SERBIAN JOURNAL OF ELECTRICAL ENGINEERING
Vol. 13, No. 1, February 2016, 111-132

UDC: 004.7:681.5]:519.218 **DOI: 10.2298/SJEE1601111S**

A Consensus-Based Distributed Calibration Algorithm for Sensor Networks

Miloš S. Stanković[1], Srdjan S. Stanković[1,2], Karl Henrik Johansson[3]

Abstract: In this paper a novel distributed algorithm for blind macro-calibration in sensor networks based on consensus is proposed. The algorithm is formulated as a set of gradient-type recursions for estimating parameters of sensor calibration functions, starting from local criteria defined as weighted sums of mean square differences between the outputs of neighboring sensors. It is proved that the algorithm achieves asymptotic agreement for sensor gains and offsets in the mean square sense and with probability one. In the case of additive measurement noise, additive inter-agent communication noise and communication outages, a modification of the original algorithm is proposed. It is proved using stochastic approximation arguments that the modified algorithm achieves asymptotic consensus for sensor gains and offsets in the mean square sense and with probability one. Special attention is paid to the situation when one sensor is selected as a reference. Illustrative simulation examples are provided.

Keywords: Calibration, Sensor Networks, Consensus, Stochastic Approximation, Convergence.

1 Introduction

Recently, *wireless sensor networks* (WSN's) have emerged as an important research area (see, *e.g.*, [3 – 5]). Diverse new applications have sparked the recognition of new classes of problems for the developers and users. *Sensor calibration* represents one of the most important challenges for the wide deployment of this new technology. Only relatively small sensor-systems can utilize micro-calibration, in which each device is individually tuned in a controlled environment. Larger WSN's demand new concepts and methods for calibration, since many devices can be in partially unobservable and dynamic environments. The so-called *macro-calibration* is based on the idea to calibrate a *network as a whole* by observing only the overall system response, thus eliminating the need to directly calibrate each and every device. The usual

[1]School of Electrical Engineering, University of Belgrade, Serbia; E-mail: milos.stankovic@ic.etf.rs
[2]Vlatacom Institute, Belgrade, Serbia; E-mail: stankovic@etf.rs
[3]K.H. Johansson is with the ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden; E-mail: kallej@kth.se

prerequisite is to frame calibration as a parameter estimation problem (*e.g.*, [6, 7]). Automatic methods for jointly calibrating WSN's, without dependence on controlled stimuli or high-fidelity groundtruth data are of significant interest. The underlying practical idea is to have sensors with homogeneous properties irrespective of the lack of an absolute reference, expecting (or imposing) dominant influence of well calibrated sensors. This problem is referred to as *blind calibration*. In [8, 9] a *centralized*, non-recursive algorithm was proposed for blind calibration assuming non-identical but correlated sensor readings. Another approach to blind WSN calibration is to assume that the deployment is relatively dense, so that neighboring nodes have nearly identical readings. In [10] this scenario was adopted, but the algorithm assumes only *pairwise* inter-node communications.

In this paper we propose a novel *blind macro-calibration method* for sensor networks based on *distributed on-line estimation* of the parameters of affine *calibration functions*. The work represents an extension of the results presented in [11]. It is assumed that the sensors form a network with *directed communication links* between neighboring nodes. By formulating the calibration problem as the problem of distributed minimization of a set of weighted sums of mean square differences between the outputs of neighboring sensors, we derive a *distributed gradient type recursive algorithm* and show that the overall network behavior can be treated as a nontrivial *consensus* problem. To the authors' best knowledge, consensus has been applied directly to the calibration problems only in [12, 13], but within different contexts. Using general arguments related to stability of diagonally dominant dynamic systems [14 – 16], we prove that the proposed basic algorithm achieves *asymptotic consensus for sensor gains and offsets* in the mean square sense and with probability one (w.p.1). The asymptotic gains and offsets depend, in general, on initial conditions, together with signal, sensor and network characteristics. The basic results are then extended by assuming: 1) additive communication noise, 2) communication outages, and 3) additive measurement noise. A modified algorithm is proposed for solving the problem posed in this case, and the achievement of the asymptotic consensus in the mean square sense and w.p.1 is proved for both gains and offsets. In case when a node is kept with fixed calibration parameters, it is proved that the algorithm provides convergence in the mean square sense and w.p.1 to the given reference.

The outline of the paper is as follows. The following subsection introduces notation and basic definitions to be used throughout the paper. In Section 2 we formulate the blind calibration problem and introduce the basic algorithm. Section 3 is devoted to the convergence analysis under different assumptions on the measured signals and network structure. A modified gradient algorithm is proposed for the general case of lossy sensor networks. Its convergence in the mean square sense and w.p.1 is proved and the convergence rate estimated.

Attention is devoted also to the convergence analysis of the proposed algorithm in the case of macro-calibration where a subset of sensors remains with fixed characteristics. In Section 4 we present some illustrative simulation results.

1.1 Notation and some definitions

$\mathbb{R}$ denotes the set of real numbers, $\mathbb{C}$ denotes the set of complex numbers, while $\mathbb{C}_+$ denotes the set of complex numbers with positive real parts. $E\{\cdot\}$ denotes the mathematical expectation. I_k denotes the identity matrix of dimension k, $0_{i\times j}$ denotes a $i\times j$ matrix with all the elements equal to zero. $\otimes$ denotes the Kronecker product. $\mathrm{diag}\{\ldots\}$ denotes a block diagonal matrix with the specified diagonal elements. $\|\cdot\|$ denotes an operator norm. $\mathrm{span}\{A\}$ denotes a linear space spanned by the columns of matrix A. $\lambda_{\min}(A)$ denotes an eigenvalue of matrix A with the smallest absolute value, while $\lambda_{\max}(A)$ denotes an eigenvalue of matrix A with the largest absolute value.

Matrix $A\in\mathbb{R}^{k\times k}$ is said to be an M-matrix if its off-diagonal entries are less than or equal to zero, and if all its principal minors are positive (see, *e.g.*, [15] for different equivalent characterizations of M-matrices).

In a given directed graph (digraph) $\mathcal{G}=(\mathcal{N},\mathcal{E})$, where $\mathcal{N}$ is the set of nodes (vertices) and $\mathcal{E}$ is the set of links (arcs), if there is a walk from the node j to the node i we say that the node i is reachable from j. A node from which every node in the digraph is reachable is called center node.

2 Problem Formulation and the Main Algorithm

Consider n distributed sensors measuring a discrete-time signal $x(t)$, $t=\ldots,-1,0,1,\ldots$, and assume that the output of the i-th sensor can be represented by

$$y_i(t)=\alpha_i x(t)+\beta_i, \tag{1}$$

where the gain α_i and the offset β_i are unknown constants.

By sensor calibration we consider the application of an affine *calibration function* which produces the overall sensor output

$$z_i(t)=a_i y_i(t)+b_i=g_i x(t)+f_i, \tag{2}$$

where a_i and b_i are the calibration parameters, $g_i=a_i\alpha_i$ is the equivalent gain and $f_i=a_i\beta_i+b_i$ the equivalent offset. In general, parameters a_i and b_i have to be chosen in such a way as to set g_i as close as possible to one and f_i as close as possible to zero.

We assume that the observed sensors form a network with a predefined structure of inter-sensor communications represented by a *directed graph* $\mathcal{G}=(\mathcal{N},\mathcal{E})$, where $\mathcal{N}$ is the set of nodes (one node corresponds to one sensor) and $\mathcal{E}$ the set of links (arcs). Define the adjacency matrix $A=[a_{ij}]$, $i,j=1,\ldots,n$, where $a_{ij}=1$ if the j-th sensor can send its message to the i-th sensor, and $a_{ij}=0$ otherwise. Let $\mathcal{N}_i$ be the set of neighboring nodes of the i-th node, *i.e.*, the set of nodes j for which $a_{ij}=1$.

