Topology-Agnostic Average Consensus in Sensor Networks with Limited Data Rate

Chang-Shen Lee, Nicolò Michelusi, and Gesualdo Scutari

Abstract—**In this paper, the distributed average consensus problem in sensor networks with limited data rate communication is studied. Unlike standard average consensus, only quantized signals with finite support are adopted for the communications among agents. To tackle this problem, a novel distributed algorithm is proposed, where each agent iteratively updates a *local* estimate based on *quantized* signals received by its neighbors. The proposed algorithm differs from the existing schemes dealing with limited data rate in the following key features: 1) each agent is not required to have information on spectral properties of the graph associated with the communication topology; 2) the initial measurements are not required to be bounded within a known interval; and 3) exact consensus to the average can be achieved asymptotically for *weight-balanced* directed topology. Thus, it is more favorable for practical implementations, especially for large networks. The proposed algorithm is proved to achieve average consensus asymptotically, almost surely and in mean square sense. The analysis of convergence rate and generalizations to random weight-balanced directed topologies and time-varying quantization are also provided. Finally, numerical results validate our theoretical findings, and demonstrate the superior performance of the proposed algorithm compared to existing topology-agnostic consensus schemes with limited data rate.**

I. Introduction

Distributed average consensus has attracted considerable attention in recent years; some representative applications include load balancing [1], vehicle formation [2], and sensor networks [3]. Since the seminal paper [4], where the first consensus scheme was proposed, several subsequent work appeared in the literature (e.g. [5], [6]) studying variations and extensions of the original protocol [4]. However, in most work it is assumed that agents can exchange precise information, which implies infinite channel capacity among agents. In practice, however, agents can only exchange *quantized* information, which motivates more recent research.

Distributed average consensus with quantized communications has been studied recently in [7]–[11]. Therein, each agent has its initial value, and the goal is to reach consensus on the average value through local quantized information exchanges. In [7], [9], the agents' local estimates are restricted to be discrete and thus the consensus can only be reached on the quantized value which is nearest to the average. In [8], the dithered quantization is introduced, which is equivalent to the probabilistic quantization, and consensus to a random variable whose expected value is the desired average is shown to be achieved almost surely. Almost sure and mean square sense convergence was proved in [10] for a similar setting. In [11], consensus to the desired average is proved both almost surely and in mean square sense for directed graphs.

Lee and Michelusi are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA; emails: <lee2495,michelus>@purdue.edu.

Scutari is with the School of Industrial Engineering, Purdue University, West Lafayette, IN, USA; email: gscutari@purdue.edu. The work of Lee and Scutari is partially supported by the NSF grant CCF 1564044 and the ONR grant N00014-16-1-2244.

The above works either assume that the number of quantization levels is infinite [11], which is impractical due to finite link capacities, or fail to converge to the average of the initial measurements [7]–[10], which is not desirable. To solve the distributed average consensus problem with finite level quantization, several algorithms based on dynamic encoding/decoding have been proposed in recent years, e.g., [12], [13]. In [12], a finite level deterministic dynamic quantizer is proposed, and convergence to the desired average is proved. In [13], a dynamic finite level quantizer is proposed which progressively reduces the quantization range to speed up consensus, but no performance guarantees are given.

However, the aforementioned dynamic encoding/decoding based algorithms require each agent to have knowledge of some spectral properties of the Laplacian matrix of the communication graph. Furthermore, each agent is required to know the bound on all initial measurements, which may not be the case when large networks are considered, since the maximum deviation of measurements will increase as the number of measurements increases (for instance, under Gaussian noise). Since the above information is used to design the quantizer and control parameters, convergence of the dynamic encoding/decoding based algorithms is no longer guaranteed without the above information. To the best of our knowledge, the distributed average consensus problem under directed topology using finite level quantization, without the need of spectral properties of the communication graph and uniform bound of the measurements at each agent, is still an open problem.

