Asynchronous Distributed Blind Calibration of Sensor Networks Under Noisy Measurements

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

***Abstract*—In this paper, a novel distributed algorithm for asynchronous blind macro-calibration in sensor networks with noisy measurements is proposed. The algorithm is formulated as a set of instrumental variable type recursions for estimating parameters of sensor calibration functions. It is proved using asynchronous stochastic approximation arguments and properties of block-diagonally dominant matrices that the algorithm achieves asymptotic consensus for sensor gains and offsets in the mean square sense and with probability one. Recommendations for system design in terms of the choice of *a priori* tunable weights are provided. Special attention is paid to the situation when a subset of sensors in the network (reference sensors) remains with fixed characteristics. In the case of only one reference sensor, convergence of the remaining sensors to its characteristics is proved. In the case of more than one reference sensor, it is proved that the calibration parameters converge to points that depend only on the characteristics of the reference sensors and the network properties.**

I. Introduction

SENSOR *calibration* represents one of the most important challenges for the wide deployment of *large wireless sensor networks* [1], [2]. Individual calibration of each sensor (*micro-calibration*) is applicable only in relatively small sensor systems. In the case of larger networks, new concepts and methods are required, having in mind that many devices can be partially unobservable and operating in dynamically changing environments. Calibration of a network as a whole is the essence of *macro-calibration*, eliminating the need to separately calibrate every device (*e.g.*, [3], [4]). Methods for calibrating sensor networks without dependence on controlled stimuli are of significant interest for practice. This problem, referred to as *blind calibration*, is a difficult one, having some similarities with the problems of blind estimation, blind deconvolution and blind equalization (*e.g.*, [5]–[7] and references therein). In [8], [9] a centralized, non-recursive blind calibration algorithm has been proposed. In [10], centralized convex optimization approaches to blind estimation of calibration gains have been proposed using sparsity. Another approach to blind sensor calibration assumes that the deployment is very dense and is based only on pairwise inter-node communications [11]. There are also methods trying to cope with situations in which sensor network deployments may not meet the density requirements, as in [12]. In [13] and [14], a methodologically different approach has been adopted, based on formulating the calibration problem as an *asymptotic consensus* problem, solved using a distributed gradient-type recursive algorithm which ensures that the corrected offsets and gains converge to the same values, assuming noiseless sensor measurements and synchronous iterations requiring a common global clock. It bears a certain resemblance with the approaches to clock synchronization in sensor networks in [15]–[21]. Extended consensus algorithms have been applied to calibration and localization problems in wireless sensor networks [22], [23], and directly to calibration [24], [25], but within different contexts.

In this paper, we focus our main attention on two important extensions of the approach in [13]:

1) the proposed algorithm operates completely *asynchronously* by using a broadcast gossip communication scheme, without requiring any type of centralized command, information or clock;

2) starting from the assumption that the measurements are corrupted by *additive measurement noise*, it is shown that the gradient algorithm proposed in [13] is not applicable; instead, a new recursive algorithm of *instrumental variable* type is proposed for solving the calibration problem.

The new algorithm is in the form of *distributed recursions* for calibration parameter estimation controlled by local clocks ticking according to the local sleeping policies. The proposed method works under the usual sleeping policies for wireless sensor networks (*e.g.*, [26] and references therein). At a local clock tick, a sensor node wakes up and sends its corrected output to its out-neighbors, initiating them to acquire measurement data and update their calibration parameters. The proposed algorithm of instrumental variable type, aimed at coping with measurement noise influence in sensor calibration, can also be considered as an important and non-trivial extension of the standard (synchronous or asynchronous) dynamic *consensus* algorithms (*e.g.*, [27]–[34]).

By using a novel methodology dealing with block-diagonally dominant systems and asynchronous stochastic approximation arguments, it is proved, starting from general and practically

Manuscript received April 10, 2016; revised September 15, 2016; accepted October 28, 2016. Date of publication December 1, 2016; date of current version March 16, 2018. This work was supported by the EU FP7 Marie Curie CIG, Knut and Alice Wallenberg Foundation, and the Swedish Research Council. Recommended by Associate Editor S. Dey.

M. S. Stanković is with the Innovation Center, School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (e-mail: milos.stankovic@ic.etf.rs).

S. S. Stanković is with the School of Electrical Engineering, University of Belgrade, Serbia and the Vlatacom Institute, Belgrade (e-mail: stankovic@etf.rs).

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

Digital Object Identifier 10.1109/TCNS.2016.2633788

nonrestrictive assumptions, that the whole algorithm converges to consensus with probability one (w.p.1) and in the mean square sense, *i.e.*, the calibration is successful in the sense that asymptotically all sensors behave identically.

The paper covers also the following two new topics, important for practice:

3) Starting from the problem of distributed minimization as in [13], a completely new analysis is provided clarifying the influence of *a priori* chosen weights to the asymptotic behavior of the algorithm. The derived conclusions can serve as an indication of how these weights are to be chosen in accordance with the system requirements.

4) Special attention is paid to the calibration problem in which a subset of sensors remains with *fixed* parameters, when thus the network becomes *pinned* to these sensors. This situation arises in practice for instance when only a newly added subset of sensors has to be calibrated. It is proved that the algorithm again converges, w.p.1 and in the mean square sense, not to a consensus, but to certain points determined by the characteristics of the fixed sensors and the network properties. In the case of one reference sensor, which can be considered to be calibrated manually to ideal characteristics, all the remaining sensors converge to the desired calibration parameter values.

A simulation study illustrates the efficiency of the algorithm in the situations treated in the theoretical part. We also numerically compare our algorithm to the one in [17], representing a typical consensus-based algorithm proposed for calibration. We apply the scheme from [17] to the (synchronous version of) blind calibration problem treated in this paper, and show the superiority of our algorithm under measurement noise.

Preliminary results presented in [35] contain some main ideas and sketches of some of the proofs presented in this work, also treating the case when the measurement noise is present but assuming completely synchronous functioning of the underlying sensor network.

The rest of the paper is organized as follows. Section II deals with the problem formulation and the derivation of the main asynchronous calibration algorithm. In Section III convergence analysis is presented, and recommendations for the choice of *a priori* tunable network weights are provided. Section IV is devoted to the case when multiple nodes with fixed calibration parameters are present. In Section V we present some illustrative simulation results and numerically compare the proposed algorithm with the algorithm from [17] adapted to the calibration problem.

II. Problem Formulation and Calibration Algorithm

Consider n distributed sensors forming a network with the topology 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 (i, j) (node i sends messages to node j). Let $\mathcal{N}_i^{\text{in}} = \{j \in \mathcal{N} | (j, i) \in \mathcal{E}\}$ be the set of in-neighbors nodes of the i-th node, and $\mathcal{N}_i^{\text{out}} = \{j \in \mathcal{N} | (i, j) \in \mathcal{E}\}$ the set of its out-neighbors. Assume that the sensors can measure a continuous-time signal $x(t)$ at discrete instants t_k, $t_k \in \Re^+$, $k = 1, 2, \ldots$, $t_{k+1} > t_k$, producing the sensor outputs

$$y_i(t_k) = \alpha_i x(t_k) + \beta_i + \xi_i(t_k) \tag{1}$$

where the *gain* α_i and the *offset* β_i are unknown constants and $\xi_i(t_k)$ the measurement noise, $i = 1, \ldots, n$. Each sensor i applies an affine *calibration function* to $y_i(t_k)$, which produces the *corrected sensor output*

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

where a_i and b_i are the *calibration parameters*, $g_i = a_i \alpha_i$ is the *corrected gain* and $f_i = a_i \beta_i + b_i$ the *corrected offset*, $i = 1, \ldots, n$. In general, it is desirable to choose parameters a_i and b_i in such a way as to set g_i and f_i as close as possible to one and zero, respectively.

Starting from the general idea of *blind macro-calibration*, the aim is to construct an algorithm for *distributed real-time estimation of the calibration parameters* a_i and b_i which asymptotically provides identical behavior of all the sensors, without the knowledge of the signal $x(t)$ and without any kind of fusion center. In order to eliminate the need for a common global clock, we assume that each node $j \in \mathcal{N}$ has its own *local clock*, and we introduce a single *virtual clock* that ticks whenever any of the local clocks ticks. The time instant t_k in (1) can now be defined as the absolute time instant of the k-th tick of the virtual clock. To have a well-posed situation, we assume that the local clocks tick independently according to some random processes such that the intervals between any two consecutive ticks are finite w.p.1 (also in the limit when $k \to \infty$). We also assume that the unconditional probability that the j-th clock ticked at an instance t_k is $q_j > 0$ which does not depend on k. These conditions are satisfied if the local clocks tick according to independent Poisson processes with rates μ_j (as in, *e.g.*, [30], [36]); in this case the virtual global clock ticks according to a Poisson process with the rate $\sum_{j=1}^n \mu_j$ (see, *e.g.*, [26] and references therein for some common sleeping policies in wireless sensor networks). For the sake of clarity of analysis, we shall adopt the last assumption throughout the paper.

Let the discrete-time instants of ticks of the clock of a node j be denoted as t_l^j, $l = 1, 2, \ldots$. At each tick of its clock, node j broadcasts its *current corrected* output $z_j(t_l^j)$ (based on the current values of a_j and b_j) to its out-neighbors $i \in \mathcal{N}_j^{\text{out}}$. We assume that each link (j, i) is subject to *random dropouts*, so that each node $i \in \mathcal{N}_j^{\text{out}}$ hears the broadcast with probability $p_{ji} > 0$. Immediately after getting the message (practically at the same instant, having in mind the time constants of the signal and the communication/computation speed of the nodes), all the nodes which have received the broadcast calculate their own current corrected outputs $z_i(t_l^j)$ and update the values of their calibration parameters a_i and b_i. The process is repeated after each tick of any node in the network; we assume that only one clock can tick at a given time. Let $j(k)$ be the index of the agent that broadcasts at instants t_k, let $J(k)$ be the subset of the set of out-neighbors $i \in \mathcal{N}_{j(k)}^{\text{out}}$ that hear the broadcast, and let $x(k) = x(t_k) = x(t_l^{j(k)})$, $y_i(k) = y_i(t_k) = y_i(t_l^{j(k)})$, $y_j(k) = y_j(t_k) = y_{j(k)}(t_l^{j(k)})$, $z_i(k) = z_i(t_k) = z_i(t_l^{j(k)})$, $z_j(k) = z_j(t_k) = z_{j(k)}(t_l^{j(k)})$, $\xi_i(k) = \xi_i(t_k) = \xi_i(t_l^{j(k)})$, and $\xi_j(k) = \xi_j(t_k) = \xi_{j(k)}(t_l^{j(k)})$ for some l.