Starting from the general concept of *blind macro calibration*, the aim of this paper is to propose an algorithm for *distributed real-time estimation of the calibration parameters* a_i and b_i which would provide asymptotically equal outputs $z_i(t)$ of all the sensors without the knowledge of the measured signal. Furthermore, the algorithm should be adjusted such that, loosely speaking, the well calibrated sensors would correct, on the basis of global *consensus*, the behavior of those that are not. Assuming that $\{x(t)\}$ is a stochastic process, the algorithm will be derived starting from the following set of *local criteria*:

$$J_i=\sum_{j\in\mathbf{N}_i}\gamma_{ij}E\{(z_j(t)-z_i(t))^2\}, \tag{3}$$

$i=1,...,n$, where γ_{ij} are nonnegative scalar weights reflecting, in general, relative importance of the neighboring nodes. Denoting $\theta_i=[a_i\ \ b_i]^T$, we obtain for the gradient of:

$$\text{grad}_{\theta_i}J_i=\sum_{j\in\mathbf{N}_i}\gamma_{ij}E\left\{(z_j(t)-z_i(t))\begin{bmatrix}y_i(t)\\1\end{bmatrix}\right\}. \tag{4}$$

The last equation gives rise to the following stochastic *gradient recursion* for estimating θ_i^* which minimizes (3)

$$\hat{\theta}_i(t+1)=\hat{\theta}_i(t)+\delta_i(t)\sum_{j\in\mathbf{N}_i}\gamma_{ij}\varepsilon_{ij}(t)\begin{bmatrix}y_i(t)\\1\end{bmatrix}, \tag{5}$$

where $\hat{\theta}_i(t)=[\hat{a}_i(t)\ \ \hat{b}_i(t)]^T$, $\delta_i(t)>0$ is a time-varying gain influencing convergence properties of the algorithm, $\varepsilon_{ij}(t)=\hat{z}_j(t)-\hat{z}_i(t)$ and $\hat{z}_i(t)=\hat{a}_i(t)y_i(t)+\hat{b}_i(t)$, with the initial conditions $\hat{\theta}_i(0)=[1\ \ 0]^T$, $i=1,\ldots,n$. Notice that each iteration of the algorithm (5) subsumes availability of the local measurement and the reception of current messages of the neighboring nodes' outputs $z_j(t)$ (communication outages will be discussed later). Local computational efforts for individual agents are minor, having in mind that only two parameters are

estimated. Communication efforts depend on the number of agents in the neighborhoods, which is kept to be small in actual WSN's.

The underlying idea of the set of recursions (5) is to ensure that the estimates of all the local gains $\hat{g}_i(t) = \hat{a}_i(t)\alpha_i$ and offsets $\hat{f}_i(t) = \hat{a}_i(t)\beta_i + \hat{b}_i(t)$ tend asymptotically to the *same values* $\overline{g}$ and $\overline{f}$, respectively, implying $\hat{z}_j(t) = \hat{z}_i(t)$, $i, j = 1, \ldots, n$. For the sake of compact notations, introduce

$$\hat{\rho}_i(t) = \begin{bmatrix} \hat{g}_i(t) \\ \hat{f}_i(t) \end{bmatrix} = \begin{bmatrix} \alpha_i & 0 \\ \beta_i & 1 \end{bmatrix} \hat{\theta}_i(t), \tag{6}$$

and

$$\varepsilon_{ij}(t) = \begin{bmatrix} x(t) & 1 \end{bmatrix} (\hat{\rho}_j(t) - \hat{\rho}_i(t)), \tag{7}$$

so that (5) becomes

$$\hat{\rho}_i(t+1) = \hat{\rho}_i(t) + \delta_i(t) \sum_{j \in N_i} \gamma_{ij} \Phi_i(t) (\hat{\rho}_j(t) - \hat{\rho}_i(t)), \tag{8}$$

where

$$\begin{aligned} \Phi_i(t) &= \begin{bmatrix} \alpha_i y_i(t) x(t) & \alpha_i y_i(t) \\ [1 + \beta_i y_i(t)] x(t) & 1 + \beta_i y_i(t) \end{bmatrix} = \\ &= \begin{bmatrix} \alpha_i \beta_i x(t) + \alpha_i^2 x^2(t) & \alpha_i \beta_i + \alpha_i^2 x(t) \\ (1 + \beta_i^2) x(t) + \alpha_i \beta_i^2 x^2(t) & 1 + \beta_i^2 + \alpha_i \beta_i x(t) \end{bmatrix}, \end{aligned} \tag{9}$$

with the initial conditions $\hat{\rho}_i(0) = [\alpha_i \;\; \beta_i]^T$, $i = 1, \ldots, n$. Recursions (8) can be represented in the following compact form

$$\hat{\rho}(t+1) = [I + (\Delta(t) \otimes I_2) B(t)] \hat{\rho}(t), \tag{10}$$

where $\hat{\rho}(t) = [\hat{\rho}_1(t)^T \cdots \hat{\rho}_n(t)^T]^T$, $\Delta(t) = \operatorname{diag}\{\delta_1(t), \ldots, \delta_n(t)\}$,

$$B(t) = \Phi(t)(\Gamma \otimes I_2),$$

$$\Phi(t) = \operatorname{diag}\{\Phi_1(t), \ldots, \Phi_n(t)\},$$

$$\Gamma = \begin{bmatrix} -\sum_{j, j \neq 1} \gamma_{1j} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & -\sum_{j, j \neq 2} \gamma_{2j} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & -\sum_{j, j \neq n} \gamma_{nj} \end{bmatrix},$$

where $\gamma_{ij} = 0$ when $j \notin \mathcal{N}_i$; the initial condition is $\hat{\rho}(0) = [\hat{\rho}_1(0)^T \cdots \hat{\rho}_n(0)^T]^T$, in accordance with (8). The asymptotic value of $\hat{\rho}(t)$, which depends on the initial conditions and matrix $B(t)$, which is, in turn, a function of the signal, sensor and network parameters, should be such that the components with odd indices and the components with even indices become equal.

Notice that, in general, the choice of the weighting coefficients γ_{ij} plays an important role in achieving successful performance of the whole scheme. If the underlying idea of the whole methodology is to achieve good absolute calibration results ($\bar{g}$ as close as possible to one and $\bar{f}$ as close as possible to zero) by exploiting sensors with good characteristics in a large sensor network, there are two main possibilities: 1) to rely on the majority of good sensors, when all γ_{ij} in any neighborhood $\mathcal{N}_i$ can take the same value, or 2) to emphasize the effect of *a priori* selected good sensors j belonging to a set $\mathcal{N}^f \subset \mathcal{N}$ by: a) attaching to them relatively high values of γ_{ij} within all $\mathcal{N}_i$ for which i is an out-neighbor of j, b) multiplying all γ_{jk}, $k = 1,\ldots,n$, $k \neq j$, by a relatively small positive number (thus preventing big changes of $\hat{\rho}_j(t)$). Section 5 will be devoted to the situation in which a set of reliable sensors in the network is left with fixed characteristics.

3 Convergence Analysis

3.1 Noiseless case

In this paragraph we assume no communication errors and no measurement errors in order to emphasize structural properties of the proposed algorithm. For this scenario we assume:

A1) $\delta_i(t) = \delta = \text{const.}$, $i = 1,\ldots,n$.

For the sake of clearer presentation, we first adopt a simplifying assumption, which will be relaxed later:

A2) $\{x(t)\}$ is i.i.d., with $E\{x(t)\} = \bar{x} < \infty$ and $E\{x(t)^2\} = s^2 < \infty$.

