Finite-Time Average Consensus based Protocol for Distributed Estimation over AWGN Channels

Alain Y. Kibangou

Abstract— This paper studies the problem of distributed estimation of a static parameter with sensors communicating through an additive white Gaussian noise (AWGN) channel. In the noiseless case, we first introduce the concept of finite-time average consensus, in which nodes can compute exactly the average in a finite number of steps for an arbitrary graph, provided the topology is time-invariant. In fact, finite-time consensus is achieved owing to joint diagonalizable matrices. By considering a linear iterations scheme, we derive closed form expressions for such matrices. Then, based on an ensemble averaging method we show how average consensus can be asymptotically reached over AWGN channels. Performance analysis of the suggested protocol is given along with comparisons with other methods in the literature.

I. Introduction

The problem of unknown parameter estimation using a sensor network that consists of a given number of distributed nodes over an area have attracted a considerable attention in the last decade. In a centralized fusion scheme, each sensor sends its measurement to a fusion center. Then, the fusion center extracts the minimum variance estimate of the parameter from aggregate measurements of the sensors. Such a scheme suffers from communication overheads and is vulnerable to fusion center failure. Several approaches suggested in the literature consider distributed fusion schemes, where, from its own measurements and those of its neighbors, each node computes a local estimation, which iteratively converges to the intended result. It is well known that, the noise variance is the same for the sensor nodes, the best linear unbiased minimum variance estimator is reduced to an averaging over the measured values of the sensors. Therefore, in such a case, distributed estimation can be solved as an average consensus problem (see [1] and references therein).

In a consensus problem, a group of agents or network nodes try to reach agreement on a given quantity of interest that depends on their states [2]. Algorithms that enable such consensus reaching belong typically to two classes: Distributed averaging based algorithms and probabilistic counting approaches, such as randomized gossip schemes [3].

Average consensus can be reached using a linear iterations scheme where each node repeatedly updates its value as a weighted linear combination of its own value and those of its neighbors. The main benefit of using a linear iterations scheme is that, at each time-step, each node only has to

transmit a single value to each of its neighbors. Based on such a scheme, several algorithms have been proposed in the literature. However, in the majority of the proposed algorithms the weights are chosen so that all the nodes asymptotically converge to the same value.

Consensus can be embedded as a step of more sophisticated distributed algorithm as it is the case for the Distributed Kalman filter [4] and the Distributed Alternating Least Squares algorithm [5]. Asymptotic convergence are not suitable for these kinds of distributed methods. Even though, speed convergence of consensus algorithm have been explored [6], [7] in order to derive fast consensus algorithms, running standard consensus in finite-time constitute a source of errors not easily quantifiable. Sometimes, bounds can be derived. Therefore, it is interesting to address the question of exact consensus in finite-time.

A number of authors have studied finite-time consensus in the framework of continuous-time systems [8], [9]. In general, the results therein cannot be directly translated to discrete-time systems.

In the discrete-time framework, in [10], finite-time average consensus was briefly discussed. However, the described method requires the graph to be fully connected for at least one time-step. A data aggregation based algorithm was proposed in [3]. Such an approach requires additional memory compared to linear iterations strategies. Another approach requiring less memory, but linear iterations, is derived in [11] by calculating weights so that the observability matrix of the network is contained in some defined space. The basic idea is that given enough time, the nodes will have observed enough to reconstruct the initial state of the system. At which time, they can compute the correct average. In [12], the same authors have shown that each node can calculate the consensus value as a linear combination of its own past values over at most D time-steps, D being the degree of the minimal polynomial of the associated weight matrix. In [13], based on properties of de Bruijn's graph and block Kronecker product, it has been shown that the average consensus problem can be reached in finite time if the number of nodes is an exact power of the maximum in-degree of the graph. Another interesting contribution is that in [14] where finite-time average consensus problem is formulated as a matrix factorization problem. However, the proposed approach is fully centralized and requires scheduling of nodes connection.

In this paper, for time-invariant topologies and in the perfect information exchange case, i.e. without channel noise nor quantization, we show that the finite-time average con-

sensus problem can be solved as a matrix factorization problem with joint diagonalizable matrices. Then, by periodically restarting the consensus algorithm, we show that, in the noisy case, average consensus can be achieved asymptotically. The remainder of the paper is organized as follows: in section II, the estimation problem using average consensus is posed. Then, in section III, we derive closed-form expressions and conditions for reaching average consensus in finite-time for ideal data exchange. In section IV, finite-time average consensus is periodically used for reaching asymptotic consensus for noisy data exchange. Before concluding the papers, some simulation results are presented in section V.

