Distributed Estimation of Macroscopic Channel Parameters in Dense Cooperative Wireless Networks

Monica Nicoli[1], Gloria Soatti[1], Stefano Savazzi[2]
[1]Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano
[2]National Research Council of Italy (CNR), IEIIT institute, Milano
E-mail: monica.nicoli@polimi.it, gloria.soatti@polimi.it, stefano.savazzi@ieiit.cnr.it

Abstract— **In peer-to-peer wireless networks, knowledge of the channel quality information of multiple links is fundamental to calibrate cooperative communication/processing techniques and design efficient resource sharing strategies. This paper is focused on distributed estimation algorithms that enable the network to self-learn key environment-dependent parameters that rule the channel quality of all links in the network. Considering an indoor scenario with fixed wireless terminals and moving objects/people in the environment, we parameterize the channel quality of each link in terms of path-loss and Rician K-factor, modelling these macro-parameters according to a site-specific stochastic model. Contribution of the paper is twofold: a measurement campaign carried out with IEEE 802.15.4 devices to validated the stochastic model; distributed algorithms to estimate the environment-dependent parameters of the model. Various schemes of weighted average consensus are proposed to enable the convergence to the equivalent global (centralized) estimate. Performance analysis is carried out in terms of convergence speed, error at convergence and communication overhead using both experimental and simulated data.**

I. Introduction

Emerging wireless cloud network paradigm [1] is characterized by a dense relay network - namely, the cloud network - where data originated by external sources are flooded via massively air-interacting nodes to the intended destinations, providing an extremely high quality of service under various application contexts [2] while keeping the air-interface to sources/destinations as simple as possible. Functionalities traditionally performed at upper layers and centrally coordinated (e.g., routing), in the cloud network are distributed over the nodes' PHY layer, by means of cooperative communications [3] and wireless network coding strategies [4], making the cloud able to self-organize and adapt in critical dynamic scenarios. Self learning of the network state (link quality, timing and location information, etc.) is a fundamental step to set-up an efficient intra-cloud connectivity.

This paper is focused on distributed estimation of key macroscopic propagation parameters that are extremely dependent on the radio environment and have a great influence on the overall network state. We consider a static peer-to-peer wireless network in a mixed line-of-sight/non-line-of-sight (LOS/NLOS) indoor environment. The channel between any two nodes is modelled according to Rician fading and it is parameterized in term of path-loss and K-factor. Multipath configuration changes rapidly over space, leading to fast variations of these two parameters from link to link. Yet, some environment-related features are slowly varying in space and can be represented by a common stochastic model [5]. Contribution of this paper is twofold: i) an indoor measurement campaign for the validation of the stochastic channel model; ii) the development of distributed estimation algorithms that enable the cloud network to self learn the model parameters in a totally decentralized way. Measurement campaigns presented in this paper and based on IEEE 802.15.4 devices [6] show that the K-factor and path-loss of each link are jointly Gaussian distributed - in the dB domain - with mean defined by space-invariant functions of the link distance and space-invariant covariance. Knowledge of these long-term propagation features is instrumental in cooperative and cognitive networks to optimize robust decentralized group-selection and resource-sharing strategies [5]. RSS-based cooperative [7]-[9] and non-cooperative [10] localization systems could also benefit from fast-converging distributed estimation of these channel features. For distributed estimation, in this work we propose to employ the consensus approach [11]-[14], based on successive refinements of local estimates at nodes with information exchange between neighbors. We apply the average-consensus principle to the least-squares (LS) estimation of the environment-dependent channel parameters. Compared to methods previously proposed for path-loss calibration [14], where consensus is performed on the correlation of the measured data, here new approaches are discussed where nodes exchange directly the local estimates (of both path-loss and K-factor parameters) rather than data, weighting the estimates in such a way to reach at convergence the same performance of an equivalent centralized estimation. In particular, a new weighted-average consensus method is proposed (Method 4) which allows to reach the optimal performance with reduced amount of information exchanged between nodes. The proposed method can be seen as an extension of the continuous-time approach [12] to the discrete-time estimation problem. Convergence of the proposed method to the global estimate is proved analytically. Network performance is analyzed on realistic indoor network scenarios with parameters drawn from the experimental measurement campaign.

II. Multi-link Channel Model

We consider a static peer-to-peer wireless network with N nodes and bidirectional communication links, as depicted in

978-1-4799-3083-8/14/$31.00 This work has been performed in the framework of the European project DIWINE (Dense Cooperative Wireless Cloud Network), funded by the European Union under its FP7 ICT Objective 1.1 - The Network of the Future.

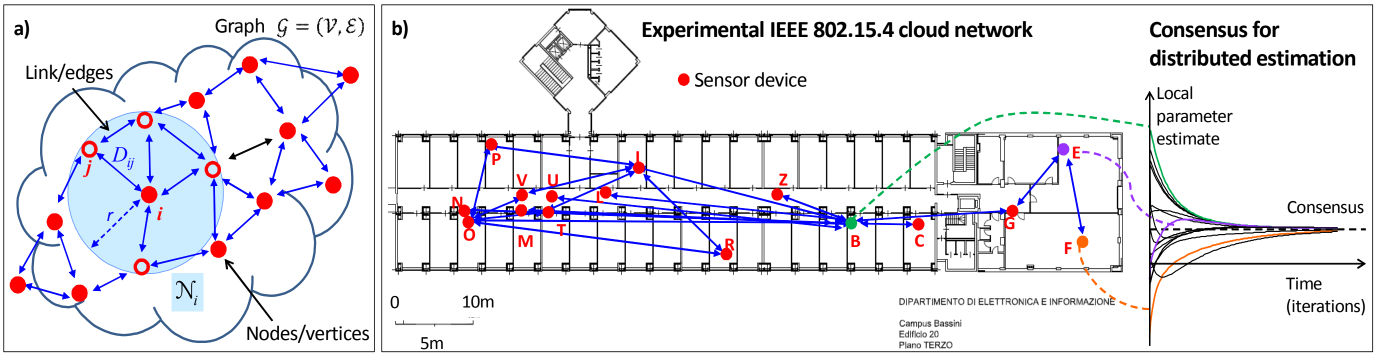

Fig. 1. (a) Cloud network graph modeling. (b) Sensor deployment for distributed consensus testing (DEIB, Politecnico di Milano) over a floor of approximately 20x110m meters rectangular area), with example of distributed consensus for a scalar parameter estimation.