Based on A1) and A2), we obtain the following recursion for the mean of the parameter estimates $\bar{\rho}(t) = E\{\rho(t)\}$

$$\bar{\rho}(t+1) = (I + \delta \bar{B})\bar{\rho}(t), \tag{11}$$

where $\bar{\rho}(0) = \rho(0)$, $\bar{B} = \bar{\Phi}(\Gamma \otimes I_2)$ and $\bar{\Phi} = E\{\Phi(t)\} = \text{diag}\{\bar{\Phi}_1 \ldots \bar{\Phi}_n\}$, with

$$\bar{\Phi}_i = \begin{bmatrix} \alpha_i\beta_i\bar{x} + \alpha_i^2 s^2 & \alpha_i\beta_i + \alpha_i^2\bar{x} \\ (1+\beta_i^2)\bar{x} + \alpha_i\beta_i s^2 & 1+\beta_i^2 + \alpha_i\beta_i\bar{x} \end{bmatrix}.$$

A closer insight into the recursion (11) shows that its properties cannot be analyzed using the known methodologies due to the block structure of $\bar{B}$ composed of specific 2×2 block matrices (see, *e.g.*, [17, 18] and many references therein). In order to cope with this problem, we shall first propose a novel methodology based on the concept of *diagonal dominance of matrices decomposed into blocks* [11, 14, 15]. We formulate several lemmas which represent basic prerequisites for the subsequent analysis.

Lemma 1 ([14, 16]**).** A matrix $A=[A_{ij}]$, where $A_{ij}\in\mathbb{C}^{m\times m}$, $i,j=1,\ldots,n$, has *quasi-dominating diagonal blocks* if the test matrix $W\in\mathbb{R}^{n\times n}$, with the elements

$$w_{ij}=1 \ (i=j); \quad w_{ij}=-\left\|A_{ii}^{-1}A_{ij}\right\| \ (i\neq j)$$

is an M-matrix; then, A is nonsingular. If $A-\lambda I$ has quasi-dominating diagonal blocks for all $\lambda\in\mathbb{C}_+$, then A is Hurwitz.

The following lemma has a very important role.

Lemma 2. If $A=[A_{ij}]$ has quasi-dominating diagonal blocks and A_{ii}, $i=1,\ldots,n$, are Hurwitz, A is also Hurwitz.

The proof is given in [11]; it is based on the basic ideas exposed in [14, 16].

The structure of the blocks of matrix $\bar{B}$ is defined by the structure of the digraph $\mathcal{G}$. We assume the following assumption, typical for consensus schemes:

A3) Graph $\mathcal{G}$ has a spanning tree.

Assumption A3) implies that matrix Γ has one eigenvalue at the origin and the other eigenvalues with negative real parts, *e.g.*, [18]. Consequently, matrix $\bar{B}$ from (11) has at least two eigenvalues at the origin. In the following, we characterize its remaining eigenvalues.

A4) $s^2-\bar{x}^2=\mathrm{var}\{x(t)\}>0$.

Assumption A4) essentially ensures sufficient excitation by the measured signal. Its important direct formal consequence is that $-\bar{\Phi}_i$ is Hurwitz, $i=1,\ldots,n$. Namely, it is straightforward to verify that $-\bar{\Phi}_i$ is Hurwitz iff

$$\alpha_i^2(s^2-\bar{x}^2)>0, \quad 2\alpha_i\beta_i\bar{x}+\alpha_i^2 s^2+1+\beta_i^2>0. \tag{12}$$

Both inequalities hold iff A4) holds.

The following lemma applies the above general results to matrix $\overline{B}$.

Lemma 3. Let Assumptions A3) and A4) be satisfied. Then, matrix $\overline{B}$ in (11) has two eigenvalues at the origin and the remaining eigenvalues have negative real parts.

The proof is given in the Appendix.

Define vectors

$$i_1 = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 \end{bmatrix}^T \in \mathbb{R}^{2n},$$

$$i_2 = \begin{bmatrix} 0 & 1 & 0 & 1 & \cdots & 0 & 1 \end{bmatrix}^T \in \mathbb{R}^{2n},$$

which represent the right eigenvectors of $\overline{B}$ corresponding to the zero eigenvalue, and let π_1 and π_2 be the corresponding normalized left eigenvectors, satisfying

$$\left[\begin{array}{c} \pi_1 \\ \hdashline \pi_2 \end{array}\right] \left[i_1 \mid i_2 \right] = I_2 .$$

The following lemma introduces a similarity transformation which will be used in the rest of the derivations.

Lemma 4 ([11]**).** Let $T = \left[i_1 \mid i_2 \mid T_{2n\times(2n-2)} \right]$, where $T_{2n\times(2n-2)}$ is an $2n \times (2n-2)$ matrix, such that span $\text{span}\left\{T_{2n\times(2n-2)}\right\} = \text{span}\left\{\overline{B}\right\}$. Then, T is nonsingular and

$$T^{-1}\overline{B}T = \left[\begin{array}{c|c} 0_{2\times 2} & 0_{2\times(2n-2)} \\ \hline 0_{(2n-2)\times 2} & \overline{B}^* \end{array}\right], \tag{13}$$

where $\overline{B}^*$ is Hurwitz.

Notice that

$$T^{-1} = \left[\begin{array}{c} \pi_1 \\ \hdashline \pi_2 \\ \hdashline S_{(2n-2)\times 2n} \end{array}\right], \tag{14}$$

where $S_{(2n-2)\times 2n}$ is defined in accordance with the definition of T.

We are now ready to prove the following theorem dealing with the asymptotic behavior of the mean of the parameter estimates generated by (11).

Theorem 1. Let Assumptions A1), A2), A3) and A4) be satisfied. Then there exists $\delta' > 0$ such that for all $\delta \le \delta'$ in (11)

$$\overline{\rho}_\infty = \lim_{t\to\infty} \overline{\rho}(t) = (i_1\pi_1 + i_2\pi_2)\overline{\rho}(0)\,,$$

implying $\overline{\rho}_\infty = [\overline{\rho}_{\infty 1}^T \cdots \overline{\rho}_{\infty n}^T]^T$ with $\overline{\rho}_{\infty 1} = \overline{\rho}_{\infty 2} = \cdots = \overline{\rho}_{\infty n}$.

Proof. Let $\tilde{\overline{\rho}}(t) = [\tilde{\overline{\rho}}_1(t)^T \cdots \tilde{\overline{\rho}}_n(t)^T]^T = T^{-1}\overline{\rho}(t)$. From (11) we obtain

$$\tilde{\overline{\rho}}(t+1)^{[1]} = \tilde{\overline{\rho}}(t)^{[1]}; \quad \tilde{\overline{\rho}}(t+1)^{[2]} = (I + \delta\overline{B}^*)\tilde{\overline{\rho}}(t)^{[2]}, \tag{15}$$

where $\dim\left\{\tilde{\overline{\rho}}(t)^{[1]}\right\} = 2$, $\dim\left\{\tilde{\overline{\rho}}(t)^{[2]}\right\} = 2n-2$. Using Lemma 3 we conclude that for δ small enough all the eigenvalues of $I + \delta\overline{B}^*$ lie within the unit circle. Therefore, $\lim\limits_{t\to\infty} = \left\|\tilde{\overline{\rho}}^{[2]}(t)\right\| = 0$, so that

$$\lim_{t\to\infty} \tilde{\overline{\rho}} = \tilde{\overline{\rho}}_\infty^T = \left[\tilde{\overline{\rho}}^{[1]T}(0) \quad 0 \quad \cdots \quad 0\right]^T.$$

Consequently,

$$\overline{\rho}_\infty = T\left[\tilde{\overline{\rho}}^{[1]T}(0) \quad 0 \quad \cdots \quad 0\right]^T = (i_1\pi_1 + i_2\pi_2)\overline{\rho}(0)\,. \tag{16}$$

Having in mind the definition of i_1 and i_2, we conclude that $\overline{\rho}_{\infty 1} = \overline{\rho}_{\infty 2} = \cdots = \overline{\rho}_{\infty n}$. ■

Now we focus on the original recursion (10). First we demonstrate the important fact, stemming from the structure of the matrices in (11), that the transformation T from Lemma 4, after being applied to the time-varying matrix $B(t)$, results in a matrix with the same structure as the transformed matrix $\overline{B}$ in (13).

