Distributed Calibration for Sensor Networks under Communication Errors and Measurement Noise

Miloš S. Stanković, Srdjan S. Stanković and Karl Henrik Johansson

Abstract— In this paper a new distributed calibration algorithm based on consensus is proposed for sensor networks. The algorithm is basically formulated as a set of stochastic gradient type recursions for estimating parameters of local sensor calibration functions, starting from local criteria defined as weighted sums of mean square errors between the outputs of neighboring sensors. It is proved that the proposed algorithm provides asymptotic consensus in the space of the sensor gains and offsets. In the case of communication dropouts and additive communication and measurement noise, a modification of the instrumental variable type of the original calibration scheme is proposed. It is proved using stochastic approximation arguments that in this case the proposed algorithm achieves asymptotic consensus in the mean square sense and with probability one. Some illustrative simulation examples are provided.

I. INTRODUCTION

Recently, wireless sensor networks (WSN) have emerged as an important research area (see, *e.g.*, [1], [2], [3]). Calibration represents one of the most important challenges, having in mind great number of sensors typical for WSN's today. 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 prerequisite is to frame calibration as a parameter estimation problem [4], [5]. Automatic methods for jointly calibrating WSN's, without dependence on controlled stimuli or high-fidelity groundtruth data, is of significant interest. This problem is referred to as *blind calibration* [6], [7]. One approach to blind WSN calibration is to assume that the deployment is very dense, so that neighboring nodes should have nearly identical readings. There are also methods trying to cope with situations in which sensor network deployments may not meet the density requirements [8], [9].

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. Assuming that the sensors form a network based on communications between neighboring nodes, it will be shown that the the convergence of the algorithm can be treated as a nontrivial *consensus* problem, to which the classical

M. S. Stanković and K. H. Johansson are with the ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, 100-44 Stockholm, Sweden; E-mail: milsta@kth.se, kallej@kth.se

S. S. Stanković is with Faculty of Electrical Engineering, University of Belgrade, Belgrade, Serbia. E-mail: stankovic@etf.rs

results (*e.g.*, [10]) are not directly applicable. Note also that, to the authors' knowledge, consensus has been applied to the calibration problems only in [11], [12], but within different contexts. Using basic arguments derived from stability of diagonally dominant dynamic systems [13], [14], [15], we prove that the proposed algorithm provides asymptotic consensus in the mean square sense and with probability one (see also [16] which deals with the noiseless case). The basic results are extended in this paper to the general case which includes: 1) measurement noise, 2) communication dropouts, and 3) additive communication noise. An algorithm of *instrumental variable* type [17] is constructed for solving the problem in the most general case, and the achievement of the asymptotic consensus in the mean square sense and with probability one (w.p. 1) is proved for both gains and offsets. Simulation results illustrate the proposed algorithm.

II. PROBLEM FORMULATION AND THE BASIC ALGORITHM

Consider n distributed sensors measuring the same discrete-time signal $x(t)$, $t = \ldots, -1, 0, 1, \ldots$, which is supposed to be a realization of a random process $\{x(t)\}$. Assume that the i-th sensor generates at its output the signal

$$y_i(t) = \alpha_i x(t) + \beta_i \qquad (1)$$

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

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

$$z_i(t) = a_i y_i(t) + b_i = a_i \alpha_i x(t) + a_i \beta_i + b_i = g_i x(t) + f_i, \quad (2)$$

where the calibration parameters a_i and b_i have to be chosen in such a way as to set the equivalent gain $g_i = a_i \alpha_i$ as close as possible to one and the equivalent offset $f_i = a_i \beta_i + b_i$ 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* $\mathcal{G} = (\mathcal{U}, \mathcal{V})$, where $\mathcal{U}$ is the set of nodes (one node corresponds to one sensor) and $\mathcal{V}$ the set of arcs. Define the adjacency matrix $C = [c_{ij}]$, $i, j = 1, \ldots, n$, such that $c_{ij} = 1$ if the j-th sensor can send its message to the i-th sensor, and $c_{ij} = 0$ otherwise.

The aim of this paper is to design an algorithm for distributed real-time estimation of the calibration parameters a_i and b_i, $i = 1, \ldots, n$, which provides asymptotically equal outputs of all the sensors in the case when no reference signal or ideal sensor is given or identified, expecting, loosely speaking, that the majority of well calibrated sensors will correct the behavior of the remaining ones. In the case of

a given reference, all the sensors should be asymptotically calibrated.

The distributed calibration algorithm is derived starting from minimization of the set of *local criteria*

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

$i = 1, \ldots, n$, where $\mathcal{N}_i$ is the set of neighboring nodes of the i-th node (the sensors sending information to the i-th sensor), and γ_{ij} are nonnegative scalar weights. Starting from (3), the following recursions of *gradient type* can be derived:

$$\hat{\theta}_i(t + 1) = \hat{\theta}_i(t) + \delta_i(t) \sum_{j \in \mathcal{N}_i} \gamma_{ij} \epsilon_{ij}(t) \begin{bmatrix} y_i(t) \\ 1 \end{bmatrix}, \quad (4)$$

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, $\epsilon_{ij}(t) = \hat{z}_j(t) - \hat{z}_i(t)$ 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 subsumes reception of the current outputs of the neighboring nodes, as well as availability of the local measurement.

