

Distributed Linear Parameter Estimation in Sensor Networks: Convergence Properties

Soumya Kar and José M. F. Moura

Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA 15213 USA
Email: soumyyak@andrew.cmu.edu, moura@ece.cmu.edu

Abstract—The paper considers the problem of distributed linear vector parameter estimation in sensor networks, when sensors can exchange quantized state information and the inter-sensor communication links fail randomly. We show that our algorithm \mathcal{LU} leads to almost sure (a.s.) consensus of the local sensor estimates to the true parameter value, under the assumptions that, a minimal global observability criterion is satisfied and the network is connected in the mean, i.e., $\lambda_2(\bar{L}) > 0$, where \bar{L} is the expected Laplacian matrix. We show that the local sensor estimates are asymptotically normal and characterize the convergence rate of the algorithm in the framework of moderate deviations.

Index Terms—Distributed Parameter Estimation, Random Link Failures, Quantization, Consistency, Asymptotic Normality.

I. INTRODUCTION

Wireless sensor network (WSN) applications generally consist of a large number of sensors which coordinate to perform a task in a distributed fashion. Unlike fusion-center based applications, there is no center and the task is performed locally at each sensor with intermittent inter-sensor message exchanges. In a coordinated environment monitoring or surveillance task, it translates to each sensor observing a part of the field of interest. With such local information, it is not possible for a particular sensor to get a reasonable estimate of the field. The sensors need to cooperate then and this is achieved by intermittent data exchanges among the sensors, whereby each sensor fuses its version of the estimate from time to time with those of other sensors with which it can communicate (in this context, see [1], [2], [3], [4], for a treatment of general distributed stochastic algorithms.) We consider the above problem in this paper in the context of distributed parameter estimation in WSNs. As an abstraction of the environment, we model it by a static vector parameter, whose dimension, M , can be arbitrarily large. We assume that each sensor receives noisy measurements (not necessarily additive) of only a part of the parameter vector. More specifically, if M_n is the dimension of the observation space of the n -th sensor, $M_n \ll M$. Assuming that the rate of receiving observations

at each sensor is comparable to the data exchange rate among sensors, each sensor updates its estimate at time index i by fusing it appropriately with the observation (innovation) received at i and the estimates at i of those sensors with which it can communicate at i . We propose the algorithm \mathcal{LU} (see also [5]) and show that under a minimum global observability constraint (required even by a centralized estimator to guarantee desirable estimation properties like consistency etc.), the estimate sequences generated at the sensors reach consensus a.s. and the limiting consensus value is the true parameter.

The inter-sensor communication is quantized with random link (communication channel) failures. This is appropriate, for example, in *digital* communication WSN when the data exchanges among a sensor and its neighbors are quantized, and the communication channels (or links) among sensors may fail at random times, e.g., as when packet dropouts occur randomly. We consider a very generic model of temporally independent link failures, whereby it is assumed that the sequence of network Laplacians, $\{L(i)\}_{i \geq 0}$ are i.i.d. with mean \bar{L} and satisfying $\lambda_2(\bar{L}) > 0$. We do not make any distributional assumptions on the link failure model. Although the link failures, and so the Laplacians, are independent at different times, during the same iteration, the link failures can be spatially dependent, i.e., correlated. This is more general and subsumes the erasure network model, where the link failures are independent over space *and* time. Wireless sensor networks motivate this model since interference among the wireless communication channels correlates the link failures over space, while, over time, it is still reasonable to assume that the channels are memoryless or independent. In particular, we do not require that the random instantiations of communication graph be connected; in fact, it is possible to have all these instantiations to be disconnected. We only require that the graph stays connected on *average*. This is captured by requiring that $\lambda_2(\bar{L}) > 0$, enabling us to capture a broad class of asynchronous communication models, as will be explained in the paper.