II. PROBLEM SETTING

Let us consider a network of N distributed nodes. Each sensor has a measurement $m(x_n)$ of a non-random parameter θ:

$$m(x_n) = \theta + v_n, \quad n = 1, 2, \cdots, N$$

where v_n's are assumed to be zero mean noise. The Maximum Likelihood (ML) is simply the average of the local measurements, i.e. $\hat{\theta} = \frac{1}{N} \sum_{n=1}^{N} m(x_n)$. The goal is to compute this averaged value distributively through the network.

Let us assume that the communication links in the network are modeled via an undirected graph $\mathscr{G} = (\mathscr{X}, \mathscr{E})$, where $\mathscr{X} = \{x_1, \cdots, x_N\}$ is the set of nodes and $\mathscr{E} \subseteq \mathscr{X} \times \mathscr{X}$ is the set of edges. We denote by $\mathscr{N}_i$ the set of nodes that can transmit information to node x_i. Its cardinality is denoted N_i. The elements of the adjacency matrix $\mathbf{A}$ are defined as $a_{ij} = 1$ if there is an edge between nodes i and j, otherwise they are zero. Since the graph is undirected, the adjacency matrix is symmetric ($\mathbf{A} = \mathbf{A}^T$). The degree matrix $\mathbf{D}$ is a diagonal matrix such that its entry d_{ii} is equal to N_i. The Laplacian matrix is then defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. It is also symmetric for undirected graphs. Its eigenvalues, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$, contain very significant information about the topology of the graph $\mathscr{G}$. In particular, we have $\lambda_1 = 0$ with $\mathbf{1}$, an all ones vector, as eigenvector.

Let us aggregate the values of all nodes at time-step k into the vector $\mathbf{x}(k)$, with $\mathbf{x}(0) = \begin{pmatrix} m(x_1) & \cdots & m(x_n) \end{pmatrix}^T$. In the linear iterations framework, each node updates its value as a linear combination of its own value and that of its neighbors:

$$\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k) \tag{1}$$

where the off-diagonal entries $w_{i,j}$ of $\mathbf{W}$ are nonzero if and only if $x_j \in \mathscr{N}_i$. The average consensus is reached if

$$\lim_{k \to \infty} \mathbf{x}(k) = \frac{1}{N} \mathbf{1}\mathbf{1}^T \mathbf{x}(0),$$

meaning that

$$\lim_{k \to \infty} \mathbf{W}^k = \frac{1}{N} \mathbf{1}\mathbf{1}^T.$$

It is now well known that consensus is achieved if and only if $\mathbf{W}$ admits 1 as a simple eigenvalue while the remaining eigenvalues have magnitude strictly less than 1, the left and right eigenvectors of $\mathbf{W}$ associated with the eigenvalue 1

being $\frac{1}{N}\mathbf{1}$ and $\mathbf{1}$ [6]. With these conditions asymptotic convergence is guaranteed. That is the case for the Perron matrix $\mathbf{W} = \mathbf{I} - \gamma\mathbf{L}$, with $0 < \gamma < 1/N_{max}$, $N_{max} = \max\{N_1, \cdots, N_N\}$.

In the following section, for a static network topology, we investigate the way for achieving average consensus in a finite number of steps.

III. CONVERGENCE IN A FINITE NUMBER OF TIME-STEPS USING JOINT DIAGONALIZABLE MATRICES

Our goal is to find a set of matrices $\{\mathbf{W}_i\}_{i=1,\cdots,D}$, consistent with the network topology[1], so that

$$\prod_{i=1}^{D} \mathbf{W}_i = \frac{1}{N} \mathbf{1}\mathbf{1}^T \tag{2}$$

Finding this set of matrices is equivalent to solve a multivariate polynomial system of equations. In general, studying the existence of solutions to such a system of equations can be untractable. However, assume that the matrices $\mathbf{W}_i$ are jointly diagonalizable. Let $\mathbf{U}$ be the orthogonal matrix that diagonalizes the matrices $\mathbf{W}_i$. We get:

$$\mathbf{W}_i = \mathbf{U}\mathbf{E}_i\mathbf{U}^T, \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_N$$

$\mathbf{E}_i$ being a diagonal matrix. We can therefore rewrite (2) as:

$$\mathbf{U}\left(\prod_{i=1}^{D} \mathbf{E}_i\right)\mathbf{U}^T = \frac{1}{N}\mathbf{1}\mathbf{1}^T \tag{3}$$

If $\mathbf{U}$ can be written as $\mathbf{U} = \begin{pmatrix} \frac{1}{\sqrt{N}}\mathbf{1} & \tilde{\mathbf{U}} \end{pmatrix}$ with $\tilde{\mathbf{U}}^T\tilde{\mathbf{U}} = \mathbf{I}_{N-1}$ and $\tilde{\mathbf{U}}^T\mathbf{1} = \mathbf{0}$ then (3) can be rewritten as:

$$\mathbf{U}\left(\prod_{i=1}^{D} \mathbf{E}_i\right)\mathbf{U}^T = \mathbf{U}diag(1 \quad 0\cdots0)\mathbf{U}^T. \tag{4}$$

