DISTRIBUTED SIGNAL SUBSPACE PROJECTION ALGORITHMS WITH MAXIMUM CONVERGENCE RATE FOR SENSOR NETWORKS WITH TOPOLOGICAL CONSTRAINTS

S. Barbarossa, G. Scutari, T. Battisti

Emails: <sergio, scutari, timothy>@infocom.uniroma1.it
INFOCOM Dpt., Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy

ABSTRACT

The observations gathered by the individual nodes of a sensor network may be unreliable due to malfunctioning, observation noise or low battery level. Global reliability is typically recovered by collecting all the measurements in a fusion center which takes proper decisions. However, centralized networks are more vulnerable and prone to congestion around the sink nodes. To relax the congestion problem, decrease the network vulnerability and improve the network efficiency, it is appropriate to bring the decisions at the lowest possible level. In this paper, we propose a distributed algorithm allowing each node to improve the reliability of its own reading thanks to the interaction with the other nodes, assuming that the field monitored by the network is a smooth function. In mathematical terms, this only requires that the useful field belongs to a subspace of dimension smaller than the number of nodes. Although fully decentralized, the proposed algorithm is globally optimal, in the sense that it performs the projection of the overall set of observations onto the signal subspace through an iterative decentralized algorithms, that requires minimum convergence time, for any given node coverage.

***Index Terms*—** Distributed projection, sensor networks.

1. INTRODUCTION

Because of limited battery level and sensor complexity, the measurements gathered by the single node of a sensor network may be highly unreliable. Improving the reliability of the individual node would require higher complexity and cost, but this would negatively affect the economy of scale, which is a fundamental concern in large scale networks. It is then particularly important to improve the accuracy of each sensor by exploiting their interaction with the other nodes. This is possible if the environment monitored by the network exhibits a spatial correlation, which is typically the case in many physical fields of interest, like in the distribution of temperatures or the concentration of a contaminant. In mathematical terms, the vector of measurements collected by a network composed of N nodes belongs, in general, to a vector space of dimension N. However, the useful signal field typically belongs to a subspace of dimension much smaller than N. In a centralized network, the accuracy of the network as a whole is recovered by collecting all the data in a fusion center, or sink node, which projects the observed vector onto the useful signal subspace, thus eliminating all the noise components lying out of the useful subspace. However, centralized networks are prone to several shortcomings, like congestion around the sink nodes and vulnerability to selected attacks or failures of hub nodes. Also from a fundamental information theoretic perspective, if the goal of the network is to compute a function of the data which has structural properties, e.g. it is a divisible function, for example, an efficient network design requires some sort of *in-network* or distributed processing [1].

The problem addressed in this paper is how to carry out the globally optimal projection operation through a fully decentralized network, with no fusion center, using a network where each node exchanges information only with its neighbors. Distributed algorithms able to reach globally optimal processing tasks are available and they are typically iterative, see e.g. [2–4]. However, the iterative nature of distributed algorithms makes them prone to delayed decision and energy consumption, besides the complexity of handling repeated exhanges of data. It is then clear that distributed algorithms become really attractive only if we are able to limit the energy consumption and complexity of each node. The goal of this paper is to propose a fully decentralized iterative algorithm allowing each node to converge to its own globally optimal value with minimum convergence time, for any given transmit power of each node. Interestingly, we show that there is a strict reationship between the minimum transmit power of each node and the dimension of the useful signal subspace onto which the iterative algorithm is going to project the observation vector: the higher is the subspace dimension, the greater is the minimum transmit power.

2. ACHIEVING GLOBALLY OPTIMAL SIGNAL SUBSPACE PROJECTION THROUGH LOCAL INTERACTIONS

Let us consider a network composed of N sensors and denote the measurement collected by the i-th sensor, located at (x_i, y_i), by $g(x_i, y_i) = z(x_i, y_i) + \xi_i$, where $z(x_i, y_i)$ is the useful field and ξ_i is the observation error, assumed to be a zero mean random variable with variance σ_ξ^2. In vector notation, a quite general observation model is

$$\mathbf{g} = \mathbf{z} + \mathbf{v} = \mathbf{U}\mathbf{s} + \boldsymbol{\xi}, \tag{1}$$