We analyze the algorithm \mathcal{LU} using stochastic approximation and show that the estimate sequence generated at each sensor is consistent, asymptotically unbiased and asymptotically normal. We explicitly characterize the asymptotic variance. The \mathcal{LU} algorithm can be regarded as a generalization of consensus algorithms (see, for example, [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]), the

This work was partially supported by the DARPA DSO Advanced Computing and Mathematics Program Integrated Sensing and Processing (ISP) Initiative under ARO grant # DAAD 19-02-1-0180, by NSF under grants # ECS-0225449 and # CNS-0428404, by the Office of Naval Research under MURI N000140710747, and by an IBM Faculty Award.

latter being a specific case of the \mathcal{LU} with no observations (innovations). We comment on the relevant recent literature on distributed estimation in WSNs. The papers [19], [20], [21], [22] consider the distributed estimation problem in non-random networks with unquantized inter-sensor transmission, mostly in the context of distributed least-mean-square (LMS) filtering. However, the goal in LMS-based approaches is not necessarily to obtain a consistent estimate of the entire parameter vector, but to minimize a certain regressed error. On the contrary, we consider random networks with quantized inter-sensor exchange and show that our algorithm leads to consistent estimates of the entire parameter vector at each sensor, whenever a minimal global observability condition is satisfied. We also show that the estimate sequences are asymptotically normal and explicitly compute the asymptotic variance, which sheds light on the convergence rate properties of our algorithm in the moderate deviations framework.

We comment briefly on the organization of the rest of the paper. In Section II we recall some concepts from spectral graph theory and statistical quantization theory. Section III sets up the problem of distributed linear parameter estimation in sensor networks, where we present the algorithm \mathcal{LU} with the assumptions. Main results regarding convergence properties of the \mathcal{LU} algorithm are stated in Section IV. Finally, Section V concludes the paper.

The proofs of the results in Section IV are omitted and can be found in [5].

II. NOTATION AND PRELIMINARIES

For completeness, this subsection sets notation and presents preliminaries on algebraic graph theory, matrices, and dithered quantization to be used in the sequel.

Preliminaries. We denote the k -dimensional Euclidean space by $\mathbb{R}^{k \times 1}$. The $k \times k$ identity matrix is denoted by I_k , while $\mathbf{1}_k, \mathbf{0}_k$ denote respectively the column vector of ones and zeros in $\mathbb{R}^{k \times 1}$. We also define the rank one $k \times k$ matrix P_k by

$$P_k = \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^T \quad (1)$$

The only non-zero eigenvalue of P_k is one, and the corresponding normalized eigenvector is $\left(\frac{1}{\sqrt{k}}\right) \mathbf{1}_k$. The operator $\|\cdot\|$ applied to a vector denotes the standard Euclidean 2-norm, while applied to matrices denotes the induced 2-norm, which is equivalent to the matrix spectral radius for symmetric matrices. We assume that the parameter to be estimated belongs to a subset \mathcal{U} of the Euclidean space $\mathbb{R}^{M \times 1}$. Throughout the paper, the true (but unknown) value of the parameter is denoted by θ^* . We denote a canonical element of \mathcal{U} by θ . The estimate of θ^* at time i at sensor n is denoted by $\mathbf{x}_n(i) \in \mathbb{R}^{M \times 1}$. Without loss of generality, we assume that the initial estimate, $\mathbf{x}_n(0)$, at time 0 at sensor n is a non-random quantity.

Spectral graph theory. We model the inter-sensor communication graph at time index i by a simple undirected graph $G(i) = (V, E(i))$, where V is the set of sensors with $|V| = N$ and $E(i)$ denotes the (random) set of edges (communication

links) active at i . The neighborhood of node n at i is

$$\Omega_n(i) = \{l \in V \mid (n, l) \in E(i)\} \quad (2)$$

Node n has degree $d_n(i) = |\Omega_n(i)|$ at i . The structure of the graph can be described by the symmetric $N \times N$ adjacency matrix, $A(i) = [A_{nl}(i)]$, $A_{nl}(i) = 1$, if $(n, l) \in E(i)$, $A_{nl}(i) = 0$, otherwise. Let the degree matrix be the diagonal matrix $D(i) = \text{diag}(d_1(i) \cdots d_N(i))$. The graph Laplacian matrix, $L(i)$, at i , is

$$L(i) = D(i) - A(i) \quad (3)$$

The Laplacian is a positive semidefinite matrix; hence, its eigenvalues can be ordered as

$$0 = \lambda_1(L(i)) \leq \lambda_2(L(i)) \leq \cdots \leq \lambda_N(L(i)) \quad (4)$$

