Technical Notes and Correspondence

Distributed Blind Calibration in Lossy Sensor Networks via Output Synchronization

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

***Abstract*—A novel distributed algorithm for blind macro-calibration of large sensor networks is introduced. The algorithm is in the form of a system of gradient-type recursions for estimating parameters of local sensor calibration functions. The method does not require any fusion center. The convergence analysis is based on diagonal dominance of the dynamical systems with block matrices. It is proved that the asymptotic consensus is achieved for all the equivalent sensor gains and offsets (in the mean square sense and with probability one) in lossy sensor networks with possible communication outages and additive communication noise. An illustrative simulation example is provided.**

***Index Terms*—Calibration, consensus, convergence, distributed estimation, sensor networks, stochastic approximation, synchronization.**

I. Introduction

Wireless sensor networks have emerged as an important research area (see, e.g., [3], [4] and the references therein). *Sensor calibration* represents one of the most important challenges for the wide deployment of this new technology, especially in the case of large networks, when many devices may be partially unobservable. *Macro-calibration* is based on the idea to calibrate a network as a whole by observing the overall system response, mostly using parameter estimation techniques (e.g., [5]). Macro-calibration without explicit dependence on controlled stimuli is referred to as *blind calibration* (e.g., [6]–[10]). The problem is a difficult one, resembling to blind deconvolution. In [6], [7] a *centralized* non-recursive algorithm has been proposed assuming restrictive signal and sensor properties. In [9], an algorithm is proposed based on pairwise inter-node calibration and subsequent centralized consistency maximization at the network level. In *decentralized strategies*, the idea is to obtain a network with uniform properties using only local data processing and communications with neighbors. In this sense, analogies with the problem of *clock synchronization* in wireless sensor networks should be pointed out (see, e.g., [11]–[16] and the references therein).

In this technical note we propose a novel concept for *distributed blind macro-calibration* for sensor networks, based on distributed gradient-based recursive parameter estimation of affine calibration functions, derived from local criterion functions. The developed algorithm is a non-trivial extension of standard *output synchronization* or *consensus* algorithms (see, e.g., [17]–[19] and references therein). The overall network behavior is such that all the equivalent sensor gains and offsets converge asymptotically to equal values, without the knowledge of the measured signal and without the need for any type of fusion center. To the authors' best knowledge, consensus techniques have been applied explicitly to the calibration problems only in [20], [21], but within different contexts. The method bears a formal resemblance to the distributed clock synchronization schemes proposed in e.g., [12], [13] (a more detailed comparison will be given later). The analysis of the proposed method starts from the development of a novel methodology based on stability of *block diagonally dominant* dynamical systems [22]–[24]. It is proved that the convergence to consensus is achieved for the equivalent gains and offsets in the mean square sense and with probability one (w.p.1), even in the presence of communication uncertainties in the form of communication outages and additive noise. Preliminary results presented in [1] contain the main ideas and sketches of some of the proofs, assuming perfect communications among the sensor nodes, while [2] deals with the case when the measurement noise is present, which has not been treated in this technical note.

The outline of the technical note is as follows. In Section II, we formulate the blind calibration problem as a distributed parameter estimation problem and introduce the basic algorithm. Section III is devoted to the algorithm's convergence analysis in both the noiseless case and the case of communication uncertainties. In Section IV we present some illustrative simulation results where a comparison with an existing algorithm is provided.

II. Problem Formulation and Algorithm Design

Consider n distributed sensors measuring a discrete-time stochastic signal $x(t)$, $t \in \mathcal{I}^+$, where $\mathcal{I}^+$ is the set of nonnegative integers, and assume that the output of the i-th sensor is represented by

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

where the gain α_i and the offset β_i are, in general, unknown. By sensor calibration we consider the application of an affine calibration transformation which produces the corrected sensor output

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

where a_i and b_i are the *calibration parameters*, while $g_i = a_i\alpha_i$ and $f_i = a_i\beta_i + b_i$ represent the *equivalent gain* and the *equivalent offset*, respectively. In general, it is desirable to choose parameters a_i and b_i

Manuscript received March 11, 2014; revised October 14, 2014 and February 23, 2015; accepted April 7, 2015. Date of publication April 24, 2015; date of current version December 1, 2015. This work was supported by the EU Marie Curie CIG, Knut and the Alice Wallenberg Foundation and the Swedish Research Council. This work was presented in part at the Mediterranean Conference on Control and Automation and at the IEEE Conference on Decision and Control. Recommended by Associate Editor S. Zampieri.

M. S. Stanković and S. S. Stanković are with the School of Electrical Engineering, University of Belgrade, 11120 Belgrade, Serbia (e-mail: milos.stankovic@ic.etf.rs; stankovic@etf.rs).

K. H. Johansson is with the ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (e-mail: kallej@kth.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2015.2426272

in such a way as to obtain g_i as close as possible to one, and f_i to zero, $i = 1, \dots, n$.

Assume that the observed sensors form a network having the structure represented by a *directed graph* $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the set of nodes and $\mathcal{E}$ the set of arcs. The *adjacency matrix* $A = [a_{ij}]$, $i, j = 1, \dots, n$, is such that $a_{ij} = 1$ if the j-th sensor can send its message to the i-th sensor, and $a_{ij} = 0$ otherwise (the corresponding arc is directed from j to i). Let $\mathcal{N}_i$ be the set of neighboring nodes of the i-th node, i.e., the set of nodes j for which $a_{ij} = 1$ [25].