Following the idea presented in [13], we attach to each node $i \in \mathcal{N}$ a *local criterion*

$$F_i(\theta_i) = \sum_{j \in \mathcal{N}_i^{\text{in}}} \gamma_{ij} E\{(z_j(k) - z_i(k))^2\} \tag{3}$$

where $\theta_i = [a_i \;\; b_i]^T$, $E\{\cdot\}$ denotes the mathematical expectation, and $\gamma_{ij} > 0$ are *a priori* chosen scalar weights defining the relative importance of the in-neighboring nodes ($\gamma_{ij} = 0$ for $j \notin \mathcal{N}_i^{\text{in}}$). Following directly the line of thought of [13], we can formulate

$$\text{grad}_{\theta_i} F_i(\theta_i) = \sum_{j \in \mathcal{N}_i^{\text{in}}} \gamma_{ij} E\left\{(z_j(k) - z_i(k)) \begin{bmatrix} y_i(k) \\ 1 \end{bmatrix}\right\} = 0 \tag{4}$$

which may be used for generation of *stochastic gradient updates* for θ_i that utilize samples of $\text{grad}_{\theta_i} F_i(\theta_i)$ (*e.g.*, [13], [37]). However, it is immediately clear from (1) and (2) that both $z_i(k)$ and $y_i(k)$ contain the same noise term $\xi_i(k)$, and that, consequently, the resulting estimates would be biased even in the white noise case, having in mind that the term $E\{\xi_i(k)^2\} \neq 0$ appears explicitly in (4) (see the discussion below). In order to cope with this problem, we shall follow a more general methodology taken from *system identification*, and replace $y_i(k)$ in (4) by an *instrumental variable* $\zeta_i(k)$, which should be, in general, uncorrelated with $\xi_i(k)$, but correlated with $x(k)$ [37], [38]. Having in mind the whole distributed calibration setting, including the assumption that $\{\xi_i(k)\}$ is white, our choice of the local instrumental variable $\zeta_i(k)$ for node i is in the form of the *delayed measurement*

$$\zeta_i(k) = y_i(d_i(k))$$

where $d_i(k)$ is the iteration number that corresponds to the instant of *immediate past local measurement* performed at the node i. Therefore, we propose the following new procedure of instrumental variable type for updating the local calibration parameters:

$$\hat{\theta}_i(k) = \hat{\theta}_i(k-1) + \delta_i(k)\gamma_{i,j(k)}\epsilon_{i,j(k)}(k) \begin{bmatrix} y_i(d_i(k)) \\ 1 \end{bmatrix} \tag{5}$$

where:

1) $\hat{\theta}_i(k) = [\hat{a}_i(k) \;\; \hat{b}_i(k)]^T$,
2) $\delta_i(k)$ is the step size defined as $\delta_i(k) = \nu_i(k)^{-c}$, where $\nu_i(k) = \sum_{m=1}^{k} I\{i \in J(m)\}$ ($I\{\cdot\}$ denotes the indicator function) represents the number of updates of node i up to the instant k, $1/2 < c \leq 1$,
3) $\epsilon_{i,j(k)}(k) = \hat{z}_{j(k)}(k) - \hat{z}_i(k)$,
4) the current corrected outputs of nodes $j(k)$ and i are given by

$$\hat{z}_{j(k)}(k) = \hat{a}_{j(k)}(k-1)y_{j(k)}(k) + \hat{b}_{j(k)}(k-1) \tag{6}$$

$$\hat{z}_i(k) = \hat{a}_i(k-1)y_i(k) + \hat{b}_i(k-1). \tag{7}$$

It will be adopted that the initial conditions are $\hat{\theta}_i(0) = [1 \;\; 0]^T$. For all $i \notin J(k)$ the calibration parameters are not updated: $\hat{\theta}_i(k) = \hat{\theta}_i(k-1)$. The algorithm is, evidently, very simple, requiring just a few arithmetic operations per iteration. Notice that at each iteration it requires locally a signal sample, the value of the local instrumental variable and the current output of one of in-neighbors. Explicit knowledge of k (or $d_i(k)$) is not required. The pseudocode of the algorithm is presented as Algorithm 1.

Algorithm 1: *A*synchronous calibration algorithm.

for All the nodes $i \in \mathcal{N}$ **do**
  Initialize $\hat{a}_i(0) = 1$, $\hat{b}_i(0) = 0$, and $y_i(0) = 0$
end for
loop
  if Tick $t_k = t_l^{j(k)}$ of the local clock of a node $j(k) \in \mathcal{N}$ **then**
    Measure the current local sensor output $y_j(t_k) = y_j(k)$
    Calculate $\hat{z}_j(k)$ using (6)
    Broadcast $\hat{z}_j(k)$ to the out-neighbours $\mathcal{N}_j^{\text{out}}$
  end if
end loop
loop
  if A message is received by a node $i \in \mathcal{N}$ (from a node $j(k) \in \mathcal{N}_i^{\text{in}}$) **then**
    Measure the current local sensor output $y_i(t_k) = y_i(k)$
    Calculate $\hat{z}_i(k)$ using (7)
    Calculate new estimates of the calibration parameters according to (5)
  end if
end loop

Remark 1: It is to be noticed that $\delta_i(k)$ depends only on the locally available data, making the algorithm completely decentralized.

Remark 2: As in the synchronous operation mode treated in [13], the role of the weights γ_{ij} in (3), and, consequently, in (5), is to take into account relative importance of in-neighbors (see the discussion below).

Remark 3: Notice also that the instrumental variables in (5) can be chosen locally in different ways. One of possibilities easily implementable in practice is to choose $\zeta_i(k) = y_i(\bar{t}_l^{j,i})$, where $\bar{t}_l^{j,i}$ is the continuous time instant of an additional measurement performed locally at node i, shortly after the instant of the last update performed at node i. This scheme will not be incorporated in the general setting, since it would require a more complicated definition of the sequence of iteration numbers.

The underlying idea of the whole procedure based on (5) is to ensure that the estimates of all the local corrected gains $\hat{g}_i(k) = \hat{a}_i(k)\alpha_i$ and offsets $\hat{f}_i(k) = \hat{a}_i(k)\beta_i + \hat{b}_i(k)$ tend asymptotically to the *same values* $\bar{g}$ and $\bar{f}$, respectively, implying that in the limit $\hat{z}_j(k) = \hat{z}_i(k)$, $i, j = 1, \ldots, n$. Therefore, we introduce, for the sake of further analysis,

$$\hat{\rho}_i(k) = \begin{bmatrix} \hat{g}_i(k) \\ \hat{f}_i(k) \end{bmatrix} = \begin{bmatrix} \alpha_i & 0 \\ \beta_i & 1 \end{bmatrix} \hat{\theta}_i(k) \tag{8}$$

and

$$\epsilon_{i,j(k)}(k) = \begin{bmatrix} x(k) & 1 \end{bmatrix} (\hat{\rho}_{j(k)}(k) - \hat{\rho}_i(k)) + \hat{a}_{j(k)}(k)\xi_{j(k)}(k) - \hat{a}_i(k)\xi_i(k). \tag{9}$$

Consequently, we have

$$\hat{\rho}_i(k) = \hat{\rho}_i(k-1) + \delta_i(k)\gamma_{i,j(k)}\{(\Phi_i(k) + \Psi_i(k)) \times (\hat{\rho}_{j(k)}(k-1) - \hat{\rho}_i(k-1)) + N_{i,j(k)}(k)\hat{\rho}_{j(k)}(k-1) - N_{ii}(k)\hat{\rho}_i(k-1)\} \quad (10)$$

where

$$\Phi_i(k) = \begin{bmatrix} \alpha_i\beta_i x(k) + \alpha_i^2 x(k)x(d_i(k)) & \alpha_i\beta_i + \alpha_i^2 x(d_i(k)) \\ (1+\beta_i^2)x(k) + \alpha_i\beta_i x(k)x(d_i(k)) & 1 + \beta_i^2 + \alpha_i\beta_i x(d_i(k)) \end{bmatrix}$$

$$\Psi_i(k) = \xi_i(d_i(k)) \begin{bmatrix} \alpha_i x(k) & \alpha_i \\ \beta_i x(k) & \beta_i \end{bmatrix}$$

$$N_{i,j(k)}(k) = \frac{\xi_{j(k)}(k)}{\alpha_{j(k)}} \begin{bmatrix} \alpha_i y_i^0(d_i(k)) & 0 \\ \beta_i y_i^0(d_i(k)) & 0 \end{bmatrix} + \begin{bmatrix} \frac{\xi_{j(k)}(k)\xi_i(d_i(k))}{\alpha_{j(k)}} & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$N_{ii}(k) = \frac{\xi_i(k)}{\alpha_i} \begin{bmatrix} \alpha_i y_i^0(d_i(k)) & 0 \\ \beta_i y_i^0(d_i(k)) & 0 \end{bmatrix} + \begin{bmatrix} \frac{\xi_i(k)\xi_i(d_i(k))}{\alpha_i} & 0 \\ 0 & 0 \end{bmatrix}$$

where $y_i^0(k) = \alpha_i x(k) + \beta_i$, with the initial conditions $\hat{\rho}_i(0) = [\alpha_i \ \ \beta_i]^T$, $i \in J(k)$.

Recursion (10) for $i = 1, \ldots, n$, can be represented in the following compact form:

$$\hat{\rho}(k) = \{I + [\Phi(k) + \Psi(k)](\Delta(k)\Gamma(k) \otimes I_2) + (\Delta(k) \otimes I_2)\tilde{N}(k)\}\hat{\rho}(k-1) \quad (11)$$

where:

1) $\hat{\rho}(k) = [\hat{\rho}_1(k)^T \ldots \hat{\rho}_n(k)^T]^T$,
2) $\Delta(k) = \mathrm{diag}\{\delta_1(k), \ldots, \delta_n(k)\}$,
3) $\Phi(k) = \mathrm{diag}\{\Phi_1(k), \ldots, \Phi_n(k)\}$,
4) $\Gamma(k) = [\Gamma(k)_{lm}]$, with $\Gamma(k)_{ll} = -\gamma_{l,j(k)}$ and $\Gamma(k)_{l,j(k)} = \gamma_{l,j(k)}$ for all $l \in J(k)$, $\Gamma(k)_{lm} = 0$, otherwise,
5) $\Psi(k) = \mathrm{diag}\{\Psi_1(k), \ldots, \Psi_n(k)\}$,
6) $\tilde{N}(k) = [\tilde{N}_{lm}(k)]$, where $\tilde{N}_{ll}(k) = -\gamma_{l,j(k)}N_{ll}(k)$ and $\tilde{N}_{l,j(k)}(k) = \gamma_{l,j(k)}N_{l,j(k)}(k)$, for all $l \in J(k)$, $\tilde{N}(k)_{lm} = 0$, otherwise.

