

Almost sure convergence of a randomized algorithm for relative localization in sensor networks

*Original*

Almost sure convergence of a randomized algorithm for relative localization in sensor networks / Ravazzi, Chiara; Frasca, Paolo; Tempo, R.; Ishii, H.. - ELETTRONICO. - (2013), pp. 4778-4783. (Intervento presentato al convegno IEEE 52nd Annual Conference on Decision and Control (CDC), 2013 tenutosi a Firenze nel 10-13 Dec. 2013) [10.1109/CDC.2013.6760638].

*Availability:*

This version is available at: 11583/2549138 since:

*Publisher:*

IEEE / Institute of Electrical and Electronics Engineers

*Published*

DOI:10.1109/CDC.2013.6760638

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

IEEE postprint/Author's Accepted Manuscript

©2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

# Almost sure convergence of a randomized algorithm for relative localization in sensor networks

Chiara Ravazzi

Paolo Frasca

Roberto Tempo

Hideaki Ishii

**Abstract**—This paper regards the relative localization problem in sensor networks. We study a randomized algorithm, which is based on input-driven consensus dynamics and involves pairwise “gossip” communications and updates. Due to the randomness of the updates, the state of this algorithm ergodically oscillates around a limit value. Exploiting the ergodicity of the dynamics, we show that the time-average of the state almost surely converges to the least-squares solution of the localization problem. Remarkably, the computation of the time-average does not require the sensors to share any common clock. Hence, the proposed algorithm is fully distributed and asynchronous.

## I. INTRODUCTION

We consider the problem of relative localization for sensor networks that can be described as follows. We assume to have a group of agents representing the nodes of a graph, and a vector, indexed over the agents and unknown to them. The agents are allowed to take noisy measurements of the differences between their vector entries and those of their neighbors in the graph. The estimation problem consists in reconstructing the original vector, up to an additive constant. While an optimal solution can be easily found by a centralized least-squares approach, we are interested in finding effective *distributed* solutions. More precisely, a solution is said to be distributed if it requires each node to use information which is available only at the node itself or from its immediate neighbors. Following this approach, we have recently proposed [1] a randomized “gossip” algorithm for distributed relative localization. This algorithm, which is inspired by a gradient descent approach, involves the activation of a randomly chosen pair of neighboring nodes at each time step. In our previous work, we have already given a convergence result for the algorithm: the mean-square error between the time-average of the states and the optimal solution asymptotically goes to zero.

In this paper, we study the algorithm using tools from ergodic theory and we obtain a related convergence result: the time-average of the states converges almost surely to

the optimal solution. Significantly, our definition of time-average does not require the agents to be aware of any global clock or of any global variable which counts the number of interactions occurring on the network.

### A. State of the art

The abstract problem of relative localization is deemed to have important applications in sensor and robotic networks [2]. The proposed applications cover synchronization of uncertain clocks [3], as well as spacial localization [4] in mobile robotic networks when no absolute position information is available. In the control literature, the problem has been popularized by [2], [5], [6]. Distributed algorithms involving synchronized updates by the nodes have been proposed in the last few years [2], [3]. The last paper uses a gradient-descent technique to solve the problem: this approach has been later followed in [7], [8], as well as in our previous work [1]. Recently, papers have started to consider randomized asynchronous and randomized algorithms to solve the localization problem [9], [10], [11]. We refer the reader to [12] for randomised asynchronous algorithms for convex optimisation and to [13] for other randomized algorithms in systems and control.

### B. Contribution

In this work, we prove an ergodic theorem for the algorithm in [1, Eq. (10)] and we show that a suitable time-averaging operation removes the persistent random oscillations which affect the “raw” estimates obtained through gossip communication and updates. The resulting time-averaged state converges *almost surely* to the optimal solution of the relative localization problem.

The ergodicity analysis of seemingly non-convergent dynamics started to attract attention in the theory of multi-agent systems and distributed control only recently. For example, we remark that the algorithm under study presents strong similarities with randomized algorithms for the PageRank computation, which have been recently proposed in several papers [14], [15], [16]. Indeed, randomized PageRank dynamics have been shown to converge modulo time-averaging: available results cover both convergence in mean square [14] and almost sure [17]. We stress, however, that our proof of almost sure convergence is based on completely different tools from [17], which uses techniques from stochastic approximation. Related applications of ergodic theory can be found in the context of social networks [18], [19], where the authors show the ergodicity of specific opinion dynamics, which extend the well-known consensus dynamics to

This work was supported in part by the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement n.257462 HYCON2 Network of Excellence and in part by JST, CREST.