In this paper, we fill the gap and propose an algorithm for the distributed average consensus problem under weight-balanced directed topology with limited data rate. In the proposed algorithm, each agent is neither required to have knowledge of the communication graph, nor of the uniform bound of all measurements. We show that consensus to the desired average can be achieved asymptotically, as long as (i) the communication topology is strongly connected and *weight-balanced*, and (ii) the *average* of the (initial) measurements fall inside the quantization range. Clearly, strong connectivity is required to reach consensus, and (ii) is a much milder and more natural requirement than the uniform boundedness of the measurements, typically assumed in most prior work (e.g., see [10], [12], [14]). Note that the proposed algorithm can be regarded as a generalization of the one in [14] to weight-balanced directed graphs. However, our analysis covers more generalizations (e.g., static and i.i.d. random weight-balabced directed graphs) than [14], see [15]. The performance superiority of the proposed algorithm stems from the fact that we tackle *saturation* of the quantizer differently from dynamic encoding/decoding based algorithms. The key idea therein is to carefully scale the input to the quantizer, so that the latter will never be saturated (or be saturated with small probability [13]).

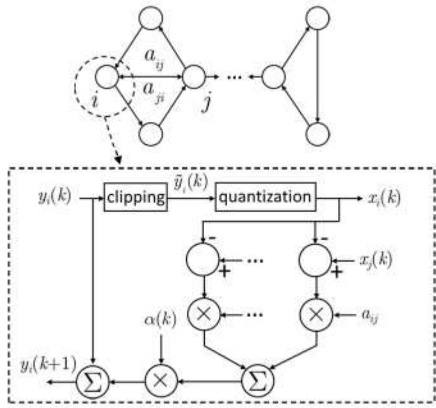

Fig. 1. System and signal model.

In contrast, our algorithm adopts a more intuitive approach: when data fall outside the range, they are simply clipped and then quantized. However, since clipping is a *non-linear* operation, the analysis of the proposed scheme is much more involved. Therefore, another major contribution of this work is that we directly tackle such *non-linearity* and show that consensus to the average can still be achieved almost surely and in mean square sense. We show that the convergence results also hold for two generalizations: i.i.d. random topologies and time-varying quantization. Finally, we numerically evaluate the effectiveness of the proposed algorithm.

The paper is organized as follows. In Section II, we introduce the system model and the proposed algorithm. In Section III, we elaborate the analytical results. In Section IV, we numerically validate our theoretical findings, followed by some final remarks in Section V. Due to the space limitations, the proofs are provided in [15]. We denote the vector with all components equal to 1, the one with all components equal to 0, and the ith canonical vector as $\mathbf{1}$, $\mathbf{0}$ and $\mathbf{e}_i$, respectively.

II. Problem Formulation and Proposed Algorithm

Consider a sensor network with N agents, depicted in Fig. 1. Each agent i has its observation s_i of a common unknown parameter $\theta \in (\theta_{\min}, \theta_{\max})$, with boundaries $\theta_{\min} < \theta_{\max}$ known at each agent. The goal of distributed average consensus is to compute, at each agent, an average of the initial measurements, that is,

$$\bar{s} = \frac{1}{N}\sum_{i=1}^{N} s_i, \tag{1}$$

using only local processing and information exchange among neighbors. For instance, $s_i = \theta + w_i$, where w_i denotes Gaussian noise, so that $\bar{s}$ represents the sample average measurement across the network. Clearly, s_i is not bounded since the Gaussian noise w_i can take any value in $\mathbb{R}$. The measurements are said to be *informative* if $\bar{s} \in (\theta_{\min}, \theta_{\max})$.

The communication network is modeled as a static directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the vertex set, with each vertex corresponding to an agent, and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the edge set, with each edge representing the communication link between two agents. Let $\mathcal{N}_i^+, \mathcal{N}_i^-$ be the out-neighbor and in-neighbor set of agent i, i.e., $j \in \mathcal{N}_i^+$ and $i \in \mathcal{N}_j^-$, iff $(i,j) \in \mathcal{E}$. In other words, agent j can receive the signal from agent i. Let $\mathbf{A} = [a_{i,j}]_{N\times N}$ be the weight matrix associated with $\mathcal{G}$, $\mathbf{D}$ be the degree matrix and $\mathbf{L}$ the Laplacian matrix, defined as $\mathbf{D} = \mathrm{diag}\{d_1, \cdots, d_N\}$, with $d_i = \sum_{j\in\mathcal{N}_i^+} a_{j,i}$, and $\mathbf{L} = \mathbf{D} - \mathbf{A}$. In the following, we state two assumptions on the graph and its weight matrix.