The initial condition is $\hat{\rho}(0) = [\hat{\rho}_1(0)^T \ldots \hat{\rho}_n(0)^T]^T = [[\alpha_1 \ \beta_1]^T \ldots [\alpha_n \ \beta_n]^T]^T$.

III. Analysis of the Algorithm

A. Convergence Proof

In the basic setting, we assume:

(A1) $\{x(k)\}$ is a stationary random sequence bounded w.p.1, satisfying the ϕ-mixing condition.

Remark 4: The ϕ-mixing condition is one of the strong mixing conditions (implying the frequently used α-mixing condition) [39]–[41]. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space, $\mathcal{F}_J^L$ the σ-field generated by $x(k)$, $J \le k \le L$ and $\phi(\nu) = \sup_M \phi(\mathcal{F}_{-\infty}^M, \mathcal{F}_{M+\nu}^\infty) = \sup_M \sup |P(B|A) - P(A)|$, $A \in \mathcal{F}_{-\infty}^M$, $B \in \mathcal{F}_{M+\nu}^\infty$, $\nu \ge 1$; for stationary sequences $M = 0$. The process is ϕ-mixing if $\lim_{\nu\to\infty} \phi(\nu) = 0$ [39].

(A2) Let $\{t^{i,l}\}$, $l = 1, 2, \ldots$ represent time instants in which node i measures the signal $x(t)$. Then, $\min_i \bar{r}_i > m^2$, where $m = E\{x(k)\}$ and $\bar{r}_i = E\{x(t^{i,l})x(t^{i,l-1})\}$, $i = 1, \ldots, n$.

Remark 5: Assumption (A2) represents an extension of the condition (A1.c) from [13], resulting from the introduction of delayed measurements $y_i(d_i(k))$ in (5). It implies that the variance of $x(k)$ is greater than zero (for all k, due to stationarity); hence, it ensures the persistence of excitation, not allowing, *e.g.*, constant signals [13], [37]. On the other hand, it does not allow, *e.g.*, white noise signals, having in mind the requirement that the instrumental variable should be correlated with the current measurement $x(k)$. Note that, due to stationarity of $\{x(t)\}$, $\bar{r}_i$ can be written as $E\{E\{x(0)x(\Delta t^i)|\Delta t^i\}\}$, where Δt^i is exponentially distributed with the mean depending on the rates μ_j and the probabilities p_{ji}, $j \in \mathcal{N}_i^{\mathrm{in}}$. Hence, for fixed rates μ_j, (A2) is always satisfied for processes $\{x(t)\}$ for which the autocovariance function $C_{xx}(\tau) = E\{x(0)x(\tau)\} - m^2$ is positive in a sufficiently wide interval about the origin. If the rates μ_j are adjustable and $C_{xx}(\tau)$ is continuous, it is always possible to choose $\mu_{\min} = \min_{j\in\mathcal{N}} \mu_j$ large enough, such that the assumption is always satisfied (since $C_{xx}(0) > 0$). Hence, the assumption is not restrictive for processes having low frequency spectrum, which are typical in practice (*e.g.*, continuous-time ARMA low pass processes [37]). We shall provide an additional insight into this subject in Remark 8.

(A3) $\{\xi_i(k)\}$, $i = 1, \ldots n$, are zero mean sequences of bounded w.p.1 and independent random variables, independent of the measured process $\{x(t)\}$, with $E\{\xi_i(k)^2\} = (\sigma_i^\xi)^2$ for all k.

(A4) Graph $\mathcal{G}$ has a spanning tree.

Remark 6: Assumptions (A3) and (A4) are standard for similar problems. Assumption (A4) implies that graph $\mathcal{G}$ has a center node from which all the remaining nodes are reachable.

Based on the given problem definition and the adopted assumptions, we are able to prove the following main result, stating that the calibration parameters will converge to the proper values. Having in mind the above discussion related to the choice of instrumental variables (Remark 3), w.l.o.g. we shall adopt, for the sake of more transparent analysis, that the estimate $\hat{\theta}_i(k-1)$ in (5) is connected to the instant $d_i(d_i(k))$, *i.e.*, to the measurement instant at node i preceding the instant $d_i(k)$.

Theorem 1: Let Assumptions (A1)–(A4) be satisfied. Then $\hat{\rho}(k)$ generated by (11) converges to $\hat{\rho}_\infty = \chi_1 v_1 + \chi_2 v_2$ in the mean square sense and w.p.1, where χ_1 and χ_2 are random variables with bounded second moments, $v_1 = \begin{bmatrix} 1 & 0 & 1 & 0 & \ldots & 1 & 0 \end{bmatrix}^T \in \mathbb{R}^{2n}$ and $v_2 = \begin{bmatrix} 0 & 1 & 0 & 1 & \ldots & 0 & 1 \end{bmatrix}^T \in \mathbb{R}^{2n}$.

Before proceeding to the proof of the theorem, we shall pay attention to several important prerequisites, related to the

asynchronous functioning of the algorithm and some structural properties of the matrices involved.

We first discuss asymptotics of the step-size matrix $\Delta(k)$. The following lemma represents a generalization of the result in [36].

Lemma 1: Let (A4) be satisfied, and let p_i be the unconditional probability of node i *to update* at the instant k, $i=1,\dots,n$. Then, for a given $q\in(0,\frac{1}{2})$, there exists such a $\bar{k}(q)$ that w.p.1 for all $k\geq\bar{k}$

$$\Delta(k)=\frac{1}{k^c}P^{-c}+\tilde{\Delta}(k) \tag{12}$$

where $P=\mathrm{diag}\{p_1,\dots,p_n\}$, $\tilde{\Delta}(k)=\mathrm{diag}\{\tilde{\delta}_1(k),\dots,\tilde{\delta}_n(k)\}$

$$|\tilde{\delta}_i(k)|\leq\bar{\delta}_i\frac{1}{k^{c+\frac{1}{2}-q}} \tag{13}$$

$0<\bar{\delta}_i<\infty$, $i=1,\dots,n$.

Proof: According to (A4), $p_i=\sum_{j\in\mathcal{N}_i^{\mathrm{in}}}q_jp_{ji}>0$, $i=1,\dots,n$. From Lemma 3 in [36] we have that w.p.1 for k large enough $|\frac{1}{\nu_i(k)}-\frac{1}{kp_i}|\leq\kappa\frac{1}{k^{\frac{3}{2}-q}}$, $\kappa>0$. Using this result, we obtain that there exist such constants $\bar{\delta}_i>0$ that w.p.1 for k large enough

$$\left|\frac{1}{\nu_i(k)^c}-\frac{1}{(kp_i)^c}\right|\leq\bar{\delta}_i\frac{1}{k^{c+\frac{1}{2}-q}}$$

which proves the lemma. ■

After applying Lemma 1 to (11), we realize that the convergence properties of the recursion essentially depend on matrix $B(k)=\Phi(k)(P^{-c}\Gamma(k)\otimes I_2)$ (both $\Psi(k)$ and $\tilde{N}(k)$ are noise terms). Therefore, we shall first pay attention to the matrix $\bar{B}=E\{B(k)\}=\bar{\Phi}(P^{-c}\bar{\Gamma}\otimes I_2)$, where $\bar{\Phi}=\mathrm{diag}\{\bar{\Phi}_1,\dots,\bar{\Phi}_n\}$, with

$$\begin{aligned}\bar{\Phi}_i&=E\{\Phi_i(k)\}\\&=\begin{bmatrix}\alpha_i\beta_im+\alpha_i^2\bar{r}_i & \alpha_i\beta_i+\alpha_i^2m\\(1+\beta_i^2)m+\alpha_i\beta_i\bar{r}_i & 1+\beta_i^2+\alpha_i\beta_im\end{bmatrix}\end{aligned}$$

and $\bar{\Gamma}=E\{\Gamma(k)\}$, which has the following form:

$$\bar{\Gamma}=\begin{bmatrix}-\sum_{j,j\neq1}\gamma_{1j}\pi_{1j} & \gamma_{12}\pi_{12} & \cdots & \gamma_{1n}\pi_{1n}\\ \gamma_{21}\pi_{21} & -\sum_{j,j\neq2}\gamma_{2j}\pi_{2j} & \cdots & \gamma_{2n}\pi_{2n}\\ & & \ddots & \\ \gamma_{n1}\pi_{n1} & \gamma_{n2}\pi_{n2} & \cdots & -\sum_{j,j\neq n}\gamma_{nj}\pi_{nj}\end{bmatrix}$$

($\gamma_{ij}=0$ when $j\notin\mathcal{N}_i$), where $\pi_{ij}=q_jp_{ji}$ is the probability that node i updates as a consequence of a tick of node j (note that $\pi_{ij}=0$ when $i\notin\mathcal{N}_j^{\mathrm{out}}$). Matrix $\bar{\Gamma}$ contains information on the network structure, including transmission probabilities, as well as on the relative importance of nodes for the whole calibration process (see Lemma 4 below).

The following lemmas follow methodologically [13]; however, the results presented here are new, since they are connected to a basically different setting, based on asynchronous communications.

Lemma 2: Let (A1), (A2) and (A4) be satisfied. Then $\bar{B}$ has two eigenvalues at the origin and the remaining ones in the left half complex plane.

Proof: We first observe that $-\bar{\Phi}_i$ is Hurwitz, because

$$\alpha_i^2(\bar{r}_i-m^2)>0,\quad 2\alpha_i\beta_im+\alpha_i^2\bar{r}_i+1+\beta_i^2>0, \tag{14}$$

under (A2). On the other hand, matrix $\bar{\Gamma}$, which has the form of a weighted Laplacian of $\mathcal{G}$, under (A4) has one eigenvalue at the origin and the remaining ones in the open left half complex plane (*e.g.*, [28]). Having in mind the block structure of $\bar{B}$, we use the basic result derived in [13], stating that a matrix $C=[C_{ij}]$, composed of the blocks $C_{ij}\in C^{m\times m}$, is Hurwitz, provided: 1) it has quasi dominating diagonal blocks, and 2) the blocks C_{ii} are Hurwitz, $i,j=1,\dots,n$ (matrices with quasi-dominating diagonal blocks are defined and discussed in, *e.g.*, [42], [43]). Thus, the result follows. ■

The following lemma enables application of the result of Lemma 2 to the analysis of the recursion (11). Note that vectors v_1 and v_2 represent right eigenvectors of $\bar{B}$ corresponding to the zero eigenvalue. Let w_1 and w_2 be the corresponding normalized left eigenvectors, satisfying $\begin{bmatrix}w_1\\\cdots\\w_2\end{bmatrix}[v_1\vdots v_2]=I_2$.