Chiara Ravazzi is with the Department of Electronics and Telecommunications (DET), Politecnico di Torino, Italy. E-mail: chiara.ravazzi@polito.it

Paolo Frasca is with the Department of Mathematical Sciences (DISMA), Politecnico di Torino, Italy; he is now with the Department of Applied Mathematics, University of Twente, Enschede, The Netherlands. E-mail: p.frasca@utwente.nl

Roberto Tempo is with CNR-IEIIT, Politecnico di Torino, Italy. E-mail: roberto.tempo@polito.it

Hideaki Ishii is with the Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Japan. E-mail: ishii@dis.titech.ac.jp.

incorporate external influences and heterogeneous (stubborn) agents. Our analysis owes much to these techniques.

Furthermore, we stress that ergodicity is a key property in enabling our framework to easily accommodate time-averages. Hence, we study an algorithm which is fully asynchronous and distributed. On the contrary, prior work on randomized algorithms assumed time-averages to be computed using a global iteration counter, see [14], [17].

### C. Organization of the paper

We formally present the problem of relative localization in Section II. Then, in Section III we define the randomized algorithm and state the main convergence results. These results are proved by the analysis presented in Section IV. We conclude the paper with some remarks on future research in the final section.

## II. THE PROBLEM OF RELATIVE LOCALIZATION

We consider a set of nodes  $\mathcal{V} = \{1, 2, \dots, N\}$ , endowed with an unknown scalar quantity  $\bar{x}_v$  for  $v \in \mathcal{V}$ . The relative localization problem consists, for each node  $u \in \mathcal{V}$ , in estimating the scalar value  $\bar{x}_u$ , based on noisy measurements of differences  $\bar{x}_u - \bar{x}_v$  with certain neighbors  $v$ . An oriented graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is used to represent the available measurements. The orientation of the pairs is conventionally assumed to be such that  $(v, u) \in \mathcal{E}$  only if  $u < v$ . We let  $A \in \{0, \pm 1\}^{\mathcal{E} \times \mathcal{V}}$  be the incidence matrix of the graph  $\mathcal{G}$ , which is defined as

$$A_{ew} = \begin{cases} +1 & \text{if } e = (v, w) \\ -1 & \text{if } e = (w, v) \\ 0 & \text{otherwise} \end{cases}$$

for every  $e \in \mathcal{E}$ . We let  $b \in \mathbb{R}^{\mathcal{E}}$  be the vector collecting the measurements

$$b = A\bar{x} + \eta,$$

where  $\bar{x}$  is the stacked vector of  $\bar{x}_i$  with  $i \in \mathcal{V}$ ,  $\eta \in \mathbb{R}^{\mathcal{E}}$  is additive noise. We define the set of the optimal estimates in a least squares sense as

$$X = \operatorname{argmin}_{z \in \mathbb{R}^N} \|Az - b\|_2^2 \quad (1)$$

where  $\|\cdot\|_2$  is the Euclidean norm. The set  $X$  is described in the following well-known lemma [2].

*Lemma 1 (Centralized solution):* Let the graph  $\mathcal{G}$  be connected and let  $L := A^\top A$  denote the Laplacian of the graph. The following facts hold:

- 1)  $x \in X$  if and only if  $A^\top Ax = A^\top b$ ;
- 2) there exists a unique  $x^* \in X$  such that  $\|x^*\|_2 = \min_{z \in X} \|z\|_2$ ;
- 3)  $x^* = L^\dagger A^\top b$ , where  $L^\dagger$  denotes the Moore-Penrose pseudo-inverse of the Laplacian  $L$ .

Note that  $x^*$  is the minimum-norm element of the affine space of solutions of (1). Indeed,  $Ax^* = A(x^* + c\mathbf{1})$  for all scalar  $c$ , where  $\mathbf{1}$  denotes the vector of ones.

In this work, we are interested in designing randomized asynchronous algorithms to solve this problem: in view of the result above, we shall assume from now on that  $\mathcal{G}$  is *connected*.

## III. ALGORITHM DESCRIPTION AND MAIN RESULTS

This section is devoted to describe the algorithm which was proposed in [1, Section V] and to state our main results.