where $\mathbf{z} = \mathbf{U}\mathbf{s}$ is the useful signal, $\mathbf{U}$ is a $N \times r$ matrix, with $r \leq N$, and $\mathbf{s}$ is a $r \times 1$ column vector. The columns of $\mathbf{U}$ are assumed to be linearly independet and thus constitute a basis spanning the useful signal subspace. In many applications, the useful signal is a smooth function. This property can be captured by choosing the columns of $\mathbf{U}$ as the low frequency components of the Fourier basis or low-order polynomials, for example. In practice, the dimension r of the useful signal subspace is typically much smaller than the dimension N of the observation space. Hence, a strong noise reduction may be obtained by projecting the observation vector onto the signal subspace. More specifically, if the noise vector is Gaussian, with zero mean and covariance $\sigma_\xi^2\mathbf{I}$, the maximum likelihood estimator of $\mathbf{z}$ is

$$\hat{\mathbf{z}} = \mathbf{U}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{g}. \tag{2}$$

If the noise pdf is unknown, the estimator (2) is still significant, as it is the so called Best Linear Unbiased Estimator (BLUE) .

This work has been funded by the WINSOC project, a Specific Targeted Research Project (Contract Number 0033914) co-funded by the INFSO DG of the European Commission within the RTD activities of the Thematic Priority Information Society Technologies.

The operation performed in (2) corresponds to the orthogonal projection of the observation vector onto the subspace spanned by the columns of $\mathbf{U}$. Assuming, without any loss of generality (w.l.o.g.), the columns of $\mathbf{U}$ to be orthonormal, the projector simplifies into

$$\hat{\mathbf{z}} = \mathbf{U}\mathbf{U}^T\mathbf{g}. \tag{3}$$

With a centralized system, the computation of (3) requires that all nodes send their measurements (vector $\mathbf{g}$) to a fusion center that computes (3). Conversely, our problem is how to compute (3) with a decentralized network, where each node exchanges information with its neighbors only. The proposed approach is based on an iterative procedure, where each node initializes a state variable with the local measurement, let us say $z_i[0] = g(x_i, y_i)$, and then it evolves by interacting with nearby nodes. Denoting by $\mathbf{z}[k]$, the $N \times 1$ vector containing the states of all nodes at iteration k, the whole system evolves according to the following linear state equation:

$$\mathbf{z}[k+1] = \mathbf{W}\mathbf{z}[k], \qquad k = 0, 1, \ldots, \quad \mathbf{z}[0] = \mathbf{g} \in \mathbb{R}^N, \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{N\times N}$ is a *sparse* (not necessarily symmetric) matrix. We assume here that $\mathbf{W}$ does not vary with time. The sparsity of $\mathbf{W}$ is what characterizes the network topology and, in particular, it makes explicit that each node interacts with a few neighbors only. In a wireless network, the network topology is dictated by the coverage radius and then, ultimately, by the transmit power of each node. Given the interaction mechanism (4), our problem is twofold: 1) guarantee that system (4) converges to the desired vector (3), although using a sparse matrix $\mathbf{W}$; 2) find the sparse matrix $\mathbf{W}$, under a topological constraint, so that the convergence time is minimized.

Let us denote by $\mathbf{P}_{\mathcal{R}(\mathbf{U})} \in \mathbb{R}^{N\times N}$ the orthogonal projector onto the r-dimensional subspace of $\mathbb{R}^N$ spanned by the columns of $\mathcal{R}(\mathbf{U})$, where $\mathcal{R}(\cdot)$ denotes the range space operator and $\mathbf{U} \in \mathbb{R}^{N\times r}$ is a full-column rank matrix, assumed, w.l.o.g., to be semi-unitary. System (4) converges to the desired orthogonal projection of the initial value vector $\mathbf{z}[0] = \mathbf{g}$ onto $\mathcal{R}(\mathbf{U})$, for *any* given $\mathbf{g} \in \mathbb{R}^N$, if and only if

$$\lim_{k\to+\infty} \mathbf{z}[k] = \lim_{k\to+\infty} \mathbf{W}^k\mathbf{g} = \mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{g}, \tag{5}$$

i.e.,

$$\lim_{k\to+\infty} \mathbf{W}^k = \mathbf{P}_{\mathcal{R}(\mathbf{U})}. \tag{6}$$

Necessary and sufficient conditions for (6) were proved in [8] and are given in the following.