Lemma 5 ([11])**.** Matrix $B(t)$ in (10) satisfies for all t

$$T^{-1}\overline{B}(t)T = \left[\begin{array}{c|c} 0_{2\times 2} & 0_{2\times(2n-2)} \\ \hline 0_{(2n-2)\times 2} & \overline{B}^*(t) \end{array}\right], \tag{17}$$

where $B(t)^*$ is an $(2n-2)\times(2n-2)$ matrix and T is given in Lemma 4.

Now we are in a position to prove the following theorem dealing with the convergence of the main recursion (1) in the mean square sense and w.p.1.

Theorem 2. Let Assumptions A1), A2), A3) and A4) be satisfied. Then there exists $\delta'' > 0$ such that for all $\delta \le \delta''$ in (10)

$$\lim_{t\to\infty} \hat{\rho}(t) = (i_1\pi_1 + i_2\pi_2)\hat{\rho}(0) \tag{18}$$

in the mean square sense and w.p.1.

Proof. Using Lemmas 4 and 5, we define $\tilde{\rho}(t) = T^{-1}\hat{\rho}(t)$ and obtain

$$\begin{aligned} \tilde{\rho}(t+1)^{[1]} &= \tilde{\rho}(t)^{[1]}, \\ \tilde{\rho}(t+1)^{[2]} &= (I + \delta B(t)^*)\tilde{\rho}(t)^{[2]}, \end{aligned} \tag{19}$$

where $\dim\left\{\tilde{\rho}(t)^{[1]}\right\} = 2$, $\dim\left\{\tilde{\rho}(t)^{[2]}\right\} = 2n-2$. Recalling that $\overline{B}^*$ is Hurwitz, we observe that there exists such a positive definite matrix R^* that

$$\overline{B}^{*T}R^* + R^*\overline{B}^* = -Q^*, \tag{20}$$

where Q^* is positive definite. Define $q(t) = E\left\{\tilde{\rho}(t)^{[2]T}R^*\tilde{\rho}(t)^{[2]}\right\}$, and let $\lambda_Q = \min_i \lambda_i\{Q^*\}$ and $k' = \max_i \lambda_i\left\{E\left\{B(t)^*B(t)^{*T}\right\}\right\}$. From (onetwo1) and A2) we obtain

$$q(t+1) = E\left\{\tilde{\rho}(t)^{[2]T}E\left\{(I+\delta B(t)^*)^T R^*(I+\delta B(t)^*)\right\}\tilde{\rho}(t)^{[2]}\right\} \tag{21}$$

and, further,

$$q(t+1) \le \left[1 - \delta\frac{\lambda_Q}{\max_i \lambda_i\{R^*\}} + \delta^2 k'\frac{\max_i \lambda_i\{R^*\}}{\min_i \lambda_i\{R^*\}}\right]q(t), \tag{22}$$

having in mind that $E\{B(t)^*\} = \overline{B}^*$. Consequently, there exists such a δ'' that for $\delta < \delta''$, $i = 1,\ldots,n$, the term in the brackets at the right hand side of (22) is less than one. Therefore, $q(t)$ tends to zero exponentially, implying that $\tilde{\rho}(t)^{[2]}$ converges to zero in the mean square sense and with probability one. The result follows from $\hat{\rho}(t) = T\tilde{\rho}(t)$, analogously with Theorem 1. ■

3.2 Communication errors

We assume that communication errors are manifested in two ways: 1) communication dropouts and 2) additive communication noise. Accordingly, we formally assume:

A5) The weights γ_{ij} in the algorithm (5) are stochastic processes represented as $\{\gamma_{ij}(t)\} = \{u_{ij}(t)\gamma_{ij}\}$, where $\{u_{ij}(t)\}$ are i.i.d. binary random sequences, such that $u_{ij}(t) = 1$ with probability p_{ij} ($p_{ij} > 0$ when $j \in \mathcal{N}_i$), and $u_{ij}(t) = 0$ with probability $1 - p_{ij}$.

A6) Instead of receiving $\hat{z}_j(t)$ from the j-th node, the i-th node receives $\hat{z}_j(t) + \xi_{ij}(t)$, where $\{\xi_{ij}(t)\}$ is an i.i.d. random sequence with $E\{\xi_{ij}(t)\} = 0$ and $E\{\xi_{ij}(t)^2\} = (\sigma_{ij}^{\xi})^2 < \infty$.

A7) Processes $x(t)$, $u_{ij}(t)$ and $\xi_{ij}(t)$ are mutually independent.

Denoting

$$\nu_i(t) = \sum_{j\in\mathcal{N}_i} \gamma_{ij}(t)\xi_{ij}(t)\begin{bmatrix} \alpha_i y_i(t) \\ 1+\beta_i y_i(t) \end{bmatrix},$$

and $\nu(t) = \begin{bmatrix} v_1(t) & \cdots & v_n(t) \end{bmatrix}$, we obtain from (10) that

$$\hat{\rho}(t+1) = [I + (\Delta(t)\otimes I_2)B'(t)]\hat{\rho}(t) + \Delta(t)\nu(t), \tag{23}$$

where $B'(t) = \Phi(t)(\Gamma(t)\otimes I_2)$, with $\Gamma(t)$ obtained from Γ by replacing constants γ_{ij} with time varying gains $\gamma_{ij}(t)$, according to A5).

We will study convergence of the recursion (23) starting from the results of the previous paragraph. Notice first that, due to mutual independence between the random variables in $B'(t)$, we have $E\{B'(t)\} = \bar{B}' = \bar{\Phi}(\bar{\Gamma}\otimes I_2)$, where $\bar{\Gamma} = E\{\Gamma(t)\}$ is obtained from Γ by replacing γ_{ij} with $\gamma_{ij}p_{ij}$. Defining $\tilde{B}'(t) = B'(t) - \bar{B}'$, we conclude that $E\{\tilde{B}'(t)\} = 0$ and $E\{\tilde{B}'(t)\,|\,\mathcal{F}_{t-1}\} = 0$. It is obvious that $\bar{B}' = \bar{\Phi}(\bar{\Gamma}\otimes I_2)$ has qualitatively the same properties as $\bar{B}$ in (11): it has two eigenvalues at the origin and the remaining eigenvalues in the left half plane.

Further, we now assume that the step sizes $\delta_i(t)$ satisfy the following assumption standard for stochastic approximation algorithms (*e.g.*, [19]):

A8) $\delta_i(t) = \delta(t) > 0$, $\sum_{t=0}^{\infty}\delta(t) = \infty$, $\sum_{t=0}^{\infty}\delta(t)^2 < \infty$, $i = 1,\ldots,n$.

Therefore, we have

$$\hat{\rho}(t+1) = (I + \delta(t)\bar{B}')\hat{\rho}(t) + \delta(t)\tilde{B}'(t)\hat{\rho}(t) + \delta(t)\nu(t). \tag{24}$$

Let

$$T' = \left[\, i_1 \;\vdots\; i_2 \;\vdots\; T'_{2n\times(2n-2)} \,\right],$$

where $T'_{2n\times(2n-2)}$ is an $2n\times(2n-2)$ matrix, such that $\text{span}\{T'_{2n\times(2n-2)}\} = \text{span}\{\bar{B}'\}$. Then

$$(T')^{-1} = \begin{bmatrix} \pi'_1 \\ \hdashline \pi'_2 \\ \hdashline S'_{(2n-2)\times 2n} \end{bmatrix},$$

where π'_1 and π'_2 are the left eigenvectors of $\bar{B}'$ corresponding to the zero eigenvalue.

Theorem 3. Let Assumptions A2)–A8) be satisfied. Then, $\hat{\rho}(t)$ generated by (24) converges to $i_1 w_1 + i_2 w_2$ in the mean square sense and w.p.1, where w_1 and w_2 are scalar random variables satisfying $E\{w_1\} = \pi_1' \hat{\rho}(0)$ and $E\{w_2\} = \pi_2' \hat{\rho}(0)$.

