Convergence Rate for Distributed Macro Calibration of Sensor Networks based on Consensus

Maja Stanković, Dragan Antić, Saša Nikolić

University of Niš, Faculty of Electronic Engineering, Department of Control Systems, Aleksandra Medvedeva 14, 18000 Niš, Republic of Serbia
majasta@telekom.rs, dragan.antic@elfak.ni.ac.rs, sasa.s.nikolic@elfak.ni.ac.rs

Abstract: In this paper asymptotic convergence rate of a blind distributed macro-calibration algorithm for sensor networks based on consensus is analyzed, assuming the presence of both communication and measurement noise. Convergence to consensus with probability one (w.p.1) and in the mean square sense (m.s.s.) is proved for the sensor gain correction algorithm using general stochastic approximation arguments. Starting from an original formulation of the convergence rate of quadratic Lyapunov functions, an expression for the asymptotic convergence rate of the gain correction algorithm is derived. Using this result, convergence w.p.1 to consensus of the offset correction algorithm is proved under an additional nonrestrictive condition on the step size of the algorithm. In the practically important case when one of the sensors is taken as a reference, convergence w.p.1 of all corrected gains and offsets to the given reference value is proved. Simulation results are provided in order to illustrate characteristic properties of the algorithm.

Keywords: sensor networks; macro calibration; stochastic consensus; stochastic approximation; convergence rate

List of acronyms: a.s. - almost surely; ARMA - autoregressive moving average; e.g. - exempli gratia, meaning for example; i.e. - id est, meaning that is; m.s.s. - mean square sense; w.p.1 - with probability 1

1 Introduction

Sensor networks represent one of very important parts of Cyber-physical systems, Internet of things and complex large scale Control systems [1, 2]. *Sensor network calibration* is one of very actual problems connected to wide deployment of large *sensor networks* in industry, robotics and diverse multidisciplinary fields [3, 4]. In this case, individual calibration of each sensor (*micro-calibration*) cannot be effectively implemented. The so-called *macro-calibration* appears as one of practically efficient and theoretically attractive approaches, treating a given sensor network as a whole [6, 7]. In this sense, special attention has been paid recently to *distributed methods* for macro-calibration, not requiring centralized measurements and/oractions; *blind distributed macro-calibration* methods represent, theoretically and practically a great challenge since they do not require information about the stimuli [3, 7, 8].

A promising and methodologically consistent approach to *blind distributed macro-calibration* has been proposed in [9-12]. Starting from a sensor model characterized by two parameters (gain and offset), the algorithm proposed therein is of *distributed gradient-type*, providing equal corrected sensor outputs by ensuring convergence to *consensus* of *all the corrected sensor gains and offsets*. In practice, identical readings of all sensors provide: 1) uniformly good measurements for all the sensors in a network, and 2) an important possibility to obtain ideal calibration of the whole network by individual micro-calibration of only one sensor selected as a reference. Convergence of the algorithm has been proved, but without any insight into its convergence rate [9, 11].

Following the line of thought specific for time-synchronization algorithms [13], in [14, 15] a new approach to distributed blind macro-calibration, different from the results presented in [9-12], has been proposed. The algorithm consists of two recursive schemes, in which one is used for independent correction of *sensor gains*, and another for correction of *sensor offsets,* relying on the current gain correction results. It has been demonstrated that this algorithm has theoretical and practical advantages over the previously proposed ones [14, 15, 16]. Having in mind that the recursions for gain and offset correction can be considered as *distributed stochastic approximation algorithms*, efforts have been done to prove convergence to consensus under different operating conditions using their general theoretical background, e.g. [17]. As a result, geometric convergence to consensus has been proved in the noiseless case [14]. In [15], the previously obtained results were substantially generalized by proposing an asynchronous algorithm, working under the presence of additive communication and measurement noises. However, the corresponding theoretical analysis was focused dominantly on the gain correction, as the most critical part. Also, in [16] a new nonlinear calibration algorithm has been proposed, providing high robustness in practice w.r.t. noise samples containing outliers of impulsive type.

As a result, *simultaneous operation* of gain and offset correction in a *stochastic environment* containing both communication and measurement noise has still remained insufficiently clarified for the algorithm proposed in [14, 15]. Such a regime has been completely understood only in the noiseless case as a consequence of the exponential convergence rate. It turned out that the convergence to consensus of the offset correction scheme in a stochastic environment depends on the *convergence rate* of the gain correction scheme, which is far from being exponential. This situation requires a methodologically new tool for the theoretical analysis, since the existing general results cannot be successfully applied, e.g. [11, 15, 17-26].

This paper treats the distributed calibration algorithm proposed in [14, 15] under stochastic disturbances in a new and methodologically original way, trying to formulate the *convergence rate to consensus* of the gain correction scheme, and to prove exactly, for the first time, *convergence to consensus* with probability one (w.p.1) for both gain and offset correction in simultaneous operation. The autonomous scheme for gain correction will be analyzed first using the methodology from [15], generalizing the deterministic approach presented in [14] (Theorem 1 below). Proof of *convergence* w.p.1 and in the mean square sense (m.s.s.) *to consensus* is derived in a novel way. After showing that the methodology from [17], Paragraph 3.1, (proposed in the context of a.s. convergence) may be applied to the analysis of appropriate Lyapunov functions (Theorem 2 below), *convergence rate to consensus* of the gain correction scheme is formulated, as an indispensable precondition for treating the problem of convergence of the offset correction algorithm. Using a selected methodology from [1, 18] and the obtained expression for the rate of convergence of the gain correction algorithm [17], a proof of the offset correction algorithm convergence w.p.1 to consensus has been obtained (Theorems 4 and 5 below).

A special section (Section 5) is devoted to an important extension of the results obtained within Section 4 to the practically important case in which one of the network nodes is selected as a reference. In this case, the whole calibration algorithm succeeds in making all the asymptotical corrected gains and offsets equal to the selected ones w.p.1 and in the m.s.s., see [14, 15, 16]. The corresponding proofs are provided, representing an additional original contribution of the paper.

All in all, a major part of the presented theoretical results can be considered new, providing a new insight into the distributed blind calibration algorithms proposed in [14, 15]. They also provide a new general methodology for determining convergence rate of quadratic Lyapunov functions (Theorem 2), extending an idea from [17]. We hope that the results will fill not only an obvious gap in the theoretical understanding of the basic calibration algorithms from [14, 15] themselves, but also an analogous gap in the general understanding of complex dynamic consensus schemes (having in mind that the algorithm from [14, 15] can be formally considered as a special form of complex dynamic consensus schemes working in a stochastic environment [15]). Notice that the theoretical results from [9-12], having a similar character as the above, cannot be applied to the algorithm treated in the paper, having in mind that: 1) the algorithm itself is basically different, 2) convergence rate has not been analyzed and 3) the analysis methodology developed therein is not applicable to the algorithm from [14, 15]. Namely, the two-way coupling between the corrected gains and the corrected offsets adopted therein implies utilization of mathematical tools (e.g., related to diagonal dominance of the matrices characterizing the system) that are not applicable in the case of the algorithms treated in this paper. On the other hand, numerous existing results have shown that the algorithm from [14, 15] has a superior convergence rate in practice.