This condition implies that $\mathbf{1}$ should be an eigenvector of $\mathbf{W}_i$. Therefore, our aim in the sequel is to define a set of topology consistent matrices $\mathbf{W}_i$ jointly diagonalizable, with $\mathbf{1}$ as eigenvector, so that solutions of equation (4) be tractable. For this purpose, we define the matrix $\tilde{\mathbf{A}} = N_{max}\mathbf{I} - \mathbf{L}$ whose properties are:

1) $\tilde{\mathbf{A}}$ is symmetric;
2) Its eigenvalues ε_i are such that $\varepsilon_i = N_{max} - \lambda_i$, where λ_i are the eigenvalues of the Laplacian matrix $\mathbf{L}$. In particular N_{max} is a simple eigenvalue.
3) The eigenvectors of the Laplacian matrix are also eigenvectors of $\tilde{\mathbf{A}}$. In particular $\mathbf{1}$ is the eigenvector associated with the eigenvalue N_{max}.

We can now state the following lemma:

Lemma 1: The matrices $\mathbf{W}_i = \alpha_i\mathbf{I} + \beta\tilde{\mathbf{A}}$ are jointly diagonalizable, consistent with the network topology, and admits $\mathbf{1}$ as eigenvector.

Proof: The eigenvalue decomposition of $\tilde{\mathbf{A}}$ yields $\tilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, with $\boldsymbol{\Delta} = diag(\varepsilon_1, \cdots, \varepsilon_N)$ and $\mathbf{U}$ containing

---

[1] A matrix is said to be consistent with the network topology if its off-diagonal (i, j)th entry equals to zero for $j \notin \mathscr{N}_i$

the eigenvectors of the Laplacian matrix. Therefore, since $\mathbf{W}_i = \alpha_i \mathbf{I} + \beta \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, we get

$$
\begin{aligned}
\mathbf{U}^T \mathbf{W}_i \mathbf{U} &= \alpha_i \mathbf{I} + \beta \mathbf{\Lambda} \\
&= diag(\alpha_i + \beta \varepsilon_1, \cdots, \alpha_i + \beta \varepsilon_N).
\end{aligned}
\tag{5}
$$

Note that the consistency with the network topology comes directly from the dependence with the Laplacian matrix. ∎ Since $\mathbf{U}$ contains the eigenvectors of the Laplacian matrix, the set of matrices $\mathbf{W}_i$ is completely determined by solving

$$
\begin{aligned}
\prod_{i=1}^{D} \mathbf{E}_i &= diag(\prod_{i=1}^{D}(\alpha_i + \beta \varepsilon_1), \cdots, \prod_{i=1}^{D}(\alpha_i + \beta \varepsilon_N)) \\
&= diag(1 \quad 0 \cdots 0).
\end{aligned}
\tag{6}
$$

or equivalently:

$$
\begin{cases}
\prod_{i=1}^{D}(\alpha_i + \beta \varepsilon_1) = 1 \\
\prod_{i=1}^{D}(\alpha_i + \beta \varepsilon_k) = 0, \quad k = 2, \cdots, N
\end{cases}
\tag{7}
$$

A solution to this problem is stated in the following theorem:

Theorem 1: Given a connected undirected graph associated with the Laplacian matrix $\mathbf{L}$, the set of matrices $\mathbf{W}_k = (\alpha_k + N_{max}\beta)\mathbf{I} - \beta \mathbf{L}$, $k = 1, \cdots, D$, $\beta \neq 0$, allows reaching the average consensus in D steps if:

1) $D + 1$ is the number of distinct eigenvalues of the Laplacian matrix;
2) the parameters β and α_i are given by

$$
\beta = \frac{1}{\prod_{i=2}^{D+1} \sqrt[D]{\lambda_i}}
\tag{8}
$$

$$
\alpha_k = \frac{\lambda_{k+1} - N_{max}}{\prod_{i=2}^{D+1} \sqrt[D]{\lambda_i}}, \quad k = 1, \cdots, D
\tag{9}
$$

λ_i, $i = 2, \cdots, D+1$ being the nonzero distinct eigenvalues of $\mathbf{L}$.

Proof: The system of equations (7) to be solved can be rewritten as follows:

$$
\prod_{k=1}^{D}(\alpha_k + N_{max}\beta) = 1
\tag{10}
$$

$$
\prod_{k=1}^{D}(\alpha_k + N_{max}\beta - \beta \lambda_i) = 0, \quad i = 2, \cdots, N
\tag{11}
$$

One can note that the second equation can be redundant for non simple eigenvalues λ_i. Let $D + 1$ be the number of the distinct eigenvalues. Therefore, we get D distinct equations $\prod_{k=1}^{D}(\alpha_k + N_{max}\beta - \beta \lambda_i) = 0$, $i = 2, \cdots, D+1$. A solution parameterized by β is given by:

$$
\alpha_k = \beta(\lambda_{k+1} - N_{max}), \quad k = 1, \cdots, D.
$$

Replacing these expressions in (10), we get:

$$
\beta^D \prod_{k=1}^{D} \lambda_{k+1} = 1
$$

Hence the solutions given above. ∎

We can note that the structure of the weighting matrix is similar to that of the Perron matrix commonly used in standard average consensus problems. However, the main difference, in the proposed scheme, is that each node modifies the weight allocated to its own value while the weights of the values coming from the neighbors remain constant.