Lemma 3: The left eigenvectors w_1 and w_2 of $\bar{B}$ corresponding to the zero eigenvalue depend only on the sensor and network parameters.

If $T=\left[v_1\vdots v_2\vdots T_{2n\times(2n-2)}\right]$, 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) $T^{-1}=\begin{bmatrix}w_1\\\cdots\\w_2\\\cdots\\S_{(2n-2)\times2n}\end{bmatrix}$, where $S_{(2n-2)\times2n}$ follows from the definition of T, and
3) $T^{-1}\bar{B}T=\begin{bmatrix}0_{2\times2} & \vdots & 0_{2\times(2n-2)}\\\cdots & \vdots & \cdots\\0_{(2n-2)\times2} & \vdots & \bar{B}^*\end{bmatrix}$, where $\bar{B}^*$ is Hurwitz.

Proof: Define $\psi_i^{[1]}=[\alpha_i\beta_i\quad 1+\beta_i^2]^T$ and $\psi_i^{[2]}=[\alpha_i^2\ \alpha_i\beta_i]^T$, so that one can write

$$\bar{\Phi}_i=[\psi_i^{[1]}m+\psi_i^{[2]}\bar{r}_i\vdots\psi_i^{[1]}+\psi_i^{[2]}m].$$

If $w=[w^{[1]},\dots,w^{[n]}]$, where $w^{[1]},\dots,w^{[n]}$ are 2-D row-vectors, the equation $w\bar{B}=0$ gives the following set of equations:

$$-w^{[i]}\bar{\Phi}_ip_i^{-c}\sum_{j=1,j\neq i}^{n}\gamma_{ij}\pi_{ij}+\sum_{l=1,l\neq i}^{n}w^{[l]}\bar{\Phi}_l\gamma_{li}\pi_{li}p_l^{-c}=0 \tag{15}$$

equivalent to

$$-w^{[i]}\psi_i^{[1]}p_i^{-c}\sum_{j,j\neq i}\gamma_{ij}\pi_{ij}+\sum_{l,l\neq i}w^{[l]}\psi_l^{[1]}\gamma_{li}\pi_{li}p_l^{-c}=0$$

$$-w^{[i]}\psi_i^{[2]}p_i^{-c}\sum_{j,j\neq i}\gamma_{ij}\pi_{ij}+\sum_{l,l\neq i}w^{[l]}\psi_l^{[2]}\gamma_{li}\pi_{li}p_l^{-c}=0$$

$i = 1, \ldots, n$, which, obviously, do not depend on $x(k)$; this proves the first part.

Assertions 1), 2), and 3) follow from the Jordan decomposition. ■

It is important to notice that

$$T^{-1}B(k)T = \left[\begin{array}{c:c} 0_{2\times 2} & B_1(k) \\ \hdashline 0_{(2n-2)\times 2} & B_2(k) \end{array}\right] \tag{16}$$

where $B_1(k)$ and $B_2(k)$ are $2 \times (2n-2)$ and $(2n-2) \times (2n-2)$ matrices, respectively; this property will be utilized below. We now proceed with the proof of Theorem 1.

Proof of Theorem 1: In order to achieve a more compact notation, we shall introduce the sequence $\{\tau_i(k)\}$, $i = -1, 0, 1, 2, \ldots$, where $\tau_{-1}(k) = k+1$, $\tau_0(k) = k$, $\tau_1(k) = d(k)$, $\tau_2(k) = d(d(k)), \ldots$, etc. and $d(k) = \min_i(d_i(k))$, also we shall w.l.o.g. replace $\hat{\rho}(k-1)$ in (11) by $\hat{\rho}(\tau_2(k))$.

Introduce $\tilde{\rho}(k) = T^{-1}\hat{\rho}(k)$, where T is defined in Lemma 3. Using (11), Lemma 3 and (16), we obtain, in a similar way as in [13], that $\tilde{\rho}(k)$ can be decomposed as $\tilde{\rho}(k) = [\tilde{\rho}(k)^{[1]T} \;\; \tilde{\rho}(k)^{[2]T}]^T$, where

$$\tilde{\rho}(k)^{[1]} = [\tilde{\rho}_1(k) \;\; \tilde{\rho}_2(k)]^T$$

and

$$\tilde{\rho}(k)^{[2]} = [\tilde{\rho}_3(k) \ldots \tilde{\rho}_{2n}(k)]^T.$$

For k large enough we have, according to Lemma 1, the following recursions:

$$\tilde{\rho}(k)^{[1]} = \tilde{\rho}(\tau_2(k))^{[1]} + \left[\frac{1}{k^c}B_1(k) + F_1(k)\right]\tilde{\rho}(\tau_2(k))^{[2]} + \frac{1}{k^c}H_1(k)\tilde{\rho}(\tau_2(k)) \tag{17}$$

$$\tilde{\rho}(k)^{[2]} = \left[I + \frac{1}{k^c}B_2(k) + F_2(k)\right]\tilde{\rho}(\tau_2(k))^{[2]} + \frac{1}{k^c}H_2(k)\tilde{\rho}(\tau_2(k)), \tag{18}$$

where $F_1(k)$ and $F_2(k)$ are $2 \times (2n-2)$ and $(2n-2) \times (2n-2)$ matrices, respectively, resulting from

$$T^{-1}\Phi(k)(\tilde{\Delta}(k)\Gamma(k) \otimes I_2)T = \left[\begin{array}{c:c} 0_{2\times 2} & F_1(k) \\ \hdashline 0_{(2n-2)\times 2} & F_2(k) \end{array}\right]$$

while $H_1(k)$ and $H_2(k)$ are $2 \times 2n$ and $(2n-2) \times 2n$ matrices, respectively, defined by

$$T^{-1}\{\Psi(k)(P^{-c}\Gamma(k) \otimes I_2) + (P^{-c} \otimes I_2)\tilde{N}(k) + k^c[\Psi(k)(\tilde{\Delta}(k)\Gamma(k) \otimes I_2) + (\tilde{\Delta}(k) \otimes I_2)\tilde{N}(k)]\}T = \left[\begin{array}{c} H_1(k) \\ \hdashline H_2(k) \end{array}\right]. \tag{19}$$

Obviously, both $H_1(k)$ and $H_2(k)$ are zero mean and uncorrelated with $\tilde{\rho}(\tau_2(k))$, having in mind (A3).

In order to present the main line of thought, we start the analysis from (18). In order to handle the correlation in the sequence $\{x(k)\}$, we iterate the recursion (18) $n_k > 0$ times backwards, and obtain

$$\tilde{\rho}(k)^{[2]} = \Pi(k, \tau_{2n_k}(k))\tilde{\rho}(\tau_{2(n_k+1)}(k))^{[2]} + \sum_{\sigma=0}^{n_k}\frac{1}{\tau_{2\sigma}(k)^c}\Pi(k, \tau_{2(\sigma-1)}(k))H_2(\tau_{2\sigma}(k)) \times \tilde{\rho}(\tau_{2(\sigma+1)}(k)) \tag{20}$$

where $\Pi(k, \tau_{2l}(k)) = \prod_{\sigma=0}^{l}(I + \frac{1}{\tau_{2\sigma}(k)^c}B_2(\tau_{2\sigma}(k)) + F_2(\tau_{2\sigma}(k)))$, with $\Pi(k, k+1) = I$.

It follows from (A4) and Lemma 2 that $\bar{B}^* = E\{B_2(k)\}$ is Hurwitz; therefore, there exists a symmetric positive definite matrix R, such that

$$R\bar{B}^* + \bar{B}^{*T}R = -Q \tag{21}$$

for any given $Q > 0$.

Denote

$$V(k) = E\{\|\tilde{\rho}(k)^{[1]}\|^2\}$$

and

$$W(k) = E\{\tilde{\rho}(k)^{[2]T}R\tilde{\rho}(k)^{[2]}\}$$

where R follows from (21) for an *a priori* chosen $Q > 0$.

Calculation of $W(k)$ from (20) is straightforward, having in mind (A3), and that $H_2(\tau_{2\sigma}(k))\tilde{\rho}(\tau_{2(\sigma+1)}(k)) = 0$ in (20). The terms linearly depending on $\frac{1}{s^c}$, $s = k, \tau_2(k), \ldots, \tau_{2n_k}(k)$, are dominant in $E\{\Pi(k, \tau_{2n_k}(k))^T R\Pi(k, \tau_{2n_k}(k))|\mathcal{F}(\tau_{2(n_k+1)}(k))\}$ following from (20) for k large enough, where $\mathcal{F}(\tau_{2(n_k+1)}(k))$ is the minimal σ-algebra generated by the measurements up to the instant $\tau_{2(n_k+1)}(k)$. After introducing $B_2(k) = \bar{B}^* + \tilde{B}_2(k)$, we obtain the following expression for these terms:

$$E\left\{\sum_{s=\tau_{2n_k}(k)}^{\tau_0(k)}\frac{1}{s^c}(B_2(s)^T R + RB_2(s))|\mathcal{F}(\tau_{2(n_k+1)}(k))\right\} = -Q\sum_{s=\tau_{2n_k}(k)}^{\tau_0(k)}\frac{1}{s^c} + E\left\{\sum_{s=\tau_{2n_k}(k)}^{\tau_0(k)}\frac{1}{s^c}(\tilde{B}_2(s)^T R + R\tilde{B}_2(s))|\mathcal{F}(\tau_{2(n_k+1)}(k))\right\}. \tag{22}$$

According to the mixing condition in (A1), it follows that

$$\max_i |\lambda_i(E(\tilde{B}_2(s)^T R + R\tilde{B}_2(s))|\mathcal{F}(\tau_{2(n_k+1)}(k)))| \leq \varphi(s - \tau_{2(n_k+1)}(k)) \tag{23}$$

where $\tau_{2n_k}(k) \leq s \leq k$, $\varphi(s) \geq 0$ and $\lim_{s\to\infty}\varphi(s) = 0$, having in mind that $E\{\tilde{B}_2(s)\} = 0$. Therefore, there exists w.p.1 for k large enough such an n_k^0 that for all $n_k \geq n_k^0$

$$\sum_{s=\tau_{2n_k}(k)}^{\tau_0(k)}\frac{1}{s^c}\left[\lambda_{\min}(Q) - \varphi(s - \tau_{2(n_k+1)}(k))\right] > \epsilon\frac{1}{k^c} \tag{24}$$

w.p.1 for some $\epsilon > 0$, since $\lambda_{\min}(Q) > 0$ by definition.