A number of simulations illustrate the behavior of the algorithm under different assumptions and operating conditions, enabling getting a better feeling of its behavior in a stochastic environment. These simulations do not serve for direct validation of the derived theoretical results, having in mind their nature - they are strictly valid asymptotically and provide, as usual [17, 24, 25], upper limits of the mean square distance w.r.t. consensus points.

2 Problem Formulation

Assume that there exist *n* distributed sensors measuring a real signal $x(t)$, where t is discrete time$t = 0,1, \ldots$ Sensor outputs are defined by the *standard linear stochastic model*

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

where the sensor *gain* α_i and the sensor *offset* β_i are unknown constants and $\zeta_i(t)$the *measurement noise* always present in real signals (its properties will be specified later), $i = 1, \ldots, n$ [14, 15].

As it is usual [9, 11, 12, 14], we will assume that each sensor implements an *affine calibration transformation* which generates the *corrected (modified) sensor output*

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

where a_i and b_i are the *calibration parameters* to be determined according to the chosen calibration goals, while $g_i = \alpha_i a_i$ represents the *corrected (modified) gain* and $f_i = \beta_i a_i + b_i$ the *corrected (modified) offset*. In general, the calibration goal is to keep g_i close to one and f_i close to zero.

It is assumed that the sensors are mutually interconnected, so that the sensor network is represented by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the set of nodes (sensors) and $\mathcal{E}$ the set of directed links (edges) (i,j) (node i sends messages to node j). Let $\mathcal{N}_i^{in}$ represent the set of *in-neighbours* of node i and $\mathcal{N}_i^{out}$ the set of its *out-neighbours* [19].

Remark 1. *In the literature, blind macro-calibration of sensor networks in the sense of finding such parameters* a_i *and* b_i *from (2) which ensure some global properties of a network as a whole, usually assumes a static model (1) and synchronous time [5-14]. Asynchronous functioning based on broadcast gossip is tretaed in [15]. Eventual communication time delays appearing as a consequence of propagation through the network are considered small w.r.t. the variations of the measured signal. In wireless sensor networks, this assumption is simply fulfilled when the nodes from* $\mathcal{N}_i^{out}$ *receive information from node i by broadcasting. Otherwise, blind macro-calibration can easily loose sense, as in the case of large and unpredictable communication delays.*

3 Gain Correction

Following the methodology presented in [14, 15], the algorithm for estimating the correction parameters a_i is derived using the *increments of the local measurement signals*:

$$\Delta y_i(t) = y_i(t) - y_i(t-1) = \alpha_i \Delta x(t) + \Delta\zeta_i(t) \quad (3)$$

where $\Delta x(t) = x(t) - x(t-1)$ and $\Delta\zeta_i(t) = \zeta_i(t) - \zeta_i(t-1)$ are the signal and measurement noise increments. The algorithm (not depending on b_i) is derived starting from the following *local criteria*, aimed at making the *sensor output increments equal*:

$$J_i^a(a_i) = \sum_{j\in\mathcal{N}_i^{in}} \gamma_{ij}\, E\{\left(\Delta z_j(t) - \Delta z_i(t)\right)^2\} \quad (4)$$

where $\gamma_{ij} > 0$ are *scalar weights reflecting the relative importance of the nodes j* from the in-neighbourhood of node i and $\Delta z_i(t) = a_i \Delta y_i(t)$, $i = 1, \dots, n$. The expression for the gradient of the criterion (4) w.r.t. a_i can be used directly as a generator of *stochastic gradient algorithms* for estimating calibration parameters a_i[14, 15]. In such a way, one obtains the following set of recursions

$$\hat{a}_i(t+1) = \hat{a}_i(t) + \delta(t) \sum_{j\in\mathcal{N}_i^{in}} \gamma_{ij}\, \varepsilon_{i,j}^{\Delta}(t) \Delta y_i(t) \quad (5)$$

where $\hat{a}_i(t)$ is an estimate of a_i, $\delta(t) > 0$ is the step size of the algorithm common for all the nodes, $\varepsilon_{i,j}^{\Delta}(t) = \Delta\hat{z}_j(t) + \xi_{ij}(t) - \Delta\hat{z}_i(t)$, $\Delta\hat{z}_i(t) = \hat{a}_i(t)\Delta y_i(t)$,and$\xi_{ij}(t)$ represents the *additive communication noise* between the nodes j and i (to be specified later). In general, one adopts that $\hat{a}_i(0) = 1$ [14]. Notice that the algorithm is distributed (based on local information), linear and simple for real time implementation.

The algorithm (5) can also be expressed by using the corrected gains $g_i(t) = \alpha_i \hat{a}_i(t)$:

$$g_i(t+1) = g_i(t) + \delta(t) \sum_{j\in\mathcal{N}_i^{in}} \gamma_{ij} \left(g_j(t) - g_i(t)\right) \alpha_i^2 \Delta x(t)^2 + \delta(t) n_i'(t), \quad (6)$$

where$n'_i(t) = \sum_j \gamma_{ij} \{[\hat{a}_j(t)\Delta\zeta_j(t) - \hat{a}_i(t)\Delta\zeta_i(t) + \xi_{ij}(t)]\left(\alpha_i \Delta x(t) + \Delta\zeta_i(t)\right) + \left(g_j(t) - g_i(t)\right)\Delta x(t)\Delta\zeta_i(t)\}$.

The last relation is important for a proper understanding of the algorithm behavior. The whole measurement noise term is composed of two parts, the first of which is zero-mean, and the second is not (having in mind that $E\{\hat{a}_i(t)\Delta\zeta_i(t)^2\} \neq 0$). An analogous phenomenon can be found in system identification algorithms when the measurement noise corrupts the dynamic system output, leading to biasedness of the parameter estimates [20]. In the case of the above calibration algorithm, this effect is much more pronounced: the algorithm may loose its fundamental capability of achieving consensus, and may become inapplicable to calibration. This problem can be efficiently overcome by applying *instrumental variables*, a very popular tool for system identification [20]. In our case, instrumental variables have to be *correlated with the measured signalx(t)*, and *uncorrelated with the random noise* $\Delta\zeta_i(t)$, so that the delayed measured noisy signals by at least two sampling instants represent a simple and efficient solution [15]. If $Z_i(t)$ denotes the instrumental variable of the i-th node at the instant t, $Z_i(t) = y_i(t-\tau)$ simply solves the problem for $\tau \geq 2$, having in mind that $\Delta y_i(t-\tau) = \alpha_i \Delta x(t-\tau) + \Delta\zeta_i(t-\tau)$ and that, evidently, $\Delta\zeta_i(t-\tau)$ is uncorrelated with $\Delta\zeta_i(t)$ for $\tau \geq 2$, assuming a realistic situation that $\Delta x(t)$and $\Delta x(t-\tau)$ are correlated.