The smallest eigenvalue $\lambda_1(L(i))$ is always equal to zero, with $\left(\frac{1}{\sqrt{N}}\right) \mathbf{1}_N$ being the corresponding normalized eigenvector. The multiplicity of the zero eigenvalue equals the number of connected components of the network; for a connected graph, $\lambda_2(L(i)) > 0$. This second eigenvalue is the algebraic connectivity or the Fiedler value of the network; see [23] for detailed treatment of graphs and their spectral theory.

We now review results from statistical quantization theory.

Quantizer: We assume that all sensors are equipped with identical quantizers, which uniformly quantize each component of the M -dimensional estimates by the quantizing function, $\mathbf{q}(\cdot) : \mathbb{R}^{M \times 1} \rightarrow \mathcal{Q}^M$. For $\mathbf{y} \in \mathbb{R}^{M \times 1}$ we have, $\mathbf{q}(\mathbf{y}) = [k_1 \Delta, \dots, k_M \Delta] = \mathbf{y} + \mathbf{e}(\mathbf{y})$, where

$$(k_m - \frac{1}{2})\Delta \leq y_i < (k_m + \frac{1}{2})\Delta, \quad 1 \leq m \leq M \quad (5)$$

$$\frac{\Delta}{2} \mathbf{1}_N \leq \mathbf{e}(\mathbf{y}) < \frac{\Delta}{2} \mathbf{1}_N \quad (6)$$

and $\mathbf{e}(\mathbf{y})$ is the quantization error and the inequalities are interpreted component-wise. The quantizer alphabet is

$$\mathcal{Q}^M = \left\{ [k_1 \Delta, \dots, k_M \Delta]^T \mid k_i \in \mathbb{Z}, \forall i \right\} \quad (7)$$

We take the quantizer alphabet to be countable because no *a priori* bound is assumed on the parameter. Conditioned on the input, the quantization error $\mathbf{e}(\mathbf{y})$ is deterministic. This strong correlation of the error with the input creates unacceptable statistical properties. In particular, for iterative algorithms, it leads to error accumulation and divergence of the algorithm (see the discussion in [24].) To avoid this divergence, we consider dithered quantization, which makes the quantization error possess nice statistical properties. We review briefly basic results on dithered quantization, which are needed in the sequel.

Dithered Quantization: Schuchman Conditions Consider a uniform scalar quantizer $q(\cdot)$ of step-size Δ , where $y \in \mathbb{R}$ is the channel input. Let $\{y(i)\}_{i \geq 0}$ be a scalar input sequence to which we added a dither sequence $\{\nu(i)\}_{i \geq 0}$ of i.i.d. uniformly distributed random variables on $[-\Delta/2, \Delta/2)$, independent of the input sequence $\{y(i)\}_{i \geq 0}$. This is a sufficient condition for the dither to satisfy the Schuchman conditions (see [25]).

Under these conditions, the error sequence for subtractively dithered systems ([26]) $\{\varepsilon(i)\}_{i \geq 0}$

$$\varepsilon(i) = q(y(i) + \nu(i)) - (y(i) + \nu(i)) \quad (8)$$

is an i.i.d. sequence of uniformly distributed random variables on $[-\Delta/2, \Delta/2)$, which is independent of the input sequence $\{y(i)\}_{i \geq 0}$. To be more precise, this result is valid if the quantizer does not overload, which is trivially satisfied here as the dynamic range of the quantizer is the entire real line. Thus, by randomizing appropriately the input to a uniform quantizer, we can render the error to be independent of the input and uniformly distributed on $[-\Delta/2, \Delta/2)$. This leads to nice statistical properties of the error, which we exploit in this paper.