Fig. 1-(a). The network is modelled as a connected undirected graph, $\mathcal{G}=(\mathcal{V},\mathcal{E})$, with vertices $\mathcal{V}=\{1,\ldots,N\}$ representing the nodes and edges $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ representing the communication links between the nodes. Let $\mathbf{A}=[a_{ij}]$ be the $N\times N$ symmetric adjacency matrix, with $a_{ij}=1$ if $(i,j)\in\mathcal{E}$ (i.e., if node j communicates with node i); $a_{ij}=0$ for any $(i,j)\notin\mathcal{E}$. The set of neighbors for node i is denoted as $\mathcal{N}_i=\{j\in\mathcal{V}:a_{ij}\neq 0\}$ and its cardinality - the node degree - as $d_i=|\mathcal{N}_i|$. The maximum degree is indicated as $\Delta=\max_i d_i$. The Laplacian matrix of the graph is $\mathbf{L}=\mathbf{D}-\mathbf{A}$ with $\mathbf{D}=\mathrm{diag}(d_1,\ldots,d_N)$ being the degree matrix of $\mathcal{G}$.

Since the network is static, the baseband flat-fading channel $h_{i,j}$ observed between any two nodes i and j, with $a_{ij}=1$, is characterized by two main components: a dominant static component $\mu_{h_{ij}}$ accounting for the effects of fixed scatterers/absorbers (the so-called static multipath); a dynamic component with standard deviation $\sigma_{h_{ij}}$ modelling the temporal fluctuations of fading due to moving scatterers/absorbers in the environment such as people/equipments. The overall channel can be reasonably characterized by the Rician fading model, as a complex Gaussian variable with mean $\mu_{h_{ij}}$ and variance $\sigma^2_{h_{ij}}$, $h_{i,j}\sim\mathcal{CN}(\mu_{h_{ij}},\sigma^2_{h_{ij}})$, with path-loss $L_{ij}=-[|\mu_{h_{ij}}|^2+\sigma^2_{h_{ij}}]_{\mathrm{dB}}$ and K-factor $K_{ij}=[|\mu_{h_{ij}}|^2/\sigma^2_{h_{ij}}]_{\mathrm{dB}}$.

Knowledge of the macro-parameters $\{L_{ij},K_{ij}\}$ is fundamental for calibration of transmission/reception algorithms and for the optimization of resource allocation during the cloud network set-up. The multipath configuration is known to change rapidly with the node locations, leading to fast variations of the parameters $\{L_{ij},K_{ij}\}$ over the links (i,j). Yet, some statistical features of the channel are slowly varying in space and can be considered as site-specific parameters that are common to all the peer-to-peer links; these environment-dependent parameters can be estimated cooperatively by the nodes of the network through a distributed processing.

The stochastic channel model herein adopted from [5] and validated on real data in Sect. IV, referred to as the bivariate Gaussian model, assumes that the channel parameters $\mathbf{y}_{ij}=[L_{ij},K_{ij}]^{\mathrm{T}}$ of each link are jointly Gaussian distributed, $\mathbf{y}_{ij}\sim\mathcal{N}(\boldsymbol{\mu}_{\mathrm{LK}}(D_{ij}),\mathbf{Q}_{\mathrm{LK}})$, with mean value $\boldsymbol{\mu}_{\mathrm{LK}}(D_{ij})$ depending on the link distance D_{ij} and covariance matrix $\mathbf{Q}_{\mathrm{LK}}$:

$$\boldsymbol{\mu}_{\mathrm{LK}}(D_{ij})=\begin{bmatrix}\mu_L(D_{ij})\\ \mu_K(D_{ij})\end{bmatrix};\ \mathbf{Q}_{\mathrm{LK}}=\begin{bmatrix}\sigma_L^2 & \rho\sigma_K\sigma_L\\ \rho\sigma_K\sigma_L & \sigma_K^2\end{bmatrix}. \tag{1}$$

The larger is the link distance, the higher is the path-loss (according to the path-loss law) and the lower is the K-factor on average. Values $\mu_L(D_{ij})$ and $\mu_K(D_{ij})$ are in fact linearly related to the distance D_{ij} according to the functions:

$$\begin{aligned}\mu_L(D_{ij}) &= L_0+10\gamma_{\mathrm{L}}\log_{10}(D_{ij}/D_0)=\mathbf{h}_{ij}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{L}} \\ \mu_K(D_{ij}) &= K_0-10\gamma_{\mathrm{K}}\log_{10}(D_{ij}/D_0)=\mathbf{h}_{ij}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{K}}\end{aligned} \tag{2}$$

with coefficients $\mathbf{h}_{ij}=[1,10\log_{10}(D_{ij}/D_0)]^{\mathrm{T}}$, $\boldsymbol{\theta}_{\mathrm{L}}=[L_0,\gamma_{\mathrm{L}}]^{\mathrm{T}}$ and $\boldsymbol{\theta}_{\mathrm{K}}=[K_0,\gamma_{\mathrm{K}}]^{\mathrm{T}}$. Parameters $\{L_0,K_0\}$ represent the path-loss and K-factor at reference distance D_0, while $\{\gamma_{\mathrm{L}},\gamma_{\mathrm{K}}\}$ are the decay/increase indexes for path-loss/K-factor. The standard deviations $\{\sigma_K,\sigma_L\}$ are expressed in dB. The cross-correlation coefficient is $\rho\leq 0$ as it has been observed experimentally that L_{ij} and K_{ij} are negatively correlated [5].

The macroscopic channel parameters that identify the network environment are the coefficients of the linear average functions $\boldsymbol{\theta}=[\boldsymbol{\theta}_{\mathrm{L}}^{\mathrm{T}},\boldsymbol{\theta}_{\mathrm{K}}^{\mathrm{T}}]^{\mathrm{T}}=[L_0,\gamma_{\mathrm{L}},K_0,\gamma_{\mathrm{K}}]^{\mathrm{T}}$ and the parameters $\{\sigma_L,\sigma_K,\rho\}$ of the covariance matrix $\mathbf{Q}_{\mathrm{LK}}$. The focus of this paper is on the estimation of $\boldsymbol{\theta}$ and $\mathbf{Q}_{\mathrm{LK}}$ for the characterization of the environment, starting from local observations $\mathbf{y}_{ij}$ made over the active device-to-device links.