The algorithm involves for each node  $v \in \mathcal{V}$  a triple of states  $(x_v, \kappa_v, \tilde{x}_v)$ , depending on a discrete time index  $k \in \mathbb{Z}_{\geq 0}$ : these three variables play the following roles:  $x_v(k)$  is the “raw” estimate of  $\bar{x}_v$  obtained by  $v$  at time  $k$  through communications with its neighbors,  $\kappa_v(k)$  counts the number of updates performed by  $v$  up to time  $k$ , and  $\tilde{x}_v(k)$  is the “smoothed” estimate obtained through time-averaging. The algorithm is defined by choosing a scalar parameter  $\gamma \in (0, 1)$  and a sequence of random variables  $\theta(k)_{k \in \mathbb{Z}_{\geq 0}}$  taking values in  $\mathcal{E}$ . At each time  $k$ , provided that  $\theta(k) = (u, v)$ , the state updates are performed according to the following rules: the estimates evolve as

$$\begin{aligned} x_u(k+1) &= (1-\gamma)x_u(k) + \gamma x_v(k) + \gamma b_{(u,v)} \\ x_v(k+1) &= (1-\gamma)x_v(k) + \gamma x_u(k) - \gamma b_{(u,v)} \\ x_w(k+1) &= x_w(k) \quad \text{if } w \notin \{u, v\}; \end{aligned} \quad (2a)$$

the local times as

$$\begin{aligned} \kappa_u(k+1) &= \kappa_u(k) + 1 \\ \kappa_v(k+1) &= \kappa_v(k) + 1 \\ \kappa_w(k+1) &= \kappa_w(k) \quad \text{if } w \notin \{u, v\}; \end{aligned} \quad (2b)$$

and the time-averages as

$$\begin{aligned} \tilde{x}_u(k+1) &= \frac{1}{\kappa_u(k+1)} (\kappa_u(k)\tilde{x}_u(k) + x_u(k+1)) \\ \tilde{x}_v(k+1) &= \frac{1}{\kappa_v(k+1)} (\kappa_v(k)\tilde{x}_v(k) + x_v(k+1)) \\ \tilde{x}_w(k+1) &= \tilde{x}_w(k) \quad \text{if } w \notin \{u, v\}. \end{aligned} \quad (2c)$$

We assume the sequence  $\{\theta(k)\}_{k \in \mathbb{Z}_{\geq 0}}$  to be i.i.d., and its probability distribution to be uniform, *i.e.*,

$$\mathbb{P}[\theta(k) = (u, v)] = \frac{1}{|\mathcal{E}|}, \quad \forall k \in \mathbb{Z}_{\geq 0}, \quad (3)$$

where  $|\mathcal{E}|$  denotes the cardinality of  $\mathcal{E}$ . Note that this choice is made without loss of generality and convenience: the same approach may accommodate other distributions, as required by the application.

It should be noted that the time index  $k$  in fact counts the number of updates which have occurred in the network, whereas for each  $u \in \mathcal{V}$  the variable  $\kappa_u(k)$  is the number of updates involving  $u$  up to the current time. Hence,  $\kappa_u$  is a local variable which is inherently known to agent  $u$ , even in case the common clock  $k$  is unavailable. This algorithm is totally asynchronous and fully distributed, in the sense that the time averaging process does not need the nodes to be aware of a common clock. This point provides a major improvement with respect to our previous work [1, Theorem 3].

The dynamics in (2a) oscillates persistently and fails to converge in a deterministic sense, as shown in Figure 1. However, the oscillations concentrate around the solution of the least squares problem, as it is formally stated in the