Assumption 1: The graph $\mathcal{G}$ is connected, and does not contain self loops, i.e., $(i,i) \notin \mathcal{E}, \forall i$.

Assumption 2: The graph is weight-balanced, i.e., $\sum_{j\in\mathcal{N}_i^+} a_{j,i} = \sum_{j\in\mathcal{N}_i^-} a_{i,j}, \forall i$, with $a_{j,i} > 0$ iff $(i,j) \in \mathcal{E}$, and $a_{j,i} = 0$ iff $(i,j) \notin \mathcal{E}$.

Let $y_i(k)$ be the local estimate of the average of the initial measurements $\bar{s}$, owned by agent i at iteration k. We assume that communication occurs at finite rate, thus the local estimate $y_i(k)$ needs to be quantized before transmission. Specifically, at time k, each agent i first clips its local estimate $y_i(k)$ within the quantization range $[q_{\min}, q_{\max}]$, generating the *clipped* estimate $\tilde{y}_i(k)$ given by

$$\tilde{y}_i(k) = \min\big\{\max\{y_i(k), q_{\min}\}, q_{\max}\big\}. \tag{2}$$

Then, it quantizes the clipped estimate $\tilde{y}_i(k)$ with B bits, building the quantized signal $x_i(k)$, given by

$$x_i(k) = \begin{cases} q_{\min} + \left\lceil \frac{\tilde{y}_i(k) - q_{\min}}{\Delta} \right\rceil \Delta, & \text{with probability } p_i(k), \\ q_{\min} + \left\lfloor \frac{\tilde{y}_i(k) - q_{\min}}{\Delta} \right\rfloor \Delta, & \text{otherwise}, \end{cases} \tag{3}$$

where $\Delta = \frac{q_{\max} - q_{\min}}{2^B - 1}$ is the quantization step size, $\lceil\cdot\rceil$, $\lfloor\cdot\rfloor$ are the ceiling and floor functions, respectively, and

$$p_i(k) = \frac{\tilde{y}_i(k) - q_{\min}}{\Delta} - \left\lfloor \frac{\tilde{y}_i(k) - q_{\min}}{\Delta} \right\rfloor$$

is the probability of the probabilistic quantizer, whose value is such that $\mathbb{E}[x_i(k)|\tilde{y}_i(k)] = \tilde{y}_i(k)$. Finally, agent i broadcasts such a quantized estimate to its neighbors $\mathcal{N}_i^+$. After receiving the signal from its neighbors, each agent i will update its local estimate $y_i(k)$ based on the received signal $x_j(k), j \in \mathcal{N}_i^-$, according to the following dynamics:

$$y_i(k+1) = y_i(k) + \alpha(k) \sum_{j\in\mathcal{N}_i^-} a_{i,j}[x_j(k) - x_i(k)], \forall k \geq 1, \tag{4}$$

where $\alpha(k) \geq 0$ is the step-size and $y_i(1) = s_i$ at initialization.

Remark 1: Since the input to the quantizer $\tilde{y}_i(k)$ always falls within the range $[q_{\min}, q_{\max}]$, the quantizer will never be saturated. By doing so, the probabilistic quantizer is equivalent to the dithered quantizer with dither noise being a random variable, uniform in $[-\Delta/2, \Delta/2)$ (cf. [8, Lemma 2]).

Remark 2: The update rule (4) has the following property: the average of the local estimates is preserved with probability 1. In contrast, the topology-agnostic consensus algorithm with limited data rate for undirected graph developed in [10], i.e.,

$$y_i(k+1) = y_i(k) + \alpha(k) \sum_{j\in\mathcal{N}_i^-} a_{i,j}[x_j(k) - y_i(k)], \tag{5}$$

does not have this property.[1] As we will show numerically in Section IV, such property leads to better performance of the proposed algorithm compared to [10].

[1]To be precise, in [10], $a_{j,i} = 1$ if $(i,j) \in \mathcal{E}$ and $a_{j,i} = 0$ if $(i,j) \notin \mathcal{E}$, so (5) can be regarded as a generalization of the algorithm in [10].

