Optimal Share Factors in the Push-Sum Algorithm for Ring and Random Geometric Graph Sensor Networks

Elma Zanaj, Marco Baldi, and Franco Chiaraluce

Original scientific paper

***Abstract*—The convergence speed of an asynchronous point-to-point version of the push-sum algorithm in sensor networks is investigated both through numerical simulations and theoretical arguments. The main contribution of the paper consists in studying the application of such algorithm in realistic scenarios, represented by non fully-meshed networks. Simulations show that, in this case, convergence may be strongly dependent on the adopted share factor, whose value should be optimized as a function of the connectivity level of the network. Optimum shares are derived for some common topologies, like the ring and the random geometric graph. The effect of possible link failures is also investigated.**

***Index Terms*—Sensor networks, averaging algorithms, push-sum algorithm, share factors.**

I. INTRODUCTION

There is a wealth of past and recent literature devoted to the analysis and performance evaluation of protocols for sensor networks. Among them, a special place is assigned to simple uncentralized protocols for distributed information processing over networks. The class of gossip algorithms, in particular, permits to compute sums, averages, random samples, quantiles, and other aggregate functions, starting from an ensemble of initial sensed values [1]-[4]. Focusing on the averaging problem, that is the object of this paper, N sensors form the nodes of a graph whose edges correspond to reliable communication links. Each node is initially given a scalar value (which could correspond to some sensor measurement, like temperature) and, through a suitable message-passing algorithm, all nodes can compute the average of all N initial values in a possibly short time.

The most attractive features of gossip-based networks are their simplicity and scalability. Each sensor node, in fact, sends a fixed number of messages (in many cases only one message), and this number is independent of the network size. On the other hand, provided that effective communication among nodes is achieved through suitable protocols, a fundamental aspect of gossip algorithms concerns convergence to the average. This topic has been discussed extensively in previous literature. Most of published papers, however, are focused on theoretical issues and system modeling, while the literature is much less rich in regard to performance evaluation and parameters optimization.

In [5], Kempe et al. proposed a gossip algorithm named push-sum. This protocol, that will be reminded in Section II, can be managed either in a synchronous or in an asynchronous way. In the latter case, that we prefer because of its simplicity, following its own clock, a node forwards a share of its values, properly defined, to one of its neighbors randomly selected, while keeping the remaining part. The neighbor of a node is linked to it by an edge in the graph. Intuitively, the amount of the share reflects on the convergence speed of the averaging algorithm, and an optimization problem arises. Actually, in the case of fully-meshed networks, where each node can reach any other node in a single hop, the choice to assume a share factor equal to 1/2 seems qualitatively acceptable on the basis of symmetry considerations. However, if the network is non fully-meshed, that is the more frequent case in practice, simulations show that the optimal shares can be no longer equal to 1/2, and the convergence speed may be significantly faster by assuming a different value.

At the best of our knowledge, in spite of its practical importance, the problem of choosing the optimal share factor has not been faced extensively in previous literature; in [5], for example, the authors limited to say that the choice of shares may be deterministic or randomized, and may or may not depend on the time. In other papers, instead, the optimization effort looked at other targets, like to optimize the neighbor selection probabilities [6] or the probability of gossip [7]. These targets have similarities with the share factors optimization, but are not of interest in the present context. In our analysis, in particular, we will assume undirected unweighted graphs, where each node, if selected for transmission, contacts one of its neighbors with the same probability. Moreover, in absence of link failures, the gossip probability is unitary, which means that, if chosen, a node certainly transmits.

More involved protocols could be most suited for other contexts, as communication networks [8] or decentralized reputation systems [9].

As mentioned above, the first evidence of the convergence time dependence on the share factors is given by simulations.

Manuscript received June 12, 2008, and revised February 21, 2009.

E. Zanaj is with the Departamenti i Elektronikes dhe Telekomunikacionit, Fakulteti i Teknologjise se Informacionit, Universiteti Politeknik i Tiranes, Tirana, Albania (e-mail: ezanaj@fti.upt.al).

M. Baldi and F. Chiaraluce are with the Dipartimento di Ingegneria Biomedica, Elettronica e Telecomunicazioni, Università Politecnica delle Marche, Ancona, Italy (e-mail: {m.baldi, f.chiaraluce}@univpm.it).

So, we will discuss simulation results referred to some typical network topologies, like the ring and the random geometric graph. Plots of optimum share factors will be given, as a function of the nodal degree or maximum coverage radius.

A more ambitious task is to derive the optimum shares through a theoretical analysis. Although to face the problem in a very general way is very difficult, a useful, though simple, analytical tool is given by the potential function. In [5] the potential function was studied by assuming the synchronous version of the push-sum algorithm and used to demonstrate the convergence of the algorithm in the case of fully-meshed networks.

Another contribution of our work will be to extend such approach to the asynchronous case and, mostly important, to non fully-meshed networks. In the latter case, we will show that even establishing convergence of the algorithm is not possible "a priori". On the other hand, at least in some cases, a rough approximation of the average potential function behavior permits to make previsions on the expected optimum share factors.

Finally, a sensor network can be obviously subjected to link failures. When this occurs, the network connectivity changes and the optimum share factor should be recalculated. If this is not possible, because of the absence of information over the failed links and the limited processing capability of the nodes, a negative effect is expected on the convergence speed. Some numerical examples in this sense will be reported, yielding the conclusion that link failures do not influence significantly the choice of optimal share factors.

The remainder of the paper is organized as follows. In Section II, after having recalled some definitions, we will focus on simulated results, thus proving the role of the share factors for some relevant topologies. Section III will be devoted to the theoretical analysis, that is mainly based on an original adaptation of the well-known concept of potential function. In Section IV we will provide some examples of the effect of link failures. Finally, Section V will conclude the paper.

II. SHARE FACTORS IN THE PUSH-SUM ALGORITHM

Let us denote by $\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_N(k)]^T$, where T stays for "transposed", the vector of the estimates after k clock ticks. In particular, $\mathbf{x}(0)$ contains the initial sensed values. The target of the averaging algorithm is to disseminate, in the shortest possible time, throughout the nodes of the network, a reliable measure of the average value $x_{\text{ave}} = \sum_{i=1}^{N} x_i(0) \Big/ N$, to be used for control or rescue purposes, for example.

A. Push-sum algorithm

The push-sum protocol proceeds as follows. At the i-th node, with $i = 1, 2, \ldots, N$, two quantities are stored and updated following interaction with the other nodes: they are named $s_i(k)$ and $w_i(k)$, respectively.