Proposition 1 *Given the dynamical system in (4) and the projection matrix $\mathbf{P}_{\mathcal{R}(\mathbf{U})}$, the vector $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ is globally asymptotically stable for* any fixed $\mathbf{z}[0] \in \mathbb{R}^N$*, if and only if the following conditions are satisfied:*

$$\mathbf{W}\,\mathbf{P}_{\mathcal{R}(\mathbf{U})} = \mathbf{P}_{\mathcal{R}(\mathbf{U})} \tag{C.1}$$

$$\mathbf{P}_{\mathcal{R}(\mathbf{U})}\,\mathbf{W} = \mathbf{P}_{\mathcal{R}(\mathbf{U})} \tag{C.2}$$

$$\rho\left(\mathbf{W} - \mathbf{P}_{\mathcal{R}(\mathbf{U})}\right) < 1 \tag{C.3}$$

where $\rho(\cdot)$ denotes the spectral radius operator. Under (C.1)-(C.3), the error vector $\mathbf{e}[k] \triangleq \mathbf{z}[k] - \mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ satisfies the following dynamics:

$$\mathbf{e}[k+1] = \left(\mathbf{W} - \mathbf{P}_{\mathcal{R}(\mathbf{U})}\right)\mathbf{e}[k], \qquad k = 0, 1, \ldots. \tag{7}$$

□

Remark 1−**Interpretation of necessary and sufficient conditions**: Interestingly, conditions C.1-C.3 have an intuitive interpretation, as described next. C.1 and C.2 state that, if system in (4) asymptotically converges, then it is guaranteed to converge to the desired value. In fact, C.1 guarantees that the projection of vector $\mathbf{z}[k]$ onto $\mathcal{R}(\mathbf{U})$ is an invariant quantity for the dynamical system, implying that the system in (4), during its evolution, keeps the component $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ of $\mathbf{z}[0]$ unaltered; whereas C.2 makes $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ a fixed point of matrix $\mathbf{W}$ and thus a potential accumulation point for the sequence $\{\mathbf{z}[k]\}_k$. However, both conditions C.1 and C.2 do not state anything about the convergence of the dynamical system; which is instead guaranteed by C.3, imposing that all the modes associated to the eigenvectors orthogonal to $\mathcal{R}(\mathbf{U})$ be asymptotically vanishing [cf. (7)].

Remark 2−**Special cases**: Observe that, as special case, our conditions C.1-C.3 contain the well-known convergence conditions of linear discrete-time dynamical systems toward the (weighted) average consensus (see, e.g., [2]). It is sufficient to set in (5), $r = 1$ and $\mathbf{U} = \mathbf{u} = \frac{1}{\sqrt{N}}\mathbf{1}_N$, where $\mathbf{1}_N$ is the N-length vector of all ones. In such a case, C.1-C.3 can be restated as following: the digraph associated to the network described by $\mathbf{W}$ must be strongly connected and balanced.

Given Proposition 1, the challenging question in our sensor network context is to find whether (6) can be satisfied using a *sparse* matrix.

3. MAXIMUM CONVERGENCE RATE UNDER TOPOLOGY CONSTRAINTS VIA SDP

Energy consumption is one of the major concerns of distributed (iterative) algorithms. Assuming, for simplicity, that all nodes use the same transmit power, the energy spent to reach decisions with the desired accuracy or reliability is the product between the transmit power and the convergence time. For any given spatial distribution of the sensors, the transmit power constraint reflects into a topology constraint, as each node interacts only with the nodes lying within its coverage radius. We propose now an optimization method that, for any given transmit power (and thus network topology), minimizes the convergence time. More specifically, we derive the optimal matrix $\mathbf{W}$, consistent with the network topology constraints, that maximizes the convergence speed of system (4), while guaranteeing the convergence of (4) to the desired final vector $\mathbf{z}^\star = \mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{g}$ (i.e., satisfying constraints C.1-C.3 in Proposition 1). There is indeed a fundamental feasibility constraint, implying that not all network topologies guarantee the convergence to any desired projection vector. For the moment, we assume that the topology constraint and the signal subspace lead to a feasible problem. Later, we will show that, to guarantee the feasibility of the problem, it is necessary to increase the transmit power of each node proportionally to the dimension of the useful signal subspace.