Proof. Let $\tilde{\rho}(t) = [\tilde{\rho}_1(t)\ \tilde{\rho}_2(t) \cdots \tilde{\rho}_{2n}(t)]^T = (T')^{-1} \hat{\rho}(t)$. Then, (24) gives

$$\tilde{\rho}(t+1)^{[1]} = \tilde{\rho}(t)^{[1]} + \delta(t) G_1(t) \tilde{\rho}(t) + \delta(t) \nu'(t), \tag{25}$$

$$\tilde{\rho}(t+1)^{[2]} = (I + \delta(t) \bar{B}'^{*}) \tilde{\rho}(t)^{[2]} + \delta(t) G_2(t) \tilde{\rho}(t) + \delta(t) \nu''(t), \tag{26}$$

where $\tilde{\rho}(t)^{[1]}$ and $\tilde{\rho}(t)^{[2]}$ are defined as in (15), $\left[\begin{array}{c} G_1(t) \\ \hline G_2(t) \end{array}\right] = (T')^{-1} \tilde{B}'(t) T'$ in such a way that $G_1(t)$ contains the first two rows, $\nu'(t) = \left[\begin{array}{c} \pi_1' \\ \hline \pi_2' \end{array}\right] \nu(t)$ and $\nu''(t) = S'_{(2n-2)\times 2n} \nu(t)$, while $\bar{B}'^{*}$ is a $(2n-2) \times (2n-2)$ Hurwitz matrix such that $(T')^{-1} \bar{B}' T' = \mathrm{diag}\{0_{2\times 2}, \bar{B}'^{*}\}$ (see Theorem 1). It is easy to verify that $E\{G_1(t)\} = 0$ and $E\{G_2(t)\} = 0$, as well as that $E\{G_1(t) \mid \mathcal{F}_{t-1}\} = 0$ and $E\{G_2(t) \mid \mathcal{F}_{t-1}\} = 0$.

Let $P^* > 0$ satisfy the Lyapunov equation $P^* \bar{B}'^{*} + \bar{B}'^{*T} P^* = -Q^*$ for some $Q^* > 0$. Denote $s(t) = E\left\{\left\|\tilde{\rho}(t)^{[1]}\right\|^2\right\}$ and $V(t) = E\left\{\tilde{\rho}(t)^{[2]T} P^* \tilde{\rho}(t)^{[2]}\right\}$. Then, directly following the methodology of [20] (Theorem 11), one obtains

$$\begin{aligned} s(t+1) &\le s(t) + C_1 \delta(t)^2 (1 + s(t) + V(t)), \\ V(t+1) &\le (1 - c_0 \delta(t)) V(t) + C_2 \delta(t)^2 (1 + s(t) + V(t)), \end{aligned} \tag{27}$$

where c_0, C_1 and C_2 are appropriately chosen positive constants. According to [20] (Lemma 12 and Theorem 11) and [21], inequalities (27) give rise to the conclusion that $\tilde{\rho}(t)^{[1]}$ tends to a vector random variable and $\tilde{\rho}(t)^{[2]}$ to zero in the mean square sense and w.p.1. Finally, the claim of the theorem follows after calculating $\lim_{t\to\infty} T' \left[\begin{array}{c} \lim_{t\to\infty} \tilde{\rho}(t)^{[1]} \\ \hline 0 \end{array}\right]$. The expressions for $E\{w_1\}$ and $E\{w_2\}$ follow from Theorems 1 and 2. This is an important property from the point of view of the global behavior of the proposed calibration scheme. ■

3.3 Measurement noise

We assume in this subsection that the signal $x(t)$ is contaminated by additive measurement noise. This is formally defined by the following assumption:

A9) Instead of $y_i(t)$ in (1), the sensors generate the signals contaminated by noise $y_i^{\eta}(t) = \alpha_i x(t) + \beta_i + \eta_i(t)$, where $\{\eta_i(t)\}$, $i = 1,\ldots,n$, are zero mean i.i.d. random sequences with $E\{\eta_i(t)^2\} = (\sigma_i^{\eta})^2$, independent of the measured signal $x(t)$.

Inserting $y_i^{\eta}(t)$ instead of $y_i(t)$ in the basic algorithm (5), we obtain, after changing the variables, the following "noisy" version of (8):

$$\begin{aligned}\hat{\rho}_i(t+1) = \hat{\rho}_i(t) + \delta_i(t)\sum_{j\in\mathcal{N}_i}\gamma_{ij}\Big\{&\big[\Phi_i(t)+\Psi_i(t)\big]\big[\hat{\rho}_j(t)-\hat{\rho}_i(t)\big]\\ &+N_{ij}(t)\hat{\rho}_j(t) - N_{ii}(t)\hat{\rho}_i(t)\Big\},\end{aligned} \tag{28}$$

where

$$\Psi_i(t) = \eta_i(t)\begin{bmatrix}\alpha_i x(t) & \alpha_i\\ \beta_i x(t) & \beta_i\end{bmatrix},$$

$$N_{ij}(t) = \frac{\eta_j(t)}{\alpha_j}\begin{bmatrix}\alpha_i y_i(t) & 0\\ \beta_i y_i x(t) & 0\end{bmatrix} + \begin{bmatrix}\dfrac{\eta_j(t)\eta_i(t)}{\alpha_j} & 0\\ 0 & 0\end{bmatrix},$$

$$N_{ii}(t) = \frac{\eta_i(t)}{\alpha_j}\begin{bmatrix}\alpha_i y_i(t) & 0\\ \beta_i y_i x(t) & 0\end{bmatrix} + \begin{bmatrix}\dfrac{\eta_i^2(t)}{\alpha_i} & 0\\ 0 & 0\end{bmatrix},$$

assuming $\alpha_i \neq 0$, $i = 1,...,n.$ Notice that $E\{\Psi_i(t)\} = 0$, $E\{N_{ij}(t)\} = 0$, but $E\{N_{ii}(t)\} = \begin{bmatrix}\dfrac{(\sigma_i^{\eta})^2}{\alpha_i} & 0\\ 0 & 0\end{bmatrix}$.

Assuming $\delta_i(t) = \delta(t)$, $i = 1,\ldots,n,$ we can write in accordance with (10)

$$\hat{\rho}(t+1) = \Big(I + \delta(t)\Big\{[\Phi(t)+\Psi(t)](\Gamma\otimes I_2) + \tilde{N}(t)\Big\}\Big)\hat{\rho}(t), \tag{29}$$

where $\Psi(t) = \mathrm{diag}\{\Psi_1(t),\ldots,\Psi_n(t)\}$ and $\tilde{N}(t) = [\tilde{N}_{ij}(t)]$ with

$$\tilde{N}_{ij}(t) = -\sum_{k,k\neq i} \gamma_{ik} N_{ii}(t)$$

for $i = j$ and $\tilde{N}_{ij}(t) = \gamma_{ij} N_{ij}(t)$ for $i \neq j$, $i, j = 1,\ldots,n$.

Applying the methodology from the previous section to the analysis of (11), we conclude that, instead of (29), we have now $\overline{\rho}(t+1) = [I + \delta(t)(\overline{B} + \Sigma_\eta)]\overline{\rho}(t)$, where $\overline{B}$ is defined in (11) and $\sum_\eta - \text{diag}\left\{\frac{(\sigma_1^\eta)^2}{\alpha_1}\sum_j \gamma_{1j}, 0, \ldots, \frac{(\sigma_n^\eta)^2}{\alpha_n}\sum_j \gamma_{nj}, 0\right\}$. Under A2) and A3) the last recursion does not have the properties of (11), due to the additional term Σ_η; the fact that all the row sums of $\overline{B} + \Sigma_\eta$ are now not equal to zero prevents the achievement of asymptotic consensus (see Theorem 1).