III. PROBLEM FORMULATION

Let $\theta^* \in \mathbb{R}^{M \times 1}$ be an M -dimensional parameter that is to be estimated by a network of N sensors. We refer to θ as a parameter, although it is a vector of M parameters. Each sensor makes i.i.d. observations of noise corrupted linear functions of the parameter. We assume the following observation model for the n -th sensor:

$$\mathbf{z}_n(i) = H_n(i)\theta^* + \zeta_n(i) \quad (9)$$

where: $\{\mathbf{z}_n(i) \in \mathbb{R}^{M_n \times 1}\}_{i \geq 0}$ is the i.i.d. observation sequence for the n -th sensor; $\{\zeta_n(i)\}_{i \geq 0}$ is a zero-mean i.i.d. noise sequence of bounded variance; and $\{H_n(i)\}_{i \geq 0}$ is an i.i.d. sequence of observation matrices with mean \bar{H}_n and bounded second moment. For most practical sensor network applications, each sensor observes only a subset of M_n of the components of θ , with $M_n \ll M$. Under such a situation, in isolation, each sensor can estimate at most only a part of the parameter. However, if the sensor network is connected in the mean sense (as will be explained later), and under appropriate observability conditions, we will show that it is possible for each sensor to get a consistent estimate of the parameter θ^* by means of quantized local inter-sensor communication.

In this subsection, we present the algorithm \mathcal{LU} for distributed parameter estimation in the linear observation model (9). Starting from some initial deterministic estimate of the parameters (the initial states may be random, we assume deterministic for notational simplicity), $\mathbf{x}_n(0) \in \mathbb{R}^{M \times 1}$, each sensor generates by a distributed iterative algorithm a sequence of estimates, $\{\mathbf{x}_n(i)\}_{i \geq 0}$. The parameter estimate $\mathbf{x}_n(i+1)$ at the n -th sensor at time $i+1$ is a function of: its previous estimate; the communicated quantized estimates at time i of its neighboring sensors; and the new observation $\mathbf{z}_n(i)$. As described in Section II, the data is subtractively dithered quantized, i.e., there exists a vector quantizer $\mathbf{q}(\cdot)$ and a family, $\{\nu_{nl}^m(i)\}$, of i.i.d. uniformly distributed random variables on $[-\Delta/2, \Delta/2)$ such that the quantized data received by the n -th sensor from the l -th sensor at time i is $\mathbf{q}(\mathbf{x}_l(i) + \nu_{nl}(i))$, where $\nu_{nl}(i) = [\nu_{nl}^1(i), \dots, \nu_{nl}^M(i)]^T$. It then follows from the discussion in Section II that the quantization error, $\varepsilon_{nl}(i) \in \mathbb{R}^{M \times 1}$ given by (8), is a random vector, whose components are i.i.d. uniform on $[-\Delta/2, \Delta/2)$ and independent of $\mathbf{x}_l(i)$.

Algorithm \mathcal{LU} Based on the current state $\mathbf{x}_n(i)$, the quantized exchanged data $\{\mathbf{q}(\mathbf{x}_l(i) + \nu_{nl}(i))\}_{l \in \Omega_n(i)}$, and the observation $\mathbf{z}_n(i)$, we update the estimate at the n -th sensor by the following distributed iterative algorithm:

$$\begin{aligned} \mathbf{x}_n(i+1) = & \mathbf{x}_n(i) - \alpha(i) \left[b \sum_{l \in \Omega_n(i)} (\mathbf{x}_n(i) \right. \\ & \left. - \mathbf{q}(\mathbf{x}_l(i) + \nu_{nl}(i))) - \bar{H}_n^T (\mathbf{z}_n(i) - \bar{H}_n \mathbf{x}_n(i)) \right] \end{aligned} \quad (10)$$