III. Analytical Results

In this section, we will first present the convergence result of the proposed algorithm (cf. Theorems 1 and 2), and then analyze its convergence rate. We will also generalize the proposed algorithm to 1) i.i.d. random weight-balanced directed graphs, and 2) time-varying quantization.

A. Convergence to the average of the initial measurements

Let $\mathbf{y}(k) = [y_1(k), \cdots, y_N(k)]^T$. The following two theorems state convergence in mean square sense (m.s.s.) and almost sure (a.s.) convergence of the local estimate y_i to the average of the initial measurements.

Theorem 1 (Convergence to the Average of Initial Measurements in m.s.s.): Given the quantized consensus protocol (4), suppose that:

(i) The measurements are informative, $\bar{s} \in (q_{\min}, q_{\max})$.
(ii) $\alpha(k)$ satisfies

$$\alpha(k) > 0, \forall k \geq 1, \sum_{k \geq 1} \alpha(k) = \infty, \lim_{k \to \infty} \alpha(k) = 0. \quad (6)$$

Then,

$$\lim_{k \to \infty} \mathbb{E}\left[\|\mathbf{y}(k) - \bar{s}\mathbf{1}\|^2\right] = 0.$$

Proof: See [15]. ∎

Theorem 2 (A.s. Convergence to the Average of the Initial Measurements): Given the quantized consensus protocol (4), suppose that:

(i) $\bar{s} \in (q_{\min}, q_{\max})$.
(ii) $\alpha(k)$ satisfies

$$(6) \text{ and } \sum_{k \geq 1} \alpha(k)^2 < \infty. \quad (7)$$

Then,

$$\mathbb{P}\left\{\lim_{k \to \infty} \mathbf{y}(k) = \bar{s}\mathbf{1}\right\} = 1.$$

Proof: See [15]. ∎

Remark 3 (Selection of parameters): In practical implementation, we can simply set $(q_{\min}, q_{\max}) = (\theta_{\min}, \theta_{\max})$. By doing so, if $\bar{s} \in (q_{\min}, q_{\max})$, i.e., if the measurements are *informative*, the algorithm is guaranteed to converge to $\bar{s}$, as desired. Herein, we are not interested in the case of non-informative measurements $\bar{s} \notin (q_{\min}, q_{\max})$, which implies that the sample average $\bar{s}$ is not a desired estimate of θ.

Conditions (6) and (7) ensure that $\mathbf{y}(k)$ is able to travel far enough to reach the desired value ($\sum_{k \geq 1} \alpha(k) = \infty$), while not traveling too far ($\lim_{k \to \infty} \alpha(k) = 0$ in (6) and $\sum_{k \geq 1} \alpha(k)^2 < \infty$ in (7)), so as to reach convergence. The use of such sequences is a common technique in stochastic optimization [16]. There are many choices of $\{\alpha(k)\}_{k \geq 1}$ satisfying (7). For instance,

(a) $\alpha(k) = \frac{c}{k^\beta}, \forall k \geq 1$ with $c > 0, 0.5 < \beta \leq 1$ [10], [17];
(b) $\alpha(k) = \alpha(k-1)\left[1 - \mu\alpha(k-1)\right], \forall k \geq 2$ with $\alpha(1) \in (0, 1], \mu \in (0, 1)$ [17].

Note that (6) is a looser condition compared to (7), i.e., all sequences satisfying (7) also satisfy (6). A simple choice of $\{\alpha(k)\}_{k \geq 1}$ satisfying (6) is

$$\alpha(k) = \frac{c}{k^\beta}, \forall k \geq 1, \text{ with } c > 0, \ 0 < \beta \leq 1.$$

Remark 4 (Convergence with any communication rate): The proposed algorithm does not require any condition on the communication rate to guarantee the asymptotic convergence, i.e., asymptotic convergence can be achieved with any *non-zero* communication rate. In Section IV, we will show numerically that, even with 1-bit communication, convergence to the average of the initial measurements can still be achieved asymptotically.