The underlying idea of (4) 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 $\bar{g}$ and $\bar{f}$, implying $\hat{z}_j(t) = \hat{z}_i(t)$, $i, j = 1, \ldots, n$. We introduce

$$\hat{\rho}_i(t) = \begin{bmatrix} \hat{g}_i(t) \\ \hat{f}_i(t) \end{bmatrix} = \begin{bmatrix} \alpha_i & 0 \\ \beta_i & 1 \end{bmatrix} \hat{\theta}_i(t), \qquad (5)$$

and

$$\epsilon_{ij}(t)(t) = \begin{bmatrix} x(t) & 1 \end{bmatrix} (\hat{\rho}_j(t) - \hat{\rho}_i(t)), \qquad (6)$$

so that (4) becomes

$$\hat{\rho}_i(t + 1) = \hat{\rho}_i(t) + \delta_i(t) \sum_{j \in \mathcal{N}_i} \gamma_{ij} \Phi_i(t)(\hat{\rho}_j(t) - \hat{\rho}_i(t)), \quad (7)$$

where $\Phi_i(t) = \begin{bmatrix} \alpha_i y_i(t) x(t) & \alpha_i y_i(t) \\ [1 + \beta_i y_i(t)] x(t) & 1 + \beta_i y_i(t) \end{bmatrix}$, with the initial conditions $\hat{\rho}_i(0) = [\alpha_i \ \beta_i]^T$, $i = 1, \ldots, n$. Recursions (7) can be represented in the following compact form

$$\hat{\rho}(t + 1) = [I + (\Delta(t) \otimes I_2) B(t)] \hat{\rho}(t), \qquad (8)$$

where $\hat{\rho}(t) = [\hat{\rho}_1(t)^T \cdots \hat{\rho}_n(t)^T]^T$, $\Delta(t) = \operatorname{diag}\{\delta_1(t), \cdots, \delta_n(t)\}$, $B(t) = \Phi(t)(\Gamma \otimes I_2)$, $\Phi(t) = \operatorname{diag}\{\Phi_1(t), \ldots, \Phi_n(t)\}$, $\otimes$ denotes the Kronecker's product, I_2 is the 2×2 unit matrix and

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

where $\gamma_{ij} = 0$ when $j \in \mathcal{N} - \mathcal{N}_i$; the initial condition is $\hat{\rho}(0) = [\hat{\rho}_1(0)^T \cdots \hat{\rho}_n(0)^T]^T$, in accordance with (7). The

desirable asymptotic value of $\hat{\rho}(t)$, which depends on the initial conditions and the matrix $B(t)$ which is, in turn, a function of the signal and the sensor parameters, should be such that the components with odd indices and the components with even indices are the same.

III. CONVERGENCE ANALYSIS - NOISELESS CASE

We assume that:
A1) $\delta_i(t) = \delta = \text{const}$, $i = 1, \ldots, n$;
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

$$\bar{\rho}(t + 1) = (I + \delta \bar{B}) \bar{\rho}(t), \qquad (9)$$

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

Asymptotic properties of (9) cannot be analyzed by applying the well known results related to the classical consensus schemes (*e.g.*, [10]) due to the specific structure of $\bar{B}$ which is composed of 2×2 block matrices. Therefore, we start our analysis with several basic lemmas derived using the results related to the *diagonal dominance of matrices decomposed into blocks* [13], [14] (for more details see [16]).

Lemma 1. [13], [15] A matrix $A = [A_{ij}]$, where $A_{ij} \in \mathcal{C}^{m \times m}$, $i, j = 1, \ldots n$, has quasi-dominating diagonal blocks and is nonsingular if the test matrix $W \in R^{n \times n}$, with the elements

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

in an M-matrix ($\|.\|$ denotes an operator norm). If $A - \lambda I$ has quasi-dominating diagonal blocks for all $\lambda \in \mathcal{C}_+$, then A is Hurwitz ($\mathcal{C}_+$ denotes the set of complex numbers with nonnegative real parts). ∎

Lemma 2. [16], [13], [15] If A has quasi-dominating diagonal blocks and A_{ii}, $i = 1, \ldots, n$, are Hurwitz, A is also Hurwitz. ∎

Coming back to the matrix $\bar{B}$ in (9), we assume:
A3) the graph $\mathcal{G}$ has a spanning tree.

This assumption implies that Γ has one eigenvalue at the origin and the other eigenvalues with negative real parts [10]. Also, if the i-th node is a center node of $\mathcal{G}$, the matrix $\Gamma' \in R^{(n-1) \times (n-1)}$, obtained from Γ by deleting its i-th row and its i-th column, is nonsingular [18], [16].

Lemma 3. [16] Let assumption A3) be satisfied and let
A4) $-\bar{\Phi}_i$ is Hurwitz, $i = 1, \ldots, n$.

Then, matrix $\bar{B}$ in (9) has two eigenvalues at the origin and the remaining eigenvalues have negative real parts. ∎

Let us define vectors $i_1 = [1 \ 0 \ 1 \ 0 \ \ldots \ 1 \ 0]^T$ and $i_2 = [0 \ 1 \ 0 \ 1 \ \ldots \ 0 \ 1]^T$, being the right eigenvectors of $\bar{B}$ corresponding to the zero eigenvalue, and let π_1 and π_2 be the corresponding left eigenvectors.