Assume that M_0 independent channel measurements can be performed over each link (i,j) (e.g., over different subcarriers): $\mathbf{y}_{ij}^{(m)}\sim\mathcal{N}(\boldsymbol{\mu}_{\mathrm{LK}}(D_{ij}),\mathbf{Q}_{\mathrm{LK}})$ for $m=1,\ldots,M_0$. Considering that node i is connected to the d_i neighbors $\mathcal{N}_i=\{j_1,\ldots,j_{d_i}\}$, the node has access to an overall set of $d_iM_0\times 2$ observations, $\mathbf{Y}_i=[\mathbf{y}_{ij_1}^{(1)}\cdots\mathbf{y}_{ij_1}^{(M_0)}\cdots\mathbf{y}_{ij_{d_i}}^{(1)}\cdots\mathbf{y}_{ij_{d_i}}^{(M_0)}]^{\mathrm{T}}$, that are here collected into the $2d_iM_0\times 1$ data vector:

$$\mathbf{y}_i=\mathrm{vec}[\mathbf{Y}_i]=\mathbf{H}_i\boldsymbol{\theta}+\mathbf{n}_i, \tag{3}$$

with $2d_iM_0\times 4$ system matrix $\mathbf{H}_i=\mathbf{I}_2\otimes\tilde{\mathbf{H}}_i$,

$$\tilde{\mathbf{H}}_i=\begin{bmatrix}\mathbf{1}_{M_0}^{\mathrm{T}}\otimes\mathbf{h}_{ij_1} & \mathbf{1}_{M_0}^{\mathrm{T}}\otimes\mathbf{h}_{ij_2} & \cdots & \mathbf{1}_{M_0}^{\mathrm{T}}\otimes\mathbf{h}_{ij_{d_i}}\end{bmatrix}^{\mathrm{T}}, \tag{4}$$

$\otimes$ denoting the Kronecker matrix product and $\mathbf{1}_{M_0}$ the $M_0\times 1$ vector with all entries equal to 1. Due to the re-arrangement of measurements, the noise term in the signal model (3) is $\mathbf{n}_i\sim\mathcal{N}(\mathbf{0},\mathbf{Q})$ with $\mathbf{Q}=\mathbf{Q}_{\mathrm{LK}}\otimes\mathbf{I}_{d_iM_0}$.

III. Estimation of channel parameters

A. Centralized LS Estimation

Assuming that measurements from all the nodes can be aggregated by a fusion center, the global LS estimate of the channel parameters $\boldsymbol{\theta}$ is obtained as follows:

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = \left(\sum\nolimits_{i=1}^{N} \mathbf{R}_{hh,i}\right)^{-1} \sum\nolimits_{j=1}^{N} \mathbf{R}_{hy,j}, \tag{5}$$

where $\mathbf{R}_{hh,i} = \mathbf{H}_i^{\mathrm{T}}\mathbf{H}_i$ and $\mathbf{R}_{hy,i} = \mathbf{H}_i^{\mathrm{T}}\mathbf{y}_i$ are the correlation matrices at node i. Notice that for data aggregation each node has to transmit to the fusion center the data $\{\mathbf{R}_{hh,i}, \mathbf{R}_{hy,i}\}$ with size of $p(p+1)$ real-valued numbers.

The covariance of the global LS estimate is:

$$\mathbf{C}_{\mathrm{LS}} = \mathrm{Cov}\left(\hat{\boldsymbol{\theta}}_{\mathrm{LS}}\right) = \left(\sum\nolimits_{i=1}^{N} \mathbf{C}_i^{-1}\right)^{-1}, \tag{6}$$

where $\mathbf{C}_i = \mathbf{R}_{hh,i}^{-1}$ denotes the covariance of the local LS estimate at node i, $\hat{\boldsymbol{\theta}}_{\mathrm{LS,i}} = \mathbf{R}_{hh,i}^{-1}\mathbf{R}_{hy,i}$.

The LS estimate of the measurement covariance is obtained by reconstructing the Gaussian deviation to the channel model mean as $\Delta\hat{\mathbf{y}}_{ij}^{(m)} = \mathbf{y}_{ij}^{(m)} - (\mathbf{I}_2 \otimes \mathbf{h}_{ij}^{\mathrm{T}})\hat{\boldsymbol{\theta}}_{\mathrm{LS}}$ and computing the sample covariance as

$$\hat{\mathbf{Q}}_{\mathrm{LS}} = \frac{1}{M_0 \sum_{i=1}^{N} d_i} \sum\nolimits_{i,j,m} \Delta\hat{\mathbf{y}}_{ij}^{(m)} \Delta\hat{\mathbf{y}}_{ij}^{(m)\mathrm{T}}. \tag{7}$$

B. Distributed LS Estimation

In distributed estimation nodes rely solely on their local data and on interactions with neighbors: each node computes a local estimate, it exchanges information with neighbors and updates the computation, repeating the operations till a consensus is reached. Different approaches are presented below, considering both consensus on data (where nodes exchange local measurements) and on parameter estimates (where node exchange directly the estimates of parameters).

1) Method 1 - Consensus on Data: The method is based on the average-consensus algorithm [11] applied to the correlation matrices $\mathbf{R}_{hh}$ and $\mathbf{R}_{hy}$. It is initialized at node i with $\mathbf{R}_{hh,i}(0) = \mathbf{R}_{hh,i}$ and $\mathbf{R}_{hy,i}(0) = \mathbf{R}_{hy,i}$, while consensus-based iterations are performed as:

$$\mathbf{R}_{hh,i}(k+1) = \mathbf{R}_{hh,i}(k) + \varepsilon \sum_{j\in\mathcal{N}_i} (\mathbf{R}_{hh,j}(k) - \mathbf{R}_{hh,i}(k)) \tag{8}$$

$$\mathbf{R}_{hy,i}(k+1) = \mathbf{R}_{hy,i}(k) + \varepsilon \sum_{j\in\mathcal{N}_i} (\mathbf{R}_{hy,j}(k) - \mathbf{R}_{hy,i}(k)) \tag{9}$$

where $\varepsilon > 0$ is the step-size. The estimate at node i is then updated as

$$\hat{\boldsymbol{\theta}}_i(k+1) = \mathbf{R}_{hh,i}^{-1}(k+1)\,\mathbf{R}_{hy,i}(k+1). \tag{10}$$

For $\varepsilon < 1/\Delta$ [11] average consensus is guaranteed, meaning that the local estimates (8)-(9) tend for $k \to \infty$ to the average of the initial node states, $\{\mathbf{R}_{hh,i}\}$ and $\{\mathbf{R}_{hy,i}\}$ respectively. It follows that (10) converges to the global LS estimate:

$$\hat{\boldsymbol{\theta}}_i(\infty) \triangleq \lim_{k\to\infty} \hat{\boldsymbol{\theta}}_i(k) = \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{R}_{hh,i}\right)^{-1} \frac{1}{N}\sum_{j=1}^{N} \mathbf{R}_{hy,j} = \hat{\boldsymbol{\theta}}_{\mathrm{LS}}. \tag{11}$$

As regards the communication overhead required for consensus implementation, each node needs to send the message $\mathrm{msg}_i(k) = \{\mathbf{R}_{hh,i}(k), \mathbf{R}_{hy,i}(k)\}$ to its neighbors at each iteration; this can be accomplished by a packet-based communication by the node i in broadcast mode (i.e., to all its neighbors), with a packet size of $p(p+1)$ real-valued numbers.