Remark 5 (Discussion on the informative measurements condition (i)): In [12], [13], the algorithms need the initial measurements to be bounded by a known constant. However, our proposed algorithm requires only the following much looser criterion: the *average* of the initial measurements must be bounded, whereas the initial measurements *need not be*. This criterion can be satisfied more broadly than the uniform boundedness of the initial measurements, especially when the number of agents is large. To explain this point, consider the following simple signal model: $s_i = \theta + w_i, \forall i = 1, \cdots, N$, where w_i is white Gaussian noise. It is easy to see that, as N gets larger, $\max\{s_i\}$ increases, $\min\{s_i\}$ decreases, and the gap $\max\{s_i\} - \min\{s_i\}$ increases, while the average of s_i, $\bar{s}$, tends to approach θ, hence to be more informative.

B. Convergence rate

After showing the convergence to the average of the initial measurements, the next interesting question is: how fast does the algorithm converge to the average? Specifically, the convergence rate of $\{\mathbf{y}(k)\}_{k \geq 1}$ is given in the following proposition.

Proposition 1: $\mathbb{E}\left[\|\mathbf{y}(k) - \bar{s}\mathbf{1}\|^2 | \mathbf{s}\right]$ is upper bounded by

$$\mathbb{E}\left[\|\mathbf{y}(k) - \bar{s}\mathbf{1}\|^2 | \mathbf{s}\right] \leq \left(\prod_{t=1}^{k-1}\left[1 - \alpha(t)c_3\right]\right)\|\mathbf{s} - \bar{s}\mathbf{1}\|^2 + \sum_{t_1=1}^{k-1}\left(\prod_{t_2=2}^{k-t_1}\left[1 - \alpha(t_2)c_3\right]\right)c_6\alpha(t_1)^2, \quad (8)$$

where c_3 is a function of $\lambda_2(\mathbf{L})$, c_6 is a function of $\lambda_N(\mathbf{L}^T\mathbf{L}), \Delta, q_{\max}, q_{\min}$ (cf. [15]).

Proof: See [15]. ∎

The upper bound (8) is obtained by considering the following two cases: 1) saturation does not occur, i.e., $\tilde{\mathbf{y}}(k) = \mathbf{y}(k)$ and 2) saturation does occur, i.e., $\tilde{\mathbf{y}}(k) \neq \mathbf{y}(k)$. Since it is hard to obtain the probability of saturation, the worst case among these two is adopted for the derivation of (8). In most scenarios, saturation (case 2) yields larger mean square error (MSE), but, since the distance between $\tilde{\mathbf{y}}(k)$ and $\mathbf{y}(k)$ decreases as k gets larger, saturation tends to occur less likely. Therefore, the upper bound (8) is conservative in most scenarios.

Remark 6 (Parameters affecting convergence rate): From (8), we can see that the asymptotic convergence rate of $\{\mathbf{y}(k)\}$ is affected by $\lambda_2(\mathbf{L})$ and $\lambda_N(\mathbf{L}^T\mathbf{L})$, which depends on the graph topology. In addition, the convergence rate also depends on the sequence $\{\alpha(k)\}$ and the weight $\{a_{i,j}\}$. Finally, the effect of the number of bits used for quantization (i.e., B) on the convergence rate is captured by (8). As B gets larger, the quantization step size Δ becomes smaller, yielding a smaller upper bound on the MSE, as expected.

C. Generalization: random weight-balanced directed graphs

In our system model, we assume that the graph is static. However, the proposed algorithm can also be applied to random weight-balanced directed graph. In the following, we will discuss this scenario.

Let $\mathcal{G}(k) = \{\mathcal{V}, \mathcal{E}(k)\}, \forall k \geq 1$, with $\mathcal{E}(k)$ i.i.d. over time (cf. [10]). In addition, we assume that $\mathcal{E}(k)$ is independent of $\mathbf{y}(k)$, for all $k \geq 1$. Let $\mathbf{A}(k) = [a_{i,j}(k)], \mathbf{L}(k) = [l_{i,j}(k)]$ be the weight and Laplacian matrix associated with $\mathcal{G}(k)$. This case models the random link failures and random connections. For this model, we state the convergence results below.