Consequently, our calibration algorithm based on instrumental variables, is defined by

$$\hat{a}_i(t+\tau) = \hat{a}_i(t) + \delta(t) \sum_{j \in \mathcal{N}_i^{in}} \gamma_{ij} \, \Delta\varepsilon_{i,j}^{\Delta}(t) Z_i(t) \tag{7}$$

where $Z_i(t) = \Delta y_i(t-\tau)$, $\tau \geq 2$. The algorithm consists, in fact, of τ interlaced recursions at the network level, all having identical asymptotic behaviour [23]. The recursion (6) for $g_i(t)$ based on instrumental variables becomes now

$$g_i(t+\tau) = g_i(t) + \delta(t) \sum_{j \in \mathcal{N}_i^{in}} \gamma_{ij} \left(g_j(t) - g_i(t)\right) \alpha_i^2 \Delta x(t) \Delta x(t-\tau) + \delta(t) n_i(t), \tag{8}$$

where $n_i(t) = \sum_j \gamma_{ij} \{ [\hat{a}_j(t)\Delta\zeta_j(t) - \hat{a}_i(t)\Delta\zeta_i(t) + \xi_{ij}(t)] (\alpha_i \Delta x(t-\tau) + \Delta\zeta_i(t-\tau)) + \left(g_j(t) - g_i(t)\right) \Delta x(t) \Delta\zeta_i(t-\tau) \}$.

A compact form of the algorithm (8) at the network level is

$$g(t+\tau) = [I + \delta(t)\Delta x(t)\Delta x(t-\tau)\alpha^2 \Gamma] g(t) + \delta(t) N(t) \tag{9}$$

where $g(t) = [g_1(t) \dots g_n(t)]^T$, $N(t) = [n_1(t) \dots n_n(t)]^T$ and $\alpha = \text{diag}\{\alpha_1 \dots . \alpha_n\}$, while,

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

is a matrix of weights having the form of the Laplacian of the underlying directed graph [19, 21, 22].

The convergence analysis given below is based on the general results related to the stochastic approximation method [17, 24, 25, 26] and the properties of dynamic consensus schemes in stochastic environments [18, 21].

We will adopt the following assumptions:

(A.1) Graph $\mathcal{G}$ possesses a central node [22], (A.2) $\delta(t) > 0; \sum_{t=0}^{\infty} \delta(t) = \infty; \sum_{t=0}^{\infty} \delta(t)^2 < \infty$,

(A.3) For some $\tau \geq 2$, $E\{\Delta x(t)\Delta x(t-\tau)|\mathcal{F}_t\} \geq c_1 > 0$, $E\{\Delta x(t)^4|\mathcal{F}_t\} \leq c_2 < \infty$, where $\mathcal{F}_t$ is the minimal σ-algebra generated by $x(t), x(t-1)$

(A.4) Signal $\Delta\hat{z}_j(t)$ (or $\hat{z}_j(t)$) sent by node j is received by node i together with the additive noise $\xi_{ij}(t)$ (or $\eta_{ij}(t)$); all random communication noise processes are zero mean and i.i.d. with finite variances.

(A.5) Measurement noise sequences $\{\zeta_i(t)\}$ are mutually independent and composed of zero mean, finite variance i.i.d. random variables.

(A.6) Communication noise, measurement noise and the signal are mutually independent.

Theorem 1 (Convergence of the gain correction algorithm). *Let assumptions (A.1) - (A.6) be satisfied. Then, the estimate $g(t)$ generated by (9) tends w.p.1 and m.s.s. to a random vector $w\mathbf{1}, \mathbf{1} = [1 \ldots 1]^T$, where w is a random variable ($E\{w^2\} < \infty$).*

Proof: Let $T = \begin{bmatrix}\mathbf{1} & T_{n\times(n-1)}\end{bmatrix}$, where $\mathrm{span}(T_{n\times(n-1)}) = \mathrm{span}(\alpha^2\Gamma)$. Having in mind properties of Γ expressed by (A.1), it follows that

$$T^{-1}\alpha^2\Gamma\, T = \begin{bmatrix} 0 & \vdots & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & \vdots & \Gamma^* \end{bmatrix}$$

where Γ^* is a Hurwitz matrix, with all the eigenvalues in the left half plane [21, 27]. Define the transformed vector $\tilde{g}(t) = T^{-1}g(t)$, where $\tilde{g}(t) = \begin{bmatrix}\tilde{g}(t)^{[1]} & \tilde{g}(t)^{[2]T}\end{bmatrix}^T$, with $\tilde{g}(t)^{[1]} = \tilde{g}_1(t)$ and $\tilde{g}(t)^{[2]} = [\tilde{g}_2(t) \cdots \tilde{g}_n(t)]^T$.

After applying transformation T to (9), one obtains

$$\tilde{g}(t+\tau) = [I + \Delta x(t)\Delta x(t-\tau)T^{-1}\alpha^2\Gamma T]\tilde{g}(t) + \delta(t)T^{-1}N(t) \tag{10}$$

or

$$\tilde{g}(t+\tau)^{[1]} = \tilde{g}(t)^{[1]} + \delta(t)N(t)^{[1]} \tag{11}$$

$$\tilde{g}(t+\tau)^{[2]} = [I + \delta(t)\Delta x(t)\Delta x(t-\tau)\Gamma^*]\tilde{g}(t)^2 + \delta(t)N(t)^{[2]} \tag{12}$$

where $N(t)^{[1])}$ and$N(t)^{[2]}$ are obtained from the vector $T^{-1}N(t)$ as its first and the remaining $(n-1)$ components, respectively, satisfying$E\{N(t)^{[1]}|F_t\} = 0$ and $E\{N(t)^{[2]}|F_t\} = 0$. It is possible to show that $N(t)^{[1]} = \pi N(t)$ and $N(t)^{[2]} = SN(t)$, where π is the left eigenvector of $\alpha^2\Gamma$ corresponding to the zero eigenvalue and S the block matrix in the representation $T^{-1} = \begin{bmatrix}\pi \\ \cdots \\ S\end{bmatrix}$, with ($\dim S = (n-1) \times n$) (see [21]).

Consequently, there exists a symmetric positive definite matrix R such that the following Lyapunov equation is satisfied

$$\Gamma^{*T}R + R\Gamma^{*} = -Q \tag{13}$$

where Q is a symmetric positive definite matrix. Define the following Lyapunov functions $s(t) = E\{(\tilde{g}(t)^{[1]})^2\}$ and $v(t) = E\{\tilde{g}(t)^{[2]T}R\tilde{g}(t)^{[2]}\}$.

Remark 2. *The applied methodology of convergence analysis of stochastic consensus schemes based on the quadratic Lyapunov functions originates from [21]. This approach starts from the stability criteria of the underlying stochastic difference equations (11) and (12), and derives the convergence results w.p.1 and in the m.s.s. under appropriate assumptions. In this sense, the result of Theorem 2 givenbelow becomes even more important, having in mind its potential application to the convergence rate analysis of the stochastic consensus schemes in general (not yet existing in literature)*

Starting from the adopted assumptions, we immediately obtain

$$s(t+\tau) \le s(t) + C_1\delta(t)^2(1+s(t)) \tag{14}$$

where $C_1 > 0$.