The aim of this technical note is to propose an algorithm for blind macro-calibration, considered as distributed real-time estimation of the calibration parameters a_i and b_i, without requiring the knowledge of the measured signal. The algorithm will be derived starting from the set of local criteria

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

$i = 1, \dots, n$, where $\gamma_{ij} > 0$, $j \in \mathcal{N}_i$, are *a priori* chosen design parameters. Denoting $\theta_i = [a_i \ b_i]^T$, we obtain for the gradient of (3)

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

which generate the following gradient recursions [26] for estimating θ_i^* which minimizes (3):

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

where $\hat{\theta}_i(t) = [\hat{a}_i(t) \ \ \hat{b}_i(t)]^T$ are the estimates of $\theta_i = [a_i \ \ b_i]^T$, $\delta_i(t) > 0$ is a time-varying step size, $\epsilon_{ij}(t) = \hat{z}_j(t) - \hat{z}_i(t)$ and $\hat{z}_i(t) = \hat{a}_i(t) y_i(t) + \hat{b}_i(t)$ is the current corrected sensor output; assuming no *a priori* information about sensor characteristics, the initial condition is set to $\hat{\theta}_i(0) = [1 \ 0]^T$, $i = 1, \dots, n$. Notice that the i-th recursion (5) subsumes availability of current corrected sensor outputs communicated only by the nodes $j \in \mathcal{N}_i$. The underlying idea is to achieve $\hat{z}_j(t) = \hat{z}_i(t)$, $i, j = 1, \dots, n$, by minimizing all the local criteria, so that all the estimates $\hat{g}_i(t) = \hat{a}_i(t)\alpha_i$ and $\hat{f}_i(t) = \hat{a}_i(t)\beta_i + \hat{b}_i(t)$, $i = 1, \dots, n$, tend asymptotically to the same values g^* and f^*, respectively.

For the sake of analysis, introduce

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

so that from (5) we obtain

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

$\Phi_i(t) = \begin{bmatrix} \alpha_i\beta_i x(t) + \alpha_i^2 x(t)^2 & \alpha_i\beta_i + \alpha_i^2 x(t) \\ (1+\beta_i^2)x(t) + \alpha_i\beta_i x(t)^2 & 1+\beta_i^2 + \alpha_i\beta_i x(t) \end{bmatrix}$ with $\hat{\rho}_i(0) = [\alpha_i \ \beta_i]^T$, $i = 1, \dots, n$. If $\hat{\rho}(t) = [\hat{\rho}_1(t)^T \cdots \hat{\rho}_n(t)^T]^T$, we obtain the following compact form:

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

where $\Delta(t) = \mathrm{diag}\{\delta_1(t), \dots, \delta_n(t)\}$, $B(t) = \Phi(t)(\Gamma \otimes I_2)$, $\Phi(t) = \mathrm{diag}\{\Phi_1(t), \dots, \Phi_n(t)\}$, $\Gamma = [\Gamma_{ij}]$, $i, j = 1, \dots, n$, $\Gamma_{ij} = \gamma_{ij}$ ($i \neq j$, $j \in \mathcal{N}_i$), $\Gamma_{ij} = 0$ ($i \neq j$, $j \notin \mathcal{N}_i$), $\Gamma_{ii} = -\sum_{j, j \neq i} \Gamma_{ij}$ ($\otimes$ denotes the Kronecker's product). Notice that Γ is a weighted Laplacian of $\mathcal{G}$ (e.g., [17]).

III. Convergence Analysis

A. Main Result

In the basic setting we assume:

A1) $\{x(t)\}$ is a stochastic process satisfying the following conditions:

a) $E\{x(t)\} = \bar{x} < \infty$ and $E\{x(t)^2\} = m < \infty$,
b) $|x(t)| \leq K < \infty$ w.p.1,
c) $m > \bar{x}^2$,
d) $\max(|E\{x(t)|\mathcal{F}_{t-\tau}\} - \bar{x}|, |E\{x(t)^2|\mathcal{F}_{t-\tau}\} - m|) = o(1)$, where $\mathcal{F}_{t-\tau}$, $\tau \geq 1$, denotes the minimal σ-algebra generated by $\{x(t-\tau), x(t-\tau-1), \dots\}$, and $o(1)$ a function that tends to zero as $|t-\tau| \to \infty$.

Remark 1: In relation with the process $\{x(t)\}$, Assumption A1.a) requires stationarity, and A1.b) boundedness, A1.c) ensures sufficient excitation, while A1.d) is a mixing condition.

A closer insight into the main recursion (8) shows that its asymptotic behavior cannot be analyzed by applying numerous known results on *consensus schemes*, because of the specific block structure of $B(t)$ (see, e.g., [17]–[19], [27] and references therein). We will first develop a methodological background based on diagonal dominance of matrices decomposed into blocks [22], [28], presented by Lemmas 1–4 (the proofs of Lemmas 2–4 are given in the Appendix).

Lemma 1—[22], [24]: A matrix $C = [C_{ij}]$, where $C_{ij} \in \mathbb{C}^{m \times m}$, $i, j = 1, \dots n$, has *quasi-dominating diagonal blocks* if the test matrix $W \in \mathbb{R}^{n \times n}$, with the elements $w_{ij} = 1$ ($i = j$) and $w_{ij} = -\|C_{ii}^{-1} C_{ij}\|_{ij}$ ($i \neq j$), is an M-matrix, where $\|.\|_{ij}$ denotes a matrix norm; then, C is nonsingular. If $C - \lambda I$ has quasi-dominating diagonal blocks for all $\lambda \in \mathbb{C}_+$, then C is Hurwitz ($\mathbb{C}_+$ denotes the closed right half complex plane).