Lemma 4. [16] Let $T = \begin{bmatrix} i_1 \vdots i_2 \vdots T_{2n \times (2n-2)} \end{bmatrix}$, where $T_{2n \times 2n-2}$ is an $2n \times (2n - 2)$ matrix, such that $\operatorname{span}\{T_{2n \times (2n-2)}\} = \operatorname{span}\{\bar{B}\}$. Then, T is nonsingular and

$$T^{-1} \bar{B} T = \begin{bmatrix} 0_{2 \times 2} & 0_{2 \times (2n-2)} \\ 0_{(2n-2) \times 2} & \bar{B}^* \end{bmatrix}, \qquad (10)$$

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

Theorem 1. Let assumptions A1), A2), A3) and A4) be satisfied. Then there exists $\delta' > 0$ such that for all $\delta \leq \delta'$ in (9) $\lim_{t\to\infty} \bar{\rho}(t) = \bar{\rho}_\infty = [\bar{\rho}_{\infty 1} \cdots \bar{\rho}_{\infty 2n}]^T$, with $\bar{\rho}_{\infty 1} = \cdots = \bar{\rho}_{\infty(2n-1)}$ and $\bar{\rho}_{\infty 2} = \cdots = \bar{\rho}_{\infty(2n)}$.

Proof: Let $\tilde{\rho}(t) = [\tilde{\rho}_1(t) \quad \tilde{\rho}_2(t) \cdots \tilde{\rho}_{2n}(t)^T]^T = T^{-1}\bar{\rho}(t)$. From (9) we obtain

$$\tilde{\rho}(t+1)^{[1]} = \tilde{\rho}(t)^{[1]}; \quad \tilde{\rho}(t+1)^{[2]} = (I + \delta\bar{B}^*)\tilde{\rho}(t)^{[2]}, \quad (11)$$

where $\tilde{\rho}(t)^{[1]} = [\tilde{\rho}_1(t) \ \tilde{\rho}_2(t)]^T$, $\tilde{\rho}(t)^{[2]} = [\tilde{\rho}_3(t) \cdots \tilde{\rho}_{2n}(t)]^T$. For δ small enough all the eigenvalues of $I + \delta\bar{B}^*$ lie within the unit circle. Therefore, $\lim_{t\to\infty} \tilde{\rho}(t)^{[2]} = 0$, so that

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

Consequently,

$$\bar{\rho}_\infty = T[\tilde{\rho}(0)^{[1]T} 0 \cdots 0]^T = (i_1\pi_1 + i_2\pi_2)\bar{\rho}(0). \quad (12)$$

Having in mind the definition of i_1 and i_2, we conclude that $\bar{\rho}_{\infty 1} = \cdots = \bar{\rho}_{\infty(2n-1)}$ and $\bar{\rho}_{\infty 2} = \cdots = \bar{\rho}_{\infty(2n)}$. ∎

Lemma 5. [16] Matrix $B(t)$ in (8) satisfies for all t

$$T^{-1}B(t)T = \begin{bmatrix} 0_{2\times 2} & \vdots & 0_{2\times(2n-2)} \\ \cdots\cdots\cdots & \vdots & \cdots\cdots\cdots \\ 0_{(2n-2)\times 2} & \vdots & B(t)^* \end{bmatrix}, \quad (13)$$

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

Theorem 2. Let assumptions A1), A2), A3) and A4) be satisfied. Then there exists $\delta'' > 0$ such that for all $\delta \leq \delta''$ in (8)

$$\lim_{t\to\infty} \hat{\rho}(t) = (i_1\pi_1 + i_2\pi_2)\hat{\rho}(0) \quad (14)$$

in the mean square sense and with probability one.

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

$$\begin{align} \tilde{\rho}(t+1)^{[1]} &= \tilde{\rho}(t)^{[1]}; & (15) \\ \tilde{\rho}(t+1)^{[2]} &= (I + \delta B(t)^*)\tilde{\rho}(t)^{[2]}, \end{align}$$

where $\tilde{\rho}(t)^{[1]} = [\tilde{\rho}_1(t) \ \tilde{\rho}_2(t)]^T$, $\tilde{\rho}(t)^{[2]} = [\tilde{\rho}_3(t) \cdots \tilde{\rho}_{2n}(t)]^T$. Recalling that $\bar{B}^*$ is Hurwitz, we observe that there exists such a positive definite matrix R^* that

$$\bar{B}^{*T}R^* + R^*\bar{B}^* = -Q^*, \quad (16)$$

where Q^* is positive definite. Define $q(t) = E\{\tilde{\rho}(t)^{[2]T}R^*\tilde{\rho}(t)^{[2]}\}$, and let $\lambda_Q = \min_i \lambda_i\{Q^*\}$ and $k' = \max_i \lambda_i\{E\{B(t)^* \ B(t)^{*T}\}\}$. From (15) we obtain

$$q(t+1) = E\{\tilde{\rho}(t)^{[2]T}E\{(I+B(t)^*)^T R^*(I+B(t)^*)\}\tilde{\rho}(t)^{[2]}\} \quad (17)$$

and, further,

$$q(t+1) \leq (1 - \delta\frac{\lambda_Q}{\max_i \lambda_i\{R^*\}} + \delta^2 k'\frac{\max_i \lambda_i\{R^*\}}{\min_i \lambda_i\{R^*\}})q(t), \quad (18)$$

having in mind that $E\{B(t)^*\} = \bar{B}^*$. 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 (18) 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, also,

with probability one, since the sequence $\{q(t)\}$ is summable. ∎

Convergence of the proposed algorithm in the case of correlated signals will be analyzed assuming:

A2') Process $\{x(t)\}$ is weakly stationary with $E\{x(t)\} = \bar{x}$, $E\{x(t)x(t-d)\} = m(d)$, $m(0) = s^2$, $|x(t)| \leq K < \infty$ (a.s.) and

$$a) \quad |E\{x(t)|\mathcal{F}_{t-\tau}\} - \bar{x}| = o(1), \ \text{(a.s.)} \quad (19)$$

$$b) \quad |E\{x(t)x(t-d)|\mathcal{F}_{t-\tau}\} - m(d)| = o(1), \ \text{(a.s.)} \quad (20)$$

for any fixed $d \in \{0, 1, 2, \ldots\}$, $\tau > d$, where $\mathcal{F}_{t-\tau}$ denotes the minimal σ-algebra generated by $\{x(t-\tau), x(t-\tau-1), \ldots, x(0)\}$ ($o(1)$ denotes a function that tends to zero as $\tau \to \infty$).