The estimate of $\mathbf{Q}$ can be obtained once the consensus on $\boldsymbol{\theta}$ has been reached, by locally reconstructing the measurement errors $\Delta\hat{\mathbf{y}}_{ij}^{(m)}$ and implementing a consensus algorithm similar to (8)-(9) for the sample covariance matrix of $\Delta\hat{\mathbf{y}}_{ij}^{(m)}$.

2) Method 2 - Consensus on LS Estimates: This method is introduced to reduce the amount of information exchange between nodes, from $p(p+1)$ to p, by sending as messages the parameter estimates, $\mathrm{msg}_i(k) = \{\hat{\boldsymbol{\theta}}_i(k)\}$, instead of the data correlation matrices. The ith node estimate is obtained by a consensus algorithm on the estimate $\hat{\boldsymbol{\theta}}$ performed as follows

$$\hat{\boldsymbol{\theta}}_i(k+1) = \hat{\boldsymbol{\theta}}_i(k) + \varepsilon \sum\nolimits_{j\in\mathcal{N}_i} \left(\hat{\boldsymbol{\theta}}_j(k) - \hat{\boldsymbol{\theta}}_i(k)\right), \tag{12}$$

with initialization $\hat{\boldsymbol{\theta}}_i(0) = \hat{\boldsymbol{\theta}}_{\mathrm{LS,i}}$.

According to this initialization, if $\varepsilon < 1/\Delta$, the estimate converges to the average of local LS estimates (ALS):

$$\hat{\boldsymbol{\theta}}_i(\infty) = \frac{1}{N}\sum\nolimits_{i=1}^{N} \hat{\boldsymbol{\theta}}_{\mathrm{LS,i}} = \hat{\boldsymbol{\theta}}_{\mathrm{ALS}}, \tag{13}$$

whose covariance can be calculated as

$$\mathbf{C}_{\mathrm{ALS}} = \frac{1}{N^2}\sum\nolimits_{i=1}^{N} \mathbf{C}_i \geq \left(\sum\nolimits_{i=1}^{N} \mathbf{C}_i^{-1}\right)^{-1} = \mathbf{C}_{\mathrm{LS}}. \tag{14}$$

The above inequality derives from the harmonic-arithmetic matrix inequality [15], and $\mathbf{C}_{\mathrm{ALS}} \geq \mathbf{C}_{\mathrm{LS}}$ stands for $\mathbf{C}_{\mathrm{ALS}} - \mathbf{C}_{\mathrm{LS}}$ positive-semidefinite. From (14), it follows that Method 2 is suboptimal; intuitively this is due to the fact that local estimates have different accuracies and fusion does not account for this unbalance. In the next section weighting matrices are introduced to overcome this limit.

3) Method 3 - Consensus on Weighted LS Estimates: The method employs the same consensus algorithm as in (12) but with weighted local LS estimates as initial states, $\hat{\boldsymbol{\theta}}_i(0) = \mathbf{W}_i\hat{\boldsymbol{\theta}}_{\mathrm{LS,i}}$, with $p\times p$ positive definite weighting matrix $\mathbf{W}_i$. In this case the estimate at convergence is the weighted average of the local LS estimates (WALS):

$$\hat{\boldsymbol{\theta}}_{\mathrm{WALS}} = \lim_{k\to\infty} \hat{\boldsymbol{\theta}}_i(k) = \frac{1}{N}\sum\nolimits_{i=1}^{N} \mathbf{W}_i\hat{\boldsymbol{\theta}}_{\mathrm{LS,i}}, \tag{15}$$

and the corresponding covariance of the estimate is

$$\mathbf{C}_{\mathrm{WALS}} = \frac{1}{N^2}\sum\nolimits_{i=1}^{N} \mathbf{W}_i\mathbf{C}_i\mathbf{W}_i^{\mathrm{T}}. \tag{16}$$

Looking at (16), we see that $\mathbf{W}_i$ can be chosen so that the estimate covariance equals the global LS performance (6), $\mathbf{C}_{\mathrm{WALS}} = \mathbf{C}_{\mathrm{LS}}$, i.e. the minimum covariance for any unbiased estimate (given that $\mathbf{Q}$ is unknown). By setting the weights:

$$\mathbf{W}_{i,\mathrm{opt}} = N\left(\sum\nolimits_{j=1}^{N} \mathbf{C}_j^{-1}\right)^{-1} \mathbf{C}_i^{-1}, \tag{17}$$

we obtain in fact that (15) equals the global LS estimate (5): $\hat{\boldsymbol{\theta}}_{\mathrm{WALS}} = \hat{\boldsymbol{\theta}}_{\mathrm{LS}}$.

Unfortunately, in general networks scenarios (without all-to-all connectivity) optimal weights (17) are not available, as each node i has access to measurements provided by a limited set of nodes, i.e. the set of neighbors in $\mathcal{N}_i$. An alternative solution for weighting is to let the nodes exchange the local LS estimates and the related covariances at the first iteration, i.e. $\mathrm{msg}_i(0) = \{\hat{\boldsymbol{\theta}}_{\mathrm{LS,i}}, \mathbf{C}_i\}$, compute the suboptimal weights

$$\mathbf{W}_{i,\mathrm{sub}} = (d_i + 1)\left(\sum\nolimits_{j\in\mathcal{N}_i\cup\{i\}}^{N} \mathbf{C}_j^{-1}\right)^{-1} \mathbf{C}_i^{-1}, \quad (18)$$

and then implement the consensus iterations (12) with $\mathrm{msg}_i(k) = \{\hat{\boldsymbol{\theta}}_i(k)\}$ for $k > 0$. The amount of information exchanged between nodes with each neighbor is now $p(p+3)/2$ at first iteration followed by p values in the subsequent iterations. The estimate at convergence is $\hat{\boldsymbol{\theta}}_{\mathrm{WALS}} = \frac{1}{N}\sum\limits_i \mathbf{W}_{i,\mathrm{sub}}\hat{\boldsymbol{\theta}}_{\mathrm{LS,i}} \neq \hat{\boldsymbol{\theta}}_{\mathrm{LS}}$ with covariance $\mathbf{C}_{\mathrm{WALS}} \geq \mathbf{C}_{\mathrm{LS}}$. Equality with the global LS estimate is guaranteed only for all-to-all connectivity (i.e., for complete graph).