The above problem can be overcame in the case when the measurement noise variances $(\sigma_i^\eta)^2$ are *a priori* known. Consequently, the following algorithm, able to achieve asymptotic consensus, is proposed as a modification of:

$$\hat{\theta}(t+1) = \hat{\theta}_i(t) + \delta(t)\left\{\sum_{j\in\mathcal{N}_i} \gamma_{ij}\varepsilon_{ij}^\eta(t)\begin{bmatrix} y_i^\eta(t) \\ 1 \end{bmatrix} + \left[\begin{array}{c|c} (\sigma_i^\eta)^2 \sum\limits_{j\in\mathcal{N}_i} \gamma_{ij} & 0 \\ \hline 0 & 0 \end{array}\right]\hat{\theta}_i(t)\right\}, \quad (30)$$

where $\varepsilon_{ij}^\eta(t) = \hat{z}_j^\eta(t) - \hat{z}_i^\eta(t)$ and $\hat{z}_i^\eta(t) = \hat{a}_i(t) y_i^\eta(t) + \hat{b}_i(t)$, $i = 1,\ldots,n$.

Theorem 4. Let Assumptions A2)–A4), A8) and A9) be satisfied. Then, $\hat{\rho}(t)$ generated by (30) converges to $i_1 w_1 + i_2 w_2$ in the mean square sense and w.p.1, where w_1 and w_2 are scalar random variables satisfying $E\{w_1\} = \pi_1' \hat{\rho}(0)$ and $E\{w_2\} = \pi_2' \hat{\rho}(0)$.

Proof. The algorithm (30) can be represented in the following compact form (compare with (29)):

$$\hat{\rho}(t+1) = \{I + \delta(t)[B(t) - \Sigma_\eta + \Psi(t)(\Gamma \otimes I_2) + \tilde{N}(t)]\}\hat{\rho}(t). \quad (31)$$

Since $B(t) = \overline{B} + \tilde{B}(t)$, where $\overline{B}$ is defined in (11) and $\tilde{B}(t)$ satisfies $E\{\tilde{B}(t)\} = 0$ and is independent from $\tilde{B}(\tau)$, $\tau < t$, we have

$$\hat{\rho}(t+1) = (I + \delta(t)\overline{B})\hat{\rho}(t) + \delta(t)E(t)\hat{\rho}(t), \quad (32)$$

where $E(t) = \tilde{B}(t) - \Sigma_\eta + \Psi(t)(\Gamma \otimes I_2) + \tilde{N}(t)$ is a zero mean white noise term. We observe that (32) represents structurally a special case of (24), not containing the last, stochastic, term. Therefore, the proof can be completed by applying the methodology of the proof of Theorem 4. ■

3.4 Convergence rate

A closer view on the asymptotic formulae in the above theorems shows that the asymptotic convergence rate of all the analyzed algorithms follows general statements related to stochastic approximation algorithms. Focusing the attention on the basic aspects of *convergence to consensus* determined by the behavior of $\tilde{\rho}(t)^{[2]}$ in the context of all the above theorems, we have the following result giving an upper bound of the mean square error:

Theorem 5. Under the assumptions of Theorem 4, together with A2') $\lim_{t\to\infty}\left(\delta^{-1}(t+1)-\delta^{-1}(t)\right)=d\geq 0$, there exists such a positive number $\sigma'<1$ that for all $0<\sigma<\sigma'$ the asymptotic consensus is achieved by the analyzed algorithms with the convergence rate

$$E\left\{\left\|\tilde{\rho}(t)^{[2]}\right\|^2\right\}=o(\delta(t)^{\sigma}). \tag{33}$$

Proof. Select the pair of inequalities related to $s(t)$ and $V(t)$ in (27). Applying Lemma 12 from [20], one obtains

$$V(t+1)\leq(1-c_0\delta(t))V(t)+C_1\delta(t)^2, \tag{34}$$

where c_0 and C_1 are generic constants.

From A2') and A2'') we have by Taylor expansion

$$\frac{\delta(t)^{\sigma}}{\delta(t+1)^{\sigma}}=1+\sigma\frac{\delta(t)-\delta(t+1)}{\delta(t+1)}+O((\frac{\delta(t)-\delta(t+1)}{\delta(t+1}})^2), \tag{35}$$

having in mind that $\dfrac{\delta(t)-\delta(t+1)}{\delta(t+1)}\to 0$. Consequently, from (34) we have

$$\frac{V(t+1)}{\delta(t+1)^{\sigma}}\leq\frac{\delta(t)^{\sigma}}{\delta(t+1)^{\sigma}}[1+c_0\delta(t)]\frac{V(t)}{\delta(t)^{\sigma}}+C_1\delta(t)^{2-\sigma}, \tag{36}$$

so that

$$\begin{aligned}W(t+1)\leq\Bigg\{1+\Bigg[&-c_0+\sigma\frac{\delta(t)-\delta(t+1)}{\delta(t+1)\delta(t)}+\sigma\frac{\delta(t)-\delta(t+1)}{\delta(t+1)}\\&+O\left(\left(\frac{\delta(t)-\delta(t+1)}{\delta(t+1}\right)^2\right)\Bigg]\delta(t)\Bigg\}W(t)+C_1\delta(t)^{2-\sigma},\end{aligned} \tag{37}$$

where $W(t)=\dfrac{V(t)}{\delta(t)^{\sigma}}$. Taking into account A2'), for t large enough one has

$$W(t+1) \le [1 + (-c_0 + \sigma d + \varepsilon)\delta(t)]W(t) + C_1\delta(t)^{2-\sigma}. \tag{38}$$

Having in mind that $c_0 > 0$, it is evident that there exists $\sigma' > 0$ such that $c_1 = -c_0 + \sigma d + \varepsilon$ is negative for all $0 < \sigma \le \sigma'$, since ε can be made arbitrarily small. Therefore, according to the well known results (*e.g.*, [19, 22]), $\lim_{t\to\infty} W(t) = 0,$ and the assertion immediately follows. ■

In practice it is difficult to make an *a priori* estimate of the value of σ'. It is, however, clear that it depends directly on $\lambda_{\min}(Q^*)$, which, in turn, depends on the sensor and network properties expressed by matrix $B(t)$. It is possible to choose $Q^* > 0$ and obtain uniquely $P^* > 0$ as a consequence of the fact that the nonzero eigenvalues of $\overline{B}$ (or $\overline{B}(d)$) are in the left half plane. Without going into details of the relationship between the eigenvalues of Q^*, P^* and $\overline{B}^*$ in general, we can notice here that in the case of undirected graphs and symmetric matrices $\overline{B}$, c_0 is proportional to the largest eigenvalue of $\overline{B}$. This fact leads to the general qualitative conclusion that the convergence rate of the proposed algorithms depends on the properties of the underlying graph, including its connectivity (see [23]). The number of nodes increases dimensionality of the parameter estimates and potentially decreases convergence rate in the same way as in stochastic approximation schemes; however, convergence rate increases by increasing connectedness of the graph. In the case when the graph is fully connected, convergence rate is high at the expense of a large number of direct communication links; a compromise should be found, as in all analogous problems in wireless sensor networks.

In general, according to [24], the choice of $\delta(t)$ should be based on the following qualitative estimate

$$E\left\{\left\|\tilde{\rho}(t)^{[2]}\right\|^2\right\} \le v_1(\delta(t)) + v_2(\delta(t))\exp\left\{-k_1\sum_{\tau}\delta(t-\tau)\right\}, \tag{39}$$

where $v_1(\delta(t)) \to 0$ following $\delta(t)$, $v_2(\cdot)$ is bounded and $k_1 > 0$. The first term in (39) depends on the noise and the second on the initial conditions. Obviously, the choice of $\delta(t)$ should be based on a compromise between these two terms. Assuming that $\delta(t)$ has the standard form $\delta(t) = m_1/(m_2 + t^{\mu}),\quad m_1, m_2 > 0,$ $1/2 < \mu \le 1$, the values of μ closer to $\frac{1}{2}$ would give faster convergence at the expense of noise immunity; the values of μ closer to one provide the opposite effect. In the noiseless case, the choice $\delta(t) = \delta = \text{const}.$ provides an exponential convergence rate, since $v_1(\cdot) = 0$.