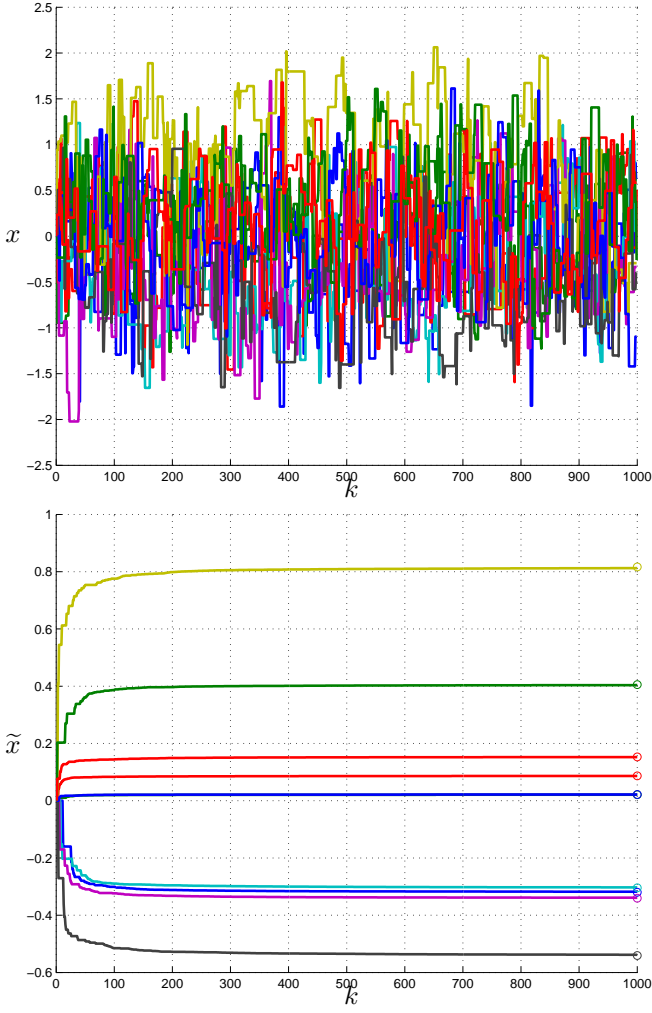


Fig. 1. Complete graph,  $N = 10$ ,  $\gamma = 0.5$ , zero initial conditions.

following two results. The first result regards the behavior of the average dynamics.

*Proposition 2 (Convergence in expectation):* Consider the dynamics (2a) with uniform edge selection (3). Then,

$$\lim_{k \rightarrow +\infty} \mathbb{E}[x(k)] = x^*.$$

The second result, instead, states that the sample dynamics is well-represented by the average one, *i.e.*, the process is ergodic.

*Theorem 3 (Almost sure convergence of ergodic means):* The dynamics in (2a), with uniform edge selection (3) and  $x(0) = 0$ , is *ergodic*. If  $\{n_\ell\}_{\ell \in \mathbb{N}}$  is a sequence of nonnegative integers, then with probability one

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \sum_{\ell=1}^k x(n_\ell) = x^*.$$

From these two facts, which are proved in the next section, we immediately deduce the following statement, which motivates the definition of the algorithm previously introduced indicates that  $\tilde{x}_u(k)$  is the “right variable” to approximate the optimal estimate  $x_u^*$ .

*Corollary 4:* The dynamics in (2c) is such that

$$\lim_{k \rightarrow +\infty} \tilde{x}(k) = x^*$$

with probability one.

#### IV. ERGODICITY ANALYSIS

In this section we study the convergence properties of the vector  $x(k)$  for the update model described in (2a). As shown in Fig. 1, the estimate of each agent oscillates persistently; on the other hand the time averages approach the optimal solution  $x^*$  in Lemma 1. Although the process  $x(k)$  almost surely fails to converge, we prove that it converges in distribution to a random variable  $x_\infty$  and is ergodic.

To begin, we rewrite the dynamics of (2a) as

$$x(k+1) = Q(k)x(k) + y(k) \quad (4)$$

where

$$Q(k) = I - \gamma(e_u - e_v)(e_u - e_v)^\top,$$

vector  $e_u$  denotes the vector of the canonical basis corresponding to  $u$ , and

$$y(k) = b_{\theta(k)}(e_u - e_v)$$

provided  $\theta(k) = (u, v)$  with  $k \in \mathbb{Z}_{\geq 0}$ . Consequently, we prove the following result either by direct computation or by using Lemma 5 in [1].

*Lemma 5:* For the distribution (3), it holds

$$\mathbb{E}[Q(k)] = I - \gamma \frac{L}{|\mathcal{E}|}, \quad \mathbb{E}[y(k)] = \gamma \frac{A^\top b}{|\mathcal{E}|}.$$

We note that for all  $k$  the matrix  $Q(k)$  is doubly stochastic and the sum of the elements in  $y(k)$  is zero, that is,

$$\mathbf{1}^\top Q(k) = \mathbf{1}, \quad Q(k)\mathbf{1} = \mathbf{1}, \quad \mathbf{1}^\top y(k) = 0. \quad (5)$$

In particular, if the vector  $x(0)$  is initialized to zero, then  $\mathbf{1}^\top x(k) = 0$  for each  $k \in \mathbb{Z}_{\geq 0}$ .

These observations imply that the dynamics of  $x(k)$  is equivalently described by the iterate

$$x(k+1) = P(k)x(k) + y(k), \quad (6)$$

where  $P(k) = Q(k)(I - \frac{1}{n}\mathbf{1}\mathbf{1}^\top) = (I - \frac{1}{n}\mathbf{1}\mathbf{1}^\top)Q(k)$  is a projection of  $Q(k)$  outside the “consensus sub-space” spanned by  $\mathbf{1}$ . This rewriting is instrumental to study the convergence behavior of the process  $\{x(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ , by asymptotic techniques of iterated random functions, which we recall from [20]. These techniques require, in order to study the random process (6), to consider the associated *backward* process  $\overleftarrow{x}(k)$ , which we define below.