The following lemma has a fundamental role in the subsequent analysis:

Lemma 2: If $C = [C_{ij}]$ has quasi-dominating diagonal blocks and C_{ii}, $i = 1, \dots, n$, are Hurwitz, then C is also Hurwitz.

Coming back to (8), we observe that matrix $\bar{B} = \bar{\Phi}(\Gamma \otimes I_2)$, where $\bar{\Phi} = E\{\Phi(t)\} = \mathrm{diag}\{\bar{\Phi}_1, \dots, \bar{\Phi}_n\}$ is essential for the analysis. It is straightforward to demonstrate that A1.c) implies that $-\bar{\Phi}_i(t) = -E\{\Phi_i(t)\}$ is Hurwitz, $i = 1, \dots, n$. We assume further:

A2) Graph $\mathcal{G}$ has a center node (a node from which all the other nodes are reachable).

Assumption A2) implies that $\mathcal{G}$ contains one strong component (strongly connected subgraph) [29]. Consequently, matrix Γ has one simple eigenvalue at the origin and the remaining ones with negative real parts (e.g., [18]).

Lemma 3: Let Assumptions A1) and A2) be satisfied. Then, matrix $\bar{B}$ has two eigenvalues at the origin and the remaining ones have negative real parts.

Define vectors $i_1 = [1 \ 0 \ 1 \ 0 \ \dots \ 1 \ 0]^T \in \mathbb{R}^{2n}$ and $i_2 = [0 \ 1 \ 0 \ 1 \ \dots \ 0 \ 1]^T \in \mathbb{R}^{2n}$, which represent right eigenvectors of $\bar{B}$ corresponding to the zero eigenvalue, and let π_1 and π_2 be the corresponding normalized left eigenvectors, satisfying $\begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} [i_1 \ \ i_2] = I_2$.

Lemma 4: The left eigenvectors π_1 and π_2 depend only on the sensor and network parameters. If $T = [\, i_1 \ | \ i_2 \ | \ T_{2n \times (2n-2)} \,]$, where $T_{2n \times (2n-2)}$ is an $2n \times (2n-2)$ matrix such that $\mathrm{span}\{T_{2n \times (2n-2)}\} = \mathrm{span}\{\bar{B}\}$, then: 1) T is nonsingular, 2) $T^{-1} = \begin{bmatrix} \pi_1 \\ \hline \pi_2 \\ \hline S_{(2n-2) \times 2n} \end{bmatrix}$, where $S_{(2n-2) \times 2n}$ follows from the definition of T, and 3) $T^{-1}\bar{B}T = \left[\begin{array}{c|c} 0_{2\times 2} & 0_{2\times(2n-2)} \\ \hline 0_{(2n-2)\times 2} & \bar{B}^* \end{array}\right]$, where $\bar{B}^*$ is Hurwitz.

Notice that Lemma 4 implies that $T^{-1}B(t)T = \left[\begin{array}{c|c} 0_{2\times 2} & 0_{2\times(2n-2)} \\ \hline 0_{(2n-2)\times 2} & B(t)^* \end{array}\right]$, where $B(t)^*$ is a $(2n-2)\times(2n-2)$ matrix.

We now introduce the following assumption on the step size:

A3) $\delta_i(t) = \delta = \text{const}, i = 1, \ldots, n$.

Based on Lemmas 1–4 we have the main convergence theorem:

Theorem 1: Let Assumptions A1)–A3) be satisfied. Then there exists $\delta' > 0$ such that for all $\delta \le \delta'$ 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 w.p.1.

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

$$\begin{aligned} \tilde{\rho}(t+1)^{[1]} &= \tilde{\rho}(t)^{[1]} \\ \tilde{\rho}(t+1)^{[2]} &= (I + \delta B(t)^*)\,\tilde{\rho}(t)^{[2]} \end{aligned} \tag{9}$$

where $\tilde{\rho}(t) = [\tilde{\rho}(t)^{[1]T}\ \tilde{\rho}(t)^{[2]T}]^T$, $\dim\{\tilde{\rho}(t)^{[1]}\} = 2$, $\dim\{\tilde{\rho}(t)^{[2]}\} = 2n-2$. Iterating (9) τ times backwards, we obtain

$$\tilde{\rho}(t+1)^{[2]} = \prod_{s=t}^{t-\tau}(I + \delta B(s)^*)\,\tilde{\rho}(t-\tau)^{[2]}. \tag{10}$$

According to Lemma 4, $E\{B(t)^*\} = \bar{B}^*$ is Hurwitz, so that there exists $R^* > 0$ such that

$$\bar{B}^{*T}R^* + R^*\bar{B}^* = -Q^* \tag{11}$$

where $Q^* > 0$. Define $q(t) = E\{\tilde{\rho}(t)^{[2]T}R^*\tilde{\rho}(t)^{[2]}\}$. After calculating $q(t+1)$ and introducing $B(t)^* = \bar{B}^* + \tilde{B}(t)^*$, with $E\{\tilde{B}(t)^*\} = 0$, we obtain that the term multiplying δ is given by

$$\begin{aligned} E\Bigg\{&\tilde{\rho}(t-\tau)^{[2]T} \\ &\times E\left\{\sum_{s=t}^{t-\tau}\left(B(s)^{*T}R^* + R^*B(s)^*\right)|\mathcal{F}_{t-\tau-1}\right\}\tilde{\rho}(t-\tau)^{[2]}\Bigg\} \\ = E\Bigg\{&\tilde{\rho}(t-\tau)^{[2]T}\Bigg[-(\tau+1)Q^* \\ &+ E\left\{\sum_{s=t}^{t-\tau}\left(\tilde{B}(s)^{*T}R^* + R^*\tilde{B}(s)^*\right)|\mathcal{F}_{t-\tau-1}\right\}\Bigg] \\ &\times\ \tilde{\rho}(t-\tau)^{[2]}\Bigg\}. \end{aligned} \tag{12}$$