In (10), $b > 0$ is a constant and $\{\alpha(i)\}_{i \geq 0}$ is a sequence of weights with properties to be defined below. Algorithm (10) is distributed because for sensor n it involves only the data from the sensors in its neighborhood $\Omega_n(i)$. Using eqn. (8), the state update can be written as

$$\begin{aligned} \mathbf{x}_n(i+1) = & \mathbf{x}_n(i) - \alpha(i) \left[b \sum_{l \in \Omega_n(i)} (\mathbf{x}_n(i) - \mathbf{x}_l(i)) \right. \\ & \left. - \bar{H}_n^T (\mathbf{z}_n(i) - \bar{H}_n \mathbf{x}_n(i)) + b\nu_{nl}(i) + b\varepsilon_{nl}(i) \right] \end{aligned} \quad (11)$$

We rewrite (11) in compact form. Define the random vectors, $\Upsilon(i)$ and $\Psi(i) \in \mathbb{R}^{NM \times 1}$ with vector components

$$\Upsilon_n(i) = - \sum_{l \in \Omega_n(i)} \nu_{nl}(i) \quad (12)$$

$$\Psi_n(i) = - \sum_{l \in \Omega_n(i)} \varepsilon_{nl}(i) \quad (13)$$

It follows from the Schuchman conditions on the dither that

$$\mathbb{E}[\Upsilon(i)] = \mathbb{E}[\Psi(i)] = \mathbf{0}, \quad \forall i \quad (14)$$

$$\sup_i \mathbb{E}[\|\Upsilon(i)\|^2] = \sup_i \mathbb{E}[\|\Psi(i)\|^2] \leq \frac{N(N-1)M\Delta^2}{12} \quad (15)$$

The iterations in (10) can be written in compact form as:

$$\begin{aligned} \mathbf{x}(i+1) = & \mathbf{x}(i) - \alpha(i) [b(L(i) \otimes I_M)\mathbf{x}(i) \\ & - \bar{D}_{\bar{H}} (\mathbf{z}(i) - \bar{D}_{\bar{H}}^T \mathbf{x}(i)) + b\Upsilon(i) + b\Psi(i)] \end{aligned} \quad (16)$$

Here, $\mathbf{x}(i) = [\mathbf{x}_1^T(i) \dots \mathbf{x}_N^T(i)]^T$ is the vector of sensor states (estimates.) The sequence of Laplacian matrices $\{L(i)\}_{i \geq 0}$ captures the topology of the sensor network. They are random to accommodate link failures, which occur in packet communications. We also define the matrices $\bar{D}_{\bar{H}}$ and $D_{\bar{H}}$ as

$$\bar{D}_{\bar{H}} = \text{diag} [\bar{H}_1^T \dots \bar{H}_N^T] \quad (17)$$

$$D_{\bar{H}} = \bar{D}_{\bar{H}} \bar{D}_{\bar{H}}^T = \text{diag} [\bar{H}_1^T \bar{H}_1 \dots \bar{H}_N^T \bar{H}_N] \quad (18)$$

Also define

$$S_q = \mathbb{E} [(\Upsilon(i) + \Psi(i)) (\Upsilon(i) + \Psi(i))^T] \quad (19)$$

We refer to the recursive estimation algorithm in eqn. (16) as \mathcal{LU} . We now summarize formally the assumptions on the \mathcal{LU} algorithm and their implications.

A.1) Observation Noise. Recall the observation model in

eqn. (9). We assume that the observation noise process, $\{\zeta(i) = [\zeta_1^T(i), \dots, \zeta_N^T(i)]^T\}_{i \geq 0}$ is an i.i.d. zero mean process, with finite second moment. In particular, the observation noise covariance is independent of i

$$\mathbb{E}[\zeta(i)\zeta^T(j)] = S_\zeta \delta_{ij}, \forall i, j \geq 0 \quad (20)$$

where the Kronecker symbol $\delta_{ij} = 1$ if $i = j$ and zero otherwise. Note that the observation noises at different sensors may be correlated during a particular iteration. Eqn. (20) states only temporal independence. The spatial correlation of the observation noise makes our model applicable to practical sensor network problems, for instance, for distributed target localization, where the observation noise is generally correlated across sensors.