For any time instant  $k$ , consider the random matrices  $P(k)$  and  $y(k)$  and define the matrix product

$$\overrightarrow{P}(\ell, m) := P(m)P(m-1) \cdots P(\ell+1)P(\ell) \quad (7)$$

with  $\ell \in \{0, \dots, m\}$ . Then, the iterated affine system in (6) can be rewritten as

$$x(k+1) = \overrightarrow{P}(0, k)x(0) + \sum_{0 \leq \ell \leq k} \overrightarrow{P}(\ell+1, k)y(\ell). \quad (8)$$

The corresponding *backward process* is defined by

$$\overleftarrow{x}(k+1) = \overleftarrow{P}(0, k)x(0) + \sum_{0 \leq \ell \leq k} \overleftarrow{P}(0, \ell-1)y(\ell), \quad (9)$$

where

$$\overleftarrow{P}(\ell, m) := P(\ell)P(\ell+1) \cdots P(m-1)P(m) \quad (10)$$

with  $\ell \in \{0, \dots, m\}$ . Crucially, the backward process  $\overleftarrow{x}(k)$  has at every time  $k \in \mathbb{Z}_{\geq 0}$  the same probability distribution of  $x(k)$ . The main tool to study the backward process is the following well-known result.

*Lemma 6 (Theorem 2.1 in [20]):* Let us consider the Markov process  $\{X(k)\}_{k \in \mathbb{N}}$  defined by

$$X(k+1) = A(k)X(k) + B(k) \quad k \in \mathbb{Z}_{\geq 0}$$

where  $A(k) \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$  and  $B(k) \in \mathbb{R}^{\mathcal{V}}$  are i.i.d. random variables. Let us assume that

$$\mathbb{E}[\log \|A(k)\|] < \infty \quad \mathbb{E}[\log \|B(k)\|] < \infty. \quad (11)$$

The corresponding backward random process  $\overleftarrow{X}(k)$  converges a.s. to a finite limit  $X_\infty$  if and only if

$$\inf_{k > 0} \frac{1}{k} \mathbb{E}[\log \|A(1) \cdots A(k)\|] < 0. \quad (12)$$

If (12) holds, the distribution of  $X_\infty$  is the unique invariant distribution for the Markov process  $X(k)$ .

This result provides conditions for the backward process to converge to a limit. Although the forward process has a different behavior compared to the backward process, the forward and backward processes have the same distribution. This fact allows us to determine, by studying the backward process  $\overleftarrow{x}(k)$ , whether the sequence of random variables  $\{x(k)\}_{k \in \mathbb{Z}_{\geq 0}}$  converges in distribution to the invariant distribution of the Markov process in (6).

This analysis is done in the following result.

*Lemma 7:* Consider the random process  $x(k)$  defined in (6), where  $P(k)$  and  $y(k)$  are i.i.d.. Then  $\overleftarrow{x}(k)$  converges a.s. to a finite limit  $x_\infty$ , and the distribution of  $x_\infty$  is the unique invariant distribution for  $x(k)$ .

*Proof:* In order to apply Lemma 6, let us compute

$$\begin{aligned} & \inf_{k \in \mathbb{N}} \frac{1}{k} \mathbb{E} \left[ \log \|\overleftarrow{P}(0, k-1)\|_1 \right] \\ & \leq \inf_{k \in \mathbb{N}} \frac{1}{k} \log \mathbb{E} \left[ \|\overleftarrow{P}(0, k-1)\|_1 \right] \\ & = \inf_{k \in \mathbb{N}} \frac{1}{k} \log \mathbb{E} \left[ \max_{w \in \mathcal{V}} \sum_{i \in \mathcal{V}} (\overleftarrow{P}(0, k-1))_{vw} \right] \\ & \leq \inf_{k \in \mathbb{N}} \frac{1}{k} \log \mathbb{E} \left[ \sum_{w \in \mathcal{V}} \sum_{v \in \mathcal{V}} (\overleftarrow{P}(0, k-1))_{vw} \right] \\ & \leq \inf_{k \in \mathbb{N}} \frac{1}{k} \log \sum_{w \in \mathcal{V}} \sum_{v \in \mathcal{V}} \mathbb{E} \left[ \overleftarrow{P}(0, k-1)_{vw} \right] \\ & \leq \inf_{k \in \mathbb{N}} \frac{1}{k} \log \left( n \left\| \mathbb{E} \left[ \overleftarrow{P}(0, k-1) \right] \right\|_\infty \right) \\ & = \inf_{k \in \mathbb{N}} \frac{1}{k} \log \left( n \left\| \prod_{\ell=0}^{k-1} \mathbb{E} [P(\ell)] \right\|_\infty \right). \end{aligned}$$