According to A1.d), one obtains

$$\begin{aligned} &\left|E\left\{\tilde{\rho}(t)^{[2]T}E\left\{\tilde{B}(s)^{*T}R^* + R^*\tilde{B}(s)^*|\mathcal{F}_{t-\tau-1}\right\}\times\tilde{\rho}(t-\tau)^{[2]}\right\}\right| \\ &\qquad\le \phi(s-t+\tau+1)q(t-\tau) \end{aligned} \tag{13}$$

for $t-\tau \le s \le t$, where $\phi(\sigma) > 0$ and $\lim_{\sigma\to\infty}\phi(\sigma) = 0$. Let $\lambda_{Q^*} = \min_i \lambda_i\{Q^*\} > 0$; then it is possible to find such $\tau' > 0$ that for all $\tau \ge \tau'$ and $t \ge \tau$, we have

$$(\tau+1)\lambda_{Q^*} - \sum_{s=t}^{t-\tau}\phi(s-t+\tau+1) > \epsilon \tag{14}$$

where $\epsilon > 0$. Therefore, according to A1.b), there exist constants k_s, $s = 2, \ldots, 2(\tau+1)$, such that

$$q(t+1) \le \left(1 - \epsilon\delta + \sum_{s=2}^{2(\tau+1)} k_s\delta^s\right) q(t-\tau). \tag{15}$$

It follows that there exists $\delta' > 0$ such that $0 < 1 - \epsilon\delta + \sum_{s=2}^{2(\tau+1)} k_s\delta^s < 1$ for all $\delta < \delta'$. Therefore, $q(t)$ tends to zero (exponentially) so that $\tilde{\rho}(t)^{[2]}$ converges to zero in the mean square sense and w.p.1 (see [30]). This implies that $\lim_{t\to\infty}\tilde{\rho}(t) = \tilde{\rho}_\infty^T = [\tilde{\rho}(0)^{[1]T}0\cdots 0]^T$, so that $\hat{\rho}_\infty = T[\tilde{\rho}(0)^{[1]T}0\cdots 0]^T = (i_1\pi_1 + i_2\pi_2)\hat{\rho}(0)$ which completes the proof. ■

Remark 2: The convergence point of $\hat{\rho}(t)$ does not depend on the signal, but on the sensor parameters α_i and β_i and the design parameters γ_{ij} in J_i, $i, j = 1, \ldots n$. For the selected initial conditions in (5), $\pi_1\hat{\rho}(0)$ and $\pi_2\hat{\rho}(0)$ are in the form of weighted sums of α_i and β_i, $1, \ldots, n$, respectively. These weighted sums can be expected to be close to one and zero, respectively, if α_i have a distribution centered around one, and β_i around zero, assuming the same weights γ_{ij} for all the nodes. If one wishes to emphasize the effect of a selected subset $\mathcal{N}^f \subset \mathcal{N}$ of "good sensors," one can attach to all $j \in \mathcal{N}^f$ relatively high values of γ_{ij} for all i for which node j is a neighbor. The same effect is achieved by multiplying all γ_{jk} by relatively small positive numbers (implying that $\hat{\rho}_j(t)$ will stay in a small neighborhood of their initial values). This conclusion indicates that one can fix calibration parameters of one selected sensor (providing the equivalent gain one, and equivalent offset zero) by multiplying the related row of matrix Γ by zero, i.e. by "pinning" the network to that node. Convergence of all equivalent sensor parameters to the fixed values dictated by the selected node is guaranteed by Theorem 1 (see also [1], [2]).

Remark 3: The existence of $\delta' > 0$ guaranteeing convergence is conceptually important, but its value resulting from (15) may be restrictive. In practice, the step size δ should be chosen carefully, on the basis of an insight into signal and network properties (as in standard recursive parameter estimation schemes [26]).

Remark 4: Assumption A1.a) is not restrictive in practice; theoretically, it is possible to prove that the result of Theorem 1 holds for time varying functions $\bar{x}(t)$ and $m(t)$ when $\max(|\bar{x}(t+1) - \bar{x}(t)|, |m(t+1) - m(t)|)$ is small enough.

B. Lossy Sensor Networks

In this section, we analyze the proposed algorithm applied to *lossy sensor networks*, characterized by *communication outages* and *additive communication noise*.

A4) Weights γ_{ij} are time varying, given by $\gamma_{ij}(t) = u_{ij}(t)\gamma_{ij}$, where $\{u_{ij}(t)\}$ are i.i.d. binary random sequences independent of $\{x(t)\}$, such that $u_{ij}(t) = 1$ with probability p_{ij} ($p_{ij} > 0$ when $j \in \mathcal{N}_i$).