4) Method 4 - Weighted Consensus on LS Estimates: In this new approach nodes perform weighted-average consensus as follows:

$$\hat{\boldsymbol{\theta}}_i(k+1) = \hat{\boldsymbol{\theta}}_i(k) + \varepsilon\mathbf{W}_i^{-1}\sum\nolimits_{j\in\mathcal{N}_i}\left(\hat{\boldsymbol{\theta}}_j(k) - \hat{\boldsymbol{\theta}}_i(k)\right), \quad (19)$$

initialized with $\hat{\boldsymbol{\theta}}_i(0) = \hat{\boldsymbol{\theta}}_{\mathrm{LS,i}}$. Unlike previous methods, here each node weights - *directly* in the consensus estimate update - its local estimate with the term $\mathbf{W}_i$ or, equivalently, the information coming from other nodes (the summation term in (19)) with $\mathbf{W}_i^{-1}$. As shown in the Appendix, this leads at convergence to the weighted average of the initial estimates:

$$\hat{\boldsymbol{\theta}}_i(\infty) = \hat{\boldsymbol{\theta}}_{\mathrm{WA}} = \left(\sum\nolimits_{n=1}^{N}\mathbf{W}_n\right)^{-1}\sum\nolimits_{i=1}^{N}\mathbf{W}_i\hat{\boldsymbol{\theta}}_i(0), \quad (20)$$

if the step size is $\varepsilon < 2/\lambda_{\max}(\mathbf{W}^{-1}(\mathbf{L}\otimes\mathbf{I}_p))$, with $\lambda_{\max}(\cdot)$ denoting the maximum eigenvalue of the argument matrix and $\mathbf{W} = \mathrm{blockdiag}(\mathbf{W}_1, \ldots, \mathbf{W}_N)$ being the $Np \times Np$ block-diagonal matrix built from the N weighting matrices.

Furthermore, if weights are optimally selected as $\mathbf{W}_i = \mathbf{C}_i^{-1}$, we get that (20) equals (5) yielding convergence to the global LS estimate. By this choice the update term in (19) based on neighboring node information is weighted by $\mathbf{C}_i$: the higher is the local estimate covariance, the higher is the reliability given to the information provided by other nodes.

The advantage of this method is that it allows to reach the global LS performance with message exchange limited to the p-parameter estimates $\mathrm{msg}_i(k) = \{\hat{\boldsymbol{\theta}}_i(k)\}$, without the need of any covariance matrix exchange.

Remark. If weights are chosen such that $\mathbf{W}_i \geq \mathbf{I}$ (i.e., with eigenvalues $\lambda(\mathbf{W}_i) \geq 1$), the bound on the step-size for convergence is simplified as it is:

$$\lambda(\mathbf{W}^{-1}(\mathbf{L}\otimes\mathbf{I}_p)) \leq \lambda((\mathbf{L}\otimes\mathbf{I}_p)) \leq 2\Delta, \quad (21)$$

being the eigenvalues of the Laplacian matrix bounded by $\lambda(\mathbf{L}) \leq 2\Delta$ [11]. Thereby, for guaranteeing convergence we can simply impose $\varepsilon < 1/\Delta$. To get normalized weights such that $\mathbf{W}_i \geq \mathbf{I}$ from the optimal ones $\tilde{\mathbf{W}}_i = \mathbf{C}_i^{-1}$ we can set: $\mathbf{W}_i = \left(\sum_n \tilde{\mathbf{W}}_n^{-1}\right)\tilde{\mathbf{W}}_i = \left(\sum_n \mathbf{C}_n\right)\mathbf{C}_i^{-1}$.

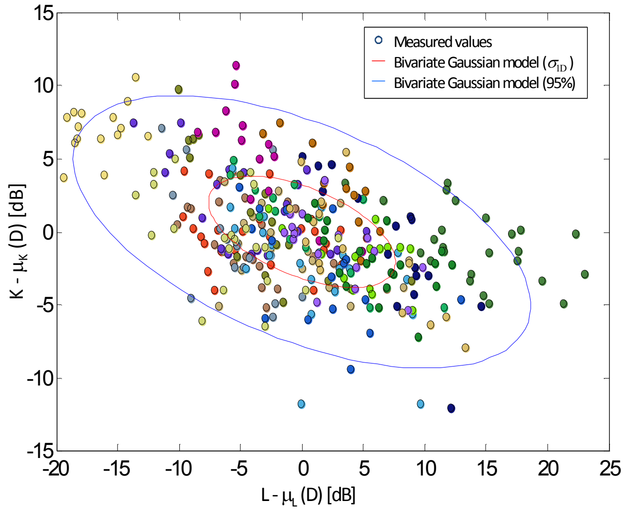

Fig. 2. Bivariate Gaussian model fitting from experiments.

IV. Channel-Model Validation on Real Data

In this section, we validate the channel model presented in Sec. II on real data collected by an indoor measurement campaign at the third floor of the department DEIB of Politecnico di Milano, with IEEE 802.15.4 compliant wireless terminals [6] deployed according to the cloud architecture of the DIWINE project [1]. Nodes were deployed in $N = 16$ fixed locations, connected over 19 links as in Fig. 1-(b). Received signal strength (RSS) measurements have been taken during day-time over distances ranging from about $D = 7$m to $D = 49$m, in mixed LOS/NLOS conditions. Devices were positioned on wooden supports, about 1m above the floor. A PC connected by serial interface to one of the devices collected all measurements. The radio module used for the experiment provides different programmable high-power modes with maximum transmit power of $P_{\mathrm{tx}} = 18$dBm; the receive power ranges from the minimum sensitivity $P_{\mathrm{rx,min}} = -98$dBm to the maximum of $P_{\mathrm{rx,max}} = -25$dBm.