Let  $q$  be the number of distinct eigenvalues of  $\mathbb{E}[P(k)]$ , denoted as  $\{\lambda_\ell\}_{\ell=1}^q$ , and consider the Jordan canonical decomposition  $\mathbb{E}[P(k)] = UJU^{-1}$ . Then  $\left\| \prod_{s=0}^{k-1} \mathbb{E}[P(s)] \right\|_\infty \leq \|U\|_\infty \|J^k\|_\infty \|U^{-1}\|_\infty$ . Since the  $k$ -th power of the Jordan block of size  $s$  is

$$\begin{aligned} & \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & & \lambda \end{bmatrix}^k \\ & = \begin{bmatrix} \lambda^k & \binom{k}{1} \lambda^{k-1} & \binom{k}{2} \lambda^{k-2} & \cdots & \binom{k}{s-1} \lambda^{k-s+1} \\ 0 & \lambda^k & \binom{k}{1} \lambda^{k-1} & \cdots & \binom{k}{s-2} \lambda^{k-s+2} \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & \lambda^k & \binom{k}{1} \lambda^{k-1} \\ 0 & 0 & \cdots & & \lambda^k \end{bmatrix}, \end{aligned}$$

we deduce that

$$\|J^k\|_\infty = \max_{v \in \mathcal{V}} \sum_{w \in \mathcal{V}} (J^k)_{vw} = \max_{\ell=1, \dots, q} \sum_{m=0}^{s_\ell-1} \lambda_\ell^{k-m} \binom{k}{m},$$

where  $s_\ell$  is the size of the largest Jordan block corresponding to  $\lambda_\ell$ . Then

$$\begin{aligned} \|J^k\|_\infty & \leq \max_{\ell=1, \dots, q} |\lambda_\ell|^k \sum_{m=0}^{s_\ell-1} |\lambda_\ell|^{-m} \binom{k}{m} \\ & \leq \max_{\ell=1, \dots, q} |\lambda_\ell|^k k^n \sum_{m=0}^{s_\ell-1} |\lambda_\ell|^{-m} \\ & \leq \chi \rho^k k^n, \end{aligned}$$

where  $\chi$  is a constant independent of  $k$  and

$$\rho = \|\mathbb{E}[P(k)]\|_2.$$

From Lemma 5, it follows that  $\rho < 1$ . We conclude that there exists a constant  $C = \|U\|_\infty \|U^{-1}\|_\infty \chi$ , independent of  $k$ , such that

$$\mathbb{E} \left[ \log \|\overleftarrow{P}(0, k-1)\|_1 \right] \leq \log (nC \rho^k k^n).$$

and, consequently,

$$\begin{aligned} & \inf_{k \in \mathbb{N}} \frac{1}{k} \mathbb{E} \left[ \log \|\overleftarrow{P}(0, k-1)\|_1 \right] \\ & \leq \lim_{k \rightarrow \infty} \frac{\log (C n k^n \rho^k)}{k} \\ & = \log \rho < 0. \end{aligned} \quad (13)$$

The claim then follows from Lemma 6.  $\blacksquare$

As a consequence, we deduce that also the (forward) random process  $x(k)$ , defined in (6), converges in distribution to a limit  $x_\infty$ , and the distribution of  $x_\infty$  is the unique invariant distribution for  $x(k)$ . The oscillations of (2a) are *ergodic*, as stated in the following result.

*Lemma 8:* The random process  $x(k)$  defined in (6) is *ergodic*: if  $\{n_\ell\}_{\ell \in \mathbb{N}}$  is a sequence of nonnegative integers, then with probability one

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{\ell=1}^k x(n_\ell) = \mathbb{E}[x_\infty].$$

Furthermore,  $\mathbb{E}[x_\infty] = L^\dagger A^\top b$ .

*Proof:* We begin by showing the ergodicity. Let  $z(0)$  be a random vector independent from  $x(0)$  with the same distribution as  $x_\infty$ . Let  $\{z(k)\}_{k \in \mathbb{Z}_{\geq 0}}$  be the sequence such that

$$z(k) = \vec{P}(0, k-1)z(0) + \sum_{0 \leq \ell \leq k-1} \vec{P}(\ell+1, k-1)y(\ell)$$

where  $\vec{P}(\ell+1, k-1)$  is defined as in (7). It should be noted that, since the process  $z(k)$  is stationary, we can apply the law of large numbers and immediately conclude that, with probability one,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{s=0}^{k-1} z(s) = \mathbb{E}[x_\infty].$$

