14], which is normalized in order to obtain a stable filter. We will called such estimator *Laplacian based estimator*, and it is defined as

$$k_{ij}(t) = h_{ij}(t) = \begin{cases} 1 - \dfrac{|\mathcal{N}_i|}{\max_i |\mathcal{N}_i|} & \text{if } i = j \\ \dfrac{|\mathcal{N}_i|}{\max_i |\mathcal{N}_i|} & \text{if } j \in \mathcal{N}_i \\ 0 & \text{otherwise}. \end{cases}$$

Two random generated networks have been considered. The first is shown in Figure 1(a), with $N = 15$ nodes, and the second with $N = 150$ nodes is reported in Figure 1(b). The signal to be tracked is $d(t) = 3\sin(2\pi t/780) - \cos(2\pi t/620)$ with a noise normally distributed around $d(t)$ with variance $\sigma^2 = 1.2$. The network is generated with sensors randomly distributed in a squared area of side $N/2$. Two nodes are connected if and only if their relative distance is less than $1.5\sqrt{N}$.

For the two network cases, all the N realizations are shown in Figure 1(c) and Figure 1(d). The first plot of the two figures shows the signal corrupted by noise, and the second refers the the realization for the arithmetic mean based estimator, the third to the Laplacian based, and the last refers to the proposed algorithm. In particular, it is possible to appreciate, visually, the improvements due to the solution proposed.

In the following table are collected the standard deviation of the mean square errors.

| Estimator	Std. dev. MSE $N = 15$	Std. dev. MSE $N = 150$
Arithmetic mean	0.296	0.209
Laplacian based	0.350	0.356
Decentralized	0.175	0.132

In the first case our algorithm yields and improvement of about 40% and of about 37% in the second case, with respect to the arithmetic mean estimator and of about 50% and of about 63% compared with the Laplacian based estimator.

VII. CONCLUSIONS

In this paper, we have presented a decentralized coopera-tive estimation algorithm for the estimation of time-varying signal using a wireless sensor network. Specifically, the algorithm allows for accurate refinements of the estimates by employing previous estimates and noisy measurements of the signal to be estimated. We provide optimal time-varying weights to be used in combining the information, where the cost function is the variance of the estimate. Numerical results shows that the proposed algorithm exhibits
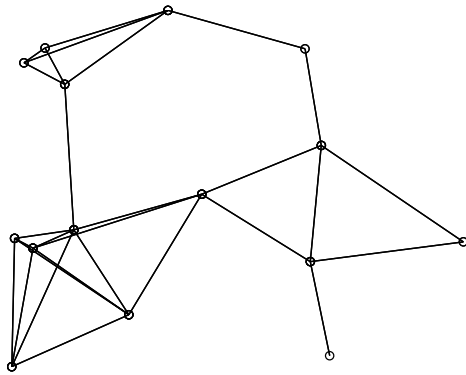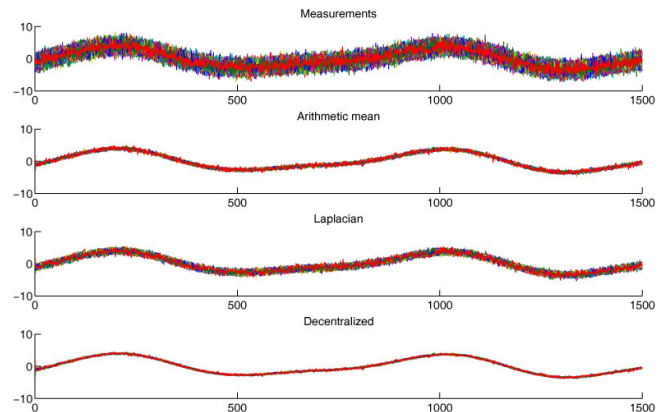
(a) Random generated network with $N = 15$. Each node has 3.6 neighbors, in average.
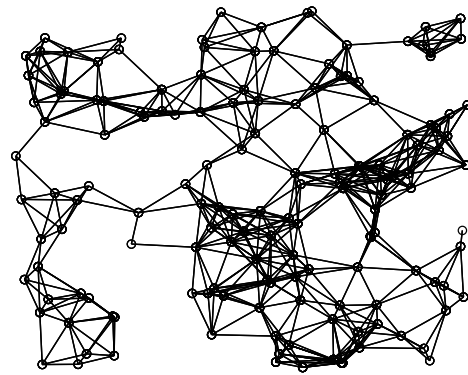
(b) Random generated network with $N = 150$. Each node has 8.3 neighbors, in average.
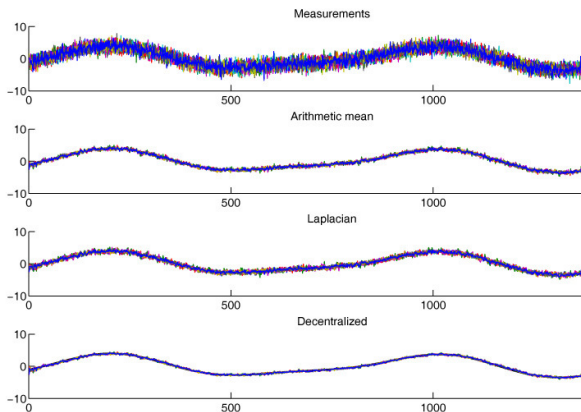
(c) Realization of the different estimators versus time. The first pl the measurements. The second shows the arithmetic mean estimator, the second the Laplacian based and the last show the proposed distributed estimator. The noise variance is of 1.2.