Example 1: Let us consider the network characterized by the following Adjacency and Laplacian matrices:

$$
\mathbf{A} = \begin{pmatrix}
0 & 1 & 0 & 0 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 1 & 0
\end{pmatrix}.
$$

We have $N_{max} = 5$ and the nonzero distinct eigenvalues of $\mathbf{L}$ are 2, 4 and 6. Therefore, $D = 3$ and $\beta = 0.2752$, $\alpha_1 = -0.8255$, $\alpha_2 = -0.2752$, and $\alpha_3 = 0.2752$.

.

A. Case of a ring topology

The Laplacian matrix associated with a ring topology is a circulant matrix given by:

$$
\mathbf{L} = \sum_{n=0}^{N-1} c_n \mathbf{J}^n, \quad c_n = \begin{cases}
2 & if \quad n = 0 \\
-1 & if \quad n = 1, N-1 \\
0 & otherwise
\end{cases}
$$

where

$$
\mathbf{J} = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & & & \ddots & \\
1 & 0 & 0 & \cdots & 0
\end{pmatrix}
$$

Its eigenvalues λ_i, $i = 1, \cdots, N$, can be computed as:

$$
\lambda_i = \sum_{n=0}^{N-1} c_n e^{\frac{2j\pi(i-1)n}{N}} = 2\left(1 - \cos\left(\frac{2\pi(i-1)}{N}\right)\right).
\tag{12}
$$

One can check that $\mathbf{L}$ has $1 + \lfloor N/2 \rfloor$ distinct eigenvalues, where $\lfloor x \rfloor$ denotes the integer part of x. Therefore, applying Theorem 1, average consensus is reached in $D = \lfloor N/2 \rfloor$ steps, with the set of matrices $\mathbf{W}_i$ parameterized by:

$$
\beta = \frac{1}{2\left(\prod_{i=1}^{D}\left(1 - \cos\frac{2\pi i}{N}\right)\right)^{1/D}}
\tag{13}
$$

and

$$
\alpha_k = \frac{-\cos\frac{2\pi k}{N}}{\left(\prod_{i=1}^{D}\left(1 - \cos\frac{2\pi i}{N}\right)\right)^{1/D}}, \quad k = 1, \cdots, D
\tag{14}
$$

Example 2: For a ring with 9 nodes. We get $D = 4$, $\beta = 0.5774$, $\alpha_1 = -0.8846$, $\alpha_2 = -0.2005$, $\alpha_3 = 0.5774$, and $\alpha_4 = 1.0851$.

B. Case of cubic lattice graphs

A cubic lattice graph is defined to be a graph $\mathscr{G}$, whose vertices are the ordered triplets on n symbols, such that two vertices are adjacent if and only if they have two coordinates in common [15]. The network has n^3 nodes and each node has $3(n-1)$ connections. The adjacency matrix has 4 distinct eigenvalues: -3, $n-3$, $2n-3$, and $3(n-1)$. As a consequence, the Laplacian matrix has 0,n,$2n$, and $3n$ as distinct eigenvalues. Applying Theorem 1, average consensus is reached in 3 steps, with the set of matrices $\mathbf{W}_i$ parameterized by:

$$\beta = \frac{1}{n\sqrt[3]{6}} \tag{15}$$

and

$$\alpha_1 = \frac{3}{n\sqrt[3]{6}}, \quad \alpha_2 = \frac{-n+3}{n\sqrt[3]{6}}, \quad \alpha_3 = \frac{-2n+3}{n\sqrt[3]{6}}. \tag{16}$$

Example 3: As stated before, for a cubic lattice graph with 27 nodes, exact average consensus is computed in 3 steps. By computing the spectral radius ρ of $\mathbf{W} - \frac{1}{N}\mathbf{1}\mathbf{1}^T$, with $\mathbf{W}$ the weights matrix in standard consensus schemes. ρ^k is the spectral radius of $\mathbf{W}^k - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ and informs us about the disagreement between $\mathbf{W}^k$ and $\frac{1}{N}\mathbf{1}\mathbf{1}^T$ [16].