3.5 Macro calibration for networks with a reference

The choice of the elements of the matrix Γ, dictated by the relative importance (precision) of sensors in a given network, plays an important role in achieving good performance of the proposed method in practice. It is possible to increase importance of the sensors from a given set $\mathcal{N}^f$ by multiplying all γ_{ij}, $i \in \mathcal{N}^f$, $j = 1,...,n$, $j \neq i$, with a small positive number. In such a way, the corresponding $\hat{\rho}_i(t)$ remains closer to its initial condition $\hat{\rho}_i(0)$, $i \in \mathcal{N}^f$, thus having more influence on the global point of convergence. In the limit, sensors from $\mathcal{N}^f$ can be left unchanged (with fixed characteristics), so that the recursions (5) are applied only to the nodes $i \in \mathcal{N} - \mathcal{N}^f$.

Take as an example the case when one of the sensors, say *k*-th, $k \in \{1,\ldots,n\}$, is taken as a reference because it has ideal (or desirable) characteristics; then, the whole calibration network can be "pinned" to that sensor. The proposed algorithm (5) can be simply applied by setting

$$\hat{\theta}_k(t+1) = \hat{\theta}_k(t), \tag{40}$$

where the initial condition $\hat{\theta}_k(0) = \hat{\theta}_{k0}$ should be appropriately chosen since the whole calibration algorithm should ensure convergence of $\hat{\rho}_i(t)$, $i = 1,\ldots,n$, $i \neq k$, to the same vector $\hat{\rho}_{k0} = \begin{bmatrix} \alpha_k & 0 \\ \beta_k & 1 \end{bmatrix} \hat{\theta}_{k0}$ (in the ideal case $\hat{\rho}_{k0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$). The corresponding modification of the general form of the algorithm consists of setting to zero all the block matrices in the *k*-th block row of $B(t)$ in (10), *e.g.*, by setting $\gamma_{kj} = 0$, $j = 1,\ldots,n$. An insight into the assumptions A1)–A9) shows that the only new aspect resulting from the application of (40) is a change in the structure of the graph $\mathcal{G}$, since the arcs leading to the node k are eliminated by (40). However, A3) still holds, since the new graph still has a spanning tree with the node k as the center node. Therefore, $\hat{\rho}_i(t)$, $i = 1,\ldots,n$, converge in the sense of all the above proven theorems to the same limit. Since $\lim_{t\to\infty} \hat{\rho}_k(t) = \hat{\rho}_{k0}$ by definition, this limit is equal to $\hat{\rho}_{k0}$.

4 Simulation Results

In order to illustrate properties of the proposed algorithms, a sensor network with ten nodes has been simulated. A fixed randomly selected communication graph satisfying A3) has been adopted, and parameters α_i and β_i have been randomly selected around one and zero, respectively, with variance 0.3.

In Fig. 1 the equivalent gains $\hat{g}_i(t)$ and offsets $\hat{f}_i(t)$ generated by the proposed algorithm (5) are presented for a preselected gain $\delta = 0.01$ in the noiseless case. It is clear that calibration is achieved, and that the asymptotic values are equal; in this case they are close to the optimal values. Fig. 2 depicts the situation when the first node is assumed to be a reference node with $\alpha_1 = 1$ and $\beta_1 = 0$. Convergence to the reference value is obvious (see Section 5).

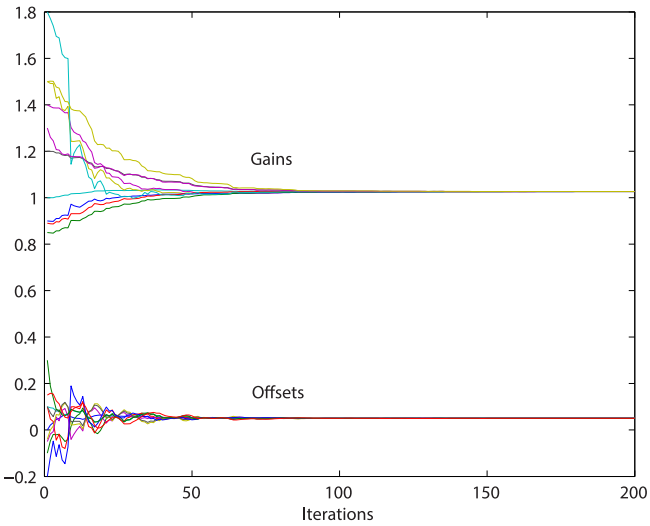

Fig. 1 – *Noiseless case: offset and gain estimates, no reference.*

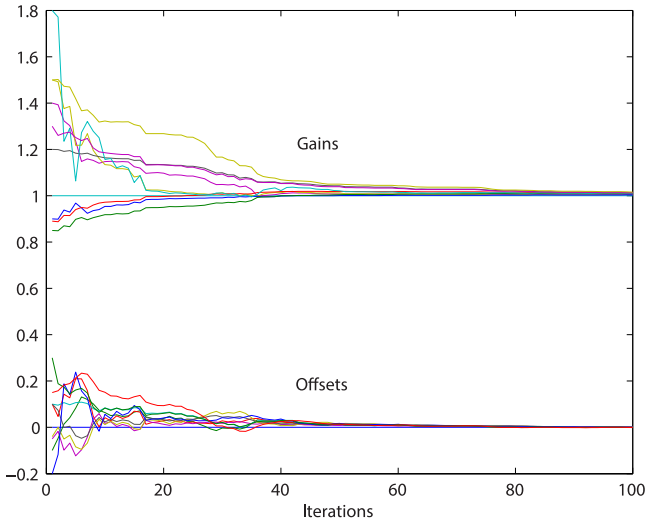

Fig. 2 – *Noiseless case: offset and gain estimates, reference included.*

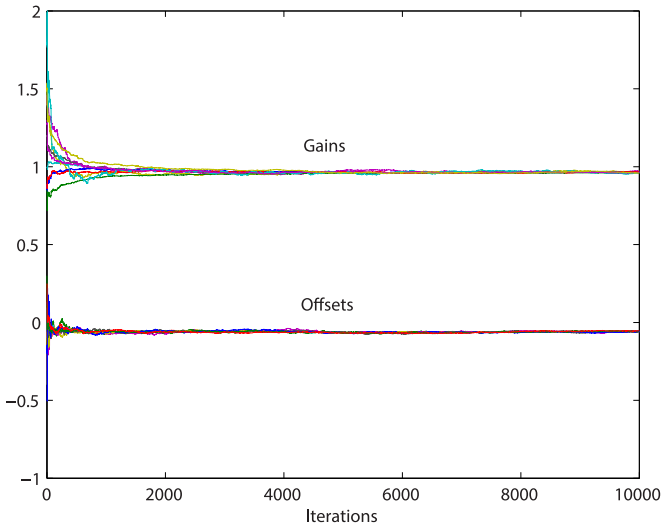

Fig. 3 – *Lossy sensor network: offset and gain estimates.*

In Fig. 3 the equivalent gains $\hat{g}_i(t)$ and offsets $\hat{f}_i(t)$ generated by the proposed algorithm (30) are presented for the sequence $\delta(t) = 0.01/t^{0.6}$. All the discussed uncertainties are included: communication outages with $p = 0.2$, communication additive noise with variance 0.1, and measurement noise with variances randomly chosen in the range $(0, 0.1)$; the signal $x(t)$ is a correlated random sequence with zero mean and variance 1. It is clear that calibration is achieved in spite of the noise existence.

5 Conclusion

In this paper a new *distributed blind calibration algorithm* of gradient type resulting in extended consensus has been proposed for sensor networks. The algorithm provides a new efficient tool for coping with the problem of calibration of large wireless sensor networks with communications limited to close neighbors, without requiring any fusion center. It is proved, after developing a novel methodology of treating *consensus schemes* on the basis of the diagonal dominance of matrices decomposed into blocks, that the algorithm achieves asymptotic agreement on all sensor gains and offsets in a given network. Convergence to consensus in the mean square sense and with probability one is proved separately for the cases of noiseless and noisy environments. In the case of imperfect inter-node communications and measurement noise, a modification of the basic gradient algorithm is proposed. Special attention is paid to the convergence rate of the proposed algorithms. The problem of distributed macro

calibration of sensor networks when a node is kept with fixed characteristics is also discussed. Some simulation results illustrate the behavior of the proposed algorithms.