Theorem 3. Let assumptions A1), A2'), A3) and A4) be satisfied. Then there exists $\delta'' > 0$ such that for all $\delta \leq \delta''$ in (8) $\lim_{t\to\infty} \hat{\rho}(t) = (i_1\pi_1 + i_2\pi_2)\hat{\rho}(0)$ in the mean square sense and with probability one.

Proof: Following Theorem 2, we first compute $\tilde{\rho}(i) = T^{-1}\hat{\rho}(t)$, and obtain the same relations as in (15). Iterating back the second one, one obtains

$$\tilde{\rho}(t+1)^{[2]} = \prod_{s=t}^{t-\tau}(I + \delta B(s)^*)\tilde{\rho}(t-\tau)^{[2]}. \quad (21)$$

After calculating $E\{\tilde{\rho}(t+1)^{[2]T}R^*\tilde{\rho}(t+1)^{[2]}\}$ using (21), we extract the term linear in δ and replace $B(t)^* = \bar{B}^* + \tilde{B}(t)^*$, where $E\{\tilde{B}(t)^*\} = 0$. According to A4'),

$$|E\{\tilde{\rho}(t-\tau)^{[2]T}E\{\tilde{B}(s)^*|\mathcal{F}_{t-\tau-1}\}\tilde{\rho}(t-\tau)^{[2]}\}| \leq \phi(s-t+\tau+1)q(t-\tau), \quad (22)$$

where $\phi(t) > 0$, $\lim_{t\to\infty} \phi(t) = 0$. Therefore, it is possible to find such $\tau_0 > 0$ that for all $\tau \geq \tau_0$

$$(\tau+1)\lambda_{min}(Q^*) - \sum_{s=t}^{t-\tau}\phi(s) > \lambda_0 > 0, \quad (23)$$

since $\lambda_{min}(Q^*) > 0$ by definition. Therefore,

$$q(t+1) \leq (1 - \lambda_0\delta + \sum_{s=2}^{2(\tau+1)} k_s\delta^s)q(t), \quad (24)$$

where $|k_s| < \infty$ due to signal boundedness. It follows from (24) that there exists a $\delta' > 0$ such that $1 - \lambda_0\delta + \sum_{s=2}^{2(\tau+1)} k_s\delta^s < 1$. The result follows now in the same way as in Theorem 2. ∎

IV. CONVERGENCE ANALYSIS: COMMUNICATION ERRORS AND MEASUREMENT NOISE

A. Communication Errors

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

A5) the weights γ_{ij} in the algorithm (4) are 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 $\{\xi_{ij}(t)\}$ are i.i.d. random sequences with $E\{\xi_{ij}(t)\} = 0$ and $E\{\xi_{ij}(t)^2\} = (\sigma_{ij}^{\xi})^2$.

We assume that the 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 $\nu(t) = \begin{bmatrix} \nu_1(t) & \dots & \nu_n(t) \end{bmatrix}$, we obtain from (8) that

$$\hat{\rho}(t+1) = [I + (\Delta(t) \otimes I_2)B'(t)]\hat{\rho}(t) + \Delta(t)\nu(t), \quad (25)$$

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 $\gamma_{ij}(t)$.

Convergence of the recursion (25) will be studied starting from the above results. Notice first that now $E\{B'(t)\} = \bar{B}' = \bar{\Phi}(\bar{\Gamma} \otimes I_2)$, 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 $E\{\tilde{B}'(t)\} = 0$, due to mutual independence between the random variables in $B'(t)$; also, $E\{\tilde{B}'(t)|\mathcal{F}_{t-1}\} = 0$. It is obvious that $\bar{B}' = \bar{\Phi}(\bar{\Gamma} \otimes I_2)$ has qualitatively the same properties as $\bar{B}$ in (9): it has two eigenvalues at the origin and the remaining eigenvalues in the left half plane.

Further, we assume

$$A7) \quad \delta_i(t) = \delta(t) > 0; \quad \sum_{t=1}^{\infty} \delta(t) = \infty; \quad \sum_{t=1}^{\infty} \delta(t)^2 < \infty,$$

$i = 1, \dots, n$, so that

$$\hat{\rho}(t+1) = (I + \delta(t)\bar{B}')\hat{\rho}(t) + \delta(t)\tilde{B}'(t)\hat{\rho}(t) + \delta(t)\nu(t). \quad (26)$$

Theorem 4. Let assumptions A2)–A7) be satisfied. Then, $\hat{\rho}(t)$ generated by (26) 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.