On the other hand,

$$\begin{aligned} &v(t+\tau) \\ &= E\left\{\tilde{g}(t)^{[2]T}E\{(I+\delta(t)x(t)x(t-\tau)\Gamma^{*})^T R(I+\delta(t)x(t)x(t-\tau)\Gamma^{*})\tilde{g}(t)^{[2]}|\mathcal{F}_t\}\right\} \\ &+ \delta(t)^2E\{\| N(t) \|^2\} \end{aligned}$$

For the linear term in $\delta(t)$ we have the following inequality

$$\begin{aligned} E\{\tilde{g}(t)^{[2]T}E\{x(t)x(t-\tau)Q|\mathcal{F}_t\}\tilde{g}(t)^{[2]}\} &\ge c_1\min_i\lambda_i(Q)\left(\max_i\lambda_i(Q)\right)^{-1} v(t) \\ &\ge c_3v(t), \end{aligned}$$

$(0 < c_3 < \infty)$, and for the quadratic term

$$E\{\tilde{g}(t)^{[2]T}\Gamma^T E\{x(t)^2x(t-\tau)^2|\mathcal{F}_t\}R\Gamma\tilde{g}(t)^{[2]}\} \le c_4v(t).$$

$(0 < c_4 < \infty)$. In such a way one obtains

$$v(t+\tau) \le (1-\delta(t)k_1)v(t) + k_2\delta(t)^2) \tag{15}$$

where k_1 and k_2 are positive constants [14].

Using the result of Huang and Manton [21], Theorem 11, (14) and (15) imply that $v(t)$ tends to zero w.p.1 and in m.s.s., that $\sup_t s(t) < \infty$ and that $\tilde{g}(t)^{[1]}$ tends to a finite random variable $\tilde{g}_\infty^{[1]}$. According to [21], it follows that $\tilde{g}(t)$ tends w.p.1 and in the m.s.s. $\left[\tilde{g}_\infty^{[1]} \;\; 0\right]^T$, where $E\{\tilde{g}_\infty^{[1]2}\} < \infty$. Using the proof of Theorem 11 in [21], one obtains that $g(t)$ tends to $\tilde{g}_\infty^{[1]}\mathbf{1}$ w.p.1 and in m.s.s. □

The following theorem is devoted to a modification of the general results from [17], Chapter 3, making them applicable to the formulation of the convergence rate of the gain correction algorithm.

Theorem 2 (Convergence rate of the quadratic Lyapunov functions). *Let*

$$V(t+1) \leq [1-\delta(t)A]V(t) + \delta(t)^2 B \tag{16}$$

where $V(t) \geq 0$*,* $A > 0$*,* $B > 0$*, and* $\delta(t)$ *satisfies (A.2), with*

(A.7) $\frac{\delta(t)-\delta(t+1)}{\delta(t)\delta(t+1)} \rightarrow_{t\rightarrow\infty} \varepsilon > 0$

(A.8) $\sum_t \delta(t)^{2(1-d)} < \infty$ *for some* $d > 0$.

Then $V(t) = o(\delta(t)^{2d})$.

Remark 3. *The result given above has its roots in the approach to the analysis of the convergence w.p.1 of stochastic approximation schemes from [17], Theorem 3.1.1.. The above theorem is focused on quadratic Lyapunov function in order to achieve full compatibility with Theorem 1.*

Proof: From (16) we directly obtain

$$\frac{V(t+1)}{\delta(t+1)^{2d}} \leq \left(\frac{\delta(t)}{\delta(t+1)}\right)^{2d} \left\{[1-\delta(t)A]\frac{V(t)}{\delta(t)^{2d}} + \delta(t)^{2(1-d)}B\right\} \tag{17}$$

As

$$\left(\frac{\delta(t)}{\delta(t+1)}\right)^{2d} = 1 + 2d\frac{\delta(t)-\delta(t+1)}{\delta(t+1)} + O\left(\left(\frac{\delta(t)-\delta(t+1)}{\delta(t+1)}\right)^2\right) \tag{18}$$

one obtains

$$W(t+1) \leq [1-\delta(t)(A-2d\varepsilon)]W(t) + \delta(t)^{2(1-d)}B \tag{19}$$

where $W(t) = V(t)/\delta(t)^{2d}$, having in mind that

$$O\left(\left(\frac{\delta(t)-\delta(t+1)}{\delta(t+1)}\right)^2\right)(\delta(t)^{-1}-A) \underset{t\rightarrow\infty}{\rightarrow} 0$$

$$2d\frac{\delta(t)-\delta(t+1)}{\delta(t+1)}(\delta(t)^{-1}-A) \underset{t\rightarrow\infty}{\rightarrow} = -(A-2d\varepsilon).$$

Therefore, according to [17], $W(t)$ tends to zero. Thus, the result follows. □

Theorem 3 (Convergence rate of the gain correction algorithm). *Let assumptions (A.1)-(A.8) be satisfied. Then, the gain correction algorithm (6) converges to consensus at the rate*

$$\| \tilde{g}(t)^{[2]} \| = o(\delta(t)^d) \tag{20}$$

Proof: Direct application of Theorem 2 gives directly that $\tilde{g}(t)^{[2]}/\delta(t)^d$ converges to zero in the m.s.s. (compare (19) and (15)). In order to prove convergence w.p.1, one has simply to take into account that $N(t)$ is a martingale difference process with bounded $E\{\| N(t \|^2\}$. Application of Theorem B.6.1 from [17] leads to the conclusion that $\sum_{t=1}^{\infty} \delta(t)^{1-d} N(t) < \infty$ if $\sum_{t=1}^{\infty} \delta(t)^{2(1-d)} < \infty$, as it is assumed by Theorem 2. Therefore, we can apply Theorem 3.1.1. from [17] to (6) and conclude that $\| \tilde{g}(t)^{[2]} \| = o(\delta(t)^d)$ w.p.1.□

Remark 4. *The practical meaning of the above theorems becomes clearer in the case when* $\delta(t) = 1/t^{\mu}$, $\mu \in (0.5,1]$. *Sufficient conditions for convergence of* $W(t)$ *are then* $A - 2d\varepsilon > 0$ *and* $1/t^{2\beta(1-d)} < \infty$, *implied by* $d < 1 - \frac{1}{2\mu}$. *For* $\mu = 1$ *(which is the maximal value) one obtains* $d < 0.5$. *The convergence rate decreases to 0.5 when* μ *decreases; for* $\mu = 0.5$ *convergence cannot be guaranteed. In the case when* $\delta(t) = 1/t^{\mu}$ *we have* $\| \tilde{g}(t)^{[2]} \| = o(1/t^{\mu d})$.