The convergence rate can be measured as the worst case convergence rate of (4) that, under constraints C.1-C.3 (Proposition 1), is given by the following asymptotic convergence exponent [11]:

$$d(\mathbf{W}) = \sup_{z[0]\neq\mathbf{z}^\star} \lim_{k\to\infty} \frac{1}{k} \ln\left(\frac{\|\mathbf{z}[k] - \mathbf{z}^\star\|}{\|\mathbf{z}[0] - \mathbf{z}^\star\|}\right), \tag{8}$$

where $\mathbf{z}^\star = \mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{g}$ is the desired final vector and $\|\cdot\|$ is any vector norm. Since, for large k

$$\|\mathbf{z}[k] - \mathbf{z}^\star\| \simeq C\, e^{d(\mathbf{W})k}, \tag{9}$$

where C is a constant that depends on the initial conditions, $d(\mathbf{W})$ gives the (*asymptotic*) number of iterations for the error to decrease by the factor $1/e$ (for the worst possible initial vector), also known as *convergence time:*

$$\tau(\mathbf{W}) = \frac{1}{\ln(1/d(\mathbf{W}))}. \tag{10}$$

Given the dynamical system (4), with $\mathbf{W}$ satisfying C.1-C.3, it follows from (7) in Proposition 1 and [11, Theorem 3.4] that

$$d(\mathbf{W}) = \rho(\mathbf{W} - \mathbf{P}_{\mathcal{R}(\mathbf{U})}). \tag{11}$$

Hence, the minimization of the convergence time in (10) while guaranteeing the convergence of the system to the desired final vector in (5), is equivalent to the minimization of $\rho(\mathbf{W}-\mathbf{P}_{\mathcal{R}(\mathbf{U})})$, under constraints C.1-C.3.

3.1. Maximization of the convergence rate

Minimizing the spectral radius of a non-symmetric matrix is a notoriously difficult problem, intractable except for small-medium values of the dimensions [7]. Some optimization problems involving the minimization of the spectral radius were indeed shown to be NP-hard [9, 10]. Since in typical sensor network problems, the dimension of $\mathbf{W}$ may be quite large and, furthermore, the advantage of using a non-symmetric matrix as opposed to a symmetric one is unclear, we will restrict our search to the class of *symmetric* matrices $\mathbf{W}$. Furthermore, we consider w.l.o.g. matrices $\mathbf{W}$ having the following structure:

$$\mathbf{W}=\mathbf{I}-\epsilon\mathbf{L}, \qquad \epsilon\in\mathbb{R}, \qquad \mathbf{L}=\mathbf{L}^T, \qquad \mathbf{L}\succeq\mathbf{0}. \tag{12}$$

We focus now on the constraints to impose in the optimization problem. First of all, we need to guarantee the convergence of (4) to the final vector $\mathbf{z}^\star=\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$, for any given $\mathbf{z}[0]\in\mathbb{R}^N$, implying the necessary (and sufficient) conditions C.1-C.3 in Proposition 1 to be satisfied by matrix $\mathbf{L}$ in (12). In addition, we have the topological constraints, resulting from the network geometry and the transmit power of each node. We rewrite those constraints in terms of constraints on matrix $\mathbf{L}$, as detailed next.

Conditions C.1 and C.2 : Since $\mathbf{L}=\mathbf{L}^T$ [see (12)], both C.1 and C.2 lead to

$$\mathbf{U}^T\mathbf{L}=\mathbf{0} \quad\Leftrightarrow\quad \mathbf{L}=\mathbf{P}_{\mathcal{R}(\mathbf{U})^\perp}\mathbf{L}\mathbf{P}_{\mathcal{R}(\mathbf{U})^\perp}. \tag{13}$$