By A4), we have now a modified form of (8) in which $B(t)$ is replaced by $B'(t) = \Phi(t)(\Gamma(t)\otimes I_2)$, where $\Gamma(t)$ is obtained from Γ by replacing γ_{ij} by $\gamma_{ij}(t)$. Now, we have $\bar{B}' = E\{B'(t)\} = \bar{\Phi}(\bar{\Gamma}\otimes I_2)$, where $\bar{\Gamma}$ is obtained from Γ by replacing γ_{ij} by $\gamma_{ij}p_{ij}$; also, $T' = \left[\begin{array}{c|c|c} i_1 & i_2 & \bar{T}'_{2n\times(2n-2)} \end{array}\right]$, where $\bar{T}'_{2n\times(2n-2)}$ is an $2n\times(2n-2)$ matrix, such that $\text{span}\{\bar{T}'_{2n\times(2n-2)}\} = \text{span}\{\bar{B}'\}$. Then, $(T')^{-1} = \left[\begin{array}{c} \pi'_1 \\ \hline \pi'_2 \\ \hline \bar{S}'_{(2n-2)\times 2n} \end{array}\right]$, where π'_1 and π'_2 are the left eigenvectors of $\bar{B}'$ corresponding to the zero eigenvalue (analogous to π_1 and π_2 derived from $\bar{B}$).

The following theorem deals with the convergence in the presence of communication outages:

Theorem 2: Let Assumptions A1)–A4) be satisfied. Then, there exists $\delta'' > 0$ such that for all $\delta \le \delta''$ vector $\hat{\rho}(t)$ generated by (8) 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: Define $\tilde{\rho}(t) = (T')^{-1}\hat{\rho}(t)$ and introduce $B'(t) = \bar{B}' + \tilde{B}'(t)$, where $\{\tilde{B}'(t)\}$ is a zero mean independent matrix sequence. Then

$$(T')^{-1}B'(t)T' = \left[\begin{array}{c|c} 0_{2\times 2} & 0_{2\times(2n-2)} \\ \hline 0_{(2n-2)\times 2} & \bar{B}'^{*} \end{array}\right] + \left[\begin{array}{c|c} 0_{2\times 2} & G_1(t) \\ \hline 0_{(2n-2)\times 2} & G_2(t) \end{array}\right] \quad (16)$$

where $\bar{B}'^{*}$ is Hurwitz, while $\{G_1(t)\}$ and $\{G_2(t)\}$ are zero mean independent random matrix sequences. Then, we have the recursions

$$\begin{aligned} \tilde{\rho}(t+1)^{[1]} &= \tilde{\rho}(t)^{[1]} + \delta G_1(t)\tilde{\rho}(t)^{[2]} \\ \tilde{\rho}(t+1)^{[2]} &= \left(I + \delta\bar{B}'^{*} + \delta G_2(t)\right)\tilde{\rho}(t)^{[2]} \end{aligned} \quad (17)$$

which differ from (9) by the terms depending on $G_1(t)$ and $G_2(t)$. Let $R'^{*} > 0$ satisfy the Lyapunov equation $R'^{*}\bar{B}'^{*} + \bar{B}'^{*T}R'^{*} = -Q^{*}$ for some $Q'^{*} > 0$ and let $q'(t) = E\{\tilde{\rho}(t)^{[2]T}R'^{*}\tilde{\rho}(t)^{[2]}\}$. After applying the majorization procedure from Theorem 1 to $q'(t+1)$ obtained from (17), we readily conclude that (15) holds for $q'(t)$ with the same value for ε but with different constants k_s, so that there exists $\delta'' > 0$ such that for all $\delta \le \delta''$, $\tilde{\rho}(t)^{[2]}$ converges to zero in the mean square sense and w.p.1. On the other hand, applying the martingale convergence theorem to the first recursion in (17), we conclude that $\tilde{\rho}(t)^{[1]}$ converges w.p.1 to a vector random variable and that $\sup_t E\{\|\tilde{\rho}(t)^{[1]}\|^2\} < \infty$, because of geometric convergence of $q'(t)$ (see, e.g., [31]). The result follows after coming back to $\hat{\rho}(t)$, as in the proof of Theorem 1. ■

Remark 5: The theoretical result of Theorem 2 has important technical implications, since it allows formulation of *asynchronous* calibration algorithms, including communication schemes of *random gossip* type (see, e.g., [32]).

Let us now treat the case where both communication outages and additive communication noise are present.

A5) 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 independent of $\{x(t)\}$ and $\{u_{ij}(t)\}$, $i, j = 1, \ldots, n$, with $E\{\xi_{ij}(t)\} = 0$ and $E\{\xi_{ij}(t)^2\} = (\sigma_{ij}^{\xi})^2 < \infty$.

For this case, instead of A3), we need the following assumption (typical in standard *stochastic approximation* algorithms, e.g., [30], [31]) ensuring convergence w.p.1:

A3') $\delta_i(t) = \delta(t) > 0$, $\sum_{t=0}^{\infty}\delta(t) = \infty$, $\sum_{t=0}^{\infty}\delta(t)^2 < \infty$, $i = 1, \ldots, n$.