4 Offset Correction

The basic local algorithm for offset correction results from the general idea from [14, 15]

$$\hat{b}_i(t+1) = \hat{b}_i(t) + \delta(t) \sum_{j \in \mathcal{N}_i^{in}} \gamma_{ij}\, \varepsilon_{ij}(t) \tag{21}$$

where $\varepsilon_{ij}(t) = \hat{a}_j(t) y_j(t) + \hat{b}_j(t) + \eta_{ij}(t) - \hat{a}_i(t) y_i(t) - \hat{b}_i(t)$, $\eta_{ij}(t)$ representing the communication noise when sending $\hat{z}_j(t) = \hat{a}_j(t) y_j(t) + \hat{b}_j(t)$ from node j to node i;oneassumes $\hat{b}_i(0) = 0$ [14, 15]. The corrected offsets$f_i(t) = \beta_i \hat{a}_i(t) + \hat{b}_i(t)$ are given by

$$f_i(t+1) = f_i(t) + \delta(t) \sum_{j \in \mathcal{N}_i^{in}} \gamma_{ij} \left[f_j(t) - f_i(t) \right] + \delta(t)[x(t) +$$

$$\alpha_i \beta_i \Delta x(t) \Delta x(t-\tau)] \sum_{j \in N_i^{in}} \gamma_{ij} \left[g_j(t) - g_i(t) \right] + \delta(t)[n_i(t) + \nu_i(t)] \tag{22}$$

where $n_i(t)$ is defined above and $\nu_i(t) = \hat{a}_j(t)\zeta_j(t) - \hat{a}_i(t)\zeta_i(t) + \eta_{ij}(t)$.

The overall global model at the network levelis given by

$$f(t+1) = (I + \delta(t)\Gamma) f(t) + \delta(t) G(t) \hat{g}(t) + \delta(t) M(t) \tag{23}$$

where$f(t) = [f_1(t) \dots f_n(t)]^T$, $G(t) = [x(t)I + \Delta x(t) \Delta x(t-\tau) \alpha\beta]\Gamma$, $\beta = \text{diag}\{\beta_1 \cdots \beta_n\}$, $M(t) = [m_1(t) \cdots m_n(t)]^T$, $m_i(t) = n_i(t) + \nu_i(t)$, $i = 1, \dots, n$.

Theorem 4 (Convergence of the offset correction algorithm). *Let the assumptions of Theorems 1 and 3 be satisfied, and let*

(A.9) $\sum_k \delta(k) o(\delta(k)^d) < \infty$.

Then, $f(t)$ *generated by (23) converges to* $w'\mathbf{1}$, *where* w' *is a random variable* *(*$E\{(w')^2\} < \infty$*).*

Remark 5. *The above theorem shows an important fact that convergence of the corrected offsets to consensus cannot be always obtained. It is important, however, that this fact does not represent a serious limitation for practice.*

Proof: As in Theorem 1, at the first step we introduce a transformation $T' = \left[\mathbf{1} \quad T'_{n\times(n-1)}\right]$, where $\text{span}\left(T'_{n\times(n-1)}\right) = \text{span}(\Gamma)$, i.e.

$$(T')^{-1}\Gamma T' = \begin{bmatrix} 0 & \vdots & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & \vdots & \Gamma' \end{bmatrix}$$

where Γ' is Hurwitz [21]. After applying transformation T' to (23), we obtain

$$\tilde{f}(t+1)^{[1]} = \tilde{f}(t)^{[1]} + \delta(t)\left[G(t)^{[1]}\tilde{g}(t)^{[2]} + M(t)^{[1]}\right] \quad (24)$$

$$\tilde{f}(t+1)^{[2]} = (I + \delta(t)\Gamma')\tilde{f}(t)^{[2]} + \delta(t)\left[G(t)^{[2]}\tilde{g}(t)^{[2]} + M(t)^{[2]}\right]. \quad (25)$$

where $G(t)^{[1]}$ and $G(t)^{[2]}$ are $1\times(n-1)$ and $(n-1)\times(n-1)$ matrices obtained from

$$(T')^{-1}\Gamma T = \begin{bmatrix} 0 & \vdots & G(t)^{[1]} \\ 0_{(n-1)\times 1} & \vdots & G(t)^{[2]} \end{bmatrix}$$

a) We shall first analyze (25) using the methodology derived from [17], Chapter 3. It is possible to start from Theorem 3.1.1 in [17] and to verify that: a) $G(t)^{[2]}\tilde{g}(t)^{[2]} \to 0$ as $t \to \infty$ and b) $E\left\{\| M(t)^{[2]} \|^2\right\} < \infty$, so that $\sum_t \delta(t)^2 E\{\| M(t)^{[2]} \|^2\} < \infty$, implying further that $\sum_t \delta(t) M(t)^{[2]} < \infty$ (according to Theorem B.6.1 in [17]). By direct inspection we realize that the assumptions of Theorem 3.1.1 are satisfied and that $\tilde{f}(t)^{[2]}$ tends to zero w.p.1.

b) Analysis of (24) does not follow strictly the lines of the general arguments from [17]. However, a direct insight shows that

$$\tilde{f}(t+1)^{[1]} = \tilde{f}(0)^{[1]} + \textstyle\sum_k \delta(k) G(k)^{[1]}\tilde{g}(k)^{[2]} + \sum_k \delta(k) M(k)^{[1]}, \quad (26)$$

For the second term we have

$$\| \sum_k \delta(k) G(k)^{[1]}\tilde{g}(k)^{[2]} \| \le k_3 \sum_k \delta(k) o(\delta(k)^d),$$

$k_3 > 0$. Boundedness of this term is ensured by (A.7). For the third term we conclude that $E\{\| M(k)^{[1]} \|^2\}$ is bound and (A.2) holds; it follows, using again Theorem B.6.1 from [17], that $\| \sum_k \delta(k) M(k)^{[1]} \| < \infty$. As a consequence, there exists a random variable $\tilde{f}_\infty^{[1]}$ (as in the case of Theorem 1) such that $\lim_{t\to\infty} = \tilde{f}_\infty^{[1]}$ w.p.1.

Thus, the result follows.□

Remark 6. *It is clear that in the standard case when* $\delta(t) = 1/t^\mu$, $0.5 < \mu \le 1$, *(A.9) becomes* $\frac{1}{t^{\mu+d}} < \infty$; *this inequality can be easily fulfilled in practice for a* *given value of d, by choosing μ close enough to 1. This is an example of how the above analysis can be applied to predict the algorithm behaviour in practice. However, one should bear in mind that all the above results are based on sufficient conditions, depending to a great extent on the chosen methodology of analysis and operating conditions.*

Convergence rate of the offset correction algorithm will be formulated using the second recursion (25) for the transformed variables, having in mind the existence of an additional exogenous term resulting from the gain correction recursion. Therefore, we shall apply Theorem 3.1.1 from [17] in its original form, dealing with the convergence w.p.1.