A.2) Observability. We assume that the observation matrices, $\{[H_1(i), \dots, H_N(i)]\}_{i \geq 0}$, form an i.i.d. sequence with mean $[\overline{H}_1, \dots, \overline{H}_N]$ and finite second moment. In particular, we have

$$H_n(i) = \overline{H}_n + \tilde{H}_n(i), \forall i, n \quad (21)$$

where, $\overline{H}_n = \mathbb{E}[H_n(i)]$, $\forall i, n$ and $\{[\tilde{H}_1(i), \dots, \tilde{H}_N(i)]\}_{i \geq 0}$ is a zero mean i.i.d. sequence with finite second moment. Here, also, we require only temporal independence of the observation matrices, but allow them to be spatially correlated. We require the following global observability condition. The matrix G

$$G = \sum_{n=1}^N \overline{H}_n^T \overline{H}_n \quad (22)$$

is full-rank. This distributed observability extends the observability condition for a centralized estimator to get a consistent estimate of the parameter θ^* . We note that the information available to the n -th sensor at any time i about the corresponding observation matrix is just the mean \overline{H}_n , and *not* the random $H_n(i)$. Hence, the state update equation uses only the \overline{H}_n 's, as given in eqn. (10).

A.3) Random Link Failure. In digital communications, packets may be lost at random times. To account for this, we let the links (or communication channels among sensors) to fail, so that the edge set and the connectivity graph of the sensor network are time varying. Accordingly, the sensor network at time i is modeled as an undirected graph, $G(i) = (V, E(i))$ and the graph Laplacians as a sequence of i.i.d. Laplacian matrices $\{L(i)\}_{i \geq 0}$. We write

$$L(i) = \overline{L} + \tilde{L}(i), \forall i \geq 0 \quad (23)$$

where the mean $\overline{L} = \mathbb{E}[L(i)]$. We do not make any distributional assumptions on the link failure model. Although the link failures, and so the Laplacians, are independent at different times, during the same iteration, the link failures can be spatially dependent, i.e., correlated. This is more general and subsumes the erasure network model, where the link failures are independent over space *and* time. Wireless sensor networks motivate this model since interference among the

wireless communication channels correlates the link failures over space, while, over time, it is still reasonable to assume that the channels are memoryless or independent.

Connectedness of the graph is an important issue. We do not require that the random instantiations $G(i)$ of the graph be connected; in fact, it is possible to have all these instantiations to be disconnected. We only require that the graph stays connected on *average*. This is captured by requiring that $\lambda_2(\overline{L}) > 0$, enabling us to capture a broad class of asynchronous communication models; for example, the random asynchronous gossip protocol analyzed in [27] satisfies $\lambda_2(\overline{L}) > 0$ and hence falls under this framework.

A.4) Persistence Condition. The weight sequence $\{\alpha(i)\}_{i \geq 0}$ satisfies

$$\alpha(i) > 0, \sum_{i \geq 0} \alpha(i) = \infty, \sum_{i \geq 0} \alpha^2(i) < \infty \quad (24)$$

This condition is commonly assumed in adaptive control and signal processing and implies, in particular, that, $\alpha(i) \rightarrow 0$. Examples include

$$\alpha(i) = \frac{1}{i^\beta}, \quad .5 < \beta \leq 1 \quad (25)$$

A.5) Independence Assumptions. The sequences $\{L(i)\}_{i \geq 0}, \{\zeta_n(i)\}_{1 \leq n \leq N, i \geq 0}, \{H_n(i)\}_{1 \leq n \leq N, i \geq 0}, \{\nu_{nl}^m(i)\}$ are mutually independent.

In the following sections we state the main results regarding the convergence properties of the \mathcal{LU} algorithm. For proofs, see [5].

IV. MAIN RESULTS

The first result concerns the consistency and asymptotic unbiasedness of the algorithm \mathcal{LU} .