These quantities satisfy the following mass conservation properties, for any k:

$$\sum_{i=1}^{N} s_i(k) = \sum_{i=1}^{N} x_i(0) = N x_{\text{ave}} \quad , \quad \sum_{i=1}^{N} w_i(k) = N. \tag{1}$$

It should be noted that x_{ave} does not change with k; it represents the initial condition, and the goal of the sensor network is to make this information available at all nodes, in the shortest possible time. This is also related with the concept of "convergence speed", that measures the latency by which this goal is achieved with satisfactory approximation. In practice, the goal to know exactly x_{ave} can be reached, in general, only after a very long (ideally infinite) time. However, to know x_{ave} with an acceptable error is possible after a limited time. In the following of this subsection, a definition of error will be given (see Eq. (3)); the convergence speed measures the time required for having an acceptably small error: the shorter this time, the faster the convergence speed.

When the protocol starts, that is, once having acquired the sensed values, we have $s_i(0) = x_i(0)$ and $w_i(0) = 1$, $\forall i$. Afterwards, if the clock of the i-th node ticks at the k-th time instant (let us remind that transmission is asynchronous in the considered system) it selects randomly one of its neighbors, say j, and sends to it a fraction $(1 - \alpha)$ of its parameters, while it retains the remaining fraction α. As a consequence, the parameters at nodes i and j are modified as follows:

$$\begin{aligned} &s_i(k) = \alpha s_i(k-1), \quad w_i(k) = \alpha w_i(k-1), \\ &s_j(k) = s_j(k-1) + (1-\alpha) s_i(k-1), \\ &w_j(k) = w_j(k-1) + (1-\alpha) w_i(k-1). \end{aligned} \tag{2}$$

The parameters at all the other nodes remain unchanged. This way, conditions (1) are certainly satisfied. A new estimate at the interacted nodes is then obtained as $x_m(k) = s_m(k)/w_m(k)$, with $m = i, j$. In this algorithm, the number of clock ticks is equal to the number of transmissions. It should be noted, however, that this is not true for other gossip algorithms [6].

In [5], where, besides point-to-point communications, also broadcast transmissions were considered, a more general mechanism was applied, where the share factor can be different for any node and even variable with time. However, this model seems too involved for practical applications. So, we prefer to consider a single and constant α, whose value should be optimized in order to achieve the fastest convergence speed.

The efficiency of the averaging algorithm can be measured by the error:

$$e(k) = \|\mathbf{x}(k) - x_{\text{ave}}\mathbf{1}\| / \|\mathbf{x}(0)\|, \tag{3}$$

where $\|\mathbf{x}\|$ denotes the l_2 norm of $\mathbf{x}$ and $\mathbf{1}$ is the vector of all ones. The algorithm converges if a value of k exists, say k^*, such that $e(k > k^*)$ becomes arbitrarily small [10]. The convergence of the push-sum protocol was already verified in

[5]. We will rediscuss the subject in Section III, after having reminded the definition of potential function.

B. The considered topologies

For explicative purposes, in this paper we consider two kinds of topologies: one regular, the ring, and one irregular, the random geometric graph. In the ring topology the nodes are equally spaced along a unit circle and each node is linked to d other nodes. d is called nodal degree and its value determines the connectivity level of the network. In the random geometric graph, instead, the nodes are randomly distributed in a unit square, according with a 2D homogeneous Poisson point process. The considered topologies are very different, and it is known that the ring topology may offer rather poor performance. On the other hand, it has the advantage of a constant nodal degree, that can be useful for the analysis purposes (see Section III). On the contrary, noting by r the maximum coverage radius, that is the maximum distance at which a node can transmit reliably, for the random geometric graph, even assuming that all nodes have the same r, the nodal degree is usually not unique. Moreover, for each value of r, the connectivity level can change from graph to graph. So, an average nodal degree, $\langle d \rangle$, is required. Two examples of these topologies are shown in Fig. 1 for $N = 16$, $d = 4$, and $N = 50$, $r = 0.4$, respectively. An example of $\langle d \rangle$, obtained by averaging over 100 random geometric graphs for each value of r, is shown in Fig. 2.

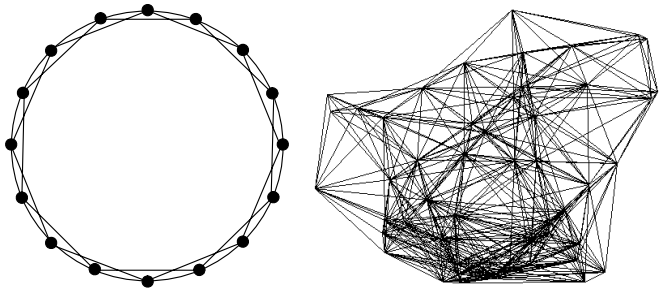


Figure 1. Examples of the considered network topologies: (a) ring with $N = 16$ and $d = 4$; (b) random geometric graph with $N = 50$ and $r = 0.4$.

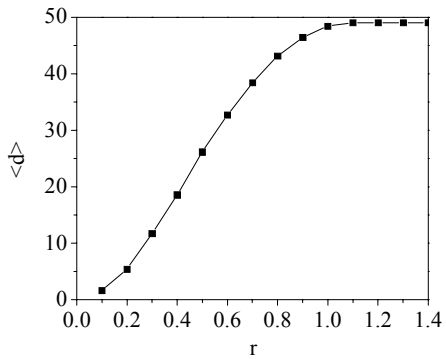


Figure 2. Average value of d as a function of r, computed over 100 random geometric graphs with $N = 50$.

C. Share factors optimization

In this subsection, the sensitivity of the convergence speed on the share factor and the consequent optimization issue will be faced through simulations. In our view, a simulation consists in assigning the initial sensed value at each node, according with a Gaussian distribution (other rules can be obviously adopted), in selecting randomly the transmitter node (TX), according with a uniform distribution and, then, in selecting randomly the destination node (RX) among the ensemble the TX can reach. After that, interaction between the two selected nodes is simulated by updating their stored values. When the link between each couple of nodes is present, we assume that their communication is reliable, and neglect lower protocols issues, as collisions and retransmissions. This is because our interest is in a simulation tool that allows fair comparison among different algorithms. In implementing such procedure, all simulations have been performed through "ad hoc" software programs written in Matlab and C++ language. As mentioned, this choice has been driven by the fact that our analysis is limited to the performance of averaging algorithms, and does not consider some issues typical of lower protocols. For this reason, we are able to adopt very simple software programs, without the need to resort to more complex network simulators.