After some additional technicalities, we obtain, for k large enough, that

$$W(k) \leq \left(1 - c_0 \frac{1}{k^c}\right) W(\tau_{2(n_k+1)}(k)) + C_1 \sum_{s=\tau_{2n_k}(k)}^{\tau_0(k)} \frac{1}{s^{1+q'}} (V(s) + W(s)) \quad (25)$$

where $c_0 > 0$, $C_1 > 0$ and $q' > 0$ are generic constants (notice that $\|F_2(k)\| = o(\frac{1}{k^{1+q}})$ w.p.1, where $q > 0$).

Similarly, from (17) we can obtain, in an analogous way

$$V(k) \leq V(\tau_2(k)) + C_2 \frac{1}{k^{1+q''}} (V(\tau_2(k)) + W(\tau_2(k))) \quad (26)$$

where $C_2 > 0$ and $q'' > 0$.

Since $\sum_{s=1}^{\infty} \frac{1}{s^c} = \infty$, $\sum_{s=1}^{\infty} \frac{1}{s^{1+q'}} < \infty$ and $\sum_{s=1}^{\infty} \frac{1}{s^{1+q''}} < \infty$, recursions (25) and (26) can be treated in a similar way as the analogous recursions analyzed in [34, Lemma 12 and Theorem 11] (see also [44]–[46]). They give rise to the conclusion that $\tilde{\rho}(k)^{[1]}$ tends to a vector random variable $[\chi_1 \ \chi_2]^T$ and $\tilde{\rho}(k)^{[2]}$ to zero in the mean square sense and w.p.1. It follows that

$$\hat{\rho}_\infty = T \begin{bmatrix} \lim\limits_{k\to\infty} \tilde{\rho}(k)^{[1]} \\ \cdots\cdots\cdots\cdots \\ 0 \end{bmatrix} = \chi_1 v_1 + \chi_2 v_2 \quad (27)$$

which proves the theorem. ■

We shall further clarify properties of χ_1 and χ_2 within the scope of three characteristic situations:

a) $\xi_i(k) = 0, i = 1, \ldots, n$. When there is no measurement noise, the results presented in [13] can be directly extended to the case of asynchronous operation. Then, the main algorithm (5) becomes a gradient scheme in which $\zeta_i(k) = y_i(k)$, and the step size becomes constant: $\delta_i(k) = \delta$. Then, it is possible to show, using some elements of the proof of Theorem 1, that for δ small enough

$$\lim_{k\to\infty} \tilde{\rho}(k)^{[1]} = \begin{bmatrix} w_1 \\ \cdots\cdots \\ w_2 \end{bmatrix} \hat{\rho}(0) + \tilde{\rho}_\infty \quad (28)$$

in the mean square sense and w.p.1, where $\tilde{\rho}_\infty = \lim_{k\to\infty} \sum_{s=1}^{k} B_1(s) \tilde{\rho}(s)^{[2]}$. The first term at the right-hand side of (28) is deterministic, and depends on the properties of the actual sensors and $\bar{B}$, and the second term stochastic, depending on on the asynchronous communication scheme and all the previous estimates $\tilde{\rho}(k)^{[2]}$. Since, in this case, $\tilde{\rho}(k)^{[2]}$ tends to zero exponentially in the mean square sense and w.p.1, $\tilde{\rho}_\infty$ remains bounded in the mean square sense and w.p.1.

b) $\tilde{\Delta}(k) = 0$. In this case, $F_1(k) = 0$; according to Theorem 1

$$\lim_{k\to\infty} \tilde{\rho}(k)^{[1]} = \begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix} \quad (29)$$

where χ_1 and χ_2 are random variables, satisfying $E\left\{\begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix}\right\} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \hat{\rho}(0) + E\{\tilde{\rho}_\infty\}$ and $E\{\chi_1^2 + \chi_1^2\} < \infty$. This case corresponds to the situation in which the step size $\delta_i(k)$ is not defined locally, but on the basis of the readings of a centralized clock specifying k (as it has been assumed in [35], in relation with the influence of additive communication noise).

c) In the most general case, the situation is more complex, since $\lim_{k\to\infty} \tilde{\rho}(k)^{[1]}$ contains an additional term depending on $F_1(s)\tilde{\rho}(\tau_2(s))^{[2]}$, $s \leq k$, expectation of which is not equal to zero, but which has finite second moment, having in mind Lemma 1, the fact that $\sum_{s=1}^{k} \frac{1}{s^{c+\frac{1}{2}-q}}$ converges and that $\tilde{\rho}(\sigma)^{[2]}$ tends to zero in the mean square sense and w.p.1.

Remark 7: Stationarity of the random process $\{x(t)\}$, assumed in (A1), is not essential in real applications. One should take into account that, formally speaking, Lemma 3 implies that a constant decoupling transformation T can be applied within the scope of the convergence analysis even when we have a time-varying matrix $\bar{B}(k)$. In addition, an insight into the given convergence proof shows that the result of Theorem 1 hold for changes of $\bar{B}(k)^*$ slow enough. For example, if $R(k) > 0$ is a unique solution of $\bar{B}(k)^{*T} R(k) + R(k)\bar{B}(k)^* = -Q(k)$ for a preselected $Q(k) > 0$, then by defining $W(k) = E\{\tilde{\rho}(k)^{[2]T} R(k) \tilde{\rho}(k)^{[2]}\}$ one obtains the result of Theorem 1 for $\|R(k) - R(k-1)\|$ small enough. Moreover, it is not difficult to prove that the result of Theorem 1 holds when the signal is only asymptotically stationary.

Remark 8: According to the proof of Theorem 1 and Lemma 1, Assumption (A2) ensures the main contraction properties of the algorithm. An idea that partially overcomes the related restriction is based on a different definition of the step-size of the algorithm. One can adopt, in accordance with the general structure of instrumental variable algorithms for system identification [37], [47], that $\delta_i(k) = \mathrm{diag}\{\delta_i^{[1]}(k), \delta_i^{[2]}(k)\}$ in (5), with

$$\delta_i^{[1]}(k) = \left(\sum_{\mu=1}^{k} y_i(\mu) y_i(d_i(\mu)) I\{i \in J(\mu)\} \right)^{-c}$$

and $\delta_i^{[2]}(k) = \nu_i(k)^{-c}$, where $\nu_i(k) = \sum_{\mu=1}^{k} I\{i \in J(\mu)\}$ (as in (5)). Having in mind that strong mixing implies ergodicity [48], we have that $\sum_{m=1}^{k} y_i(m) y_i(d_i(m)) I\{i \in J(m)\} = \sum_{\mu=1}^{l} y_i^{[i]}(\mu) y_i^{[i]}(\mu - 1) = p_i k E\{y_i^{[i]}(\mu) y_i^{[i]}(\mu - 1)\} + o(1) = p_i k \bar{\gamma}_i + o(1)$, where $\bar{\gamma}_i = \alpha_i^2 \bar{r}_i + 2\alpha_i \beta_i m + \beta_i^2$, $y_i^{[i]}(\mu)$ is defined in the same way as $x^{[i]}(\mu)$, and $o(1)$ tends to zero when k tends to infinity. According to Lemma 2, we have now matrix $-\mathrm{diag}\{\bar{\gamma}_i^{-1}, 1\}\bar{\Phi}_i$ instead of $-\bar{\Phi}_i$; this matrix is Hurwitz when either

$$\alpha_i^2(\bar{r}_i - m^2) > 0, \quad \bar{\gamma}_i > 0$$
$$2\alpha_i\beta_i m + \alpha_i^2 \bar{r}_i + \bar{\gamma}_i + \beta_i^2 > 0 \quad (30)$$

or

$$\alpha_i^2(\bar{r}_i - m^2) < 0, \quad \bar{\gamma}_i < 0$$
$$2\alpha_i\beta_i m + \alpha_i^2 \bar{r}_i + \bar{\gamma}_i + \beta_i^2 < 0. \quad (31)$$

It is easy to show that (A2) implies the first set of inequalities (30). In the case when $\bar{r}_i - m^2 < 0$, inequalities (31) give the following inequalities: $\alpha_i^2 \bar{r}_i + 2\alpha_i\beta_i m + \beta_i^2 < 0$ and $\bar{r}_i < -\frac{\beta_i}{\alpha_i}(2m + \frac{\beta_i}{\alpha_i})$. These inequalities enlarge the

theoretically permissible region for $\bar{r}_i$. In any case, Lemma 2 and this remark deal with sufficient conditions. In practice these conditions hold for moderately slow signal changes with respect to the sampling rate.

Remark 9: From the asymptotic equations derived in the proof of the above theorem we can conclude that the asymptotic *convergence rate* of the analyzed algorithm follows general statements related to the stochastic approximation algorithms (*e.g.*, [44]), *i.e.*, it can be shown that the best achievable convergence rate to consensus is $o(1/k^{c\eta})$ (see (12)), with $0 < \eta < 1$.

B. Parameter Tuning

Roughly speaking, the underlying idea of the whole method is to achieve the calibration goal by exploiting sensors with good characteristics in a large sensor network. There are two main possibilities based on the choice of the weights γ_{ij}: 1) to rely on the majority of good sensors, when all γ_{ij} in any neighborhood of node i can take the same value, or 2) to emphasize the effect of *a priori* selected good sensors belonging to a set $\mathcal{N}^f \subset \mathcal{N}$ by setting appropriate values of γ_{ij}. In this section we will give a more rigorous insight into parameter tuning problem, while Section IV will be devoted to the situation in which a set of reliable sensors has fixed sensor and calibration parameters.

The following theorem formally deals with properties of the limit values χ_1 and χ_2 in (27) and their dependence on the preselected elements of $\bar{B}$.