Proof: Let

$$T' = \begin{bmatrix} i_1 \vdots i_2 \vdots T'_{2n \times (2n-2)} \end{bmatrix},$$

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, $(T')^{-1} = \begin{bmatrix} \pi_1' \\ \hline \pi_2' \\ \hline S'_{(2n-2) \times 2n} \end{bmatrix}$, where π_1' and π_2' are the left eigenvectors of $\bar{B}'$ corresponding to the zero eigenvalue. Let $\tilde{\rho}(t) = [\tilde{\rho}_1(t) \ \tilde{\rho}_2(t) \cdots \tilde{\rho}_{2n}(t)^T]^T = (T')^{-1}\hat{\rho}(t)$; then, (26) gives

$$\tilde{\rho}(t+1)^{[1]} = \tilde{\rho}(t)^{[1]} + \delta(t)G_1(t)\tilde{\rho}(t) + \delta(t)\nu'(t), \quad (27)$$
$$\tilde{\rho}(t+1)^{[2]} = (I + \delta(t)\bar{B}'^*)\tilde{\rho}(t)^{[2]} + \delta(t)G_2(t)\tilde{\rho}(t) + \delta(t)\nu''(t), \quad (28)$$

where $\tilde{\rho}(t)^{[1]}$ and $\tilde{\rho}(t)^{[2]}$ are defined as in (11), $\begin{bmatrix} G_1(t) \\ \hline G_2(t) \end{bmatrix} = (T')^{-1}\tilde{B}'(t)T'$ in such a way that $G_1(t)$ contains the first

two rows, $\nu'(t) = \begin{bmatrix} \pi_1' \\ \hline \pi_2' \end{bmatrix} \nu(t)$ and $\nu''(t) = S'_{(2n-2) \times 2n}\nu(t)$, while $\bar{B}'^*$ is a $(2n-2) \times (2n-2)$ Hurwitz matrix such that $(T')^{-1}\bar{B}'T' = \text{diag}\{0_{2 \times 2}, \bar{B}'^*\}$ (see Theorem 1). It is easy to verify that $E\{G_1(t)\} = 0$ and $E\{G_2(t)\} = 0$, as well as that $E\{G_1(t)|\mathcal{F}_{t-1}\} = 0$ and $E\{G_2(t)|\mathcal{F}_{t-1}\} = 0$.

Let $P^* > 0$ satisfy the Lyapunov equation $P^*\bar{B}'^* + \bar{B}'^{*T}P^* = -Q^*$ for some $Q^* > 0$. Denote $s(t) = E\{\|\tilde{\rho}(t)^{[1]}\|^2\}$ and $V(t) = E\{\tilde{\rho}(t)^{[2]T}P^*\tilde{\rho}(t)^{[2]}\}$. Then, directly following the methodology of [19] (Theorem 11), one obtains

$$s(t+1) \leq s(t) + C_1\delta(t)^2(1 + s(t) + V(t)) \quad (29)$$
$$V(t+1) \leq (1 - c_0\delta(t))V(t) + C_2\delta(t)^2(1 + s(t) + V(t)),$$

where c_0, C_1 and C_2 are appropriately chosen positive constants. According to [19] (Lemma 12 and Theorem 11), inequalities (29) give rise to the conclusion that $\tilde{\rho}(t)^{[1]}$ tends to a random variable $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ and $\tilde{\rho}(t)^{[2]}$ to zero in the mean square sense and w.p. 1. The result follows after calculating $\lim_{t \to \infty} \hat{\rho}(t) = T' \begin{bmatrix} \lim_{t \to \infty} \tilde{\rho}(t)^{[1]} \\ \hline 0 \end{bmatrix}$. ∎

B. Measurement Noise

We assume in this section that the signal $x(t)$ is measured with additive noise.

A8) Instead of $y_i(t)$ in (1), the sensors generate the signals $y_i^n(t) = \alpha_i x(t) + \beta_i + \eta_i(t)$, where $\{\eta_i(t)\}$, $i = 1, \dots n$, are zero mean i.i.d. random sequences with $E\{\eta_i(t)^2\} = (\sigma_i^{\eta})^2$, independent from the measured signal $x(t)$.

Inserting $y_i^n(t)$ instead of $y_i(t)$ in the basic algorithm (4), we obtain, after changing the variables, the following "noisy" version of (7):

$$\hat{\rho}_i(t+1) = \hat{\rho}_i(t) + \delta_i(t) \sum_{j \in \mathcal{N}_i} \gamma_{ij}\{[\Phi_i(t) + \Psi_i(t)] \times [\hat{\rho}_j(t) - \hat{\rho}_i(t)] + N_{ij}(t)\hat{\rho}_j(t) - N_{ii}(t)\hat{\rho}_i(t)\}, \quad (30)$$

where $\Psi_i(t) = \eta_i(t) \begin{bmatrix} \alpha_i x(t) & \alpha_i \\ \beta_i x(t) & \beta_i \end{bmatrix}$, $N_{ij}(t) = \frac{\eta_j(t)}{\alpha_j} \begin{bmatrix} \alpha_i y_i(t) & 0 \\ \beta_i y_i(t) & 0 \end{bmatrix} + \begin{bmatrix} \frac{\eta_j(t)\eta_i(t)}{\alpha_j} & 0 \\ 0 & 0 \end{bmatrix}$ and $N_{ii}(t) = \frac{\eta_i(t)}{\alpha_i} \begin{bmatrix} \alpha_i y_i(t) & 0 \\ \beta_i y_i(t) & 0 \end{bmatrix} + \begin{bmatrix} \frac{\eta_i(t)^2}{\alpha_i} & 0 \\ 0 & 0 \end{bmatrix}$. Notice that $E\{\Psi_i(t)\} = 0$, $E\{N_{ij}(t)\} = 0$, but $E\{N_{ii}(t)\} = \begin{bmatrix} \frac{(\sigma_i^{\eta})^2}{\alpha_i} & 0 \\ 0 & 0 \end{bmatrix}$.