The value of α in (2) has an impact on the convergence speed. Actually, in the case of fully-meshed networks, where each node is able to reach any other node in a single hop (this case is rather unrealistic but can be usefully adopted as a benchmark [11]) $\alpha \approx 0.5$ can be considered a good choice. This is confirmed, through an example, in Fig. 3, that refers to the same node distribution shown in Fig. 1(b) but where each node is linked to any other node (fully-meshed network). In practice, starting from $\alpha = 0.1$, the convergence speed increases for increasing α, up to 0.5, and then decreases again for $\alpha > 0.5$. We conclude that, for this example, $\alpha = 0.5$ provides the best result.

On the contrary, the optimum value is usually different for reduced network connectivity. An example is shown in Fig. 4, for a ring topology with $N = 49$ and $d = 12$, and in Fig. 5, for a random geometric graph with $N = 50$ and $r = 0.4$ (the same network as in Fig. 1(b)). Other values of d and r could be considered: for example, $d = 4$, as in Fig. 1(a). In general, smaller values of these parameters cause reduced connectivity, so that convergence is slower. Qualitative considerations, however, remain basically the same.

The curves in Figs. 3-5 have been obtained by using the software simulator outlined above. In these cases, for each value of α, a single simulation has been considered. The goal of these figures, in fact, is to introduce the problem, while generalization able to take into account the statistical nature of the quantities involved will be introduced afterwards. The experiment consists in choosing randomly the sequence of TX/RX nodes (one for each clock tick), and in updating the error (3) and the local values (2), as a consequence of their interaction.

From Fig. 4, we see that the optimum share factor, for the considered ring topology with non fully-meshed network, is $\alpha = 0.2$, and a significant amount of time can be saved by adopting this optimum value in place of $\alpha = 0.5$: as an example, $e(k) = 10^{-5}$ requires 8352 transmissions in the case of $\alpha = 0.5$ and "only" 5991 transmissions in the case of $\alpha = 0.2$

(that is, a reduction of about 28% in the convergence time). The gap is even more evident with $\alpha > 0.5$. Similarly, in Fig. 5, the optimum α value is 0.24; $e(k) = 10^{-6}$, for example, requires 5866 transmissions with $\alpha = 0.5$ and "only" 4175 transmissions with $\alpha = 0.24$ (that is, a reduction of about 29% in the convergence time).

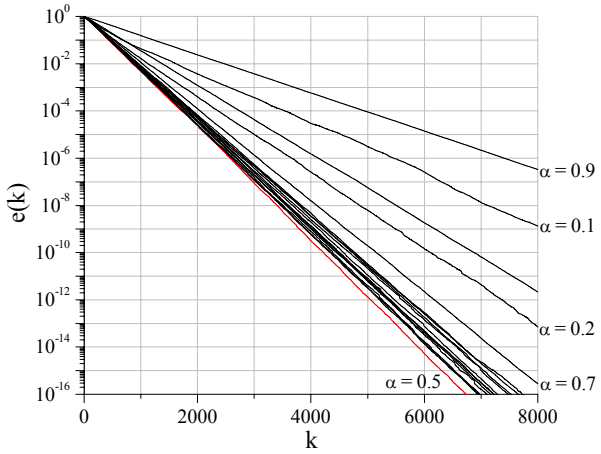


Figure 3. Simulated error for a fully-meshed random geometric graph with $N = 50$, assuming different values of α.

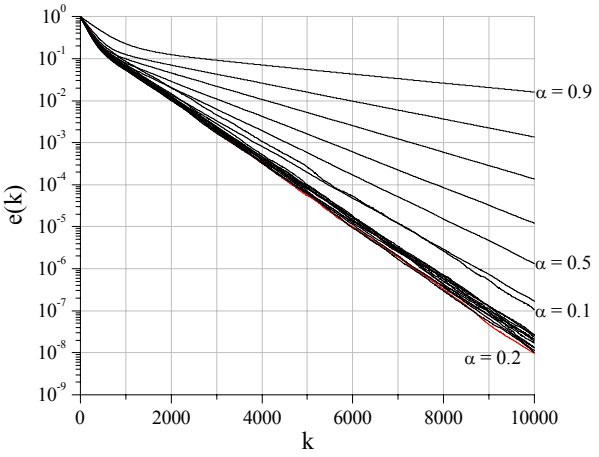


Figure 4. Simulated error for a ring topology with $N = 49$ and $d = 12$, assuming different values of α.

Obviously, to consider a specific simulation gives an idea of the framework, but it is substantially meaningless because of the random nature of the quantities involved. More significantly, from a statistical viewpoint, we can repeat the simulations R times (with R sufficiently high), finding, at any attempt, an optimum value α_{opt}^{m}, with $m = 1, \ldots, R$, and then computing an average optimum share factor as:

$$\left\langle \alpha_{\text{opt}} \right\rangle = \frac{1}{R} \sum_{m=1}^{R} \alpha_{\text{opt}}^{m} . \tag{4}$$

The procedure can be applied for any value of the nodal degree d (or the maximum coverage radius r). Some results are shown in Table I. They have been obtained by averaging the results of 200 simulations for any value of d and r. In the case of the random geometric graph, only values of $r \geq 0.3$ have been considered since, for $r = 0.2$, as an example, the network is no longer connected.

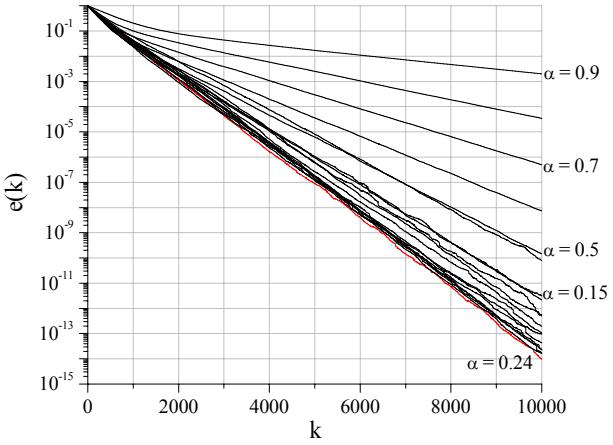


Figure 5. Simulated error for a random geometric graph with $N = 50$ and $r = 0.4$, assuming different values of α.

TABLE I

SIMULATED $\left\langle \alpha_{\text{opt}} \right\rangle$ FOR THE RING NETWORK WITH $N = 49$ AND THE RANDOM GEOMETRIC GRAPH WTH $N = 50$ BY CONSIDERING DIFFERENT CONNECTIVITY LEVELS.