	ρ	ρ^D	ρ^{10D}	ρ^{20D}
Best Perron matrix	0.5	0.125	9.3×10^{-10}	8.7×10^{-19}
Local degree	0.571	0.187	5.1×10^{-8}	2.6×10^{-15}

The best Perron matrix [6] is given by $\mathbf{I} - \frac{2}{\lambda_2 + \lambda_N}\mathbf{L}$. In this case, it is equivalent to the maximum-degree policy. We can note than more than $10D$ steps are needed for being close to the exact average value, D being the number of steps for finite-time average consensus.

IV. ASYMPTOTIC CONSENSUS WITH NOISY DATA EXCHANGE

The results derived in the previous section allow reaching an average consensus in a finite number of steps when perfect data exchange is considered. Now, we will study the case of noisy data exchange:

$$\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k) + \mathbf{v}(k+1),$$

where $\mathbf{v}$ denotes the additive noise.

In such a case, standard average consensus algorithms are known to suffer from a linear increase of noise variance with time [16]. What happens is that deviation between the values achieved by the sensors maintains a bounded variance, but the average value follows a random walk [17]. In order to limit or circumvent this undesired effect, alternative methods for designing the weights of average consensus algorithms have been proposed in the literature. For example, in [16], it was shown how to derive the weighting matrix in order to minimize the deviation between the nodes values. Starting from a Perron matrix, decreasing step-size methods have been proposed in the literature (see [18], [19] and references therein). A first order difference equation approach, which is resilient to noise, has been also proposed in [20]. In this section, we suggest a new protocol based on finite-time average consensus matrices.

The suggested protocol is as follows:

$$\mathbf{x}(k+(n-1)D) = (1-\delta_{k1})\mathbf{W}_k\mathbf{x}(k-1+(n-1)D) + \delta_{k1}\mathbf{W}_k\mathbf{x}(0) \tag{17}$$

$$\mathbf{y}(k+(n-1)D) = (1-\frac{\delta_{kD}}{n})\mathbf{y}(k-1+(n-1)D) + \frac{\delta_{kD}}{n}\mathbf{x}(k-1+(n-1)D) \tag{18}$$

where $k = 1, \cdots, D$, $n = 1, \cdots, \infty$, δ_{ij} denotes the Kronecker delta, i.e. $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ elsewhere. In this protocol, each node keeps in memory its initial value and re-send it periodically to its neighbors. With a period equals to D, in the noiseless case, one can note that

$$\mathbf{x}(nD) = \prod_{i=D}^{i}\mathbf{W}_i\mathbf{x}(0) = \bar{\mathbf{x}} = \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0)$$

Periodically the local values are all equal to the exact averaged value. When the data exchange is noisy, we get $\mathbf{x}(nD) = \bar{\mathbf{x}} + \tilde{\mathbf{v}}(nD)$ where $\tilde{\mathbf{v}}(nD)$ accounts for the noise effect. The idea represented by equation (18) is to average these noisy estimates of $\bar{\mathbf{x}}$. Therefore each node keeps in memory a value y, initialized with zero, and updates it at each period D. The price to be paid with such a scheme concerns the additional memory requirement for the initial value and for the consensus value y. With this scheme the consensus is reached on $\mathbf{y}$ and not on $\mathbf{x}$.

For noisy data exchange, equation (17) can be rewritten as:

$$\mathbf{x}(k+(n-1)D) = (1-\delta_{k1})\mathbf{W}_k\mathbf{x}(k-1+(n-1)D) + \delta_{k1}\mathbf{W}_k\mathbf{x}(0) + \mathbf{v}(k+(n-1)D),$$
$$k = 1, \cdots, D, \quad n \geq 1 \tag{19}$$

Theorem 2: Considering the consensus protocol (17)-(18), suppose that the noise sequence $\{\mathbf{v}(k+(n-1)D)\}$ is independent with zero mean and covariance $\sigma^2\mathbf{I}_N$. Then the sequence $\mathbf{y}(k+(n-1)D)$ asymptotically converges to the average value $\bar{\mathbf{x}}$ of the initial state.

Proof: We have to prove that $\lim_{n\to\infty}\mathbf{y}(k+(n-1)D) = \bar{\mathbf{x}}$. Since by construction, $\mathbf{y}(k+(n-1)D)$ is stepwise constant, it suffices to prove the above condition for a given value of k, $k = D$ for instance. We can show that $\mathbf{y}(nD) = \frac{1}{n}\sum_{m=1}^{n}\mathbf{x}(mD)$, meaning that, at each node $\mathbf{y}_i(nD) = \frac{1}{n}\sum_{m=1}^{n}\mathbf{x}_i(mD)$ acts as an unbiased estimator of a constant parameter observed in a zero mean noise. Therefore $\mathbb{E} = \lim_{n\to\infty}\mathbf{y}(nD) = \bar{\mathbf{x}}$. Now, let us compute the mean square error $\mathbb{E}\left[(\mathbf{y}(k+(n-1)D) - \bar{\mathbf{x}})^T (\mathbf{y}(k+(n-1)D) - \bar{\mathbf{x}})\right]$. From (19),

we get:

$$\mathbf{x}(nD) = \prod_{j=D}^{1} \mathbf{W}_j \mathbf{x}(0) + \sum_{i=1}^{D-1} \prod_{j=D}^{i+1} \mathbf{W}_j \mathbf{v}(i+(n-1)D) + \mathbf{v}(nD)$$

$$= \bar{\mathbf{x}} + \sum_{i=1}^{D-1} \prod_{j=D}^{i+1} \mathbf{W}_j \mathbf{v}(i+(n-1)D) + \mathbf{v}(nD).$$