Theorem 5 (Convergence rate of the offset correction algorithm). *Let assumptions (A.1) - (A.9) be satisfied. Then ,* $\|\tilde{f}(t)^{[2]}\| = o(\delta(t)^d)$ *w.p.1.*

Proof: The methodology is based on a modification of the approach in Theorem 2. We first have for t large enough that

$$\frac{\tilde{f}(t+1)^{[2]}}{\delta(t+1)^d} = \left(\frac{\delta(t)}{\delta(t+1)}\right)^d \left\{\frac{1}{\delta(t)^d}(I+\delta(t)\Gamma')\tilde{f}(t)^{[2]} + \right.$$

$$\left. +\delta(t)\left[G(t)^{[2]}\tilde{g}(t)^{[2]} + M(t)^{[2]}\right]\right\} \quad (27)$$

After applying consecutively the steps analogous to those in the proof of Theorem 2, we obtain

$$h(t+1) = [I+\delta(t)\Gamma']h(t) + \delta(t)\left[G(t)^{[2]}\tilde{g}(t)^{[2]} + M(t)^{[2]}\right] \quad (28)$$

where $h(t) = \tilde{f}(t)^{[2]}/\delta(t)^d$. As$\Gamma'$ is Hurwitz and (A.8) holds, $E\{\| M(t)^{[2]} \|^2\} < \infty$ and the proposition of Theorem 3 holds. Therefore, we can apply Theorem 3.1.1 from [17] and directly conclude that $h(t)$ tends to zero w.p.1. Thus, the result follows.□

5 Calibration with Reference

A modification of the basic algorithm is obtained when one fixed sensor (e.g., sensor k) is preliminarily micro-calibrated using the fixed values $\hat{a}_k$ and $\hat{b}_k$; practical importance of such a modification has been pointed out in [14,15,16]. Formally, we have the relations $\hat{a}_k(t+1) = \hat{a}_k(t) = \hat{a}_k$ and $\hat{b}_k(t+1) = \hat{b}_k(t) = \hat{b}_k$ for some k (we say that the network is *pinned to* node k).Assuming that the sensor parameters α_k and β_k are known (e.g., as a consequence of the micro-calibration done preliminarily for sensor k), we also have $g_k(t+1) = g_k(t) = g_k = \alpha_k\hat{a}_k$ and $f_k(t+1) = f_k(t) = f_k = \beta_k\hat{a}_k + \hat{b}_k$. We will show below that under these conditions the proposed calibration algorithms (9) and (23) ensure convergence of all the corrected gains and offsets exactly as desired to the given reference values g_k and f_k.

Assume without loss of generality that $k = 1$ and $g_k = g$. Then, (9) gives, after simple technical transformations,

$$r(t+\tau) = [I + \delta(t)\Delta x(t)\Delta x(t-\tau)\bar{\alpha}^2\bar{\Gamma}]r(t) + \delta(t)\bar{N}(t) \quad (29)$$

where $r(t)$ is an $(n-1)$-vector defined as $r(t) = \bar{g}(t) - g\mathbf{1}$, with $\bar{g}(t) = [g_2(t) \cdots g_n(t)]^T$, while $\bar{\alpha}^2$ and $\bar{\Gamma}$ are $(n-1) \times (n-1)$ matrices obtained from α^2 and Γ after deleting their first row and the first column; $(n-1)$-vector $\bar{N}(t) = [N_2(t) \cdots N_n(t)]^T$ is obtained from $N(t)$ after deleting its first component. Convergence properties of the recursion (29) can be easily analyzed by applying the methodology of Theorem 1, in relation with the recursion for $\tilde{g}(t)^{[2]}$ in (12). Namely, it is possible to show that the matrix $\bar{\alpha}^2\bar{\Gamma}$ is Hurwitz, making possible direct application of the procedure used in Theorem 1 for analysing the Lyapunov function $v(\cdot)$ (having in mind that $\bar{\Gamma}$ is quasi-diagonally dominant and therefore Hurwitz [28]). Consequently, we have:

Theorem 6 (Convergence to the reference gain). *Let the assumptions (A.1) - (A.6) be satisfied and let node 1 be a center node of the graph $\mathcal{G}$. Then $r(t)$ generated by (29) tends to zero w.p.1 and in the m.s.s. for any given value of the reference g.*

In the same way we have a modification of Theorem 3, dealing with the corresponding rate of convergence:

Theorem 7 (Convergence rate to the reference gain) *Let assumptions (A.1)-(A.8) be satisfied and let node 1 be a center node of the graph $\mathcal{G}$. Then, $r(t)$ generated by (29) converges to zero w.p.1 at the rate $\| r(t) \| = o(\delta(t)^d)$*

Similarly as in the case of gain correction in the last subsection, we can start from (23) and obtain for $f_1 = f$ the following recursion for offset correction

$$s(t+1) = [I + \delta(t)\bar{\Gamma}]s(t) + \delta(t)\bar{G}(t)s(t) + \delta(t)\bar{M}(t) \quad (30)$$

where $s(t)$ is an $(n-1)$-vector defined as $s(t) = \bar{f}(t) - f\mathbf{1}$, with $\bar{f}(t) = [f_2(t) \cdots f_n(t)]^T$, $\bar{G}(t)$ is an $(n-1) \times (n-1)$ matrix obtained from $G(t)$ by deleting its first row and its first column, while $\bar{M}(t)$ is an $(n-1)$-vector obtained from $M(t)$ bydeleting its first element. Theorem 4 can be directly applied in the following way:

Theorem 8 (Convergence to the reference offset). *Let assumptions (A.1)-(A.9) hold and let node 1 be a center node of the graph $\mathcal{G}$. Then, $s(t)$ tends to zero w.p.1. for any given value of the reference f.*

The proof follows from the proof of Theorem 4. Notice here only that the rate of convergence of (29) defined by Theorem 7 is not a prerequisite for obtaining the result of Theorem 8. Additional restrictions related to the behaviour of (24) in Theorem 4 do not exist in the case of calibration with reference.

Convergence rate for the corrected offset follows simply from Theorem 5.

Notice here only that the results of this section have great practical implications: a combination of the micro calibration procedure of a selected sensor and the distributed blind macro calibration algorithm applied to the network pinned to the same sensor seems to be a practically very efficient tool for macro calibration of large sensor networks. Some simulation results will be presented below.

6 Simulation Results

In spite of the fact that the paper has a predominantly theoretical character, some simulation results are presented in order to provide a reader with a feeling about noise immunity and achievable performance of the analyzed distributed macro-calibration algorithm in the case of stochastic disturbances. The role of the numerical results in the paper is, in fact, to provide an illustration of the functioning of the algorithm in the presence of communication and measurement noise, rather than to directly validate the main theoretical results related to the convergence rate. This is clear from the character of the main theoretical results themselves. Also, one should bear in mind that we are dealing in this paper with the *asymptotic convergence rate*, which is characterized by the asymptotic mean square value of the error, and cannot be directly observed through finite time transients w.r.t. the initial conditions (the form of these transients represents a completely different phenomenon).