Ring		**Random geometric graph**	
d	$\left\langle \alpha_{\text{opt}} \right\rangle$	r	$\left\langle \alpha_{\text{opt}} \right\rangle$
2	0.29313	0.3	0.2193
6	0.23139	0.4	0.26856
8	0.2202	0.5	0.35428
12	0.22408	0.6	0.43463
16	0.26697	0.7	0.46254
20	0.31667	0.8	0.47164
24	0.37214	0.9	0.46965
28	0.41522	1	0.47672
34	0.45667	1.1	0.47731
48	0.47945	1.2	0.471
		1.3	0.47244
		1.4	0.47289
		1.5	0.48184

On the other hand, (4) provides an average value, while the optimum share for a specific simulation may be more or less different, with a probability that depends on the dispersion around the mean.

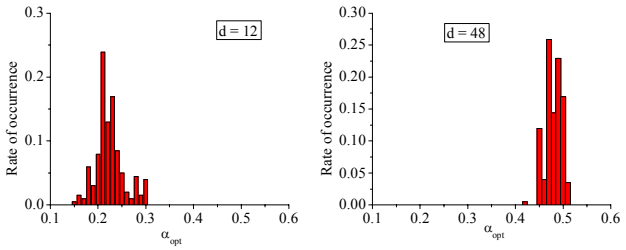
As an example, from Table I we derive that the average optimum share for the case of $d = 12$ is $\left\langle \alpha_{\text{opt}} \right\rangle\Big|_{d=12} \approx 0.224$ while for the simulation in Fig. 4 the optimum was $\alpha = 0.2$.

On the other hand, assuming $\left\langle \alpha_{\text{opt}} \right\rangle\Big|_{d=12}$ in place of the actual optimum value implies that $e(k) = 10^{-6}$ would have been reached after 7516 transmissions with a limited penalty of less than 2.2% with respect to the optimum value, that is 7356 (see Fig. 4).

In general, the reliability of the average value (4) depends on the dispersion of α_{opt} around the mean. This is shown in Fig. 6, for the ring with $N = 49$, by considering two values of d, and in Fig. 7, for the random geometric graph with $N = 50$ and two values of r.

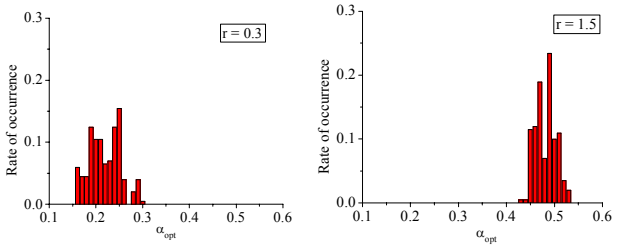
More specifically, the statistical analysis permits us to obtain the normalized standard deviation (nsd), defined as

$nsd = \sqrt{\left\langle \left(\alpha_{\text{opt}} - \langle \alpha_{\text{opt}} \rangle \right)^2 \right\rangle} / \langle \alpha_{\text{opt}} \rangle$, that is reported in Table II. We see that, except for the case of very low connectivity, the *nsd* is rather small; so, we can conclude that most networks require an α_{opt} that is not significantly different from the average value.



(a) (b)

Figure 6. Dispersion around the mean of the simulated α_{opt} for the ring network with N = 49 nodes: (a) d = 12; (b) d = 48 (fully-meshed).



(a) (b)

Figure 7. Dispersion around the mean of the simulated α_{opt} for the random geometric graph with N = 50 nodes: (a) r = 0.4; (b) r = 1.5 (fully-meshed).

III. THEORETICAL ANALYSIS

Besides the simulation work described in the previous section, a theoretical approach has been developed, based on the well known concept of "potential function".

Let us consider a network of N nodes described by a connected graph $G(V, E)$, where V is the vertex set containing the nodes and E is the edge set. Given a distribution of values $\mathbf{v} = [v_1, v_2, ..., v_N]^T$, where v_i is the value of node i, the potential Φ of the graph can be defined, in general, as follows:

$$\Phi = \left\| \mathbf{v} - v_{\text{ave}} \mathbf{1} \right\|^2 = \sum_{i \in V} \left(v_i - v_{\text{ave}} \right)^2 . \tag{5}$$

where v_{ave} is the average value over the whole network. Evidently, Φ is a measurement of the variance of the value distribution. Note that $\Phi = 0$ if and only if $\mathbf{v} = v_{\text{ave}} \mathbf{1}$. In the following we will denote by $\Phi(k)$ the potential function after the k-th clock tick ($\Phi(0)$ is the initial value).

TABLE II
SIMULATED *NSD* FOR THE RING NETWORK WITH $N = 49$ AND THE RANDOM GEOMETRIC GRAPH WTH $N = 50$ BY CONSIDERING DIFFERENT CONNECTIVITY LEVELS.

Ring		**Random geometric graph**	
d	*nsd*	*r*	*nsd*
2	0.25551	0.3	0.15276
6	0.19318	0.4	0.07522
8	0.18211	0.5	0.05165
12	0.13477	0.6	0.04257
16	0.06892	0.7	0.05902
20	0.07137	0.8	0.03986
24	0.05294	0.9	0.03449
28	0.04287	1	0.04216
34	0.04949	1.1	0.04735
48	0.03546	1.2	0.04416
		1.3	0.04572
		1.4	0.05329
		1.5	0.04441

Definition (5) could be applied, in principle, directly to the vector of the estimates $\mathbf{x}(k)$, this way obtaining, apart from the normalization by $\mathbf{x}(0)$, the square of the error $e(k)$ as defined in (3). For the push-sum algorithm, however, a slightly more complex expression is more favorable and it is described next.

In order to define the potential function for the push-sum algorithm, let us consider a vector $\mathbf{v}_i(k)$ (that does not contain any measured quantity, but is only introduced for analysis purposes), whose components, $v_{ij}(k)$, are such that:

$$s_i(k) = \sum_{j=1}^{N} v_{ij}(k) x_j(0) . \tag{6}$$

The following condition is satisfied:

$$w_i(k) = \sum_{j=1}^{N} v_{ij}(k) . \tag{7}$$

It is clear that, if $\mathbf{v}_i(k)$ is nearly proportional to the all-one vector, then $x_i(k) = s_i(k)/w_i(k)$ is close to the true average.

The potential function for the push-sum algorithm is defined as follows [5]:

$$\Phi(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} \left[v_{ij}(k) - \frac{w_i(k)}{N} \right]^2 . \tag{8}$$

So, in the limit case of all nodes perfectly aware of the true average, the potential function is null. Based on this evidence, evaluation of the mean potential function, for any k, should permit to estimate the convergence speed of the algorithm.