Defining $\mathbf{z}(nD) = \mathbf{y}(nD) - \bar{\mathbf{x}}$ yields:

$$\mathbf{z}(nD) = \frac{1}{n} \sum_{m=1}^{n} \sum_{i=1}^{D-1} \prod_{j=D}^{i+1} \mathbf{W}_j \mathbf{v}(i+(n-1)D) + \frac{1}{n} \sum_{m=1}^{n} \mathbf{v}(mD)$$

We have then to compute $\mathbb{E}\left[\mathbf{z}^T(nD)\mathbf{z}(nD)\right] = trace(\mathbb{E}\left[\mathbf{z}(nD)\mathbf{z}^T(nD)\right])$. Owing to the independence property, the matrix $\mathbb{E}\left[\mathbf{z}(nD)\mathbf{z}^T(nD)\right]$ can be written as:

$$\mathbb{E}\left[\mathbf{z}(nD)\mathbf{z}^T(nD)\right] = \frac{\sigma^2}{n^2} \sum_{m=1}^{n} \sum_{i=1}^{D-1} \prod_{j=D}^{i+1} \mathbf{W}_j \prod_{j=i+1}^{D} \mathbf{W}_j^T + \frac{\sigma^2}{n} \mathbf{I}_N$$

$$= \frac{\sigma^2}{n} \mathbf{Q} + \frac{\sigma^2}{n} \mathbf{I}_N$$

where $\mathbf{Q} = \sum_{i=1}^{D-1} \prod_{j=D}^{i+1} \mathbf{W}_j \prod_{j=i+1}^{D} \mathbf{W}_j^T$. As a consequence, we get:

$$\mathbb{E}\left[\mathbf{z}^T(nD)\mathbf{z}(nD)\right] = \frac{\sigma^2}{n}(trace(\mathbf{Q}) + N) \qquad (20)$$

that yields $\lim_{n\to\infty} \mathbb{E}\left[\mathbf{z}^T(nD)\mathbf{z}(nD)\right] = 0$. ∎

It is important to notice that the proposed approach differs from the Monte-Carlo approach in [19]. Herein, each D time steps we get a noisy observation of the actual average value. Then $\mathbf{y}$ reaches a strong consensus instead of a mean square consensus.

In figures 1 and 2, we have plotted the theoretical mean square error (20) normalized by the noise power, i.e. $\mathbb{E}\left[(\mathbf{y}(k+(n-1)D) - \bar{\mathbf{x}})^T (\mathbf{y}(k+(n-1)D) - \bar{\mathbf{x}})\right]/N\sigma^2$ respectively for a cubic lattice graph and for a ring. In both case, as expected, we get a decreasing error with approximately the same rate for any number of nodes N. We can note that one interesting property of the cubic lattice graph is to be more scalable than the ring topology. Indeed, the number of steps D is the same for any N and the mean square error does not significantly change.

V. SIMULATION RESULTS

We consider the estimation of a parameter $\theta = 3$ through a network of N sensors. Each observation is corrupted by a white Gaussian noise with variance 10^{-2}. The nodes communicate through an AWGN with variance 10^{-4}. The performance are evaluated by means of the Normalized Mean Square Error (NMSE) defined as $NMSE_k = \|\mathbf{x}(k) - \bar{\mathbf{x}}\|^2 / \|\bar{\mathbf{x}}\|^2$, $\bar{\mathbf{x}}$ being the average of the initial values of the nodes. The results below are averaged values over 100 independent runs.

Figures 3 and 4 depict the NMSE respectively for a ring Topology with 10 nodes and a cubic lattice graph with 27 nodes. We have compared the proposed scheme with two decreasing step-size ones and with a first order difference
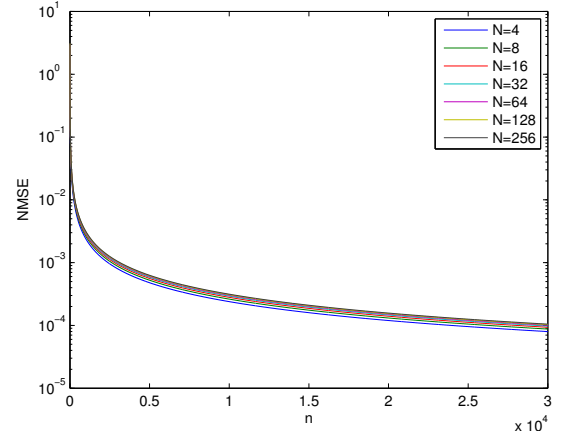