For each link (i, j), RSS measurements have been recorded over 670 time instants (with sampling $\Delta t = 1$s) and $M_0 = 15$ IEEE 802.15.4 channels (with carrier spacing $\Delta f = 5$MHz) in the 2.4 GHz band, for a total number of approximately 10^4 RSS samples per link. Each RSS sample is an average over 8 PHY symbol periods of 128μs each. The path-loss $L_{ij}^{(m)}$ has been evaluated on each frequency, $m = 1, \ldots, M_0$, as the loss with respect to the nominal maximum receive power, $L_{ij}^{(m)} = P_{\mathrm{rx,max}} - P_{\mathrm{rx}}^{(m)}$, by computing the static channel component $P_{\mathrm{rx}}^{(m)}$ as the average of the 670 RSS measurements of the considered link. The static channel component has been then removed from the RSS measurements, to highlight the dynamic one. The Rician K-factor $K_{ij}^{(m)}$ has been finally estimated by minimizing the L_2 norm of the difference between the experimental and theoretical probability density function (pdf) of the instantaneous power of the dynamic component $|h_{i,j}|^2$, expressed in dBm.

In Fig. 2, the measured values $\{L_{ij}^{(m)} - \mu_L(D_{ij}^{(m)}\}$ and

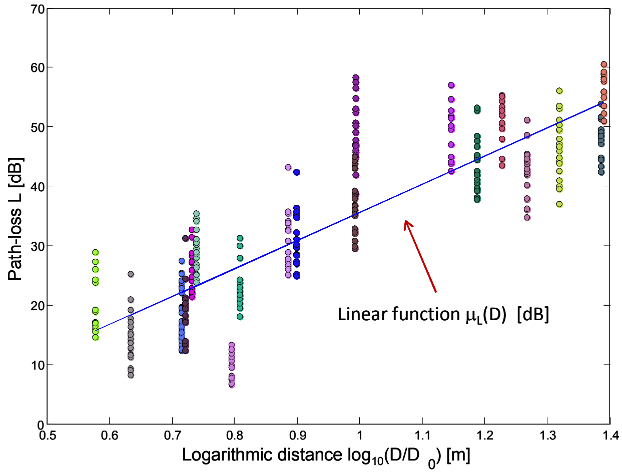

Fig. 3. Linear least-square regression of path-loss $\{L_{ij}^{(m)}\}$ from data and estimated mean function $\mu_L(D)$ vs. distance ($D = 7$m to $D = 49$m).

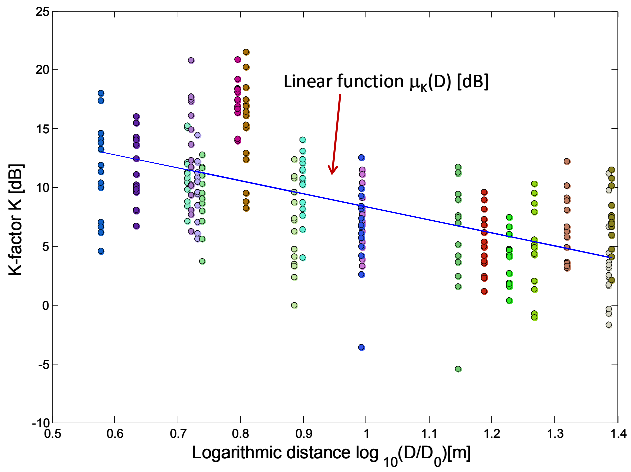

Fig. 4. Linear least square regression of K-factor $\{K_{ij}^{(m)}\}$ from data and estimated mean function $\mu_K(D)$ vs. distance ($D = 7$m to $D = 49$m).

$\{K_{ij}^{(m)} - \mu_K(D_{ij}^{(m)})\}$ are shown for all links and frequencies, together with the equidensity contours of the zero-mean bivariate Gaussian distribution whose parameters have been obtained as described below. The error ellipses are associated to 39% ($\sigma_{1\text{D}}$) and 95% of confidence level ($2.45\sigma_{1\text{D}}$).

The mean functions $\mu_L(D_{ij})$ and $\mu_K(D_{ij})$ have been estimated by performing a linear LS regression of $\{L_{ij}^{(m)}\}$ and $\{K_{ij}^{(m)}\}$ (see (2)) over the logarithmic distances $\log(D_{ij}/D_0)$ with $D_0 = 1$m. The resulting linear functions are illustrated in Fig. 3 and Fig. 4. As expected, increasing the distance leads to an increase of path-loss and a decrease of the Rician K-factor on average. The covariance matrix $\mathbf{Q}_{LK}$ has been obtained by computing variances and covariances of $\{L_{ij}^{(m)} - \mu_L(D_{ij}^{(m)})\}$ and $\{K_{ij}^{(m)} - \mu_K(D_{ij}^{(m)})\}$ from the available set of data. Since the path-loss and the Rician K-factor exhibit anti-correlated spatial variations, this leads to a strong negative correlation ρ. The radio-environment parameters (1)-(2) obtained by this analysis are: $L_0 = 15.6$ dB, $\gamma_L = 4.74$, $K_0 = 12.54$ dB, $\gamma_K = 1.04$, $\rho = -0.54$, $\sigma_L = 7.6$dB, $\sigma_K = 3.8$dB.

V. Consensus Performance Analysis

The performance of consensus methods are now analyzed for a simulated scenario with $N = 10$ nodes uniformly distributed within a 70×70m area as exemplified in Fig. 5-(a). The parameters $\boldsymbol{\theta}$ and $\mathbf{Q}_{\text{LK}}$ are chosen according to the results of the calibration procedure carried out on real data in the previous section. Connectivity is simulated assuming a radio coverage of radius 45m at each node. Measurements at node i are simulated using the model in (1)-(3), with a single path-loss and Rician K-factor measurement per link, for a total number of $M_i = 2d_i$ observations available at node i. Our purpose is to evaluate the performance of the proposed distributed LS methods based on the consensus approach for the estimation of the channel parameters $\boldsymbol{\theta}$ and $\mathbf{Q}_{LK}$.