Corollary 1 (a.s. and m.s.s. Convergence to The Average for i.i.d. Random Weight-Balanced Directed Graph): For the random graph defined in Section III-C, suppose that the assumptions of Theorem 1 hold, and $\lambda_2(\bar{\mathbf{L}}) > 0$, where $\lambda_2(\bar{\mathbf{L}})$ is the second smallest eigenvalue of $\bar{\mathbf{L}} = \mathbb{E}[\mathbf{L}(k)]$. Then,

$$\lim_{k\to\infty} \mathbb{E}\left[\|\mathbf{y}(k) - \bar{s}\mathbf{1}\|^2\right] = 0 \text{ (m.s.s. convergence)}, \quad (9)$$

$$\mathbb{P}\left\{\lim_{k\to\infty} \mathbf{y}(k) = \bar{s}\mathbf{1}\right\} = 1 \text{ (a.s. convergence)}. \quad (10)$$

Proof: See [15]. ■

It is easy to see that the results of convergence rate in Section III-B can also be generalized to the i.i.d. random weight-balanced directed graph, as (8).

Remark 7: Corollary 1 shows that, under the i.i.d. random weight-balanced directed graph model, although the graph may not be connected at some time instants, convergence to the initial average can still be achieved as long as the graph is strongly connected *on average*.

D. Generalization: time-varying quantization

In our system, we assume that the number of bits used for quantization is fixed, and we have shown that consensus to the average of the initial measurements can be achieved with 1-bit quantization (cf. Remark 4). Because 1-bit quantization can be regarded as the worst case scenario, intuitively, one may think that the convergence results also hold for time-varying quantization. The above conjecture is correct, as shown in the following Corollary.

Corollary 2 (a.s. and m.s.s. Convergence to The Average for Time-Varying Quantization): Let $B_i(k)$ be the number of bits adopted for quantization at agent i and time k. Suppose the assumption of Theorem 1 is satisfied, and $B_i(k) > 0, \forall i, k$. Then,

$$\lim_{k\to\infty} \mathbb{E}\left[\|\mathbf{y}(k) - \bar{s}\mathbf{1}\|^2\right] = 0 \text{ (m.s.s. convergence)}, \quad (11)$$

$$\mathbb{P}\left\{\lim_{k\to\infty} \mathbf{y}(k) = \bar{s}\mathbf{1}\right\} = 1 \text{ (a.s. convergence)}. \quad (12)$$

Proof: See [15]. ■

IV. Numerical Results

In this section, numerical results are presented to validate the theoretical findings. We adopt the 1-bit quantizer with $(q_{\min}, q_{\max}) = (0, 1)$, i.e., each agent can only transmit 0 or 1 at each iteration. The 0-1 weight is adopted for the weight matrix $\mathbf{A}$, i.e., $a_{j,i} = 1$ if $j \in \mathcal{N}_i^+$ and $a_{j,i} = 0$ if $j \notin \mathcal{N}_i^+$. The graph adopted is undirected, which is a special case of weight-balanced directed graph. The initial measurements $\{s_i\}$ are assumed to be i.i.d. random variables uniformly distributed in $(q_{\min}, q_{\max})$. We adopt the MSE with respect to the average as performance metric, defined as

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i(k) - \bar{s})^2, \quad (13)$$

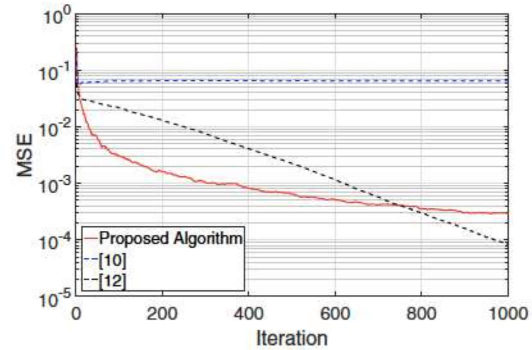

Fig. 2. MSE performance of distributed average consensus algorithms when 1-bit quantizer is adopted in the network of 6 agents, ring topology, with $(q_{\min}, q_{\max}) = (0, 1), s_i \sim \text{Uniform}(q_{\min}, q_{\max}), \forall i$, and $\alpha(k) = 1/k$.

