Rate-Constrained Distributed Estimation in Wireless Sensor Networks

Junlin Li, *Student Member, IEEE*, and Ghassan AlRegib, *Member, IEEE*

***Abstract*—In this paper, we consider the distributed parameter estimation in wireless sensor networks where a total bit rate constraint is imposed. We study the optimal tradeoff between the number of active sensors and the quantization bit rate for each active sensor to minimize the estimation mean-square error (MSE). To facilitate the solution, we first introduce a concept of equivalent 1-bit MSE function. Next, we present an optimal distributed estimation algorithm for homogeneous sensor networks based on minimizing the equivalent 1-bit MSE function. Then, we present a quasi-optimal distributed estimation algorithm for heterogeneous sensor networks, which is also based on the equivalent 1-bit MSE function, and the upper bound of the estimation MSE of the proposed algorithm is addressed. Furthermore, a theoretical nonachievable lower bound of the estimation MSE under the total bit rate constraint is stated and it is shown that our proposed algorithm is quasi-optimal within a factor 2.2872 of the theoretical lower bound. Simulation results also show that significant reduction in estimation MSE is achieved by our proposed algorithm when compared with other uniform methods.**

***Index Terms*—Best linear unbiased estimator (BLUE), collaborative signal processing, distributed estimation, distributed signal processing, wireless sensor networks.**

I. Introduction

WIRELESS sensor networks (WSNs) is an emerging technology that has many current and future envisioned applications ranging from environment monitoring, battlefield surveillance, health care, home automation, and so on [1]. A wireless sensor network is composed of a large number of geographically distributed sensor nodes. Though each sensor is characterized by low power constraint and limited computation and communication capabilities, potentially powerful networks can be constructed to accomplish various high-level tasks with sensor collaboration [2], such as distributed estimation, distributed detection, and target localization and tracking. In this paper, we study distributed estimation of an unknown deterministic parameter by a set of distributed sensor nodes and a fusion center. Sensors collect real-valued data, perform a local data compression, and send the resulting discrete messages to the fusion center that combines the received messages to produce a final estimate of the observed parameter.

Manuscript received September 20, 2005; revised May 29, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jonathan H. Manton.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: lijunlin@ece.gatech.edu; gregib@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2006.890823

Given the power constraint in sensors, one of the major objectives of the current research is to design energy-efficient devices, protocols, and algorithms [1]. Compared with sensing and computation, communication is the most energy-consuming operation in wireless sensor networks, therefore reducing the communication requirement from the sensors to the fusion center is essential in saving energy. On the other hand, the communication capacity of wireless sensor networks is limited because the wireless channel is shared across the whole network. Therefore, a total communication rate constraint is necessary to avoid communication collision that wastes energy. In this paper, we address the rate-constrained distributed estimation in wireless sensor networks by imposing a constraint on the total bit rate, that is, only B bits are used and optimally allocated among the local observations to minimize the estimation distortion at the fusion center.

The problem of distributed parameter estimation has been well studied, where most of the early works [3]–[7] assume that the joint distribution of sensors' observations is known and that the real-valued messages can be sent from the sensors to the fusion center without distortion, which are unrealistic for practical sensor networks because of the high communication cost. Recently, several rate-constrained distributed estimation algorithms have been investigated [8]–[16]. The work of [8]–[10] addressed various design and implementation issues to digitize the transmitted signal into one or several binary bits using the joint distribution of sensors' data. In [11] and [12], the focus is on finding a class of maximum-likelihood estimators (MLEs) to attain a variance that is close to the clairvoyant estimator when the observations are quantized to one bit. Without the knowledge of noise distribution, the work of [13] and [14] proposed to use a training sequence to aid the design of local data quantization strategies, and the work of [15] and [16] proposed several universal (pdf-unaware) decentralized estimation systems for distributed parameter estimation in the presence of unknown, identically distributed, additive sensor noises. In addition, minimal energy decentralized estimation in sensor networks is also studied in [17].