More precisely, assuming that, at instant k, node l is selected as the transmitter and node m as the receiver, the following difference between the potential functions at time instant $k - 1$ and k can be easily derived:

$$\delta\Phi = \Phi(k-1) - \Phi(k) = 2\alpha(1-\alpha) \sum_{j=1}^{N} \left[v_{lj}(k-1) - \frac{w_l(k-1)}{N} \right]^2$$

$$-2(1-\alpha) \sum_{j=1}^{N} \left[v_{lj}(k-1) - \frac{w_l(k-1)}{N} \right] \cdot \left[v_{mj}(k-1) - \frac{w_m(k-1)}{N} \right] . \tag{9}$$

For simplifying the notation, in the following we will omit to indicate that the quantities at the right hand side are computed at the time instant $k-1$.

We wish to compute the average of (9) over the possible choices of the transmitting and receiving nodes. Let us consider the case of a ring network, where all nodes are characterized by the same degree d. Taking into account that both the choices of transmitter and receiver are made following a uniform law, the former on the ensemble of the N nodes in the network, and the latter, for each transmitter, on the ensemble of the d nodes it is linked to, also having in mind definition (8), we find:

$$\begin{aligned}\left\langle \delta\Phi \right\rangle &= \frac{2\alpha(1-\alpha)}{N}\Phi + \frac{2(1-\alpha)}{dN}\sum_{j=1}^{N}\sum_{l=1}^{N}\left(v_{lj}-\frac{w_l}{N}\right)^2 \\ &-\frac{2(1-\alpha)}{dN}\sum_{j=1}^{N}\sum_{l=1}^{N}\sum_{m\in C_l}\left(v_{lj}-\frac{w_l}{N}\right)\left(v_{mj}-\frac{w_m}{N}\right)\end{aligned} \tag{10}$$

where $\Phi = \Phi(k-1)$, and C_l is the subgroup of nodes that includes node l and the nodes it is linked to. It is easy to see that, in the case of a fully-meshed network, when C_l is the entire network and $d = N-1$, the last term at the right side of (10) is null because of the mass conservation properties (1). In the case of a non fully-meshed network, that is $d < N-1$, it represents instead an additional contribution that, in general, could be not simple to determine. However, some elaboration is possible by reminding the definition of Laplacian matrix of a graph [12].

The Laplacian matrix $\mathbf{Q}(G)$ of the graph $G(V,E)$ is an $N\times N$ matrix whose elements are defined as follows:

$$Q_{ij} = \begin{cases} d_i & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } (i,j)\in E \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where d_i represents the nodal degree of node i. In the case of a ring topology, we have $d_i = d$ for any i. The Laplacian matrix can be also written as $\mathbf{Q} = \mathbf{\Delta} - \mathbf{A}$, where $\mathbf{\Delta}$ is the diagonal matrix with elements $\Delta_{ii} = d_i$, and $\mathbf{A}$ is the adjacency matrix of the considered graph. The eigenvalues of $\mathbf{Q}$ are called the Laplacian eigenvalues. They are all real and non negative, and satisfy the condition: $0 = \lambda_1 \le \lambda_2 \le \ldots \le \lambda_N$. The second smallest eigenvalue, λ_2, is also known as the algebraic connectivity, and it is particularly important. λ_2 is equal to zero only if G is disconnected. Other properties of matrix $\mathbf{Q}$ and its eigenvalues can be found in the literature (see [13], for example).

Let $y_{ij} = v_{ij} - w_i / N$, $i = 1, \ldots, N$, be the components of a vector $\mathbf{y}_j$. Through simple algebra, using (8) and (11), Eq. (10) can be rewritten as follows:

$$\left\langle \delta\Phi \right\rangle = -\frac{2(1-\alpha)^2}{N}\Phi + \frac{2(1-\alpha)}{dN}\sum_{j=1}^{N}\mathbf{y}_j{}^T\mathbf{Q}\mathbf{y}_j \,. \tag{12}$$

Let us define another vector $\mathbf{z} = (\mathbf{y}_1{}^T, \mathbf{y}_2{}^T, \ldots, \mathbf{y}_N{}^T)^T$ having N^2 components; it is evident that $\mathbf{z}^T\mathbf{z} = \Phi$. Moreover, let us consider a block matrix $\mathbf{L}$, with size $N^2\times N^2$, having N repetitions of $\mathbf{Q}$ along the main diagonal and all the other blocks equal to the null matrix. Also $\mathbf{L}$ can be interpreted as a Laplacian matrix, whose eigenvalues coincide with those of $\mathbf{Q}$, but each appears with multiplicity N.

Using these further definitions, Eq. (12) can be rewritten as**:**

$$\begin{aligned}\left\langle \delta\Phi \right\rangle &= \left[-\frac{2(1-\alpha)^2}{N} + \frac{2(1-\alpha)}{dN}\frac{\mathbf{z}^T\mathbf{L}\mathbf{z}}{\mathbf{z}^T\mathbf{z}}\right]\Phi \\ &= \left[-\frac{2(1-\alpha)^2}{N} + \frac{2(1-\alpha)}{dN}\mathrm{RQ}\right]\Phi\end{aligned} \tag{13}$$

having denoted by $\mathrm{RQ} = \mathbf{z}^T\mathbf{L}\mathbf{z} / \mathbf{z}^T\mathbf{z}$ the so-called Rayleigh quotient. By applying the Courant-Fischer Minimax Theorem [14] it is possible to say that:

$$\lambda_2 \le \mathrm{RQ} \le \lambda_N \tag{14}$$

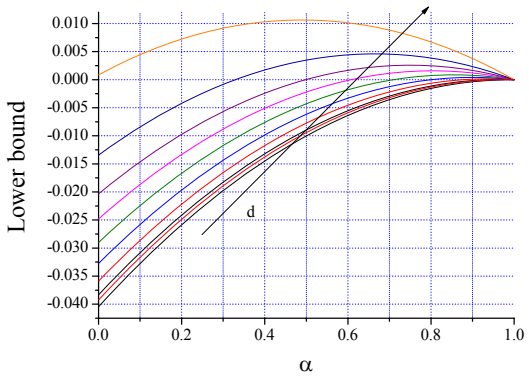
and, consequently:

$$\frac{2(1-\alpha)}{N}\left[-(1-\alpha)+\frac{\lambda_2}{d}\right] \le \left\langle\frac{\delta\Phi}{\Phi}\right\rangle \le \frac{2(1-\alpha)}{N}\left[-(1-\alpha)+\frac{\lambda_N}{d}\right]. \tag{15}$$

In practice, the above analysis permits us to find a lower bound (*lb*) and an upper bound (*ub*) for the mean variation of the potential function conditioned on a starting value Φ. If the network is fully-meshed, we have *lb* = *ub*, as all eigenvalues, except $\lambda_1 = 0$, are coincident. In this case, the value of $\left\langle \delta\Phi / \Phi \right\rangle$ is maximized assuming $\alpha = \alpha_{\mathrm{opt}} = (N-2)/(2(N-1)) \approx 0.5$ [11]. To have a maximum $\left\langle \delta\Phi / \Phi \right\rangle$ seems a good criterion to make convergence as much faster as possible, and, in fact, simulations confirm that $\alpha \approx 0.5$ is the best choice for fully-meshed network. Additionally, for any $\alpha < 1$, $\left\langle \delta\Phi / \Phi \right\rangle > 0$, and this result can be used to demonstrate convergence of the algorithm.