Theorem 3: Let Assumptions A1), A2), A3'), A4), and A5) be satisfied. Then, vector $\hat{\rho}(t)$ generated by (8) 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: In the case of communication errors of both types, (8) becomes

$$\hat{\rho}(t+1) = \left[I + (\Delta(t) \otimes I_2)\,B(t)'\right]\hat{\rho}(t) + \Delta(t)\nu(t) \quad (18)$$

where $\nu(t) = [\nu_1(t)^T \ \ldots \ \nu_n(t)^T]^T$ and

$$\nu_i(t) = \sum_{j\in\mathcal{N}_i}\gamma_{ij}(t)\xi_{ij}(t)\begin{bmatrix}\alpha_i y_i(t) \\ 1 + \beta_i y_i(t)\end{bmatrix} \quad (19)$$

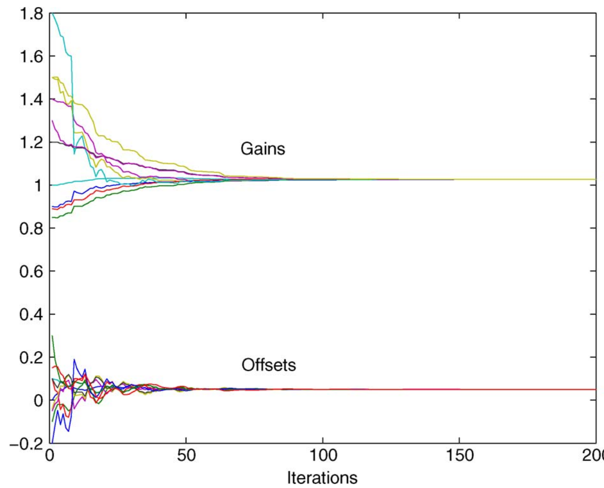

Fig. 1. Noiseless case: offset and gain estimates without reference included.

so that the term $\delta(t)\nu'(t)$ should be added to the first, and $\delta(t)\nu''(t)$ to the second relation in (17), where $\nu'(t) = \begin{bmatrix}\pi_1' \\ \pi_2'\end{bmatrix}\nu(t)$ and $\nu''(t) = S'_{(2n-2)\times 2n}\nu(t)$. According to Theorem 2 and A3') it is possible to derive (based on a similar methodology as in e.g., [30]) that

$$\begin{aligned} s'(t+1) &\le s'(t) + c_1\delta(t)^2\left(1 + s'(t) + q'(t)\right) \\ q'(t+1) &\le \left(1 - c_0\delta(t)\right)q'(t) + c_2\delta(t)^2\left(1 + q'(t)\right) \end{aligned} \quad (20)$$

where $s'(t) = E\{\|\tilde{\rho}(t)^{[1]}\|^2\}$ and c_0, c_1, and c_2 are appropriately chosen positive constants. According to [30] (Lemma 12 and Theorem 11) and [33], (20) implies 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. Hence, the result follows similar to the proof of Theorem 2. ■

Remark 6: In the case when the measurements $y_i(t)$ themselves are corrupted by additive measurement noise, the proposed algorithm cannot achieve the desired asymptotic behavior, due to correlation between the noise terms (typical for gradient based prediction error identification schemes [26], [34]). A modified version of the algorithm based on the introduction of *instrumental variables* enables consensus even in this case [2].

IV. Simulation Results

In order to illustrate characteristic properties of the proposed algorithm, a sensor network with ten nodes has been simulated. A fixed randomly selected directed communication graph satisfying A2) has been adopted, and parameters α_i and β_i have been randomly selected uniformly around one and zero, respectively, with standard deviation 0.2. A correlated random signal $x(t)$ with standard deviation equal to one, generated by a second order system driven by zero mean white noise, has been simulated.

Fig. 1 presents typical equivalent gains $\hat{g}_i(t)$ and offsets $\hat{f}_i(t)$ generated by the proposed algorithm (5) for $\delta = 0.01$ in the noiseless case. It is clear that the asymptotic consensus is achieved at the exponential rate. 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 evident (see Remark 2).

Two lines denoted by (A1) and (A2) in Fig. 3 present the mean squared error (MSE) with respect to the asymptotic mean values for the equivalent offset and equivalent gain, respectively, obtained using all nodes and 50 signal realizations with $\delta(t) = 0.03/t^{0.6}$. Communication outages with $p = 0.3$ and additive zero mean communication

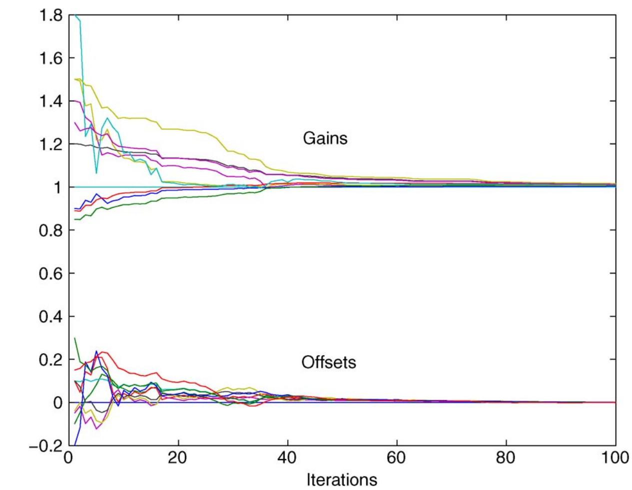

Fig. 2. Noiseless case: offset and gain estimates with reference included.