Theorem 2: Let (A1), (A2) and (A4) be satisfied. Let $\lambda\bar{\Phi}(P^{-c}\bar{\Gamma} \otimes I_2) = 0$, where (according to the notation in Lemma 3), $\lambda = [\lambda_1 \dots \lambda_{2n}] = [\lambda^{[1]} \dots \lambda^{[n]}]$, where $\lambda^{[i]} = [\lambda_{2i-1}\ \lambda_{2i}]$, $i = 1, \dots, n$, and let $\lambda'\bar{\Phi}(P^{-c}\bar{\Gamma}' \otimes I_2) = 0$, where, similarly, $\lambda' = [\lambda'_1 \dots \lambda'_{2n}] = [\lambda'^{[1]} \dots \lambda'^{[n]}]$, $\lambda'^{[i]} = [\lambda'_{2i-1}\ \lambda'_{2i}]$. Assume that $\bar{\Gamma}'$ is defined in one of the following two ways:

1) $\bar{\Gamma}' = D\bar{\Gamma}$, where $D = \mathrm{diag}\{d', 1, \dots, 1\}$, $0 < d' < 1$;
2) $\bar{\Gamma}'$ is obtained from $\bar{\Gamma}$ by replacing γ_{j1} with $d''\gamma_{j1}$, $j = 2, \dots, n$, $d'' > 1$.

Then

$$\frac{\lambda'_k}{\lambda'_j} < \frac{\lambda_k}{\lambda_j} \tag{32}$$

$j = 1, 2$, $k = 3, \dots, 2n$.

Proof: In case 1) we have that $\lambda\bar{\Phi}(D^{-1}DP^{-c}\bar{\Gamma} \otimes I_2) = 0$, wherefrom it follows that $\lambda' = \lambda(D^{-1} \otimes I_2)$, and (32) directly follows.

In case 2), define $\lambda_d = [\lambda^{[2]} \dots \lambda^{[n]}]$ and $\lambda'_d = [\lambda'^{[2]} \dots \lambda'^{[n]}]$; also, let $\bar{\Phi}_d$ be the $(2n-2) \times (2n-2)$ submatrix of $\bar{\Phi}$ with indices $i, j = 2, \dots, 2n$ and let P_d, $\bar{\Gamma}_d$ and $\bar{\Gamma}'_d$ be $(n-1) \times (n-1)$ submatrices of P, $\bar{\Gamma}$ and $\bar{\Gamma}'$, respectively, with indices $i, j = 2, \dots, n$. Then, we write $\lambda'_d = \lambda_d + \Delta\lambda'_d$ and

$$\bar{\Gamma}'_d = \bar{\Gamma}_d + (1 - d'')\mathrm{diag}\{\gamma_{21}\pi_{21}, \dots, \gamma_{n1}\pi_{n1}\}$$

[the last relation follows from the adopted assumption for case 2)]. After adopting $\lambda^{[1]}\bar{\Phi}_1 = \lambda'^{[1]}\bar{\Phi}'_1 = [1\ 1]$ (having in mind that $\mathrm{rank}\{\bar{B}\} = 2n - 2$, according to Lemma 3), from $\lambda'\bar{\Phi}(P^{-c}\bar{\Gamma}' \otimes I_2) = 0$ we obtain

$$\begin{aligned}(\lambda_d + \Delta\lambda'_d)\bar{\Phi}_d[(\bar{\Gamma}_d &+ (1 - d'')\mathrm{diag}\{\gamma_{21}\pi_{21}, \dots, \gamma_{n1}\pi_{n1}\}) \otimes I_2] \\ &= -([\gamma_{12}\pi_{12} \cdots \gamma_{1n}\pi_{1n}] \otimes [1\ 1]).\end{aligned} \tag{33}$$

Since $\lambda_d\bar{\Gamma}'_d = -([\gamma_{12}\pi_{12} \cdots \gamma_{1n}\pi_{1n}] \otimes [1\ 1])$ we obtain

$$\Delta\lambda'_d = -(1 - d'')\lambda_d[(\mathrm{diag}\{\gamma_{21}\pi_{21}, \dots, \gamma_{n1}\pi_{n1}\}(\bar{\Gamma}'_d)^{-1}) \otimes I_2]$$

wherefrom we conclude that all elements of $\Delta\lambda'_d$ are strictly negative, because of the properties of M-matrices, which imply that all the elements of $(\bar{\Gamma}'_d)^{-1}$ are non-positive [42]. This completes the proof. ■

According to the previous subsection, the dominant component of the random variables $[\chi_1\ \chi_2]$ (see (28)) is represented by a weighted sum of the sensor parameters α_i and β_i, with positive weights defined by the elements of the left eigenvectors w_1 and w_2 of $\bar{B}$ corresponding to the zero eigenvalue (see Lemma 3). These weights are functions of *a priori* chosen weights γ_{ij}, as well as of the probabilities π_{ij} and p_i, $i, j = 1, \dots, n$. In this sense, it is clear that the influence of a selected sensor i on the limit $[\chi_1\ \chi_2]$ can be increased by increasing the corresponding elements of w_1 and w_2 (components with indices $2i$ and $(2i+1)$). According to Theorem 2, this can be achieved by either:

1) decreasing all the elements of the i-th row of $\bar{\Gamma}$, or
2) increasing the elements γ_{ji}, $j \neq i$, from the i-th column (preserving, at the same time, that $\bar{\Gamma}$ is row stochastic).

Obviously, γ_{ij} can be selected accordingly. Probabilities π_{ij} depend on the clock rate of sensor j and the transmitting probabilities p_{ij}; the higher clock rate of the clock j, the higher its influence on the parameter values at consensus. In addition, influence of P^{-c} is such that the lower is the updating probability of a selected node, the higher is its influence on the limit value. Consequently, the design of the whole calibration process is flexible and can be adapted to the desired convergence points of the algorithm.

IV. Macro Calibration With Fixed Characteristics Sensors

According to Section III, the choice of the elements of the matrix $\bar{\Gamma}$ plays an important role in achieving the desired performance of the proposed method in practice. Moreover, in the limit, "good" sensors ("leaders") from a set $\mathcal{N}^f \subset \mathcal{N}$ can be left unchanged, *i.e.*, with fixed characteristics, so that the recursions (5) are applied only to the nodes $i \in \mathcal{N} - \mathcal{N}^f$. This situation can often arise in practice when a sensor network has to be enlarged, *i.e.*, when a set of uncalibrated sensors has to be added to a set of sensors already functioning in a satisfactory way. In this section, we shall present a rigorous analysis of the behavior of the proposed calibration algorithm in this case.

We shall first pay attention to the important special case when $|\mathcal{N}^f| = 1$, *i.e.*, when one sensor is well calibrated using any preselected method, and when one wishes to calibrate all the remaining sensors in such a way as to achieve the same values of the corrected parameters. Assume, without loss of generality,

that the first node is selected as a reference node ("leader"), and that it represents a center node of the underlying graph $\mathcal{G}$; this node has to be left unchanged within the calibration recursions (11). The whole calibration network becomes "pinned" to that sensor by simply setting

$$\hat{\theta}_1(k) = \hat{\theta}_1(k-1) \tag{34}$$

with $\hat{\theta}_1(0) = \hat{\theta}_{1,0}$; in the ideal case, $\hat{\rho}_{1,0} = \begin{bmatrix} \alpha_1 & 0 \\ \beta_1 & 1 \end{bmatrix} \hat{\theta}_{1,0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, where $\hat{\theta}_{1,0}$ and $\hat{\rho}_{1,0}$ are the corresponding initial conditions. According to (34), the general form of the algorithm (11) should be modified by setting to zero all the block matrices in the first block row of $B(k)$, or by setting $\gamma_{1j} = 0, j = 1, \ldots, n$. It is easy to observe that the resulting communication graph still has a spanning tree with the first node as a center node, implying that (A4) holds. Consequently, Theorem 1 can be applied, leading to the conclusion that, in the limit, the corrected calibration parameters $\hat{\rho}_i(k)$, $i = 1, \ldots, n$, should be the same, and equal to the value $\hat{\theta}_{1,0}$, imposed by the "leader".

In general, assume w.l.o.g. that $\mathcal{N}^f = \{1, 2, \ldots, n_f\}$, $n_f = |\mathcal{N}^f| > 0$, is the subset of "leaders", *i.e.*, of the sensors with arbitrary, but fixed characteristics, defined by $\rho_i^f = \begin{bmatrix} g_i^f \\ f_i^f \end{bmatrix}, i \in \mathcal{N}^f$; let $\bar{\rho}^f = [\rho_1^{fT} \ \ldots \ \rho_{n_f}^{fT}]^T$. The calibration algorithm proposed above can be applied in this situation by introducing $\hat{\theta}_i(k) = \hat{\theta}_i(k-1)$ for all $i \in \mathcal{N}^f$. Let $\mathcal{N} - \mathcal{N}^f = \{n_f + 1, \ldots, n\}$ and let $\hat{\rho}^v(k) = [\hat{\rho}_{n_f+1}(k)^T \ \ldots \ \hat{\rho}_n(k)^T]^T$ represent the vector of all the parameters to be adjusted. When $n_f = 1$, we have one "leader", and the above derived conclusions can be applied. However, when $n_f > 1$, these conclusions do not hold any more, since the graph resulting from the deletion of the arcs leading to all the nodes in $\mathcal{N}^f$ does not necessarily satisfy (A4). The next theorem treats convergence of the basic algorithm (5) in the general case of arbitrary $n_f > 1$.

Theorem 3: Let Assumptions (A1)–(A4) be satisfied and let all the nodes from $\mathcal{N} - \mathcal{N}^f$ be reachable from all the nodes in $\mathcal{N}^f$. Then the algorithm (5) in which $\gamma_{ij} = 0$ for all $i \in \mathcal{N}^f$ provides convergence of $\hat{\rho}^v(k)$ in the mean square sense and w.p.1 to the limit defined by

$$\hat{\rho}_\infty^v = -(\bar{\Gamma}^v \otimes I_2)^{-1}(\bar{\Gamma}^{f,v} \otimes I_2)\bar{\rho}^f \tag{35}$$

where matrices $\bar{\Gamma}^v$ and $\bar{\Gamma}^{f,v}$ are $(n - n_f) \times (n - n_f)$ and $(n - n_f) \times n_f$ submatrices of matrix $P^{-c}\bar{\Gamma}$, with indices $i, j = n_f + 1, \ldots, n$ and $i = n_f + 1, \ldots, n$, $j = 1, \ldots, n_f$, respectively.