with simulation results averaged over 100 iterations. We also evaluate the performance of [10], which implements the update rule (5). We adopt $\alpha(k) = 1/k, \forall k \geq 1$ for our algorithm and (5). We also evaluate the performance of [12], which is the state-of-the-art distributed average consensus algorithm with finite level quantization which does *need* the spectral properties of the topology, with the following parameters: $h = 0.99h_K^*(\epsilon_0), \epsilon_0 = 0.95, C_x = C_\delta = \max\{|q_{\max}|, |q_{\min}|\}, g_0 = 1.001g_{0,\text{lb}}$, where $g_{0,\text{lb}}$ is the lower bound of g_0 given by [12, Theorem 3.1].

Fig. 2 shows the performance of the distributed consensus algorithms under ring topology. Clearly, consensus to the average can be achieved with the proposed algorithm and [12], consistently with the analytical results. In contrast, there is an error between the final value of the algorithm in [10] and the average of the initial measurements. From Remark 2, when using the update rule (5), the average of $\mathbf{y}(k)$ can be regarded as a random variable with mean $\bar{s}$ and non-zero variance. In this way, the error floors of the algorithm in [10] in Fig. 2 is the variance of $\bar{y}_f$, where $\bar{y}_f$ is the limit of the average of $\mathbf{y}(k)$ as $k \to \infty$.

Comparing the algorithm in [12] and the proposed algorithm, it is shown in Fig. 2 that, initially, the proposed algorithm converges faster. However, since the algorithm in [12] has exponential convergence rate, it will eventually converge faster. The performance superiority of the algorithm in [12] comes from the facts that 1) it exploits information on $\max_i\{s_i\}, \lambda_2, \lambda_N$ to design the dynamic encoding/decoding parameters, 2) it uses additional memory for each agent to track the encoder and decoder states, and 3) the quantizer it adopts can represent more values than the proposed algorithm (3 instead of 2 in this specific case).

Fig. 3 shows the performance of the distributed consensus algorithms under random undirected graphs. At time k, $\mathcal{E}(k)$ is generated according to i.i.d. Bernoulli random variables with probability p, which can be regarded as the level of connectivity of the graph. Since the algorithm in [12] is

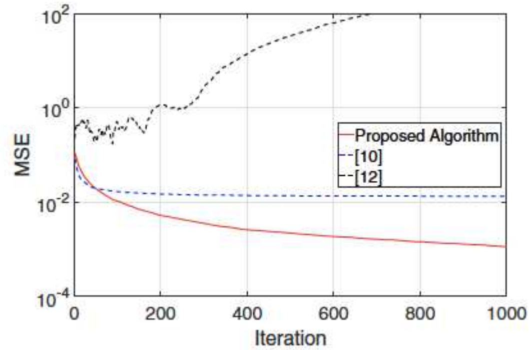

Fig. 3. MSE performance of distributed average consensus algorithms under random communication graph when 1-bit quantizer is adopted in the network of 6 agents, with $p = 0.1, (q_{\min}, q_{\max}) = (0, 1), s_i \sim \text{Uniform}(q_{\min}, q_{\max}), \forall i$, and $\alpha(k) = 1/k$.

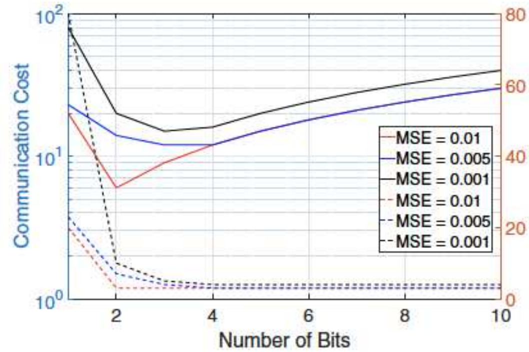

Fig. 4. Communication cost and delay for different target MSE under random communication graph when different number of bits is adopted for quantizer at different agents in the network of 6 agents, with $(q_{\min}, q_{\max}) = (0, 1), s_i \sim \text{Uniform}(q_{\min}, q_{\max}), \forall i$, and $\alpha(k) = 1/k$, where solid and dashed curves represent the communication costs and delay, respectively.

designed for fixed graph, we modify it such that the parameters are determined by the average graph, i.e., $\bar{\mathbf{A}}, \bar{\mathbf{L}}$. In this case, it is shown in Fig. 3 that the (modified) algorithm in [12] fails to converge to the desired average, while the proposed algorithm can still achieve average consensus.