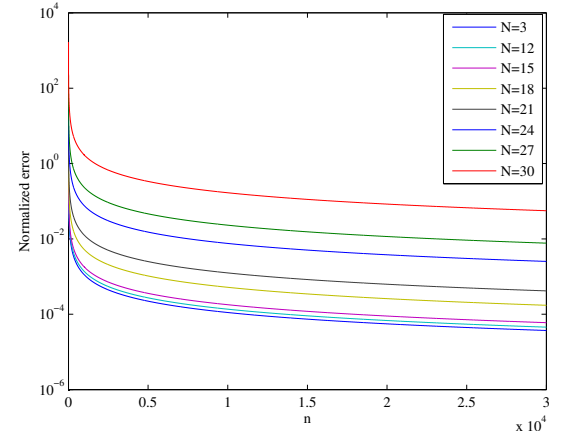Fig. 1. Theoretical mean square error normalized by the noise variance (case of a cubic lattice graph).

Fig. 2. Theoretical mean square error normalized by the noise variance (case of a ring).

based algorithm. Recall that decreasing step-size policies consist in varying the weighting matrix as follows $\mathbf{W}(k) = \mathbf{I} - \varepsilon(k)\mathbf{L}$, with $\varepsilon(k) = \frac{a}{k^b}$ for Decreasing step-size 1 and $\varepsilon(k) = \frac{1}{\lambda_N(k+1)^\mu}$ for the second one. We set $a = 0.4$, $b = 0.6$, and $\mu = 0.9$. The first order difference scheme is such that

$$\Phi(k+1) = \mathbf{x}(0) + (\mathbf{I} - \gamma\mathbf{L})\Phi(k)$$
$$\mathbf{x}(k) = \Phi(k) - \Phi(k-1)$$

Its noise resilience comes from the difference operation which counteracts the effect of the eigenvalue 1 of the weight matrix. We set $\gamma = 1/2N_{max}$.

From Figures 3 and 4 we can note that the proposed method outperforms the three others in terms of precision. Its convergence speed can be dramatically lower than that obtained with the first order difference scheme. In figure 5 where the nodes trajectory are depicted, we can see that after the D first iterations the local values are very close to the actual average value.

VI. CONCLUSION

In this paper, we have derived closed form expressions for designing weight matrices in order to reach the average consensus in a finite number of steps given by the number of
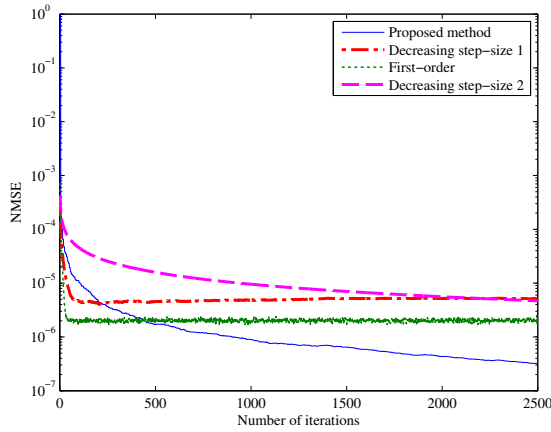
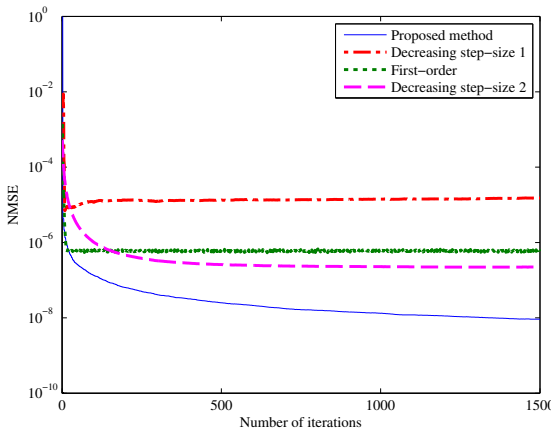Fig. 3. Evaluation of the NMSE for a ring topology with 10 nodes.

Fig. 4. Evaluation of the NMSE for a cubic lattice graph with 27 nodes.

Fig. 5. Nodes trajectory for the compared methods in the case of a cubic lattice graph.

distinct nonzero eigenvalues of the graph Laplacian matrix. The proposed finite set of weight matrices is completely characterized by the nonzero eigenvalues of the graph Laplacian matrix. By periodically restarting the finite-time consensus protocol, we have derived a new scheme for achieving asymptotic consensus in a noisy network. We have shown that strong consensus is achieved in a mean square sense. Our approach necessitates a time-invariant topology and a network configuration step for node counting and Laplacian matrix evaluation. For asymptotic convergence in the noisy configuration, our approach is very promising but the speed convergence should be improved. In addition, robustness of the proposed procedure with respect to perturbations arising from link failures and the way of computing the weighting matrices in a fully distributed scheme are to be investigated.