In order to make the computational aspects of the algorithm more transparent, Algorithm 1 below is introduced, containing a corresponding pseudo-code which can be easily programmed using any language (the results given below are obtained using a Matlab program directly derived from Algorithm 1). The given pseudo code represents also a nucleus for any type of real-time implementation.

ALGORITHM 1: **Computation of the calibration parameters**

Initial values$\hat{a}_i(0), \hat{b}_i(0), Z_i(0)$, step sizes$\delta(t), t \geq 0$, weights γ_{ij}

Initialize the iteration counter $t \leftarrow 0$

Repeat

For all $i \in \mathcal{N}$ do

Observe measurements $\Delta y_i(t)$and $y_i(t)$ according to (1) and (3)

Compute $\Delta \hat{z}_i(t)$and $\hat{z}_i(t)$ using (5) and (21)

Send $\Delta \hat{z}_i(t)$and $\hat{z}_i(t)$ to all the nodes $j \in \mathcal{N}_i^{out}$

Get data $\Delta \hat{z}_j(t) + \xi_{ij}(t)$ and $\hat{z}_j(t) + \eta_{ij}(t)$ from all $j \in \mathcal{N}_i^{in}$

Update $\hat{a}_i(t+\tau) \leftarrow \hat{a}_i(t) + \delta(t) \sum_{j\in\mathcal{N}_i^{in}} \gamma_{ij}\, \Delta\varepsilon_{i,j}^{\Delta}(t) Z_i(t)$ using (7)

Update $\hat{b}_i(t+1) \leftarrow \hat{b}_i(t) + \delta(t) \sum_{j\in\mathcal{N}_i^{in}} \gamma_{ij}\, \varepsilon_{ij}(t)$ using (21)

end for

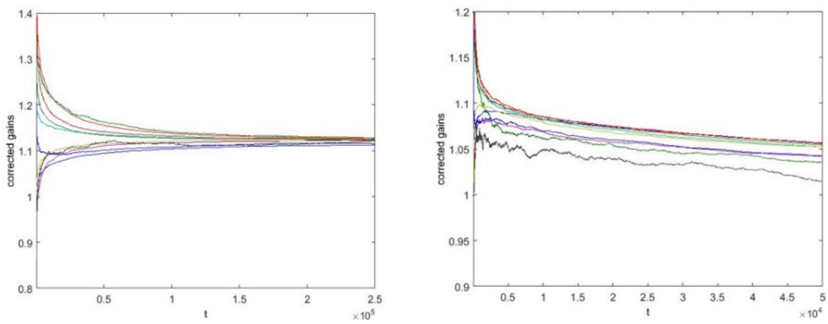
Update the iteration counter$t$←t+1

Until convergence

It is obvious that the algorithm is extremely simple, requiring a small number of operations and small memory, and, thus, allowing implementation by using very cheap processing units.

Some basic properties of the proposed algorithm are illustrated by Monte Carlo simulations related to a sensor network with ten nodes, under the presence of both communication and measurement noise. Digraphs satisfying (A.1) and parameters γ_{ij} have been selected at random, as well as the local sensor parameters α_i and β_i about one and zero, respectively, with standard deviations 0.2. The signal $x(t)$ (supposed to be unknown) has been generated by a second order ARMA process, with standard deviation 1. It has been adopted that the step size is defined by the decreasing function $\delta(t) = k/t^{\mu}$, $\mu \in (0.5,1]$.

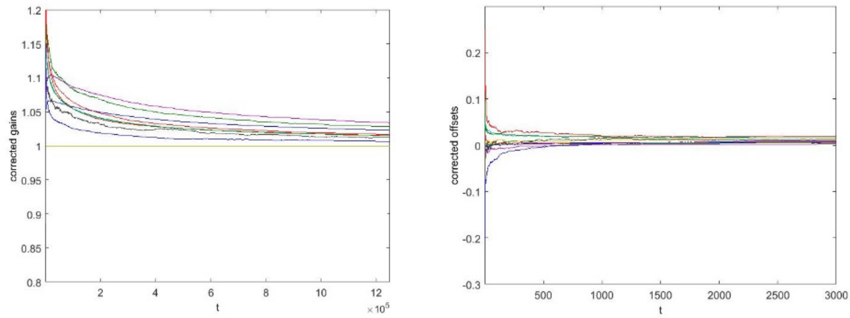
Figure 1 corresponds to the gain correction alone (offset is set to zero), with k=0.1 and $\mu = 0.6$, and noise standard deviation equal to 0.3. Figure 1a) is obtained by the algorithm with the instrumental variables defined above; the quality of the estimates is evidently high. It is obvious that the algorithm can handle the measurement noise efficiently. Figure 1b) illustrates the situation in which the instrumental variables are not applied (noisy gradient is applied directly); the corrected gains diverge, as a consequence of the correlated noise terms. Figure 2a) shows the behavior of the algorithm in the case when the network is pinned to node 1 ($g_1 = 1$ is taken as a reference). Convergence is visible, but the initial convergence rate is lower than in the case of the maximal number of degrees of freedom in adjusting calibration parameters (see also [14, 15, 16]).



a) b)

Figure 1

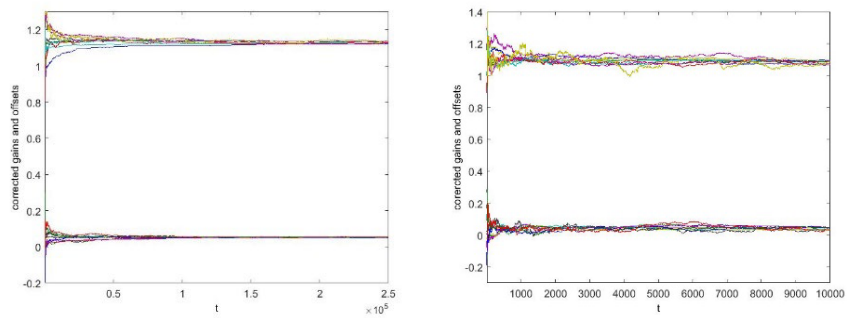
Corrected gains: a) algorithm with instrumental variables, b) gradient scheme



a) b)

Figure 2

a) corrected gains (node 1 taken as a reference), b) corrected offsets



a) b)

Figure 3

Simultaneously corrected gains and offsets: a) low level noise, b) high level noise

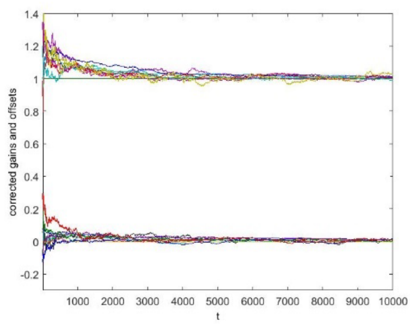


Figure 4

Simultaneously corrected gains and offsets: high-level noise, node 1 taken as reference

Figure 2b) corresponds to offset correction, when the corrected gain is fixed to one, under the same noise characteristics as in the case of Figure 1.