Assuming $\delta_i(t) = \delta(t)$, $i = 1, \dots, n$, we can write in accordance with (8)

$$\hat{\rho}(t+1) = (I + \delta(t)\{[\Phi(t) + \Psi(t)](\Gamma \otimes I_2) + \tilde{N}(t)\})\hat{\rho}(t), \quad (31)$$

where $\Psi(t) = \text{diag}\{\Psi_1(t), \dots, \Psi_n(t)\}$, $\tilde{N}(t) = [\tilde{N}_{ij}(t)]$, where $\tilde{N}_{ij}(t) = -\sum_{k,k \neq i} \gamma_{ik}N_{ii}(t)$ for $i = j$ and $\tilde{N}_{ij}(t) = \gamma_{ij}N_{ij}(t)$ for $i \neq j$, $i, j = 1, \dots, n$.

Applying the methodology from the the previous section to the analysis of (31), we conclude that, instead of (9), we have now $\bar{\rho}(t+1) = [I + \delta(t)(\bar{B} + \Sigma_\eta)]\bar{\rho}(t)$, where $\bar{B}$ is defined in (9) and $\Sigma_\eta = -\text{diag}\{\frac{(\sigma_1^\eta)^2}{\alpha_1} \sum_j \gamma_{1j}, 0, \ldots, \frac{(\sigma_n^\eta)^2}{\alpha_n} \sum_j \gamma_{nj}, 0\}$. Under A2) and A3) the last recursion does not have the properties of (9), due to the additional term Σ_η. Then, all the row sums of $\bar{B} + \Sigma_\eta$ are not equal to zero, which prevents the achievement of asymptotic consensus (see Theorem 1).

Assuming more realistically that $\{x(t)\}$ is a correlated sequence, we adopt A2') instead of A2), and construct the following general calibration algorithm of *instrumental variable* type [17]:

$$\hat{\theta}_i(t+1) = \hat{\theta}_i(t) + \delta(t) \sum_{j \in \mathcal{N}_i} \gamma_{ij} \epsilon_{ij}^n(t) \begin{bmatrix} y_i^n(t-d) \\ 1 \end{bmatrix}, \quad (32)$$

where $d \geq 1$. The motivation for the introduction of delay $d > 0$ is the elimination of correlation between the noise terms in $\epsilon_{ij}^n(t)$ and $y_i^n(t-d)$. According to Section II, one obtains from (32) the following relations involving explicitly $x(t)$ and the noise terms:

$$\hat{\rho}_i(t+1) = \hat{\rho}_i(t) + \delta(t) \sum_{j \in \mathcal{N}_i} \gamma_{ij}\{(\Phi_i(t,d) + \Psi_i(t,d))$$
$$\times (\hat{\rho}_j(t) - \hat{\rho}_i(t)) + N_{ij}(t,d)\hat{\rho}_j(t) - N_{ii}(t,d)\hat{\rho}_i(t)\},$$

where $\Phi_i(t,d)$ is easily obtained in the same way as $\Phi_i(t)$ ($\Phi_i(t) = \Phi_i(t,0)$),

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

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

and

$$N_{ii}(t,d) = \frac{\eta_i(t)}{\alpha_i} \begin{bmatrix} \alpha_i y_i(t-d) & 0 \\ \beta_i y_i(t-d) & 0 \end{bmatrix} + \begin{bmatrix} \frac{\eta_i(t)\eta_i(t-d)}{\alpha_i} & 0 \\ 0 & 0 \end{bmatrix}.$$

In the same way as in (31), we have

$$\hat{\rho}(t+1) = (I + \delta(t)\{[\Phi(t,d) + \Psi(t,d)](\Gamma \otimes I_2) + \tilde{N}(t,d)\})\hat{\rho}(t), \quad (33)$$

where $\Phi(t,d) = \text{diag}\{\Phi_1(t,d), \ldots, \Phi_n(t,d)\}$, $\Psi(t,d) = \text{diag}\{\Psi_1(t,d), \ldots, \Psi_n(t,d)\}$, $\tilde{N}(t,d) = [\tilde{N}_{ij}(t,d)]$, where $\tilde{N}_{ij}(t,d) = -\sum_{k,k \neq i} \gamma_{ik} N_{ii}(t,d)$ for $i = j$ and $\tilde{N}_{ij}(t,d) = \gamma_{ij} N_{ij}(t,d)$ for $i \neq j$, $i, j = 1, \ldots, n$.

Instead of A4), we introduce:

A4') $-\bar{\Phi}(d) = -E\{\Phi_i(t,d)\}$ is Hurwitz for some $d = d_0 > 0$.

Theorem 5. Let assumptions A2'), A3), A4'), A7) and A8) be satisfied. Then $\hat{\rho}(t)$ generated by (33) with $d = d_0$ 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.