The performance of each distributed algorithm is evaluated in terms of root mean square error (RMSE) of the estimate by averaging over 60 graphs and 170 measurement outcomes. The step size has been set to $\varepsilon = 0.94\varepsilon_{\max}$, with $\varepsilon_{\max} = 1/\Delta$ for Methods 1-3 and $\varepsilon_{\max} = 2/\lambda_{\max}(\mathbf{W}^{-1}(\mathbf{L} \otimes \mathbf{I}_p))$ for Method 4. The RMSE is shown vs. the number of iterations in Fig. 5-(a) (for $\boldsymbol{\theta}$) and 5-(b) (for $\mathbf{Q}_{LK}$), where the global LS estimate performance is shown as reference. It can be observed that Method 1, 3 with optimal weights, and 4 converge to the global LS estimate performance. Method 1 requires the smallest time for convergence, but it also needs to exchange the largest set of data between nodes ($O(p^2)$) and the largest complexity for consensus implementation ($O(p^3)$ multiplications per iteration). Method 3, on the other hand, is unfeasible in practical systems as it assumes in (17) the knowledge at all nodes of all estimate covariances (i.e., all-to-all connectivity and $O(p^2)$ data exchange over each link). Method 4 has slower convergence but it has the advantage of easy implementation, requiring only $O(p^2)$ multiplications per iteration and communication overhead of p parameters per link. The remaining methods, 2 and 3 with sub-optimal weights, do not converge to the global LS estimate. In fact, in Method 2 each node estimates the p parameters by simply applying consensus on the local LS estimates without exchanging any information about the estimate covariance. On the other hand, in Method 3 with sub-optimal weights each node exchanges local estimates and also covariances at first iteration, but weights are sub-optimal having each node access to a limited set of data. To conclude, the best solutions are Method 1 and 4, as they both reach optimal performance, with comparison - in terms of overall complexity and communication overhead - depending in general on the number p of parameters to be estimated.

VI. Concluding Remarks

Consensus based algorithms have been investigated for distributed identification of the macroscopic channel parameters that characterize the wireless propagation in an indoor wireless ad-hoc network. Various techniques have been considered, including both consensus on data correlation and consensus on parameter estimates. In the latter case, weight factors have been introduced to account for different measurement reliability at nodes and to allow convergence to the global LS estimate. Based on numerical results, Method 1 and 3

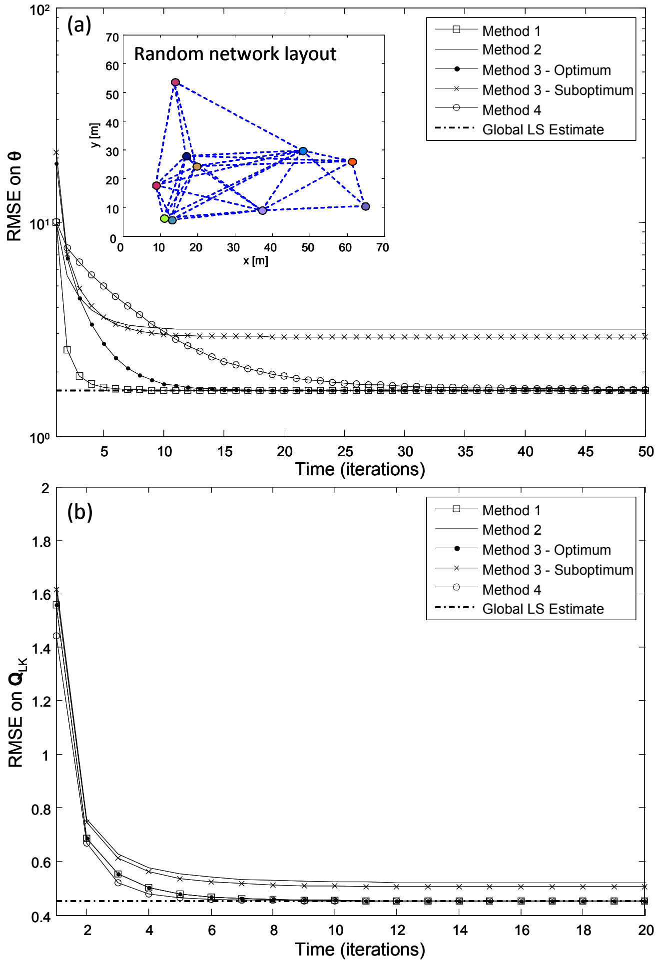

Fig. 5. RMSE performance of distributed consensus-based algorithms (Methods 1-4) for parameter estimation $\boldsymbol{\theta}$ (a) and $\mathbf{Q}_{LK}$ (in dB scale) (b). The network layout is superimposed on top subfigure.

(with optimal weights) offer faster convergence and optimal performance, but with higher amount of data exchange among sensors. Method 3 becomes suboptimal for reduced connectivity. On the other hand, Method 4 allows to reduce the number of parameters to be shared still guaranteeing convergence - with a slower rate - to the optimal performance. Improvement of the convergence rate will be considered in future work.