Introducing the semi-unitary matrix $\mathbf{U}^\perp\in\mathbb{R}^{N-r\times N-r}$ such that $\mathbf{U}^T\mathbf{U}^\perp=\mathbf{0}$, condition (13) states that every feasible $\mathbf{L}$ must belong to the range space of $\mathbf{U}^\perp$, implying the following structure for $\mathbf{L}$:

$$\mathbf{L}=\mathbf{U}^\perp\overline{\mathbf{L}}\mathbf{U}^{\perp^T}, \tag{14}$$

with

$$\overline{\mathbf{L}}\in\mathcal{L}\triangleq\left\{\mathbf{X}\in\mathbb{R}^{(N-r)\times(N-r)}\,:\,\mathbf{X}=\mathbf{X}^T,\quad \mathbf{X}\succeq\mathbf{0}\right\}, \tag{15}$$

where, in the definition of $\mathcal{L}$, we used (12). Thus, there is no loss of generality in looking for the optimal matrices $\mathbf{L}$ assuming the form (14).

Condition C.3: Using (14), we impose now constraint C.3 on $\mathbf{L}$, as given next.

$$\rho\left(\mathbf{W}-\mathbf{P}_{\mathcal{R}(\mathbf{U})}\right)=\rho\left(\mathbf{U}^\perp\left(\overline{\mathbf{I}}-\epsilon\overline{\mathbf{L}}\right)\mathbf{U}^{\perp^T}\right)=\rho\left(\overline{\mathbf{I}}-\epsilon\overline{\mathbf{L}}\right)<1, \tag{16}$$

where in the first equality we used (14) and $\overline{\mathbf{I}}$ denotes the $N-r$ identity matrix. Using (14), condition (16) is equivalent to the following:

$$\epsilon\lambda_i(\overline{\mathbf{L}})>0 \qquad\text{and}\qquad \epsilon\lambda_i(\overline{\mathbf{L}})<2,\quad\forall i, \tag{17}$$

where $\{\lambda_i(\overline{\mathbf{L}})\}_{i=1}^{N-r}$ denotes the set of eigenvalues of $\overline{\mathbf{L}}$. It follows from (12) that conditions in (17) can be rewritten as:

$$\overline{\mathbf{L}}\succ\mathbf{0} \qquad\text{and}\qquad 0<\epsilon<\frac{2}{\lambda_i(\overline{\mathbf{L}})}. \tag{18}$$

Sparsity constraints: The constraint imposing that each node interacts only with its neighbors can be formulated by setting the appropriate values of $\mathbf{W}$, and then $\mathbf{L}$, equal to 0, i.e., [see (14)]

$$[\mathbf{L}]_{ij}=\left[\mathbf{U}^\perp\overline{\mathbf{L}}\mathbf{U}^{\perp^T}\right]_{ij}=0\quad\forall i,j=0\quad\forall i,j\in\mathcal{B}, \tag{19}$$

where $\mathcal{B}$ is the set containing all the pairs of nodes connected by an active link. Using (14), (18), and (19), the optimization problem can thus be formulated as follows:

$$\begin{aligned}
\underset{\overline{\mathbf{L}},\,\epsilon}{\text{minimize}}\quad & \rho\left(\overline{\mathbf{I}}-\epsilon\overline{\mathbf{L}}\right)\\
\text{subject to}\quad & \overline{\mathbf{L}}\succ\mathbf{0},\quad \overline{\mathbf{L}}=\overline{\mathbf{L}}^T,\quad \mathrm{Tr}(\overline{\mathbf{L}})=1,\\
& \left[\mathbf{U}^\perp\overline{\mathbf{L}}\mathbf{U}^{\perp^T}\right]_{ij}=0\quad\forall i,j\in\mathcal{B},\\
& 0<\epsilon<\tfrac{2}{\lambda_i(\overline{\mathbf{L}})},\quad\forall i\in\{1,\ldots,N-r\}.
\end{aligned} \tag{20}$$

Observe that the maximization of the convergence rate as formulated in (20) differs from the approaches proposed in the literature to accelerate classical *consensus* algorithms (see, e.g., [5, 6]), since we solve a much more general problem than consensus and, in our formulation, we consider the *joint* optimization of the step size ϵ and the weight matrix $\mathbf{L}$, including also sparsity constraints on $\mathbf{W}$.