For non fully-meshed networks, instead, *lb* and *ub* show behaviors like those in Fig. 8 for the case of the ring network with $N = 49$. The values of the nodal degree considered are: d = 2, 6, 8, 12, 16, 20, 24, 28, 34, 48. From the figure we see that *lb* becomes lower and lower when reducing d, while *ub* has an opposite behavior. Fig. 8 (a), in particular, shows that convergence of the algorithm cannot be established just on the basis of the *lb* on $\left\langle \delta\Phi / \Phi \right\rangle$: for $d < N-1$, in fact, there is a large range of α's where $\left\langle \delta\Phi / \Phi \right\rangle < 0$. On the other hand, Fig. 8 (b) shows that reasoning on the *ub* would be too optimistic, since it provides $\left\langle \delta\Phi / \Phi \right\rangle > 0$ everywhere, and even greater and greater for decreasing d.

Based on these preliminary results, it is evident that an in depth analysis should weight the contributions of λ_2 and λ_N and, even more important, should take into account the role of the other eigenvalues, λ_i, with $i \in [3, N-1]$.



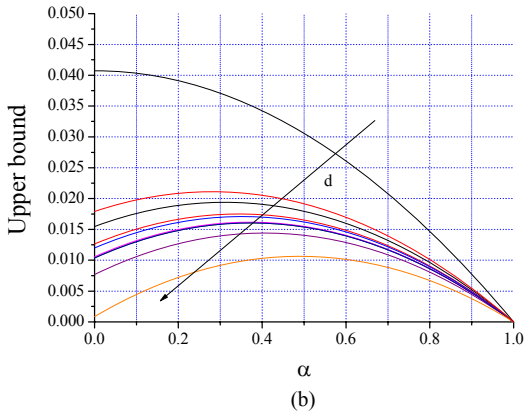


Figure 8. (a) Lower and (b) upper bounds for $\langle \delta\Phi/\Phi \rangle$ in a ring network with $N = 49$ nodes.

With the goal to develop further the theoretical analysis, a first, though coarse, approximation can consist in considering the average of the eigenvalues. An example of the estimate of α_{opt} computed this way is shown in Fig. 9 for the random geometric graph with $N = 50$, where the nodes are located as in Fig. 10; in this case, in Eq. (10) (and in the following ones) $1/d$ has been replaced by $\langle 1/d_i \rangle$. The figure shows how theory can predict the need to reduce the optimum share factor for decreasing $\langle d \rangle$, but the approximations used do not permit a satisfactory agreement between the numerical values. The difference between the theoretical and the simulated results, particularly for small $\langle d \rangle$, can be ascribed to some limits in the analytical approach. In (10), in fact, the subgroup C_l is unconstrained, and the missing links, when the network is non fully-meshed, can be found everywhere in the network. On the contrary, when reduced connectivity is due to the limited maximum coverage radius, the subgroup C_l is constrained and this introduces an effect that is not easy to capture by means of a formula. To confirm such statement, in Section IV, the comparison will be repeated by assuming that missing links due to failures are randomly distributed; and, in that case, we will show that the agreement between the theoretical curve and the simulated results can be better.

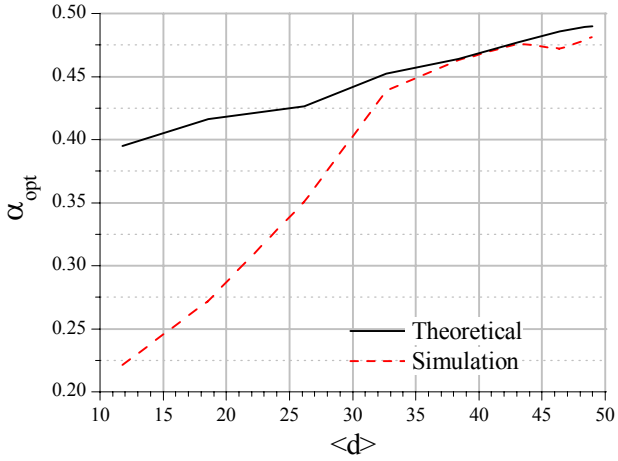


Figure 9. α_{opt} as a function of the average nodal degree for the random geometric graph with $N = 50$ nodes: simulated average vs. analytical estimate.

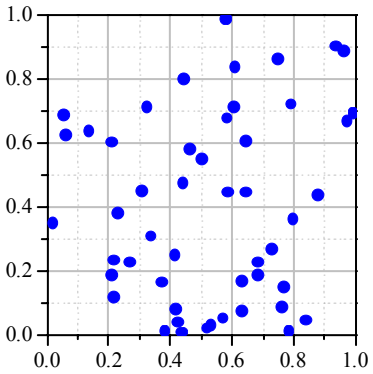


Figure 10. Example of random geometric graph for a network with $N = 50$.

Similar considerations hold for the ring topology. Starting from (13) and maximizing $\langle \delta\Phi/\Phi \rangle$ we find (remind that d is constant for the ring):

$$\alpha_{opt} = 1 - \frac{RQ}{2d}. \tag{16}$$

In this case, it can be numerically proved that the average of the eigenvalues is almost equal to d. So, the approximate optimum share factor is $\alpha_{opt} \approx 0.5$. Actually, as shown in Fig. 11, the actual optimum value, derived through simulations, can be significantly smaller, particularly for small d. It also shows a minimum, at $d = 8$ in the considered set of values. Even $d = 2$ is acceptable in the figure as, thanks to its structure, the ring network remains connected. Also in this case, the difference between the theory and the simulated results is due to the assumption of constrained links, that is even more evident for the ring topology. By removing this hypothesis, which means to distribute uniformly the missing links, that is a reasonable hypothesis in the case of link failures, the agreement between (16) and the simulated results becomes much better.