Fig. 4 shows the performance when different number of bits is adopted for the quantizer, under random communication graph. Here, we adopt the same random graph model as in Fig. 3 with $p = 0.5$. The communication cost is defined as the product of the number of bits adopted for quantization and the number of iterations, which is the total number of bits transmitted by an agent. From the three solid curves, we can see that the communication cost increases as the target MSE decreases, as expected. In addition, this figure shows that for a given network, there exists an optimum number of bits for MSE minimization in terms of communication cost, which depends on the target MSE. For instance, if the target MSE is set as 0.01 (red solid curve), then two bits should be adopted for quantization at each agent, in order to minimize the communication cost. On the other hand, we can see that the delay also increases as the target MSE decreases. Unlike the communication cost, delay is a decreasing function of the number of bits adopted for quantization, i.e., generally, more bits adopted for quantization yield better performance. However, the marginal improvement becomes smaller as the number of bits increases. For instance, if the target MSE is set as 0.001 (black dashed curve), then using more than four bits for quantization yields negligible improvement.

V. Conclusion

In this paper, we have studied the distributed average consensus problem in sensor networks with limited communication data rate, and proposed a simple but powerful algorithm to solve this problem. Unlike existing schemes for limited data rate, the proposed algorithm has three properties which makes it more favorable for practical implementation: 1) each agent is not required to have information oen th spectral properties of the communication topology; 2) each agent does not need to know the uniform bound on the initial measurements; 3) exact consensus to the average can be achieved asymptotically, both almost surely and in mean square sense. We have proved the convergence results, provided the analysis of convergence rate, and extended the analytical results to 1) i.i.d. random weight-balanced directed networks and 2) time-varying communication rate. Numerical result shows that the proposed algorithm outperforms state-of-the-art topology-agnostic consensus schemes with limited data rate.

References

[1] G. Cybenko, "Dynamic load balancing for distributed memory multiprocessors," *J. Parallel and Distrib. Comput.*, vol. 7, no. 2, pp. 279–301, 1989.

[2] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465–1476, Sep. 2004.

[3] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc wsns with noisy links—part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, Jan. 2008.

[4] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Mass. Inst. Technol. (MIT), Cambridge, 1984.

[5] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Contr. Lett.*, vol. 53, no. 1, pp. 65–78, 2004.

[6] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[7] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, May 2007.

[8] T. C. Aysal, M. J. Coates, and M. G. Rabbat, "Distributed average consensus with dithered quantization," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 4905–4918, Oct. 2008.

[9] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Trans. Autom. Control*, vol. 54, no. 11, pp. 2506–2517, Oct. 2009.

[10] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1383–1400, Mar. 2010.

[11] S. Zhu, Y. C. Soh, and L. Xie, "Distributed parameter estimation with quantized communication via running average," *IEEE Trans. Signal Process.*, vol. 63, no. 17, pp. 4634–4646, Sep. 2015.

[12] T. Li, M. Fu, L. Xie, and J.-F. Zhang, "Distributed consensus with limited communication data rate," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 279–292, Feb. 2011.

[13] D. Thanou, E. Kokiopoulou, Y. Pu, and P. Frossard, "Distributed average consensus with quantization refinement," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 194–205, Jan. 2013.

[14] R. Rajagopal and M. J. Wainwright, "Network-based consensus averaging with general noisy channels," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 373–385, Jan 2011.

[15] C.-S. Lee, N. Michulusi, and G. Scutari, "Topology-agnostic average consensus in sensor networks with limited data rate," Purdue University, West Lafayette, IN, Tech. Rep., May 2017. [Online]. Available: https://engineering.purdue.edu/~michelus/Asilomar17.pdf

[16] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation:Numerical Methods*. Belmont, MA, USA: Athena Scientific, 1989.

[17] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J. S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.