To the best of our knowledge, most of the past works on rate-constrained distributed estimation are usually posed for a given number of sensors (one observation per sensor) [8]–[16], but for rate-constrained problems, a more meaningful scenario is to constraint the total number of available bits B, that is, given the network, we are allowed to transmit up to B bits that have to be allocated among the observations. Using this concept, there exists an interesting tradeoff between the number of active sensors and the quantization precision of each active sensor. In this paper, we address the optimal tradeoff and design the optimal distributed estimation mechanism to minimize the estimation

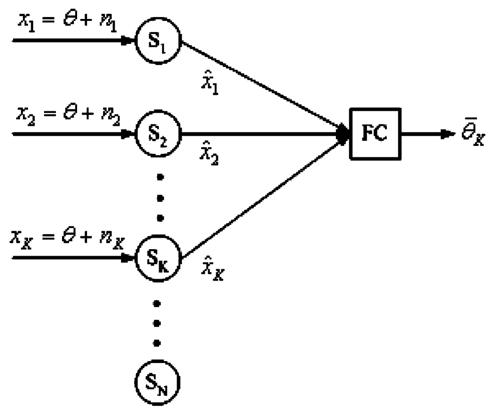

Fig. 1. Distributed estimation system under the total bit rate constraint.

distortion for a given total bit rate B. The goals in this optimization problem are 1) selecting a subset of sensors to observe the phenomenon and 2) selecting the quantizer for each active sensor to quantize the real-valued observation.

The rest of the paper is organized as follows. Section II states the distributed estimation problem under the total bit rate constraint. Section III introduces a concept of equivalent 1-bit MSE function. Then, in Sections IV and V, we propose an optimal distributed estimation algorithm for homogeneous sensor networks and a quasi-optimal distributed estimation algorithm for heterogeneous sensor networks, respectively. Further, the upper bound of the estimation mean-square error (MSE) of our proposed algorithm is addressed, and a theoretical lower bound of the estimation MSE under the total bit rate constraint is proved. Section VI shows some simulation results to demonstrate the performance of the proposed algorithms. Finally, the conclusions are given in Section VII.

II. Problem Statement and Background

Consider a dense sensor network that includes N distributed sensors and each sensor can observe, quantize, and transmit its observation to the fusion center, which will estimate the parameter θ based on the received messages. Because of the total bit rate constraint, there is a tradeoff between the number of active sensors and the quantization bit rate at each active sensor, that is to say, only a subset of the sensors will be active at each task period. Without loss of generality, we assume that the first K sensors are active. This can be accomplished as follows (Fig. 1).

First, each sensor makes an observation on the unknown parameter θ. The observations are corrupted by additive noises and are described by

$$x_k = \theta + n_k, \quad k = 1, \ldots, K. \tag{1}$$

We assume that the noises $n_k (k = 1, \ldots, K)$ are zero mean, spatially uncorrelated with variance σ_k^2, but otherwise unknown. Second, each of the K active sensors performs a local quantization $\hat{x}_k = Q_k(x_k)$, where $Q_k(x_k)$ is a quantization function, and the quantization message $\hat{x}_k$ is then transmitted to the fusion center where all the quantization messages are combined to produce a final estimation of θ using a fusion function f: $\bar{\theta}_K = f(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_K)$. The quality of an estimation for θ is measured by the MSE criterion. So the goal is to optimize the following problem:

$$\begin{aligned} \min \quad & E(\bar{\theta}_K - \theta)^2, \\ \text{s.t.} \quad & \sum_{k=1}^{K} b_k \le B, \quad b_k \in \mathbb{Z}^+, \quad k = 1, \ldots, K \end{aligned} \tag{2}$$

where B is the total bit rate constraint, K is the number of the active sensors, and b_k is the quantization bit rate of the sensor k. That is to say, adaptively select the active subset of the sensors and the b_k-bit quantizer for each active sensor to minimize the estimation MSE at the fusion center.

If the fusion center has the knowledge of the sensor noise variance $\sigma_k^2 (k = 1, \ldots, K)$ and the sensors can perfectly send their observations $x_k (k = 1, \ldots, K)$ to the fusion center, the best linear unbiased estimator (BLUE) [18] for θ is known to be as

$$\bar{\theta}_K = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2} \right)^{-1} \sum_{k=1}^{K} \frac{x_k}{\sigma_k^2} \tag{3}$$

and the estimation MSE of the BLUE estimator is

$$E(\bar{\theta}_K - \theta)^2 = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2} \right)^{-1}. \tag{4}$$

But the BLUE scheme is impractical for wireless sensor networks because of the high communication cost. Instead of sending the real-valued observations to the fusion center directly, quantization at the local sensors is essential to reduce the communication cost (bandwidth and energy). In this paper, we adopt a probabilistic quantization scheme [17] as well as a quasi-BLUE estimation scheme, based on which the optimal tradeoff between the number of active sensors and the quantization bit rate of each active sensor is addressed.