6 Appendix: Proof of Lemma 3

Proof. Let $W^d = [w_{ij}^d]$, where $w_{ij}^d = 0$ for $i = j$, and $w_{ij}^d = (\sum_{j=1,\, j\neq i}^{n} \gamma_{ij})^{-1}\gamma_{ij}$ for $i \neq j$. This matrix is row stochastic and cogredient (amenable by permutation transformations) to

$$W_c^d = \begin{bmatrix} W_1^d & 0 \\ W_2^d & W_0^d \end{bmatrix}, \tag{41}$$

where $W_1^d \in \mathbb{R}^{n_1 \times n_1}$ is an irreducible matrix, $W_2^d \in \mathbb{R}^{n_2 \times n_1} \neq 0$ and $W_0^d \in \mathbb{R}^{n_2 \times n_2}$ is such that $\max_i |\lambda_i\{W_0^d\}| < 1$. Select one node of the graph $\mathcal{G}$ from the set of center nodes and delete the corresponding row and column from Γ; the resulting matrix is cogredient to $W_c^{d-} = \begin{bmatrix} W_1^{d-} & 0 \\ W_2^{d-} & W_0^d \end{bmatrix}$, where $W_1^{d-} \in \mathbb{R}^{(n_1-1)\times(n_1-1)}$ and $W_2^{d-} \in \mathbb{R}^{n_2 \times (n_1-1)}$. As W_1^d corresponds to a closed strong component of the inverse graph of $\mathcal{G}$, deleting one node from it (together with the corresponding edges) results in a graph containing, in general, $\kappa \geq 1$ closed strong components; each of the nonnegative matrices to which these strong components are associated have at least one row with the sum of all the elements strictly less than one. Consequently, $I - W_c^{d-}$ is an M-matrix [15]. Deleting one row and column from Γ means deleting two consecutive rows and columns from $\bar{B}$; let $\bar{B}^- \in \mathbb{R}^{(2n-2)\times(2n-2)}$ be the resulting matrix. According to Lemmas 1 and 2 and assumption A4) which guarantees that $\bar{\Phi}_i$, $i = 1,\ldots,n$ is Hurwitz, $\bar{B}^-$ has quasi-dominating diagonal blocks, which directly implies that $\bar{B}^-$ has all the eigenvalues with negative real parts. Thus the result.

Acknowledgement

This work was supported by the EU Marie Curie CIG, Knut and Alice Wallenberg Foundation and the Swedish Research Council.

8 References

[1] M.S. Stankovic, S.S. Stankovic, K.H. Johansson: Distributed Calibration for Sensor Networks under Communication Errors and Measurement Noise, IEEE Conference on Decision and Control, Maui, HI, USA, 10-13 Dec. 2012, pp. 1380 – 1385.

[2] M.S. Stankovic, S.S. Stankovic, K.H. Johansson: Distributed Macro Calibration in Sensor Networks, 20th Mediterranean Conference on Control and Automation, Barcelona, Spain, 03-06 July 2012, pp. 1049 – 1054.

[3] Special Issue on Sensor Networks and Applications, Proceedings of the IEEE, Vol. 91, No. 8, Aug. 2003, pp. 1151 – 1259.

[4] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci: Wireless Sensor Networks: A Survey, Computer Networks, Vol. 38, No. 4, March 2002, pp. 393 – 422.

[5] A. Speranzon, C. Fischione, K.H. Johansson: Distributed and Collaborative Estimation over Wireless Sensor Networks, IEEE Conference on Decision and Control, San Diego, CA, USA, 13-15 Dec. 2006, pp. 1025 – 1030.

[6] K. Whitehouse, D. Culler: Calibration as Parameter Estimation in Sensor Networks, 1st ACM International Workshop on Wireless Sensor Networks and Applications, Atlanta, GA, USA, 28 Sept. 2002, pp. 59 – 67.

[7] K. Whitehouse, D. Culler: Macro-calibration in Sensor/Actuator Networks, Mobile Networks and Application, Vol. 8, No. 4, Aug. 2003, pp. 463 – 472.

[8] L. Balzano, R. Nowak: Blind Calibration, Technical Report TR-UCLA-NESL-200702-01, Networked and Embedded Systems Laboratory, University of California, Los Angeles, CA, USA, Feb. 2007.

[9] L. Balzano, R. Nowak: Blind Calibration in Sensor Networks, International Conference Information Processing in Sensor Networks, Cambridge , MA, USA, 25-27 April 2007, pp. 79 – 88.

[10] V. Bychkovskiy, S. Megerian, D. Estrin, M. Potkonjak: A Collaborative Approach to In-place Sensor Calibration, 2nd International Workshop on Information Processing in Sensor Networks, Palo Alto, CA, USA, 22-23 April 2003, pp. 301 – 316.

[11] M.S. Stankovic, S.S. Stankovic, K. H. Johansson: Distributed Blind Calibration in Lossy Sensor Networks via Output Synchronization, IEEE Transaction on Automatic Control, Vol. 60, No. 12, Dec. 2015, pp. 3257 – 3262.

[12] 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, Jan. 2010, pp. 176 – 193.

[13] E. Miluzzo, N.D. Lane, A.T. Campbell, R. Olfati-Saber: Calibree: A self-calibration System for Mobile Sensor Networks, 4th IEEE International Conference on Distributed Computing in Sensor Systems, Santorini Island, Greece, 11-14 June 2008, pp. 314 – 331.

[14] Y. Ohta, D.D. Siljak: Overlapping Block Diagonal Dominance and Existence of Lyapunov Functions, Journal of Mathematical Analysis and Application, Vol. 112, No. 2, Dec. 1985, pp. 396 – 410.

[15] D.D. Siljak: Decentralized Control of Complex Systems, Academic Press, New York, NY, USA, 1991.

[16] I.F. Pearce: Matrices with Dominating Diagonal Blocks, Journal of Economic Theory, Vol. 9, No. 2, Oct. 1974, pp. 159 – 170.

[17] R. Olfati-Saber, J.A. Fax, R. Murray: Consensus and Cooperation in Networked Multi-agent Systems, Proceedings of the IEEE, Vol. 95, No. 1, Jan. 2007, pp. 215 – 233.

[18] W. Ren, R.W. Beard: Consensus Seeking in Multi-agent Systems under Dynamically Changing Interaction Topologies, IEEE Transaction on Automatic Control, Vol. 50, No. 5, May 2005, pp. 655 – 661.

[19] H.F. Chen: Stochastic Approximation and its Applications, Kluwer Academic Dordrecht, Netherlands, 2002.

[20] M. Huang, J.H. Manton: Stochastic Consensus Seeking with Noisy and Directed Inter-agent Communications: Fixed and Randomly Varying Topologies, IEEE Transaction on Automatic Control, Vol. 55, No. 1, Jan. 2010, pp. 235 – 241

[21] T. Li, J.F. Zhang: Consensus Conditions of Multi-agent Systems with Time Varying Topologies and Stochastic Communication Noises, IEEE Transaction on Automatic Control, Vol. 55, No. 9, Sept. 2010, pp. 2043 – 2057.

[22] B.T. Polyak: Convergence and Rate of Convergence of Iterative Stochastic Algorithms - II – The Linear Case, Automation and Remote Control, No. 4, 1977, pp. 101 – 107. (In Russian).

[23] C. Godsil, G. Royle: Algebraic Graph Theory, Springer Verlag, New York, NY, USA, 2001.

[24] V.S. Borkar, S.P. Meyn: The O.D.E. Method for Convergence of Stochastic Approximation and Reinforcement Learning, SIAM Journal on Control and Optimization, Vol. 38, No. 2, 2000, pp. 447 – 469.