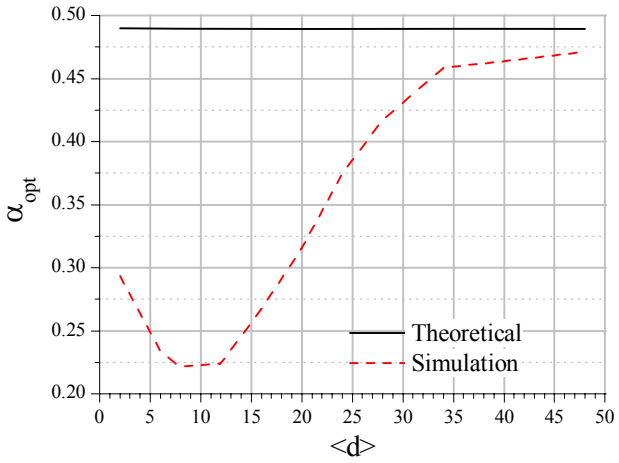


Figure 11. α_{opt} as a function of the nodal degree for the ring topology with N = 49 nodes: simulated average vs. analytical estimate.

IV. THE EFFECT OF FAILURES

When averaging algorithms, like the considered push-sum protocol, are adopted in actual wireless sensor networks, the presence of obstacles or other kinds of radio impairments, as noise and multipath, could prevent some links from being used, due their poor quality in terms of signal-to-noise ratio. In such cases, at least in principle, the network should update the choice of the share factors in order to recover optimal efficiency even in presence of link failures.

However, updating of the share factors could be quite unpractical to implement. Furthermore, the network could be unaware of the number of link failures occurred; so the value of its connectivity degree could not be available for calculating the optimal value of the share factor in the unexpected scenario.

Under these circumstances, it is reasonable to assume that the network will continue to adopt the optimal share factor found for the case without failures. On the other hand, it can be interesting to estimate the performance loss (if any) that results from the adoption of the non optimal value.

Let us consider a fully-meshed random geometric graph; for it, as discussed before, the optimal share, in absence of failures, is about 0.5. Let us suppose that, because of failures, at the beginning of the node interaction, a fraction x of its links, randomly distributed, is missing. While in absence of failures the network connectivity is $N-1$, in the new situation the average value of d becomes:

$$\langle d \rangle = (N-1)(1-x). \tag{17}$$

Fig. 12 shows the values of $\langle d \rangle$, calculated through (17), as a function of x, for different choices of N.

If we focus on the curve for N = 50 (i.e., the same considered in the numerical examples of the previous sections), we can notice that, for a fraction of link failures $\leq$ 0.5, the average nodal degree is still ≥ 25 and, from Fig. 9, we conclude that this implies $\alpha_{opt} \geq 0.35$. With similar arguments, we can derive that, when the fraction of link failures is ≤ 0.3, it is $\alpha_{opt} \geq 0.43$.

Based on this simple analysis, we can guess that, except when the percentage of link failures with respect to the total number of possible links is very high, the presence of obstacles or other impairments in a wireless sensor network based on the push-sum algorithm scarcely affects the choice of the optimal share factor. On the other hand, when link failures occur, the gossip probability becomes smaller than 1, because a node could waste its communication attempt when trying to exploit a failed link.

So, a negative effect on the convergence time is expected, in terms quite similar to those observed in [15] for another version of the gossip algorithm. Our analysis shows that such a penalty cannot be compensated through a different choice of the share factor.

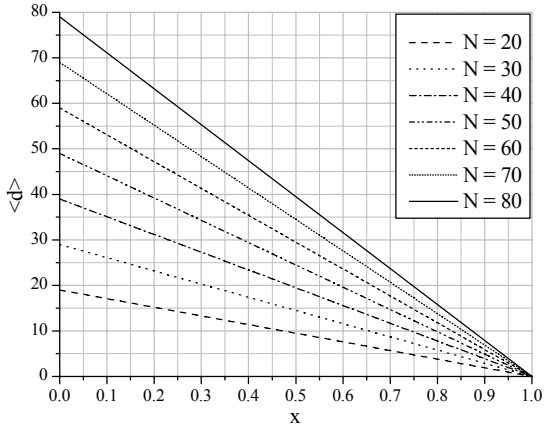


Figure 12. Average value of d for a fully-meshed network in presence of a link failure rate x, for different values of N.

The above considerations do not prevent us to estimate the optimum share factors in presence of link failures for some specific nodes configuration, like that depicted in Fig. 10. This is shown in Fig. 13, where the theoretical curve derived through the approach in Section III is also reported, for the sake of comparison. Fig. 13 is interesting as it demonstrates that the agreement between the analytical approach and the simulated results improves when the missing links are distributed uniformly, which is the hypothesis adopted for simulating link failures. This is because uniform distribution fits well the assumptions implicit in Eq. (10). On the contrary, when reduced connectivity is due to a limited maximum coverage radius, a larger difference is expected, that in fact has been confirmed in Section III.

For the sake of completeness, in Fig. 14 we have reported some values of optimum share factors found for a ring network with N = 49 in presence of link failures. Because of the random distribution of failures, the nodal degree is no longer constant for the ring too, and an average value must be determined. Similarly to Fig. 13, we see that, when missing links are randomly distributed, there is a better correspondence between theoretical expectations and simulations results.

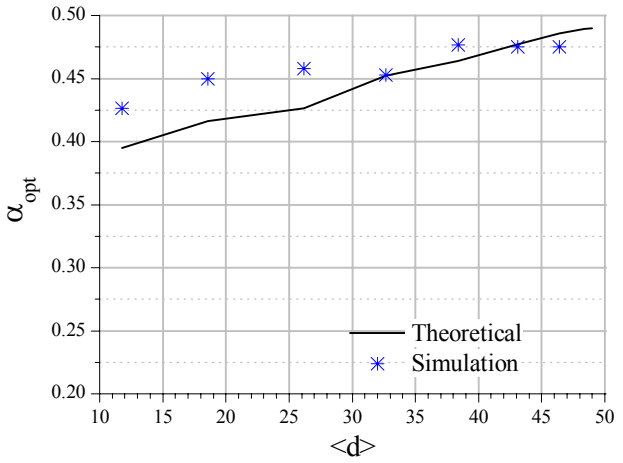


Figure 13. α_{opt} as a function of the average nodal degree for the random geometric graph with N = 50 nodes in the presence of link failures: simulated average vs. analytical estimate.

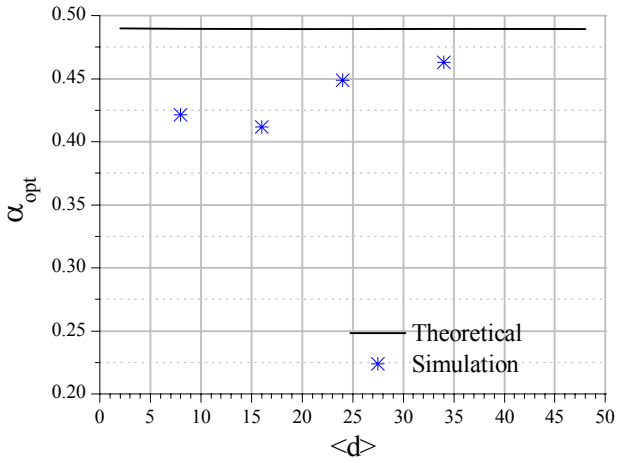


Figure 14. α_{opt} as a function of the average nodal degree for the ring network with N = 49 nodes in the presence of link failures: simulated average vs. analytical estimate.