Proof: The proof starts from the demonstration that $T^{-1}B(t,d)T = \text{diag}\{0_{2\times2}, B(t,d)^*\}$, where $B(t,d) = \Phi(t,d)(\Gamma \otimes I_2)$ and $B(t,d)^*$ is an $(2n-2) \times (2n-2)$ matrix. Then, we compute $\tilde{\rho}(i) = T^{-1}\hat{\rho}(t)$, where T is chosen according to Lemma 5, and obtain, similarly as in Theorem 2, that

$$\tilde{\rho}(t+1)^{[1]} = \tilde{\rho}(t)^{[1]} + \delta(t)H_1(t,d)\tilde{\rho}(t) \quad (34)$$
$$\tilde{\rho}(t+1)^{[2]} = (I + \delta(t)B(t,d)^*)\tilde{\rho}(t)^{[2]}$$
$$+ \delta(t)H_2(t,d)\tilde{\rho}(t), \quad (35)$$

where $\tilde{\rho}(t)^{[1]} = [\tilde{\rho}_1(t) \ \tilde{\rho}_2(t)]^T$, $\tilde{\rho}(t)^{[2]} = [\tilde{\rho}_3(t) \cdots \tilde{\rho}_{2n}(t)]^T$, $H(t) = \begin{bmatrix} H_1(t) \\ \hdashline H_2(t) \end{bmatrix} = T^{-1}[\Psi_1(t,d)(\Gamma \otimes I_2) + \tilde{N}(t,d)]T$, so that $H_1(t)$ contains the first two rows; notice that $\{H(t)\}$ is i.i.d. and zero mean, with $E\{H(t)|\mathcal{F}_{t-1} = 0\}$.

After iterating the second relation in (34) τ times backwards, one obtains

$$\tilde{\rho}(t+1)^{[2]} = \Pi(t, t-\tau, d)\tilde{\rho}(t-\tau)^{[2]}$$
$$+ \sum_{\sigma=t-\tau}^{t} \Pi(t, \sigma+1, d)\delta(\sigma)H_2(\sigma)\tilde{\rho}(\sigma), \quad (36)$$

where $\Pi(t, s, d) = \prod_{\sigma=s}^{t}(I + \delta(\sigma)B(\sigma,d)^*)$, with $\Pi(t, t+1, d) = I$.

Having in mind A4'), we conclude that $\bar{B}(d)^* = E\{B(t,d)\}$ is Hurwitz for $d = d_0$; therefore, there exist such symmetric positive definite matrices R^* and Q^* that $R^*\bar{B}(d_0)^* + \bar{B}(d_0)^{*T}R^* = -Q^*$. Denote $s(t) = E\{\|\tilde{\rho}(t)^{[1]}\|^2\}$ and $V(t) = E\{\tilde{\rho}(t)^{[2]T}R^*\tilde{\rho}(t)^{[2]}\}$, as in Theorem 4. Calculation of $V(t)$ from (36) is straightforward, because $E\{H_2(t,d)\tilde{\rho}(\sigma)\} = 0$ for all σ. The crucial term in the final expression is the linear part of $E\{\tilde{\rho}(t-\tau)^{[2]T}\Pi(t, t-\tau, d)^T\Pi(t, t-\tau, d)\tilde{\rho}(t-\tau)^{[2]}\}$ with respect to $\delta(\sigma)$, $\sigma = \{t-\tau, \ldots, t\}$, analogously with the case of time invariant gains δ in Theorem 3. According to A2'), we obtain

$$E\{\tilde{\rho}(t-\tau)^{[2]T}E\{R^*\tilde{B}(\sigma, d_0)^* + \tilde{B}(\sigma, d_0)^{*T}R^*|\mathcal{F}_{t-\tau-1}\}$$
$$\times \tilde{\rho}(t-\tau)^{[2]}\} \leq \varphi(\sigma - t + \tau)V(t-\tau), \quad (37)$$

where $\varphi(t) > 0$ and $\lim_{t\to\infty} \varphi(t) = 0$, $t - \tau \leq \sigma \leq t$. Therefore, it is possible to find such $\tau_0 > 0$ that for all $\tau \geq \tau_0$

$$\lambda_{min}(Q^*) \sum_{\sigma=t-\tau}^{t} \delta(\sigma) - \sum_{\sigma=t-\tau}^{t} \varphi(\sigma)\delta(\sigma) > \lambda_0 > 0, \quad (38)$$

since $\lambda_{min}(Q^*) > 0$ by definition. Therefore, for t large enough, we have

$$V(t+1) \leq (1 - c_0\delta(t))V(t-\tau) \quad (39)$$
$$+ C_1 \sum_{\sigma=t-\tau}^{t} \delta(\sigma)^2(1 + s(\sigma) + V(\sigma)),$$

where $c_0 > 0$ and $C_1 > 0$ are generic constants. Since from the first relation in (34) we have directly