Appendix

We first observe that the consensus algorithm (19) can be equivalently written as $\boldsymbol{\theta}(k+1) = \mathbf{P}\boldsymbol{\theta}(k)$ with $\boldsymbol{\theta}(k) = \left[\boldsymbol{\theta}_1^{\mathrm{T}}(k)\cdots\boldsymbol{\theta}_N^{\mathrm{T}}(k)\right]^{\mathrm{T}}$, Perron matrix $\mathbf{P} = \mathbf{I}_{Np} - \varepsilon\mathbf{W}^{-1}\tilde{\mathbf{L}}$ and Laplacian matrix $\mathbf{W}^{-1}\tilde{\mathbf{L}}$ with $\tilde{\mathbf{L}} = \mathbf{L}\otimes\mathbf{I}_p$. We consider the eigenvalue decomposition (EVD) of the Laplacian matrix, $\mathbf{W}^{-1}\tilde{\mathbf{L}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}$, where $\mathbf{U} = [\mathbf{u}_1\cdots\mathbf{u}_{Np}]$ and $\mathbf{V} = [\mathbf{v}_1\cdots\mathbf{v}_{Np}]$ are the $Np\times Np$ matrices of, respectively, the left and right eigenvectors, whereas $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1,\ldots,\lambda_{Np})$ is the diagonal matrix of the corresponding eigenvalues sorted in non-decreasing order. Considering that $\mathbf{L}$ has a trivial eigenvalue equal to 0 and recalling the Kronecker structure of $\tilde{\mathbf{L}}$, it follows that the first p eigenvalues of $\mathbf{W}^{-1}\tilde{\mathbf{L}}$ are $\lambda_1 = \ldots = \lambda_p = 0$. It can be shown that the associated left and right eigenvectors are: $\mathbf{U}_0 = \frac{1}{\sqrt{N}}\mathbf{1}_N\otimes\mathbf{I}_p$ and $\mathbf{V}_0 = \sqrt{N}\mathbf{W}(\mathbf{1}_N\otimes\mathbf{I}_p)(\sum_m\mathbf{W}_m)^{-1}$, with $\mathbf{U}_0$ scaled such that $\mathbf{U}_0^T\mathbf{U}_0 = \mathbf{I}$. The last statement can be proven by few algebraic passages, exploiting the doubly stochastic property of $\mathbf{L}$ (i.e., $\mathbf{1}^T\mathbf{L} = \mathbf{0}$ and $\mathbf{L1} = \mathbf{0}$) and showing that $\{\mathbf{U}_0,\mathbf{V}_0\}$ satisfy the following conditions: i) $\mathbf{V}_0^{\mathrm{T}}(\mathbf{W}^{-1}\tilde{\mathbf{L}}) = (\sum_m\mathbf{W}_m)^{-1}(\mathbf{1}^T\mathbf{L}\otimes\mathbf{I}_p) = \mathbf{0}_{p\times pN}$; ii) $(\mathbf{W}^{-1}\tilde{\mathbf{L}})\mathbf{U}_0 = \mathbf{W}^{-1}(\mathbf{L}\mathbf{1}_N\otimes\mathbf{I}_p) = \mathbf{0}_{Np\times p}$; iii) $\mathbf{V}_0^T\mathbf{U}_0 = \mathbf{I}_p$. We can thus write the EVD of $\mathbf{P}$ as $\mathbf{P} = \mathbf{U}_0\mathbf{V}_0^{\mathrm{T}} + \sum_{n=p+1}^{Np}\mu_n\mathbf{u}_n\mathbf{v}_n^{\mathrm{T}}$ with $\mu_n = 1-\varepsilon\lambda_n$. Using the above result on the first p eigenvectors, consensus iterations become:

$$\boldsymbol{\theta}(k+1) = \mathbf{P}^{k+1}\boldsymbol{\theta}(k) = \mathbf{U}_0\mathbf{V}_0^T\boldsymbol{\theta}(0) + \sum_{n=p+1}^{Np}\mu_n^{k+1}\mathbf{u}_n\mathbf{v}_n^{\mathrm{T}}\boldsymbol{\theta}(0)$$

with $\mathbf{U}_0\mathbf{V}_0^T\boldsymbol{\theta}(0) = [\hat{\boldsymbol{\theta}}_{\mathrm{WA}}^{\mathrm{T}}\cdots\hat{\boldsymbol{\theta}}_{\mathrm{WA}}^{\mathrm{T}}]^{\mathrm{T}}$ with $\hat{\boldsymbol{\theta}}_{\mathrm{WA}}$ defined as in (20). Thereby, the N node estimates converge to the weighted average of the initial states, $\hat{\boldsymbol{\theta}}_{\mathrm{WA}}$, if $\mu_n^{k+1}\to 0$, i.e. for $|\mu_n| = |1-\varepsilon\lambda_n| < 1$ or equivalently: $0 < \varepsilon < 2/\lambda_{\max}(\mathbf{W}^{-1}(\mathbf{L}\otimes\mathbf{I}_p))$.

References

[1] H. Mark, "Dense Cooperative Wireless Cloud Network (DIWINE)," *Future Network & Mobile Summit*, Lisbon, Portugal, July 2013.
[2] S. Savazzi, U. Spagnolini, L. Goratti, S. Galimberti, "Wireless cloud networks for critical industrial quality control," *Proc. IEEE 10th International Symposium on Wireless Communication Systems (ISWCS)*, Ilmenau, Germany, 27-30 Aug. 2013, pp. 592-596.
[3] J. N. Laneman, D. N. C. Tse, G. W. Womell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Information Theory*, Vol. 50, No. 12, pp. 3062-3080, Dec. 2004.
[4] P.A. Chou, Y. Wu, "Network coding for the internet and wireless networks," *IEEE Sig. Proc. Mag.*, Vol. 24, No. 5, pp. 77-85, Sep. 2007.
[5] P. Castiglione, S. Savazzi, M. Nicoli, T. Zemen, "Partner selection in indoor-to-outdoor cooperative networks: an experimental study," *IEEE Journal on Selected Areas in Communications*, special issue on Theories and Methods for Advanced Wireless Relays, Vol. 31, No. 6, Aug. 2013.
[6] Data-sheet JN-DS-JN5148-001, IEEE 802.15.4 Wireless Microcontroller JN5148-001, NXP Laboratories, UK, (2012).
[7] N. Patwari, et al., "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, Vol. 22, No. 4, pp. 54-69, July 2005.
[8] M. Nicoli, D. Fontanella, "Fundamental performance limits of TOA-based cooperative localization," *Proc. IEEE International Conference on Communications Workshops*, Dresden, Germany, June 2009, pp. 1-5.
[9] D. Fontanella, M. Nicoli, L. Vandendorpe, "Bayesian localization in sensor networks: Distributed algorithm and fundamental limits", *Proc. IEEE International Conference on Communications (ICC)*, CapeTown, South Africa, 2010, pp. 1-5.
[10] S. Savazzi, M. Nicoli, F. Carminati, M. Riva, "A Bayesian approach to device-free localization: modeling and experimental assessment," *IEEE Journal on Sel. Topics in Signal Proc.*, Vol. 8, No. 1, Feb. 2014.
[11] R. Olfati-Saber, J. A. Fax, R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, Vol. 95, No. 1, pp. 215-233, Jan. 2007.
[12] D.P. Spanos, R.M. Murray, "Distributed sensor fusion using dynamic consensus," *Proc. 16th IFAC World Congress*, Prague, Czech, 2005.
[13] I.D. Schizas, A. Ribeiro, G.B. Giannakis, "Consensus in ad hoc WSNs with noisy links - Part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Processing*, Vol. 56, No. 1, pp. 350-364, Jan. 2008.
[14] S. Bolognani, S. D. Favero, L. Schenato, D. Varagnolo,"Consensus based distributed sensor calibration and least-square parameter identification in WSNs," *International Journal of Robust and Nonlinear Control*, Vol. 20, No. 2, pp. 176-193, Jan. 2010.
[15] J. Pečarić, B. Mond, "The arithmetic mean, the geometric mean and related matrix inequalities," *ISNM International Series of Numerical Mathematics*, Vol. 123, pp. 77-91, 1997.