Theorem 1 (\mathcal{LU} : Consistency and Unbiasedness) Consider the \mathcal{LU} algorithm with the assumptions **A.1)-A.5)**. Then,

$$\mathbb{P}\left[\lim_{i \rightarrow \infty} \mathbf{x}_n(i) = \theta^*, \forall n\right] = 1 \quad (26)$$

In other words, the estimate sequence $\{\mathbf{x}_n(i)\}_{i \geq 0}$ at a sensor n , is a consistent estimate of the parameter θ^* .

Also, we have

$$\lim_{i \rightarrow \infty} \mathbb{E}[\mathbf{x}_n(i)] = \theta^*, \quad 1 \leq n \leq N \quad (27)$$

Thus, the estimate sequences generated at the sensors reach consensus a.s. and the limiting consensus variable is the true parameter value θ^* .

We now show that for appropriate choice of the weight sequence $\{\alpha(i)\}_{i \geq 0}$ the estimate sequence at each sensor is asymptotically normal. We explicitly characterize the asymptotic variance, which establishes convergence properties of the algorithm \mathcal{LU} (see [5] for detailed explanation.)

Theorem 2 (\mathcal{LU} : Asymptotic normality) Consider the \mathcal{LU} algorithm under **A.1)-A.5)** with link weight sequence, $\{\alpha(i)\}_{i \geq 0}$ given by:

$$\alpha(i) = \frac{a}{i+1}, \quad \forall i \quad (28)$$

for some constant $a > 0$. Let $\{\mathbf{x}(i)\}_{i \geq 0}$ be the state sequence generated. Then, if $a > \frac{1}{2\lambda_{\min}(\overline{bL} \otimes I_M + D_{\overline{H}})}$, we have

$$\sqrt{i}(\mathbf{x}(i) - \mathbf{1}_N \otimes \theta^*) \implies \mathcal{N}(\mathbf{0}, S(\theta^*)) \quad (29)$$

where

$$S(\theta^*) = a^2 \int_0^\infty e^{\Sigma v} S_0 e^{\Sigma v} dv \quad (30)$$

$$\Sigma = -a [\overline{bL} \otimes I_M + D_{\overline{H}}] + \frac{1}{2} I \quad (31)$$

$$S_0 = S_H + \overline{D}_{\overline{H}} S_\zeta \overline{D}_{\overline{H}}^T + b^2 S_q \quad (32)$$

In particular, at any sensor n , the estimate sequence, $\{\mathbf{x}_n(i)\}_{i \geq 0}$ is asymptotically normal:

$$\sqrt{i}(\mathbf{x}_n(i) - \theta^*) \implies \mathcal{N}(\mathbf{0}, S_{nn}(\theta^*)) \quad (33)$$

where, $S_{nn}(\theta^*) \in \mathbb{R}^{M \times M}$ denotes the n -th principal block of $S(\theta^*)$.

Here \implies denotes convergence in distribution or weak convergence and

$$S_H = \mathbb{E} \left[\left(\overline{D}_{\overline{H}} \begin{bmatrix} \tilde{H}_1(i) \\ \cdot \\ \cdot \\ \tilde{H}_N(i) \end{bmatrix} \mathbf{1}_N \theta^* \right) \left(\overline{D}_{\overline{H}} \begin{bmatrix} \tilde{H}_1(i) \\ \cdot \\ \cdot \\ \tilde{H}_N(i) \end{bmatrix} \mathbf{1}_N \theta^* \right)^T \right] \quad (34)$$

and S_ζ, S_q are defined in eqns. (20,19).

V. CONCLUSION

In this paper we study the \mathcal{LU} algorithm for linear distributed parameter estimation in sensor networks, when inter-sensor communication links fail randomly and inter-sensor communication is quantized. We show that the algorithm \mathcal{LU} is consistent, i.e., it leads to a.s. consensus of the estimate sequences generated at the sensors and the limiting consensus variable is the true parameter value. We establish asymptotic normality of the estimate sequences generated at the sensors (under appropriate assumptions on the weight sequence $\{\alpha(i)\}_{i \geq 0}$) and compute the asymptotic variances explicitly.