Problem (20) is not convex and it might not be feasible. Our goal now is to convert (20) into a convex Semidefinite Programming (SDP), so that its solution can be computed efficiently [12]. Here after, we assume that the network topology constraints are such that the feasible set of (20) is nonempty. Later on, we will show that this condition is valid if the transmit power of each node exceeds a value depending on the dimension of the signal subspace.

3.2. SDP reformulation

Problem (20) is separable in the variables $\overline{\mathbf{L}}$ and ϵ, since the constraints in (20) depend only on $\overline{\mathbf{L}}$ or ϵ . Thus, in the following, we solve (20) by first minimizing the objective function over ϵ for a given feasible $\overline{\mathbf{L}}$, and then minimizing the resulting objective function over $\overline{\mathbf{L}}$.

Minimizing over ϵ: Denoting by $\left\{\lambda_{(i)}(\overline{\mathbf{I}}-\epsilon\overline{\mathbf{L}})\right\}$ and $\left\{\lambda_{(i)}(\overline{\mathbf{L}})\right\}$ the set of eigenvalues of $\overline{\mathbf{I}}-\epsilon\overline{\mathbf{L}}$ and $\overline{\mathbf{L}}$, respectively, arranged in increasing order [i.e., $\lambda_{(i)}(\cdot)\le\lambda_{(i+1)}(\cdot)$, for all $i\in\{1,\ldots,N-r\}$], the objective function in (20) can be rewritten as:

$$\rho\left(\overline{\mathbf{I}}-\epsilon\overline{\mathbf{L}}\right)=\max_{k\in\{1,\ldots,N-r\}}\left\{1-\epsilon\lambda_{(1)}(\overline{\mathbf{L}}),\,\epsilon\lambda_{(N-r)}(\overline{\mathbf{L}})-1\right\}, \tag{21}$$

which is a piecewise-linear convex function in the variable $0<\epsilon<\frac{2}{\lambda_i(\overline{\mathbf{L}})}$. It follows that, for any given $\overline{\mathbf{L}}$, the global minimum of (21) over $\epsilon\in(0,\,2/\lambda_i(\overline{\mathbf{L}}))$ is achieved when $1-\epsilon\lambda_{(1)}(\overline{\mathbf{L}})=\epsilon\lambda_{(N-r)}(\overline{\mathbf{L}})-1$ (recall that $0<\lambda_{(1)}(\overline{\mathbf{L}})\le\lambda_{(N-r)}(\overline{\mathbf{L}})$, for any feasible $\overline{\mathbf{L}}$), implying the following optimal (feasible) value of ϵ:

$$\epsilon^\star=\frac{2}{\lambda_{(1)}\left(\overline{\mathbf{L}}\right)+\lambda_{(N-r)}\left(\overline{\mathbf{L}}\right)}. \tag{22}$$

Using (22), (21) can be rewritten as

$$\rho\left(\overline{\mathbf{I}}-\epsilon^\star\overline{\mathbf{L}}\right)=\frac{\lambda_{(N-r)}(\overline{\mathbf{L}})-\lambda_{(1)}(\overline{\mathbf{L}})}{\lambda_{(1)}(\overline{\mathbf{L}})+\lambda_{(1)}(\overline{\mathbf{L}})}\triangleq\frac{\kappa(\overline{\mathbf{L}})-1}{\kappa(\overline{\mathbf{L}})+1}, \tag{23}$$

where in the last equality we used $\overline{\mathbf{L}}\succ\mathbf{0}$ and introduced the condition number $\kappa(\overline{\mathbf{L}})\triangleq\lambda_{(N-r)}(\overline{\mathbf{L}})/\lambda_{(1)}(\overline{\mathbf{L}})$ of matrix $\overline{\mathbf{L}}$. We can now find the optimal $\overline{\mathbf{L}}$ minimizing $\rho\left(\overline{\mathbf{I}}-\epsilon^\star\overline{\mathbf{L}}\right)$, under constraints in (20).