REFERENCES

[1] F. Garin and L. Schenato. A survey on distributed estimation and control applications using linear consensus algorithms. In Alberto Bemporad, Maurice Heemels, and Mikael Johansson, editors, *Networked Control Systems*, volume 406 of *Lecture Notes in Control and Information Sciences*, pages 75–107. Springer, 2011.

[2] R. Olfati-Saber, A. Fax, and R. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. of the IEEE*, 95(1):215–233, January 2007.

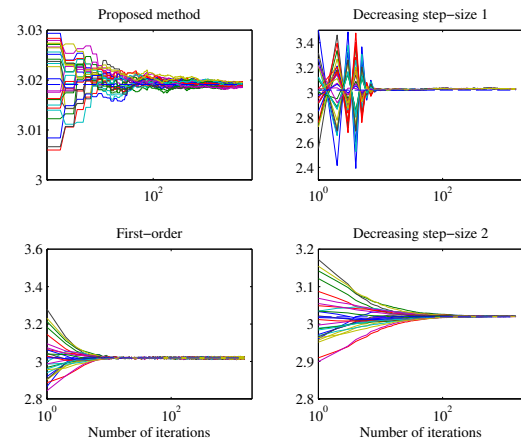[3] N. Lechevin, C.A. Rabbath, and Y. Zhang. Information broadcasting algorithm for finite-time reaching-at-risk consensus with application to weapon-target assignment. In *Proc. of American Control Conference (ACC)*, pages 3286–3291, St. Louis, MO, USA, June 10-12 2009.

[4] R. Olfati-Saber. Distributed kalman filtering for sensor networks. In *Proc. of the 46th IEEE Conf. on Decision and Control*, New Orleans, LA, USA, December 12–14 2007.

[5] A.Y. Kibangou and A.L.F. de Almeida. Distributed PARAFAC based DS-CDMA blind receiver for wireless sensor networks. In *Proc. of the IEEE Workshop SPAWC*, Marrakech, Morocco, June 20-23 2010.

[6] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems Control Lett.*, 53:65–78, 2004.

[7] E. Kokiopoulou and P. Frossard. Accelerating distributed consensus using extrapolation. *IEEE Signal Proc. Letters*, 14(10):665–668, October 2007.

[8] J. Cortes. Finite-time convergent gradient flows with application to network consensus. *Automatica*, 42(11):1993–2000, 2006.

[9] Q. Hui, W.M. Haddad, and S.P. Bhat. Finite-time semi-stability theory with applications to consensus protocols in dynamical networks. In *Proc. of American Control Conference (ACC)*, New York City, USA, 2007.

[10] D.B. Kingston and R.W. Beard. Discrete-time average consensus under switching network topologies. In *Proc. of American Control Conference (ACC)*, Minneapolis, Minenesota, USA, 2006.

[11] S. Sundaram and C. Hadjicostis. Distributed function calculation and consensus using linear iterative strategies. *IEEE Journal on Selected Areas in Communications*, 26(4):650–660, May 2008.

[12] S. Sundaram and C. Hadjicostis. Finite-time distributed consensus in graphs with time-invariant topologies. In *Proc. of American Control Conference (ACC)*, New York City, USA, July 11-13 2007.

[13] J.-C. Delvenne, R. Carli, and S. Zampieri. Optimal strategies in the average consensus problem. *Systems and Control Letters*, 58:759–765, 2009.

[14] C.-K. Ko. *On matrix factorization and scheduling for finite-time average consensus*. PhD thesis, California Institue of Technology, Pasadena, California, USA, January 2010.

[15] R. Laskar. Eigenvalues of the adjacency matrix of cubic lattice graphs. *Pacific Journal of Mathematics*, 29(3):623–629, 1969.

[16] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *J. Parallel Distrib. Comput.*, 67:33–46, 2007.

[17] L. Pescosolido, S. Barbarossa, and G. Scutari. Average consensus algorithms robust against channel noise. In *Proc. of the IEEE Workshop SPAWC*, pages 261–265, Recife, Brazil, 2008.

[18] C. Mosquera, R. Lopez-Valcarce, and S.K. Jayaweera. Distributed estimation with noisy exchanges. In *Proc. of SPAWC*, pages 236–240, Recife, Brazil, July 2008.

[19] S. Kar and J.M.F. Moura. Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise. *IEEE Trans. on Signal Proc.*, 57(1):355–369, 2008.

[20] S. Barbarossa and G. Scutari. Decentralized maximum likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems. *IEEE Trans. on Signal Proc.*, 55(7):3456–3470, 2007.