On the other hand, we compute

$$\begin{aligned} & \mathbb{P}(\|x(k) - z(k)\|_1 \geq \varepsilon^k) \\ & \leq \frac{\mathbb{E}[\|\vec{P}(0, k-1)(z(0) - x(0))\|_1]}{\varepsilon^k} \\ & \leq \frac{\mathbb{E}[\|\vec{P}(0, k-1)\|_1 \|z(0) - x(0)\|_1]}{\varepsilon^k} \\ & \leq \frac{\mathbb{E}[\|\vec{P}(0, k-1)\|_1] \mathbb{E}[\|z(0) - x(0)\|_1]}{\varepsilon^k} \\ & \leq \frac{Cnk^n \rho^k}{\varepsilon^k} \mathbb{E}[\|z(0) - x(0)\|_1], \end{aligned}$$

where we have used (13). If we choose  $\varepsilon \in (\rho, 1)$ , then the Borel-Cantelli Lemma [21, Theorem 1.4.2] implies that with probability one  $\|x(k) - z(k)\|_1 < \varepsilon^k$  for all but finitely many values of  $k \geq 0$ . Therefore, almost surely  $\frac{1}{k} \sum_{s=0}^{k-1} \|x(s) - z(s)\|_1$  converges to zero as  $k$  goes to infinity,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{s=0}^{k-1} x(s) = \mathbb{E}[x_\infty].$$

The statement follows when we observe that the same arguments hold if we replace all summations over the nonnegative integers with summations over a subsequence of nonnegative integers.

To complete the proof, we only need to compute the expectation of  $x_\infty$ . Using the independence among  $P(k)$ s and  $y(k)$ s, we obtain

$$\mathbb{E}[x(k)] = \mathbb{E}[P]^k x(0) + \sum_{0 \leq \ell \leq k-1} \mathbb{E}[P(k)]^{k-\ell-1} \mathbb{E}[y(k)]$$

and we conclude from Proposition 2 (see also [1, Proposition 6]) that

$$\mathbb{E}[x_\infty] = \lim_{k \rightarrow +\infty} \mathbb{E}[x(k)] = L^\dagger A^\top b.$$

## V. CONCLUDING REMARKS

In this paper, we have studied a randomized gossip algorithm for the relative localization problem. The states of this algorithm exhibit oscillations, which can be smoothed out by a suitable time-averaging. These time-averages converge almost surely to the optimal solution. We stress that our definition of time-average (2) does not require the nodes to be aware of any global clock or any other global variable. This feature should be contrasted against our previous contribution in [1], where we defined time-averages as  $\bar{x}(k) := \frac{1}{k} \sum_{\ell=1}^k x(\ell)$ . Figure 2 compares the evolution

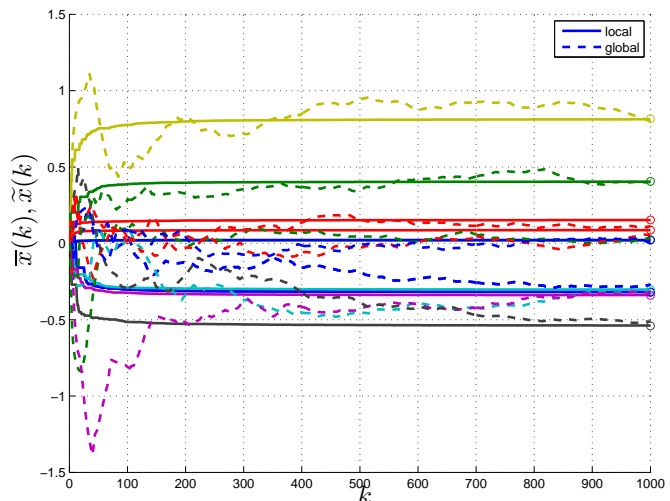


Fig. 2. Comparing local and global clocks, under the same conditions as in Figure 1. Dashed lines represent global time-averages  $\bar{x}(k)$ , solid lines represent local time-averages  $\tilde{x}(k)$ .

of  $\bar{x}(k)$  (dashed lines) and  $\tilde{x}(k)$  (solid lines). Remarkably, solid lines are smoother than dashed lines and approach the optimal solution faster. This observation is common to our simulations, and suggests that the use of local time-averaging is not only an advantage from the point of view of the implementation, but actually allows for an improved performance of the algorithm. A formal proof of this fact asks for further investigation.

In view of the effectiveness of the proposed time-averaging approach, future research will seek broader applications of these techniques to ergodic dynamics in multi-agents systems. It is also worth mentioning that different solutions have been proposed in the literature to smooth the oscillations of a dynamical system and guarantee the almost sure convergence of the system. This goal can be achieved using diminishing step-sizes, which damp the input in the long run, but maintaining the input itself 'active' for a sufficiently long time: their analysis is generally based on tools from stochastic approximation [22] or semi-martingale theory [23]. A thorough comparison of the latter approach with ours is also an interesting topic for future research.