Simultaneous correction of both gains and offsets is theoretically in the focus of the whole paper. Figure 3a) corresponds to this situation, under the same noise characteristics as in the case of Figure 1. The estimation quality is high in spite of the noise presence. In Figure 3b) the noise standard deviation is increased ten times w.r.t. the conditions related to Figure 3a). The estimation quality is still completely acceptable, although the noise level surpasses all that can be expected in reality (having in mind the signal level). Figure 4 depicts the situation corresponding to Figure 3b), but when node 1 is taken as a reference. Notice that the estimates can be additionally smoothed by choosing parameters k and μ, but then the convergence rate may become much lower.

The highly acceptable character of the estimates in Figures 3 and 4 is a consequence of an achieved high asymptotic convergence rate in the theoretical sense.

Conclusion

In this paper, the distributed blind macro-calibration algorithm based on consensus proposed in [14, 15] is analysed in detail in the case of the presence of communication and measurement noise; the existing literature covers the convergence properties of this algorithm only partially, e.g. [15, 21].

The main original contribution of the paper is the formulation of the asymptotic convergence rate of the algorithm for gain correction, not yet existing in the literature. Convergence analysis of the gain correction parameters is done at the first step, showing that under very general conditions the *corrected gains become asymptotically equal w.p.1and in the m.s.s*. The expression for the *convergence rateof the gain correction algorithm* is derived using the general results of H.F. Chen in [17], adapted in an original way to the application to quadratic Lyapunov functions, The obtained formulation of the *convergence rate of the gain correction algorithm* enables obtaining a proof of convergence w.p.1 of the *offset correction algorithm*, including the formulation of the corresponding convergence rate. To our knowledge, such a result is not available in the literature.

The paper also contains an extension of the obtained general results to the case when the *observed sensor network contains a reference node*. An original proof is provided that the corrected gains and offsets converge in this case to the given reference values (under very general conditions).

A number of simulation results serve as an illustration of the typical behaviour of the proposed calibration algorithm under stochastic disturbances.

Further research could be directed to the generalization of the obtained results to more general sensor models involving nonlinearities; this is especially important in the case of the existence of a reference node, having in mind that the error due to nonlinearities may prevent ideal convergence to consensus. As stated above, it would be a challenge to apply the methodology used in the paper to the convergence rate analysis of the calibration algorithms proposed earlier in [9-12].

References

[1] Greer, C., Burns, M., Wollman, D., and E. Griffor, E.: *Cyber-Physical Systems and Internet of Things*. NIST Special Publication, 2019

[2] Wu, F. J., Kao,Y. F. and Tseng, Y. C.: From wireless sensor networks towards cyber physical systems," Pervasive and Mobile Computing, Vol. 7, pp. 397-413, 2011

[3] Gharavi, H. and Kumar, S., Eds., Proceedings of the IEEE: Special Isssue on Sensor Networks and Applications, Vol. 91, August 2003

[4] Akyildiz, T. F., Su, W. and Cayirci, E.: Wireless sensor networks: a survey, Computer Networks, Vol. 38, pp. 393-422, 2002

[5] Barcelo-Ordinas, J. M., Doudou, M., Garcia-Vidal, J. and Badache, N.: Self-calibration methods for uncontrolled environments in sensor networks: A reference survey, Ad Hoc Networks, Vol. 88, pp. 142-159, 2019

[6] Barcelo-Ordinas, J. M, Chanet, J.- P., Hou, K.-M. and Garcia-Vidal, J.: A survey of wireless sensor technologies applied to precision agriculture, Precision Agriculture, Vol. 13, pp. 801-808, 2013

[7] Whitehouse, K. and CullerD.: Macro-calibration in sensor/actuator networks,Mobile Network Applications, Vol. 8, pp. 463-472, 2003

[8] Balzano, L. and Nowak, R.: Blind calibration in sensor networks, in Proc. Intern. Conf. Inf. Proc. in Sensor Networks, April 2007

[9] Stanković, M. S., Stanković, S. S. and Johansson, K. J.: Distributed blind calibration in lossy sensor networks via output synchronization, IEEE Trans. Aut. Control, Vol. 60, pp. 3257-3262, 2015

[10] Stanković, M. S., Stanković, S. S. and Johansson, K. J.: Distributed calibration of sensor networks, in Proc. 20th Mediterranean Conference on Control and Automation, 2012

[11] Stanković, M. S., Stanković, S. S. and Johansson, K. J.: Asynchronous distributed blind calibration of sensor networks under noisy measurements, IEEE Transactions on Control of Network Systems, Vol. 5, No. 1, pp. 571-582, 2018

[12] Stanković, M. S., Stanković, S. S., Johansson, K. J., Beko, M. and Camarinha-Matos, L. M.: On consensus-based distributed blind calibration of sensor networks, Sensors, Vol. 18, No. 11, 2018

[13] Schenato, L. and Fiorentin, F.: Average timesynch: a consensus-based protocol for time synchronization in wireless sensor networks, Automatica, Vol. 47, No. 9, pp. 1878-1886, 2011

[14] Stanković, M. and Antić, D. S.: Distributed consensus-based calibration of networked control systems, Facta Universitatis, Aut. Control Robotics, Vol. 18, pp. 95-111, 2019

[15] Stanković, M.: Distributed asynchronous consensus-based algorithm for blind calibration of sensor networks with autonomous gain correction, IET Control Theory Appl., Vol. 12, pp. 2287-2293, 2018

[16] Stanković, M. and Antić, D. S.: Distributed non-linear robust consensusbased sensor calibration for networked control systems, IET Control Theory Appl., Vol. 14, pp. 1200-1208, 2020

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

[18] Huang, M. and Manton, J. H.: Stochastic approximation for consensus seeking: mean square and almost sure convergence, in *Proc. 46th IEEE Conf. Decision and Control*, 2007, pp. 306-311

[19] Godsil, C. and Royle, G.: Algebraic Graph Theory. New York: Springer Verlag, 2001

[20] Ljung, L.: System Identification - theory for the user. Englewood Cliffs, N. J.: Prentice Hall International, 1989

[21] Huang, M. and Manton, J. H.: 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

[22] Lin, Z.: Coupled dynamic systems: from structure towards stability and stabilizability, Ph.D. dissertation, University of Toronto, 2006

[23] Saridis, G.: Stochastic approximation methods for identification and control - a survey, IEEE Trans. Autom. Control, Vol. 19, pp. 798-809, 1974

[24] Polyak, B. T. and Tsypkin, Y. Z.: Pseudo gradient algorithms of adaptation and learning, Automation and Remote Control, No. 3, pp. 45-68, 1973

[25] Polyak, B. T.: Convergence and rate of convergence of iterative stochastic algorithms, I General case, Automation and Remote Control, No. 12, pp. 83-94, 1976

[26] Borkar, V. S.: Stochastic Approximation: a Dynamical Systems Viewpoint. Springer, 2009

[27] Horn, R. A. and Johnson, C. A.: Matrix Analysis. Cambridge Univ. Press, 1985

[28] Siljak, D.: Large Scale Dynamic Systems: Stability and Structure. North Holland, 1978