Minimizing over $\overline{\mathbf{L}}$: Given the optimal $\epsilon^\star$ in (22) and the resulting objective function $\rho\left(\overline{\mathbf{I}}-\epsilon^\star\overline{\mathbf{L}}\right)$ in (23), the optimization problem (20) can be rewritten in epigraph form as:

$$\begin{aligned} \underset{\overline{\mathbf{L}},\,\gamma}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & \frac{\lambda_{(N-r)}(\overline{\mathbf{L}})}{\lambda_{(1)}(\overline{\mathbf{L}})} \leq \gamma, \\ & \overline{\mathbf{L}} \succ \mathbf{0}, \quad \overline{\mathbf{L}} = \overline{\mathbf{L}}^T, \quad \mathrm{Tr}(\overline{\mathbf{L}}) = 1, \\ & \left[\mathbf{U}^{\perp}\overline{\mathbf{L}}\mathbf{U}^{\perp\mathbf{T}}\right]_{ij} = 0 \quad \forall i,j \in \mathcal{B}. \end{aligned} \tag{24}$$

where to write (24) from (20) and (23) we used the fact that $\rho\left(\overline{\mathbf{I}} - \epsilon^{\star}\overline{\mathbf{L}}\right)$ in (23) is an increasing function of $\kappa(\overline{\mathbf{L}})$. For any given $\overline{\mathbf{L}} \succ \mathbf{0}$ and $\gamma > 0$, it is not difficult to prove the following equivalence:

$$\frac{\lambda_{(N-r)}(\overline{\mathbf{L}})}{\lambda_{(1)}(\overline{\mathbf{L}})} \leq \gamma \quad \Leftrightarrow \quad \exists \mu > 0 \text{ such that } \mu\mathbf{I} \preceq \overline{\mathbf{L}} \preceq \gamma\mu\mathbf{I}. \tag{25}$$

Using (25) and introducing the following change of variables:

$$\widetilde{\mathbf{L}} = \overline{\mathbf{L}}\frac{1}{\mu} \qquad \widetilde{\mu} = \frac{1}{\mu}, \tag{26}$$

problem (24) can be rewritten as:

$$\begin{aligned} \underset{\widetilde{\mathbf{L}},\,\gamma,\,\widetilde{\mu}}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & \mathbf{I} \preceq \widetilde{\mathbf{L}} \preceq \gamma\mathbf{I}, \quad \widetilde{\mu} > 0, \\ & \widetilde{\mathbf{L}} = \widetilde{\mathbf{L}}^T, \quad \mathrm{Tr}(\widetilde{\mathbf{L}}) = \widetilde{\mu}, \\ & \left[\mathbf{U}^{\perp}\widetilde{\mathbf{L}}\mathbf{U}^{\perp\mathbf{T}}\right]_{ij} = 0 \quad \forall i,j \in \mathcal{B}. \end{aligned} \tag{27}$$

The eigenvalue problem in (27) can be efficiently solved using classical SDP tools [12]. Once an optimal solution $\left(\widetilde{\mathbf{L}}^{\star}, \gamma^{\star}, \widetilde{\mu}^{\star}\right)$ to (27) is computed, the optimal original $\mathbf{L}^{\star}$ can be obtained through (26) and (14): $\mathbf{L}^{\star} = \widetilde{\mu}^{\star -1}\mathbf{U}^{\perp}\widetilde{\mathbf{L}}^{\star}\mathbf{U}^{\perp^T}$.

4. NUMERICAL RESULTS AND CONCLUSION

As an example, in Fig. 1, we show the minimum convergence time obtained for a network of 25 sensors distributed over a unit square, as a function of the square coverage radius (assumed to be the same for every node). Assuming a power attenuation law $p_R = p_T/r^2$, the abscissa is proportional to the transmit power necessay to induce a unit receive power p_R. We considered both cases of sensors uniformly spaced (solid line) and randomly distributed (dashed line). Fig. 1 a) refers to the projection onto a signal subspace spanned by two-dimensionl polynomials of degree $d = 0, 1$, and 2. The minimum number of neighbors guaranteeing the problem feasibility is equal to the number of two-dimensional independent polynomials of degree up to d. In this example, the number of neighbors is then $1, 3$, and 6, respectively. The transmit power has to be chosen to ensure this minimum number of neighbors. Fig. 1 b) refers to the projection onto a signal subspace spanned by two-dimensional Fourier bases including up to the harmonic of degree $d = 0$ and 1. Also in this case, the minimum number of neighbors is equal to the number of independent sinusoids of degree up to d. In this example, the minimum number of neighbors is 1 and 5, respectively. In both cases, as expected, as the coverage area increases, the convergence time decreases. However, this entails a greater transmit power to cover a larger area. On the other hand, the convergence time increases if, for a given number of neighbors, the dimension of the kernel space increases. In conclusion, the distributed algorithm proposed in this paper is able to project the whole observation vector onto the useful signal subspace using a totally distributed approach, using minimum convergence time, for any given topology ensuring the problem feasibility.