$$s(t+1) \leq s(t) + C_1\delta(t)^2(1 + s(t) + V(t)), \quad (40)$$

having in mind that $E\{H_1(t)|\mathcal{F}_{t-1}\} = 0$, recursions (39) and (40) can be treated like the recursions in (29), giving rise to the conclusion that $\tilde{\rho}(t)^{[1]}$ tends to a random variable $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ and $\tilde{\rho}(t)^{[2]}$ tends to zero in the mean square sense and w.p. 1. ∎
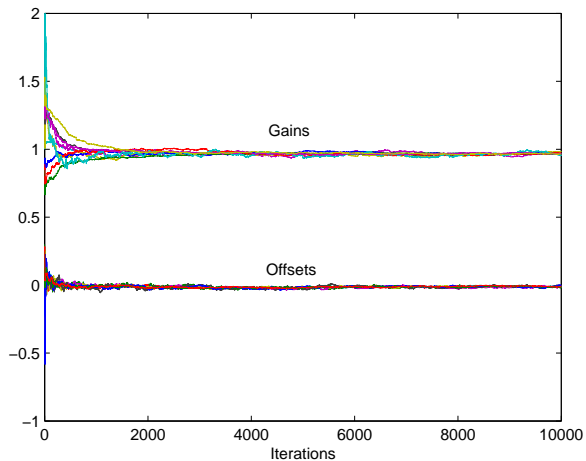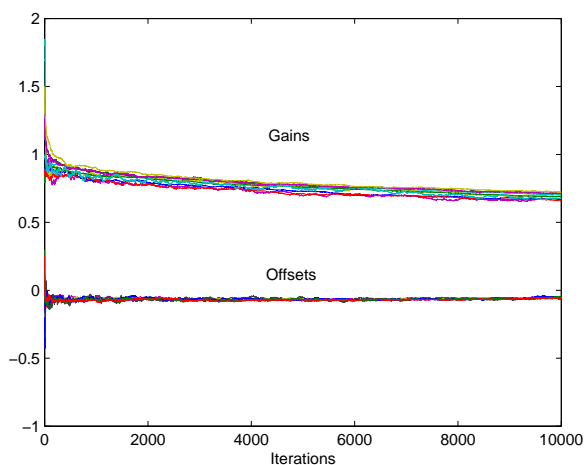
Fig. 1. Offset and gain estimates: $d = 1$

Fig. 2. Offset and gain estimates: $d = 0$

V. Simulation Results

In order to illustrate properties of the proposed algorithm, a sensor network with ten nodes has been simulated. A fixed randomly selected communications structure has been adopted, as well as parameters α_i and β_i randomly selected around one and zero, 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 (32) with $d = 1$ are presented for the sequence $\delta(t) = 0.01/t^{0.6}$. All the types of the discussed uncertainties are included: communications dropouts 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 is a correlated random sequence with zero mean and variance 1. It is clear that the consensus is achieved, and that the asymptotic values are not far from the optimal ones. Fig. 2 illustrates the necessity of the introduced instrumental variable modification. The estimates are depicted for $d = 0$ (no instrumental variable).

It is obvious that consensus is not achieved and that gains converge to zero in this case.

VI. Conclusion

In this paper a distributed calibration algorithm based on consensus has been proposed for sensor networks. It is proved, on the basis of a novel methodology of treating higher order consensus schemes using the results related to diagonal dominance of matrices decomposed into blocks, that the algorithm achieves asymptotic consensus for sensor gains and offsets in the mean square sense and with probability one, under both communication errors and measurement noise. The obtained results can be used to prove convergence to a given reference sensor characteristics (see [16]). The optimal choice of the weights γ_{ij} deserves further attention.

References

[1] *Proceedings of the IEEE, Special issue on sensor networks and applications*, August 2003, vol. 91.
[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, pp. 393–422, 2002.
[3] A. Speranzon, C. Fischione, and K. H. Johansson, "Distributed and collaborative estimation over wireless sensor networks," in *Proc. IEEE Conf. on Decision and Control*, 2006, pp. 1025–1030.
[4] K. Whitehouse and D. Culler, "Calibration as parameter estimation in sensor networks," in *Proceedings of the 1st ACM International Workshop on Wireless sensor networks and applications*, 2002, pp. 59–67.
[5] ——, "Macro-calibration in sensor/actuator networks," *Mobile Netw. Applicat.*, vol. 8, pp. 463–472, 2003.
[6] L. Balzano and R. Nowak, "Blind calibration," Networked and Embedded Systems Laboratory, UCLA, Tech. Rep. TR-UCLA-NESL-200702-01, 2007.
[7] ——, "Blind calibration in sensor networks," in *Proc. Intern. Conf. Inf. Proc. in Sensor Networks*, April 2007.
[8] M. Takruri, S. Challa, and R. Yunis, "Data fusion techniques for auto calibration in wireless sensor networks," July 2009.
[9] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak, "A collaborative approach to in-place sensor calibration," in *In Proceedings of the Second International Workshop on Information Processing in Sensor Networks (IPSN)*, 2003, pp. 301–316.
[10] R. Olfati-Saber, A. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, 2007.
[11] S. Bolognani, S. D. Favero, L. Schenato, and 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, January 2010.
[12] E. Miluzzo, N. D. Lane, A. T. Campbell, and R. Olfati-Saber, "Calibree: A self-calibration system for mobile sensor networks." in *DCOSS'08*, 2008, pp. 314–331.
[13] Y. Ohta and D. Siljak, "Overlapping block diagonal dominance and existence of Lyapunov functions," *J. Math. Analysis Appl.*, vol. 112, pp. 396–410, 1985.
[14] D. D. Šiljak, *Decentralized Control of Complex Systems*. New York: Academic Press, 1991.
[15] I. F. Pierce, "Matrices with dominating diagonal blocks," *Journ. of Economic Theory*, vol. 9, pp. 159–170, 1974.
[16] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed macro calibration in sensor networks," in *Proc. 20th Mediterranean Conference on Control and Automation*, 2012.
[17] T. Söderström and P. Stoica, *System Identification*. Hemel Hempstaed, UK: Prentice Hall International, 1989.
[18] S. S. Stanković, M. S. Stanković, and D. M. Stipanović, "Consensus based overlapping decentralized estimator," *IEEE Trans. Autom. Control*, vol. 54, pp. 410–415, 2009.
[19] M. Huang and J. H. Manton, "Stochastic consensus seeking with noisy and directed inter-agent communications: fixed and randomly varying topologies," *IEEE Trans. Autom. Control*, vol. 55, pp. 235–241, 2010.