V. CONCLUSIONS

The role of the share factor in the push-sum algorithm has been investigated through simulations and theoretical arguments. Simulations have confirmed that the adoption of an optimal share factor, depending on the network configuration and connectivity level, can improve significantly the convergence speed. Because of the statistical nature of the quantities involved, results have also been given in terms of mean values and standard deviations. We have also shown that the optimal share factor is not significantly influenced by the possible appearance of link failures, at least when the original network has a high connectivity level and the percentage of link failures is not too large.

We have also discussed a first approach, based on the concept of potential function, for a theoretical derivation of the optimal shares. Actually, the possibility to compute analytically the optimum values is very attractive but, at present, only qualitative and approximate information is achievable through simple mathematical arguments. In general, the analytical approach is suitable to describe situations where missing links are uniformly distributed, that is a realistic assumption in the case of random link failures. On the contrary, at least in the current version, it is not able to describe efficiently the constraint induced on links by a limited maximum coverage radius. Improving approximation can consist in overcoming such limitation, as well as in finding a better way for weighting the contributions of different eigenvalues of the Laplacian matrix for the considered graph. This further topic could be the subject of future work.

REFERENCES

[1] F. Bénézit, A. G. Dimakis, P. Thiran, and M. Vetterli, ‘‘Gossip along the way: order-optimal consensus through randomized path averaging,” in *Proc. Forty-Fifth Annual Allerton Conference*, Allerton House, IL, September 2007, pp. 1226-1233.

[2] J.-Y. Chen, G. Pondurangan, and D. Xu, “Robust computation of aggregates in wireless sensor networks: distributed randomized algorithms and analysis,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 9, pp. 987-1000, September 2006.

[3] E. Zanaj, M. Baldi, and F. Chiaraluce, “Efficiency of unicast and broadcast gossip algorithms for wireless sensor networks,” *Journal of Communications Software and Systems*, vol. 4, no. 2, pp. 105-112, June 2008.

[4] A. G. Dimakis, A. D. Sarwate, and M. J. Wainwright, “Geographic gossip: efficient averaging for sensor networks,” *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1205-1216, March 2008.

[5] D. Kempe, A. Dobra, and J. Gehrke, ‘‘Gossip-based computation of aggregate information,” in *Proc. IEEE Conference on Foundation of Computer Science*, Cambridge, MA, October 2003, pp. 482–491.

[6] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, ‘‘Randomized gossip algorithms,” *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, June 2006.

[7] Z. J. Haas, J. Y. Halpern, and L. Li, “Gossip-based ad hoc routing,” *Proc. IEEE Infocom 2002*, New York, NY, June 2002, pp. 1707-1716.

[8] M. Mehyar, D. Spanos, J. Pongsajapan, S. H. Low, and R. M. Murray, “Asynchronous distributed averaging on communication networks,” *IEEE/ACM Transactions on Networking*, vol. 15, no. 3, pp. 512–520, June 2007.

[9] Y. Bachrach, A. Parnes, A. D. Procaccia, and J. S. Rosenschein, “Gossip-based aggregation of trust in decentralized reputation systems,” *Autonomous Agents and Multi-Agent Systems*, vol. 19, no. 2, pp. 153–172, May 2009.

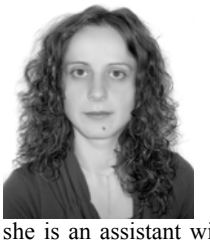
[10] A. Olshevsky and J. N. Tsitsiklis, “Convergence speed in distributed consensus and averaging,” Jan. 2009, [Online]. Available: http://arxiv.org/abs/math.OC/0612682.

[11] M. Baldi, F. Chiaraluce, and E. Zanaj, “Comparison of averaging algorithms for wireless sensor networks,” *Proc. ICTTA’08*, Damascus, Syria, April 2008, Paper TEL05_7.

[12] R. Merris, “A survey of graph Laplacians,” *Linear and Multilinear Algebra*, vol. 39, no. 1 & 2, pp.19-31, July 1995.

[13] B. Mohar, “The Laplacian spectrum of graphs,” included in *Graph Theory, Combinatorics, and Applications*, Vol. 2, Y. Alavi, G. Chartrand, O. R. Oellermann, A. J. Schwenk Eds., Wiley, pp. 871-898, 1991.

[14] J. Shawe-Taylor and N. Cristianini, "Kernel Methods for Pattern Analysis," Cambridge University Press, 2004.
[15] E. Zanaj, M. Baldi, F. Chiaraluce, "Fault tolerance in wireless sensor networks based on the gossip algorithm," in *Proc. IEEE Conference on Wireless Rural and Emergency Communications, WRECOM 2007*, Rome, Italy, 1-2 October 2007.



Elma Zanaj was born in Valona (Albania) in 1980. She received her Laurea Degree in Electronics Engineering, with a specialization in Telecommunications, from the Polytechnic University of Tirana in 2003, and the Doctoral degree in Electronics, Informatics and Telecommunications Engineering from the Polytechnic University of Marche, Ancona, Italy, in 2008. Since 2003 she cooperates with the Polytechnic University of Tirana, where she is an assistant within the group of Fundamentals in Electronics. Her main research interests are in the field of the analysis and modeling of wireless sensor networks.

Marco Baldi was born in Macerata, Italy, in 1979. He received the "Laurea" degree (summa cum laude) in Electronics Engineering in 2003, and the Doctoral degree in Electronics, Informatics and Telecommunications Engineering in 2006 from the Polytechnic University of Marche, Ancona, Italy. At present, he is a post-doctoral researcher and adjunct Professor at the same university. His main research activity is in channel coding, with particular interest in linear block codes for symmetric and asymmetric channels, low-density parity-check (LDPC) codes and their application in cryptography. He is member of IEEE.

Franco Chiaraluce was born in Ancona, Italy, in 1960. He received the "Laurea in Ingegneria Elettronica" (summa cum laude) from the University of Ancona in 1985. Since 1987 he joined the Department of Electronics and Automatics of the same university. At present, he is an Associate Professor at the Polytechnic University of Marche. His main research interests involve various aspects of communication systems theory and design, with special emphasis on error correcting codes, sensor networks, cryptography and multiple access techniques. He is co-author of more than 200 papers and two books. He is member of IEEE and IEICE.