(d) Realization of the different estimators versus time. The first plot show the measurements. The second shows the arithmetic mean estimator, the second the Laplacian based and the last show the proposed distributed estimator. The noise variance is of 1.2.

very good performance in terms of standard deviation of the measurement errors and that it outperforms other common solutions.

REFERENCES

[1] Shang, Y. and Ruml, W. and Zhang, Y. and Fromherz, M., "Localization from connectivity in sensor networks," *IEEE Transaction on Parallel and Distributed System, Vol. 15, No. 11*, November 2004.

[2] Patwari, N. and Ash, J. N. and Kyperountas, S., "Locating the nodes," *IEEE Signal Processing Magazine*, July 2005.

[3] Mills, D. L., "Precision synchronization of computer network clock," *ACM Computer Communication Review*, 1994.

[4] Di Stefano, G. and Graziosi, F. and Santucci, F., "Distributed positioning algorithm for ad-hoc networks," *International Workshop on Ultra Wideband Systems (IWUWBS), Oulu*, 2003.

[5] Savvides, A. and Garber, W. L. and Moses, R. L. and Srivastava, M. B., "An analysis of error inducing parameters in multihop sensor node localization," *IEEE Transaction on Mobile Computing, Vol. 4, No. 6*, November/December 2005.

[6] Xiao, L. and Boyd, S. and Lall, S., "A scheme for robust distributed sensor fusion based on average consensus," *In Proc. of IEEE IPSN*, 2005.

[7] Xiao, L. and Boyd, S. and Kim, S. J., "Distributed average consensus with least-mean-square deviation," *Submitted to Journal of Parallel and Distributed Computing*, 2006.

[8] L. Xiao, S. Boyd, "Fast linear iterations for distributed averaging," *System Control Letter*, 2004.

[9] Olfati-Saber, R. and Shamma, J. S., "Consensus filters for sensor networks and distributed sensor fusion," *In Proc. of IEEE CDC*, 2005.

[10] Gustafsson, F. and Gunnarsson, F., "Mobile positioning using wireless networks," *IEEE Signal Processing Magazine*, July 2005.

[11] Y. C. Eldar, "Minimum variance in biased estimation: Bounds and asymptotically optimal estimators," *IEEE Transactions on Signal Processing*, vol. 54, 2004.

[12] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, ser. Studies in Applied Mathematics. Philadelphia: SIAM, 1994.

[13] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Athena Scientific, 2003.

[14] C. Godsil and G. Royle, *Algebraic Graph Theory*. Springer, 2001.