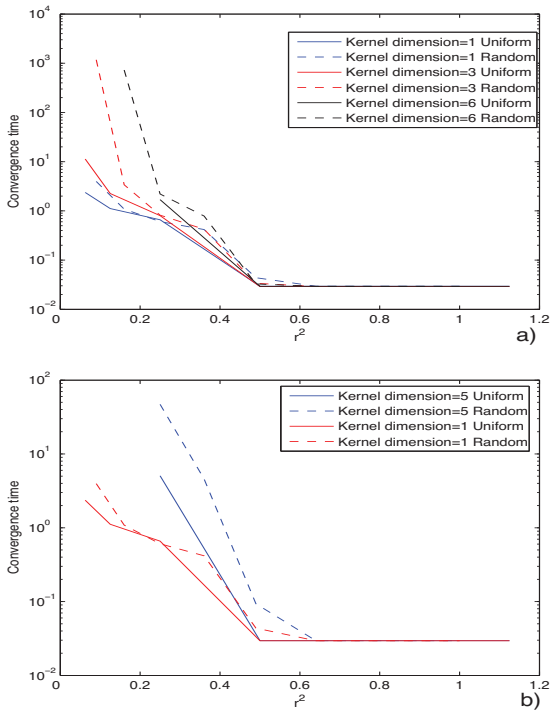

Fig. 1. Minimum convergence time vs. number of neighbors: a) polynomial basis; b) Fourier basis.

5. REFERENCES

[1] A. Giridhar, P.R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Communications Magazine*, pp. 98 – 107, April 2006.

[2] J. N. Tsitsiklis, D. P. Bertsekas, M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. on Automatic Control*, pp. 803–812, Sep. 1986.

[3] S. Barbarossa and G. Scutari, "Bio-inspired Sensor Network Design: Distributed Decision Through Self-synchronization," *IEEE Signal Processing Magazine*, vol. 24, no. 3, May 2007, pp. 26–35.

[4] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, A. Ribeiro, "Consensus in Ad Hoc WSNs With Noisy LinksPart II: Distributed Estimation and Smoothing of Random Signals", *IEEE Trans. on Signal Proc.*, April 2008, pp. 1650–1666.

[5] L. Xiao and S. Boyd, "Fast Linear Iterations for Distributed Averaging," in *Systems and Control Letters*, 53:65-78, 2004.

[6] L. Xiao, S. Boyd, S.J. Kim, "Distributed average consensus with Least Mean Square Deviations," *Journal of Parallel and Distributed Computing*, vol. 67, Jan. 2007, pp. 33–46.

[7] C. D. Meyer, R. J. Plemmons, "Convergence Powers of Matrix with Applications to Iterative Methods for Singular Linear Systems," *SIAM Jour. on Numerical Analysis*, Vol. 14, No. 4, pp. 699-705, 1977.

[8] S. Barbarossa, G. Scutari, T. Battisti, "Distributed projection of observed data onto signal subspace in wireless sensor networks using local exchange algorithms", submitted to *IEEE Trans. on Signal Processing*, (May 2008).

[9] V. Blondel, J.N. Tsitsiklis, "NP-hardness of some linear control design problems," *SIAM Jour. on Control and Optimization*Vol. 35, pp. 2118–2127, 1997.

[10] A. Nemirovskii, "Several NP-hard problems arising in robust stability analysis," *Mathematics of Control, Signal and Systems*, vol. 6, pp. 99–105, 1993.

[11] R. Varga, *Matrix Iterative Analysis*, Springer-Verlag, 2th Ed., 2000.

[12] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.