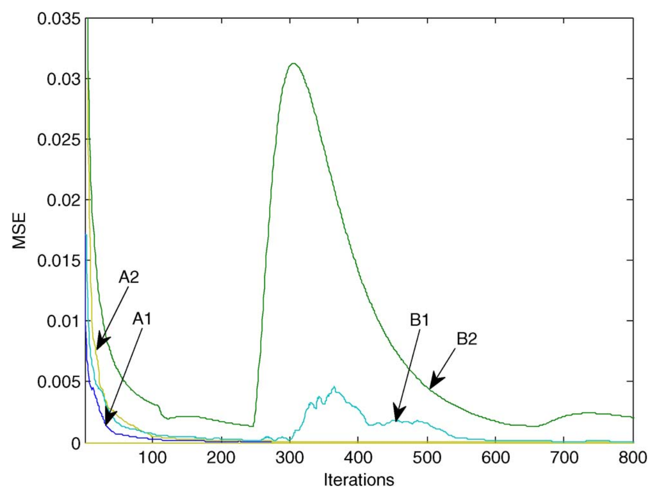

Fig. 3. Lossy sensor network: MSE for equivalent offsets (A1 and B1) and gains (A2 and B2) of the proposed algorithm (A1 and A2) and the algorithm proposed in [13] (B1 and B2).

noise with standard deviation 0.1 are included. Calibration is again achieved, but at a lower rate, as expected.

In order to clarify the relationship with calibration schemes which can be constructed starting from analogous distributed clock synchronization algorithms, we have constructed a calibration algorithm based on the algorithm proposed in [13] as a representative. It consists of two parts: 1) the first one calculates the filtered ratio of the signal increments $(y_j(t) - y_j(t-1))/(y_i(t) - y_i(t-1))$ for all neighbors $j \in N_i$ and uses this ratio in a recursion for $\hat{a}_i(t)$, 2) the second one generates $\hat{b}_i(t)$ using a recursion identical to the second recursion in (5). Without noise, this scheme provides the results comparable to the ones presented in Fig. 1. However, the algorithm is very sensitive to additive communication noise, due to the incorporated division by signal increments (this effect is not pronounced in clock synchronization). The lines denoted by (B1) and (B2) in Fig. 3 show the MSE for the equivalent offset and gain, respectively, in the case of the zero mean noise with standard deviation 0.005 (20 times lower than in the compared simulation of the proposed scheme). The advantage of our scheme is evident.

V. Conclusion

In this technical note, a new simple and easily implementable distributed blind macro-calibration algorithm of gradient type based on extended consensus with respect to equivalent gains and offsets has been proposed. The algorithm provides an efficient tool for coping with the problem of calibration of large wireless sensor networks, requiring neither the knowledge of the signal nor any fusion center. It has been proved, after developing a novel methodology based on diagonal dominance of matrices decomposed into blocks, that the calibration algorithm converges in the mean square sense and with probability one in the case of lossy networks. The behavior of the proposed algorithm has been illustrated through simulations.

Appendix

A. Proof of Lemma 2

If C_{ii} is Hurwitz, then there exists matrix $D_i > 0$ such that $C_{ii}D_i + D_iC_{ii}^* = -Q_i$, where $Q_i > 0$. Define the following *operator norm* of a matrix $X \in \mathbb{C}^{m\times m}$: $\|X\|_{ij} = \|X\|_i = \sup_{x\neq 0} \|Xx\|_{D_i}/\|x\|_{D_i}$, where $x \in \mathbb{C}^m$ and $\|x\|_{D_i} = (x^*D_i^{-1}x)^{1/2}$. Therefore, we obtain

$$\begin{aligned} w_{ij} &= -\left\|C_{ii}^{-1}C_{ij}\right\|_i \\ &= -\sqrt{\lambda_{\max}\left(C_{ij}^*C_{ii}^{-1*}D_i^{-1}C_{ii}^{-1}C_{ij}\right)} \\ &= -\max_{x\neq 0}\frac{x^*C_{ij}C_{ij}^*x}{x^*C_{ii}D_iC_{ii}^*x}. \end{aligned} \tag{21}$$

According to Lemma 1, C is Hurwitz, since $x^*(C_{ii} - \lambda I)D_i(C_{ii} - \lambda I)^*x \geq x^*C_{ii}D_iC_{ii}^*x$ for all $\lambda \in \mathbb{C}_+$, because $x^*(C_{ii}D_i - D_iC_{ii}^*)x = 0$, $i = 1,\ldots,n$. Hence, the result follows.

B. Proof of Lemma 3

The proof can be derived starting from the properties of matrix Γ, using Lemmas 1 and 2.

C. Proof of Lemma 4

Define $\psi_i^{[1]} = [\alpha_1\beta_i\ 1+\beta_i^2]^T$ and $\psi_i^{[2]} = [\alpha_i^2\ \alpha_i\beta_i]^T$, so that one can write $\bar{\Phi}_i = [\psi_i^{[1]}\bar{x} + \psi_i^{[2]}m \vdots \psi_i^{[1]} + \psi_i^{[2]}\bar{x}]$. If $\pi = [\pi^{[1]},\ldots,\pi^{[n]}]$, where $\pi^{[i]}$ are 2-dimensional row-vectors, the equation $\pi\bar{B} = 0$ (defining π) gives $-\pi^{[i]}\bar{\Phi}_i\sum_{j=1,j\neq i}^n\gamma_{ij} + \sum_{l=1,l\neq i}^n\pi^{[l]}\bar{\Phi}_l\gamma_{li} = 0$, $i = 1,\ldots,n$, and the following equivalent set of equations: $-\pi^{[i]}\psi_i^{[1]}\sum_{j=1,j\neq i}^n\gamma_{ij} + \sum_{l=1,l\neq i}^n\pi^{[l]}\psi_l^{[1]}\gamma_{li} = 0$ and $-\pi^{[i]}\psi_i^{[2]}\sum_{j=1,j\neq i}^n\gamma_{ij} + \sum_{l=1,l\neq i}^n\pi^{[l]}\psi_l^{[2]}\gamma_{li} = 0$, $i = 1,\ldots,n$, which do not depend on $x(t)$; this proves the first part. Assertions 1), 2), and 3) in the second part follow directly from the Jordan decomposition.