## REFERENCES

- [1] C. Ravazzi, P. Frasca, H. Ishii, and R. Tempo, "A distributed randomized algorithm for relative localization in sensor networks," in

- European Control Conference*, Zürich, Switzerland, Jul. 2013, pp. 1776–1781.
- [2] P. Barooah and J. P. Hespanha, “Estimation from relative measurements: Algorithms and scaling laws,” *IEEE Control Systems Magazine*, vol. 27, no. 4, pp. 57–74, 2007.
  - [3] A. Giridhar and P. R. Kumar, “Distributed clock synchronization over wireless networks: Algorithms and analysis,” in *IEEE Conference on Decision and Control*, San Diego, CA, USA, Dec. 2006, pp. 4915–4920.
  - [4] D. Moore, J. Leonard, D. Rus, and S. Teller, “Robust distributed network localization with noisy range measurements,” in *International Conference on Embedded Networked Sensor Systems*, ser. SenSys '04, New York, NY, USA, 2004, pp. 50–61.
  - [5] P. Barooah and J. P. Hespanha, “Error scaling laws for linear optimal estimation from relative measurements,” *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5661–5673, 2009.
  - [6] —, “Estimation from relative measurements: Electrical analogy and large graphs,” *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2181–2193, 2008.
  - [7] S. Bolognani, S. D. Favero, L. Schenato, and D. Varagnolo, “Consensus-based distributed sensor calibration and least-square parameter identification in WSNs,” *International Journal of Robust and Nonlinear Control*, vol. 20, no. 2, pp. 176–193, 2010.
  - [8] W. S. Rossi, P. Frasca, and F. Fagnani, “Transient and limit performance of distributed relative localization,” in *IEEE Conference on Decision and Control*, Maui, HI, USA, Dec. 2012, pp. 2744–2748.
  - [9] R. Carli and L. Schenato, “Exponential-rate consensus-based algorithms for estimation from relative measurements,” University of Padova, Tech. Rep., Sep. 2012. [Online]. Available: <http://automatica.dei.unipd.it/people/schenato/publications.html>
  - [10] R. Carli, E. D’Elia, and S. Zampieri, “A PI controller based on asymmetric gossip communications for clocks synchronization in wireless sensors networks,” in *IEEE Conf. on Decision and Control and European Control Conference*, Orlando, FL, USA, 2011.
  - [11] N. M. Freris and A. Zouzias, “Fast distributed smoothing of relative measurements,” in *IEEE Conference on Decision and Control*, Maui, HI, USA, Dec. 2012, pp. 1411–1416.
  - [12] A. Nedic, “Asynchronous broadcast-based convex optimization over a network,” *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1337–1351, 2011.
  - [13] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems, with Applications*. Springer, 2005.
  - [14] H. Ishii and R. Tempo, “Distributed randomized algorithms for the PageRank computation,” *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 1987–2002, 2010.
  - [15] H. Ishii, R. Tempo, and E. W. Bai, “A web aggregation approach for distributed randomized PageRank algorithms,” *IEEE Transactions on Automatic Control*, vol. 57, no. 11, pp. 2703–2717, 2012.
  - [16] —, “Pagerank computation via a distributed randomized approach with lossy communication,” *Systems & Control Letters*, vol. 61, no. 12, pp. 1221–1228, 2012.
  - [17] W.-X. Zhao, H.-F. Chen, and H.-T. Fang, “Almost sure convergence of distributed randomized PageRank computation,” *IEEE Transactions on Automatic Control*, 2013, to appear.
  - [18] G. Como and F. Fagnani, “Scaling limits for continuous opinion dynamics systems,” *The Annals of Applied Probability*, vol. 21, no. 4, pp. 1537–1567, 2011.
  - [19] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, “Opinion fluctuations and disagreement in social networks,” *Mathematics of Operations Research*, vol. 38, no. 1, pp. 1–27, 2013.
  - [20] P. Diaconis and D. Freedman, “Iterated random functions,” *SIAM Review*, vol. 41, pp. 45–76, 1999.
  - [21] V. Borkar, *Probability Theory: An Advanced Course*. Springer, 1995.
  - [22] H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, pages 400–407, 1951.
  - [23] S.W. He, J.G. Wang, and J.A. Yan, *Semimartingale theory and stochastic calculus*, Science Press, CRC Press Inc., 1992.