Proof: Recursion (10) holds, in general, for $i = 1, \ldots, n$; in particular, for $\hat{\rho}^v(k)$, we have explicitly

$$\begin{aligned} \hat{\rho}^v(k) = \{&I + [\Phi^v(k) + \Psi^v(k)](\Delta^v(k)\Gamma^v(k) \otimes I_2) \\ &+ \tilde{N}^v(k)\}\hat{\rho}^v(k-1) \\ &+ [\Phi^v(k) + \Psi^v(k)](\Gamma^{f,v}(k) \otimes I_2)\bar{\rho}^f \end{aligned} \tag{36}$$

where $\Phi^v(k)$, $\Psi^v(k)$, and $\tilde{N}^v(k)$ are $2(n - n_f) \times 2(n - n_f)$ submatrics of $\Phi(k)$, $\Psi(k)$ and $\tilde{N}(k)$ corresponding to the indices $n_f + 1, \ldots, n$, matrices $\Gamma^v(k)$ and $\Gamma^{f,v}(k)$ are obtained from $P^{-c}\Gamma(k)$ in the same way as $\bar{\Gamma}^v$ and $\bar{\Gamma}^{f,v}$ are obtained from $P^{-c}\bar{\Gamma}$, and $\Delta^v(k)$ is the $(n - n_f) \times (n - n_f)$ submatrix of $\Delta(k)$, with indices $i, j = n_f + 1, \ldots, n$. After introducing $r(k) = \hat{\rho}^v(k) - \hat{\rho}_\infty^v$, we obtain

$$\begin{aligned} r(k) = \{&I + [\Phi^v(k) + \Psi^v(k)](\Delta^v(k)\Gamma^v(k) \otimes I_2) \\ &+ \tilde{N}^v(k)\}r(k-1) - \tilde{N}^v(k)\hat{\rho}_\infty^v. \end{aligned} \tag{37}$$

Analysis of (37) can be entirely based on the methodology of analyzing the recursion (18) in the proof of Theorem 1. Therefore, we shall pay the main attention to matrix

$$B^v(k) = \Phi^v(k)((P^v)^{-c}\Gamma^v(k) \otimes I_2)$$

where P^v is an $(n - n_f) \times (n - n_f)$ submatrix of P, with indices $i, j = n_f + 1, \ldots, n$; furthermore, we can write $B^v(k) = \bar{B}^v + \tilde{B}^v(k)$, where $\bar{B}^v = E\{B^v(k)\}$. From the assumptions and Lemma 2, we realize that $\bar{B}^v$ is Hurwitz. Following the methodology of Theorem 1, we iterate (37) n_k steps backwards and calculate the Lyapunov function $W^v(k) = E\{r(k)^T R^v r(k)\}$, where R^v is a positive definite matrix satisfying the Lyapunov equation $R^v\bar{B}^v + \bar{B}^{vT}R^v = -Q^v$, $Q^v > 0$. Because of the specific structure of (37), we obtain, for k large enough, that

$$\begin{aligned} W^v(k) \leq &\left(1 - c_0\frac{1}{k^c}\right) W^v(\tau_{2(n_k+1)}(k)) \\ &+ C_1 \sum_{s=\tau_{2n_k}(k)}^{\tau_0(k)} \frac{1}{s^{2c}}(1 + \|\bar{r}(s)\|) \end{aligned} \tag{38}$$

where $\bar{r}(k) = E\{r(k)\}$, while c_0 and C_1 are generic constants. In order to estimate $\bar{r}(k)$, we iterate (37) back to some initial time $\tau_{2(n'_k+1)}(k)$, and directly find out that

$$\bar{r}(k) = \prod_{s=\tau_{2n'_k}(k)}^{k} \left(I + \frac{1}{s^c}\bar{B}^v + \bar{F}^v(s)\right) \bar{r}(\tau_{2(n'_k+1)}(k)) \tag{39}$$

where $\|\bar{F}^v(s)\| = o(\frac{1}{s^c})$. From (39), we simply conclude that $\|\bar{r}(k)\| \to_{k\to\infty} 0$. Coming back to (38), and using, *e.g.*, [34], [44], we conclude that $r(k)$ in (37) tends to zero in the mean square sense and w.p.1, which proves the theorem. ■

It can be directly observed that in the case when all the parameters ρ_i^f in $\mathcal{N}^f$ are equal to ρ^f, the rest of the sensor parameters will also converge to ρ^f. This statement holds approximately for small distances between ρ_i^f and ρ_j^f, $i, j \in \mathcal{N}^f$.

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 communication graph satisfying (A4) has been adopted, and parameters α_i and β_i have been randomly selected around one and zero, respectively, with standard deviation 0.3. It is assumed that all the local clocks tick according to Poisson processes having the same rates.

In Fig. 1, the corrected gains $\hat{g}_i(k)$ and offsets $\hat{f}_i(k)$ generated by the proposed instrumental variable algorithm (5) are

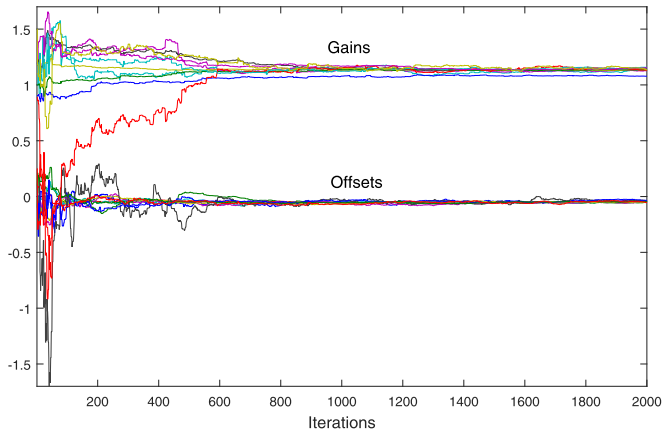


Fig. 1. Proposed algorithm based on instrumental variables without reference sensors: convergence to consensus is achieved for corrected gains and corrected offsets.

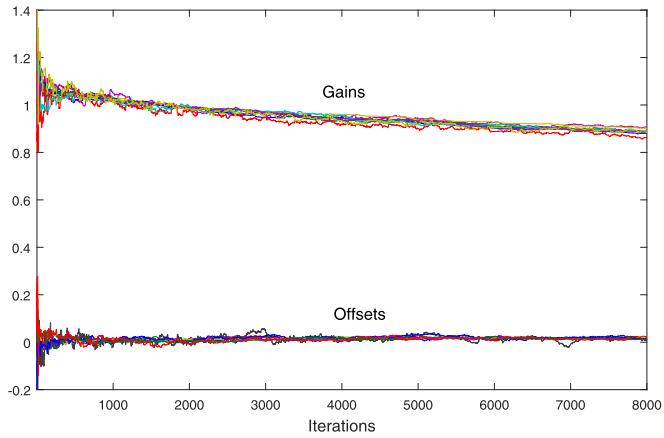


Fig. 2. Stochastic gradient algorithm: convergence to consensus is not achieved.

presented for the sequence $\delta(k) = 0.01/k^{0.6}$; communication dropouts are assumed with $p_{ij} = 0.2$ for each communication link and the measurement noises are generated with standard deviations randomly chosen in the interval $(0, 0.1)$; the signal $x(k)$ is a correlated random sequence with zero mean and variance one, generated by a second order linear system with white noise at the input. It is clear that successful calibration is achieved in spite of the noise existence.

Fig. 2 illustrates the necessity of introducing instrumental variables: the basic gradient algorithm in which $\zeta_i(k) = y_i(k)$ (instead of $y_i(d_i(k))$) has been simulated. Convergence of the network to consensus is not achieved in this case: all the corrected gains $\hat{g}_i(k)$ slowly converge to zero, instead of to consensus.

In Fig. 3, the network has been pinned to one preselected node, and in Fig. 4 to two of them. In the first case, convergence of all the corrected gains and offsets to the reference values (indicated by horizontal straight lines) takes place, as expected. However, in the second case, consensus is not achieved; instead, according to Theorem 3, all the parameters converge to different values, determined by (35).

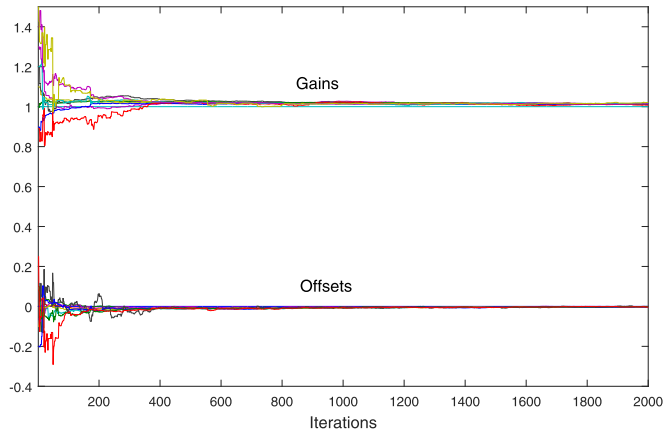


Fig. 3. Proposed algorithm with one reference sensor: all the corrected gains converge to one and all the corrected offsets to zero.

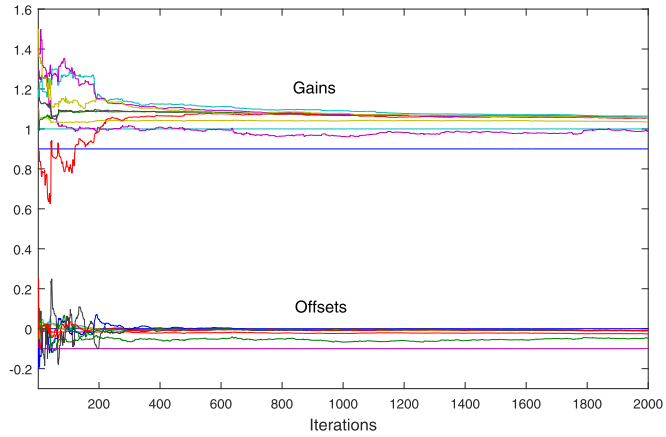


Fig. 4. Proposed algorithm with two reference sensors having different characteristics: the corrected gains and the corrected offsets converge to different values determined by (35).

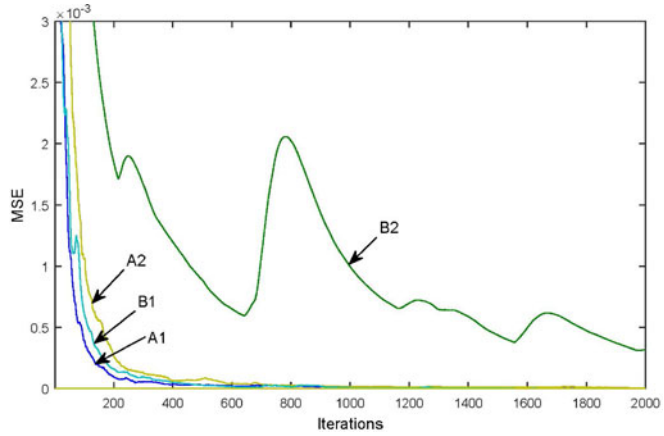


Fig. 5. MSE for corrected offsets (A1 and B1) and gains (A2 and B2) of the proposed algorithm (A1 and A2) and the algorithm proposed in [17] (B1 and B2).