REFERENCES

- [1] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D., Massachusetts Institute of Technology, Cambridge, MA, 1984.
- [2] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-31, no. 9, pp. 803–812, 1986.
- [3] D. Bertsekas, J. Tsitsiklis, and M. Athans, "Convergence theories of distributed iterative processes: A survey," *Technical Report for Information and Decision Systems, Massachusetts Inst. of Technology, Cambridge, MA*, 1984.
- [4] H. Kushner and G. Yin, "Asymptotic properties of distributed and communicating stochastic approximation algorithms," *Siam J. Control and Optimization*, vol. 25, no. 5, pp. 1266–1290, Sept. 1987.

- [5] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: nonlinear observation models and imperfect communication," August 2008, submitted for publication, 51 pages. [Online]. Available: <http://arxiv.org/abs/0809.0009>
- [6] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.
- [7] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Automat. Contr.*, vol. AC-48, no. 6, pp. 988–1001, June 2003.
- [8] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Contr. Lett.*, vol. 53, pp. 65–78, 2004.
- [9] S. Kar and J. M. F. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3315–3326, July 2008.
- [10] —, "Distributed consensus algorithms in sensor networks with communication channel noise and random link failures," in *41st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2007.
- [11] —, "Distributed average consensus in sensor networks with quantized inter-sensor communication," in *Proceedings of the 33rd International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, USA, April 1-4 2008.
- [12] Y. Hatano and M. Mesbahi, "Agreement over random networks," in *43rd IEEE Conference on Decision and Control*, vol. 2, Dec. 2004, pp. 2010–2015.
- [13] T. C. Aysal, M. Coates, and M. Rabbat, "Distributed average consensus using probabilistic quantization," in *IEEE/SP 14th Workshop on Statistical Signal Processing Workshop*, Maddison, Wisconsin, USA, August 2007, pp. 640–644.
- [14] M. E. Yildiz and A. Scaglione, "Differential nested lattice encoding for consensus problems," in *ACM/IEEE Information Processing in Sensor Networks*, Cambridge, MA, April 2007.
- [15] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, pp. 1192–1203, July 2007.
- [16] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri, "Average consensus on networks with quantized communication," *Submitted to the Int. J. Robust and Nonlinear Control*, 2008.
- [17] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *Technical Report 2778, LIDS-MIT*, Nov. 2007.
- [18] M. Huang and J. Manton, "Stochastic approximation for consensus seeking: mean square and almost sure convergence," in *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, LA, USA, Dec. 12-14 2007.
- [19] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, July 2008.
- [20] S. Stankovic, M. Stankovic, and D. Stipanovic, "Decentralized parameter estimation by consensus based stochastic approximation," in *46th IEEE Conference on Decision and Control*, New Orleans, LA, USA, 12-14 Dec. 2007, pp. 1535–1540.
- [21] I. Schizas, G. Mateos, and G. Giannakis, "Stability analysis of the consensus-based distributed lms algorithm," in *Proceedings of the 33rd International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, USA, April 1-4 2008, pp. 3289–3292.
- [22] S. Ram, V. Veeravalli, and A. Nedic, "Distributed and recursive parameter estimation in parametrized linear state-space models," *Submitted for publication*, April 2008.
- [23] F. R. K. Chung, *Spectral Graph Theory*. Providence, RI : American Mathematical Society, 1997.
- [24] S. Kar and J. Moura, "Distributed consensus algorithms in sensor networks: Quantized data," November 2007, submitted for publication, 30 pages. [Online]. Available: <http://arxiv.org/abs/0712.1609>
- [25] L. Schuchman, "Dither signals and their effect on quantization noise," *IEEE Trans. Commun. Technol.*, vol. COMM-12, pp. 162–165, December 1964.
- [26] S. P. Lipshitz, R. A. Wannamaker, and J. Vanderkooy, "Quantization and dither: A theoretical survey," *J. Audio Eng. Soc.*, vol. 40, pp. 355–375, May 1992.
- [27] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Trans. Netw.*, vol. 14, no. SI, pp. 2508–2530, 2006.