References

[1] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed macro calibration in sensor networks," in *Proc. Mediterr. Conf. Control and Automation*, 2012, pp. 1049–1054.

[2] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed calibration for sensor networks under communication errors and measurement noise," in *Proc. IEEE Conf. Decision and Control*, 2012, pp. 1380–1385.

[3] J. Chen, K. H. Johansson, S. Olariu, I. C. Paschalidis, and I. Stojmenovic, "Guest Editorial: Special Issue On Wireless Sensor and Actuator Networks," *IEEE Trans. Autom. Control*, vol. 56, no. 10, pp. 2244–2246, Oct. 2011.

[4] *Proc. IEEE, Special Issue on Sensor Networks and Applications*, vol. 91, Aug. 2003.

[5] K. Whitehouse and D. Culler, "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] L. Balzano and R. Nowak, "Blind calibration of sensor networks," in *Proc. Int. Conf. Information Processing in Sensor Networks*, Apr. 2007, pp. 79–88.

[8] K. Römer, P. Blum, and L. Meier, *Time Synchronization and Calibration in Wireless Sensor Networks*. New York, NY, USA: Wiley, 2005, pp. 199–237.

[9] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak, "A collaborative approach to in-place sensor calibration," in *Proc. Int. Conf. Information Processing in Sensor Networks*, 2003, pp. 301–316.

[10] M. Takruri, S. Challa, and R. Yunis, "Data fusion techniques for auto calibration in wireless sensor networks," in *Proc. Int. Conf. Information Fusion*, 2009, pp. 132–139.

[11] D. Giridhar and P. R. Kumar, "Distributed clock synchronization over wireless networks: Algorithms and analysis," in *Proc. IEEE Conf. Decision and Control*, 2006, pp. 263–270.

[12] P. Sommer and R. Wattenhofer, "Gradient clock synchronization in wireless sensor networks," in *Proc. Int. Conf. Information Processing in Sensor Networks*, 2009, pp. 37–48.

[13] L. Schenato and F. Fiorentin, "Average timesynch: A consensus-based protocol for time synchronization in wireless sensor networks," *Automatica*, vol. 47, no. 9, pp. 1878–1886, 2011.

[14] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Optimal synchronization for networks of noisy double integrators," *IEEE Trans. Autom. Control*, vol. 56, no. 5, pp. 1146–1152, May 2008.

[15] C. Liao and P. Barooah, "Distributed clock skew and offset estimation from relative measurements in mobile networks with Markovian switching topologies," *Automatica*, vol. 49, pp. 3015–3022, 2013.

[16] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed time synchronization in lossy wireless sensor networks," in *Proc. IFAC Workshop on Estimation and Control of Networked Systems*, 2012, pp. 25–30.

[17] R. Olfati-Saber, A. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[18] W. Ren and R. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.

[19] G. Seyboth, G. Schmidt, and F. Allgöwer, "Output synchronization of linear parameter-varying systems via dynamic couplings," in *Proc. IEEE Conf. Decision and Control*, 2012, pp. 5128–5133.

[20] S. Bolognani, S. D. Favero, L. Schenato, and D. Varagnolo, "Consensus-based distributed sensor calibration and least-square parameter identification in WSNs," *Int. J. Robust Nonlin. Control*, vol. 20, no. 2, pp. 176–193, 2010.

[21] E. Miluzzo, N. D. Lane, A. T. Campbell, and R. Olfati-Saber, "Calibree: A self-calibration system for mobile sensor networks," in *Proc. DCOSS'08*, 2008, pp. 314–331.

[22] Y. Ohta and D. Siljak, "Overlapping block diagonal dominance and existence of Lyapunov functions," *J. Math. Anal. Appl.*, vol. 112, pp. 396–410, 1985.

[23] D. D. Šiljak, *Large Scale Dynamic Systems: Stability and Structure*. Amsterdam, The Netherlands: North-Holland, 1978.

[24] I. F. Pierce, "Matrices with dominating diagonal blocks," *J. Econom. Theory*, vol. 9, pp. 159–170, 1974.

[25] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York, NY, USA: Springer-Verlag, 2001.

[26] L. Ljung and T. Söderström, *Theory and Pracice of Recursive Identification*. Cambridge, MA: MIT Press, 1983.

[27] S. S. Stanković, M. S. Stanković, and D. M. Stipanović, "Decentralized parameter estimation by consensus based stochastic approximation," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 531–543, Mar. 2011.

[28] D. D. Šiljak, *Decentralized Control of Complex Systems*. New York, NY, USA: Academic, 1991.

[29] ONZ. Lin, "Coupled Dynamic Systems: From Structure Towards Stability and Stabilizability," Ph.D. dissertation, University of Toronto, Toronto, Canada, 2006.

[30] 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, no. 1, pp. 235–241, Jan. 2010.

[31] H. F. Chen, *Stochastic Approximation and Its Applications*. Dordrecht, The Netherlands: Kluwer, 2002.

[32] T. C. Aysal, M. E. Yildriz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2748–2761, Jul. 2009.

[33] T. Li and J. F. Zhang, "Consensus conditions of multi agent systems with time varying topologies," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2043–2056, Sep. 2010.

[34] T. Söderström and P. Stoica, *System Identification*. Hemel Hempstead, U.K.: Prentice-Hall, 1989.