In order to compare our algorithm to similar algorithms existing in the literature, we have chosen a representative consensus-based time synchronization algorithm, proposed in [17], and applied it to the distributed calibration problem (it can be easily adapted to the problem treated in this paper). The results are

shown in Fig. 5. The lines denoted by (B1) and (B2) show the mean squared error (MSE) for the corrected offsets and gains, respectively, obtained using [17], in the case of the measurement noise with standard deviation 0.005 (more than 10 times lower than the value adopted in the simulation of our algorithm). The advantage of our scheme is evident. The algorithm from [17] is very sensitive to additive communication noise, due to the incorporated division by measured signal increments.

VI. Conclusion

In this paper, a new distributed asynchronous blind calibration algorithm of instrumental variable type resulting in extended consensus has been proposed for sensor networks under measurement noise. The algorithm provides a new and computationally efficient tool for coping with the problem of calibration of large wireless sensor networks with communications limited to neighboring nodes, without any type of central coordination. It has been proved using asynchronous stochastic approximation arguments that the algorithm achieves asymptotic agreement on all corrected sensor gains and offsets. An analysis has been provided, indicating flexibility of system design at the network level by proper choice of *a priori* tunable weighting parameters. Attention has been devoted to the problem of distributed macro calibration of sensor networks when characteristics of a subset of nodes are kept invariant. When the network is pinned to one reference sensor, all the remaining sensors converge to its characteristics. It has also been proved that in the general case of multiple sensors with fixed characteristics, the proposed decentralized asynchronous algorithms ensure convergence to different parameters depending on sensor and network properties. Some simulation results illustrate characteristic behavior of the proposed algorithm.

References

[1] 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, 2011.

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

[3] K. Whitehouse and D. Culler, "Calibration as parameter estimation in sensor networks," in *Proc. 1st ACM Int. Workshop on Wireless Sensor Networks and Applications*, 2002, pp. 59–67.

[4] K. Whitehouse and D. Culler, "Macro-calibration in sensor/actuator networks," *Mobile Netw. Applicat.*, vol. 8, pp. 463–472, 2003.

[5] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 1964–1971.

[6] C. Yu and L. Xie, "On recursive blind equalization in sensor networks," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 662–672, 2015.

[7] A. Nandi, *Blind Estimation Using Higher-Order Statistics*. Boston, MA: Kluwer, 1999.

[8] L. Balzano and R. Nowak, "Blind calibration," Networked and Embedded Systems Laboratory, UCLA, Tech. Rep. TR-UCLA-NESL-200702-01, 2007.

[9] L. Balzano and R. Nowak, "Blind calibration of sensor networks," in *Proc. Int. Conf. Inform. Process. in Sensor Networks*, Apr. 2007, pp. 79–88.

[10] C. Bilen, G. Puy, R. Gribonval, and L. Daudet, "Convex optimization approaches for blind sensor calibration using sparsity," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4847–4856, 2014.

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

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

[13] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed blind calibration in lossy sensor networks via output synchronization," *IEEE Trans. Autom. Control*, vol. 60, pp. 3257–3262, 2015.

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

[15] 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.

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

[17] 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.

[18] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Optimal synchronization for networks of noisy double integrators," *IEEE Trans. Autom. Control*, vol. 56, pp. 1146–1152, 2008.

[19] 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.

[20] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed time synchronization in lossy wireless sensor networks," in *Proc. 3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, 2012 pp. 25–30.

[21] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed drift estimation for time synchronization in lossy networks," in *Proc. Mediterranean. Conf. Control and Automation*, 2016, pp. 779–784.

[22] C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii, "Ergodic randomized algorithms and dynamics over networks," *IEEE Trans. Control Netw. Syst.*, vol. 2, pp. 78–87, 2015.

[23] A. Carron, M. Todescato, R. Carli, and L. Schenato, "An asynchrnous consensus-based algorithm for estimation from noisy relative measurements," *IEEE Trans. Control Netw. Syst.*, vol. 1, pp. 283–295, 2014.

[24] 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, 2010.

[25] E. Miluzzo, N. D. Lane, A. T. Campbell, and R. Olfati-Saber, "Calibree: A self-calibration system for mobile sensor networks," in *Proc. 4th IEEE Int. Conf. Distributed Computing in Sensor Systems*, 2008, pp. 314–331.

[26] P. Park, C. Fischione, A. Bonivento, K. H. Johansson, and A. Sangiovanni-Vincentelli, "Breath: An adaptive protocol for industrial control applications using wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 10, pp. 821–838, 2011.

[27] R. Olfati-Saber, A. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, pp. 215–233, 2007.

[28] W. Ren and R. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, pp. 655–661, 2005.

[29] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton, NJ: Princeton University Press, 2010.

[30] T. C. Aysal, M. E. Yildriz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Trans. Signal Process.*, vol. 57, pp. 2748–2761, 2009.

[31] 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.

[32] L. Fang and P. Antsaklis, "Asynchronous consensus protocols using nonlinear paracontractions theory," *IEEE Trans. Autom. Control*, vol. 53, no. 10, pp. 2351–2355, 2008.

[33] A. Dimakis, S. Kar, J. Moura, M. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[34] 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.

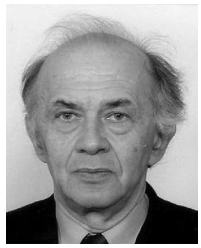
[35] M. S. Stanković, S. S. Stanković, and K. H. Johansson, "Distributed calibration for sensor networks under communication errors and measurement noise," in *Proc. IEEE Conf. Decision and Control*, 2012 pp. 1380–1385.

[36] A. Nedić, "Asynchronous broadcast-based convex optimization over a network," *IEEE Trans. Autom. Control*, vol. 56, pp. 1337–1351, 2011.

[37] L. Ljung and T. Söderström, *Theory and Pracice of Recursive Identification*. Cambridge, MA: MIT Press, 1983.
[38] T. Söderström and P. Stoica, *System Identification*. Hemel Hempstead, U.K.: Prentice-Hall Int., 1989.
[39] R. C. Bradley, "Basic properties of strong mixing conditions a survey and some open questions," *Probabil. Surv.*, vol. 2, pp. 107–144, 2005.
[40] I. Ibragimov, "Some limit theorems for stochastic processes stationary in the strict sense," *Dokl. Akad. Nauk SSSR*, vol. 125, pp. 711–714, 1959.
[41] M. Rosenblatt, "A central limit theorem and a strong mixing condition," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 42, pp. 43–47, 1956.
[42] Y. Ohta and D. Siljak, "Overlapping block diagonal dominance and existence of Lyapunov functions," *J. Math. Anal. Appl.*, vol. 112, pp. 396–410, 1985.
[43] I. F. Pierce, "Matrices with dominating diagonal blocks," *J. Econom. Theory*, vol. 9, pp. 159–170, 1974.
[44] H. F. Chen, *Stochastic Approximation and its Applications*. Dordrecht, The Netherlands: Kluwer, 2002.
[45] T. Li and J. F. Zhang, "Consensus conditions of multi agent systems with time varying topologies," *IEEE Trans. Autom. Control*, vol. 55, pp. 2043–2056, 2010.
[46] V. Borkar, "Asynchronous stochastic approximation," *SIAM J. Control Optimiz.*, vol. 36, pp. 840–851, 1998.
[47] T. Söderström and P. Stoica, *Instrumental Variable Methods for System Identification*. New York: Springer-Verlag, 1983.
[48] B. E. Fristedt and L. F. Gray, *A Modern Approach to Probability Theory*. Boston, MA: Birkhauser, 1996.

Miloš S. Stanković received the Dipl. Ing. (M.Sc.) degree from the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, in 2002, and the Ph.D. degree in systems and entrepreneurial engineering from the University of Illinois at Urbana-Champaign (UIUC) in 2009.

He was a Research and Teaching Assistant in the Control and Decision Group of the Coordinated Science Laboratory, UIUC, during 2006–2009. From 2009 to 2012, he was a postdoctoral researcher in the Automatic Control Laboratory and the ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden. In 2012, he joined the Innovation Center, School of Electrical Engineering, University of Belgrade, as an EU Marie Curie Research Fellow. His research interests include decentralized decision making, networked control systems, dynamic game theory, optimization, and machine learning, with applications to mobile sensor networks, cyber-physical systems and the Internet of Things.



Srdjan S. Stanković received the Dipl.Ing., Mgr.Sc. and Ph.D. degrees from the Faculty of Electrical Engineering, University of Belgrade, Belgrade, Yugoslavia, in 1968, 1972 and 1975, respectively.

He was with the Institute for Nuclear Sciences, Vinča, Belgrade, Yugoslavia, from 1968 to 1972. Since 1973, he has been with the Faculty of Electrical Engineering, University of Belgrade, where he is currently Emeritus Professor of Automatic Control. He is also President of the Research Council of the Vlatacom Institute, Belgrade. He held visiting positions at the Eindhoven University of Technology, Eindhoven, The Netherlands, and at the Santa Clara University, Santa Clara, CA. He has published numerous scientific papers in diverse fields, including estimation and identification, adaptive systems, large-scale systems, decentralized control and neural networks. He has also been leader of numerous scientific projects for government and industry.

Dr. Stanković was President of the National Council for Higher Education of the Republic of Serbia, as well as President of the Serbian Society for Electronics, Communications, Computers, Control and Nuclear Engineering.

Karl H. Johansson (F'13) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1992 and 1997, respectively.

He is Director of the Stockholm Strategic Research Area ICT–The Next Generation, and a Professor at the School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. He has held visiting positions at the University of California at Berkeley, California Institute of Technology, Nanyang Technological University, the Institute of Advanced Studies at Hong Kong University of Science and Technology, and the Norwegian University of Science and Technology. His research interests are in networked control systems, cyber-physical systems, and applications in transportation, energy, and automation systems.

Dr. Johansson is a member of the IEEE Control Systems Society Board of Governors and the European Control Association Council. He has received several best paper awards and other distinctions, including a ten-year Wallenberg Scholar Grant, a Senior Researcher Position with the Swedish Research Council, and Future Research Leader Award from the Swedish Foundation for Strategic Research. He is an IEEE Control Systems Society Distinguished Lecturer.