A. Probabilistic Quantization and Quasi-Blue Estimator

Suppose the observed signal of each sensor is bounded to $[-W, W]$, that is, $x = \theta + n \in [-W, W]$, where W is decided by the sensor's dynamic range, θ is the unknown parameter to be estimated, and n is the zero mean noise with variance σ^2. Regardless of the probability distribution of x, the probabilistic quantization with b bits is summarized as follows: Uniformly divide $[-W, W]$ into intervals of length $\Delta = (2W)/(2^b - 1)$, and round x to the neighboring endpoints of these small intervals in a probabilistic manner. More specifically, suppose $-W + i\Delta \le x \le -W + (i+1)\Delta$, where $0 \le i \le 2^b - 2$, then x is quantized to $\hat{x}(b)$ according to

$$\begin{aligned} P\{\hat{x}(b) = -W + i\Delta\} &= 1 - r \\ P\{\hat{x}(b) = -W + (i+1)\Delta\} &= r \end{aligned} \tag{5}$$

where $r = (x + W - i\Delta)/(\Delta)$. As shown in [17], the quantized message $\hat{x}(b)$ is an unbiased estimator of θ with a variance

$$\begin{aligned} E(|\hat{x}(b) - \theta|^2) &\le \sigma^2 + \frac{W^2}{(2^b - 1)^2} \\ &= \sigma^2 + \delta^2, \quad \text{for all } b \ge 1 \end{aligned} \tag{6}$$

where $\delta^2 = (W^2)/(2^b - 1)^2$ denotes the upper bound of the quantization noise variance.

Now, suppose all the observations of the K active sensors x_k $(k = 1, \ldots, K)$ are quantized into b_k-bits discrete messages $\hat{x}_k(b_k)$ with the above probabilistic quantization scheme, and treat all the quantized messages $\hat{x}_k$ as the new observations for the fusion center. Then, the quasi-BLUE estimator based on the quantized messages has the following form:

$$\bar{\theta}_K = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \sum_{k=1}^{K} \frac{\hat{x}_k}{\sigma_k^2 + \delta_k^2}. \tag{7}$$

Notice that $\bar{\theta}_K$ is an unbiased estimator of θ since every $\hat{x}_k$ is unbiased. Moreover, the estimation MSE of the quasi-BLUE estimator [17] is

$$E(\bar{\theta}_K - \theta)^2 \leq \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1}. \tag{8}$$

With the probabilistic quantization scheme and the quasi-BLUE fusion rule, instead of the original problem in (2), we turn to the following modified problem:

$$\begin{aligned} \min \quad & \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \frac{W^2}{(2^{b_k}-1)^2}}\right)^{-1}, \\ \text{s.t.} \quad & \sum_{k=1}^{K} b_k \leq B, \quad b_k \in \mathbb{Z}^+, \quad k = 1, \ldots, K \end{aligned} \tag{9}$$

which minimizes the bound of the estimation MSE. In the following sections, we will address this modified problem for the homogeneous and heterogeneous sensor networks, respectively. To facilitate the solution, we first define an equivalent 1-bit MSE function.

III. Equivalent 1-Bit MSE Function

As shown in Section II-A, the b-bit quantized message from a sensor with observation noise variance σ^2 is an unbiased estimation of the estimated parameter θ with the estimation MSE $D \leq \sigma^2 + (W^2)/((2^b - 1)^2)$. We denote the estimation MSE bound as

$$f(\sigma^2, b) = \sigma^2 + \frac{W^2}{(2^b - 1)^2}. \tag{10}$$

Definition 1 (Equivalent 1-Bit MSE Function): For a sensor with b-bit quantization and observation noise variance σ^2, the equivalent 1-bit MSE function is defined as

$$g(\sigma^2, b) = b \cdot f(\sigma^2, b) = b\left(\sigma^2 + \frac{W^2}{(2^b - 1)^2}\right). \tag{11}$$

With this definition, the estimation MSE of the quasi-BLUE estimator in (8) can be rewritten as (12), shown at the bottom of the page. From the estimation MSE aspect, a b-bit quantization sensor with the estimation MSE $f(\sigma^2, b)$ can be treated as b equivalent 1-bit quantization sensors, each with the same estimation MSE $g(\sigma^2, b)$ defined as above. Further, the rate-constrained distributed estimation system with K sensors under the total bit rate constraint B can be treated as another equivalent distributed estimation system with B equivalent 1-bit quantization sensors, where B is a constant for the given problem, while K is a variable.

IV. Distributed Estimation in Homogeneous Sensor Networks

In this section, we address the distributed estimation under the total bit rate constraint for homogeneous sensor networks, where every sensor has the same observation noise variance, that is, $\sigma_k^2 = \sigma^2 (k = 1, \ldots, N)$. It is natural and easily verified that each active sensor should quantize its observation with the same bit rate $b_k = b$ to minimize the estimation MSE. As a result, the number of active sensors is B/b and the estimation MSE function is simplified to

$$\begin{aligned} E(\bar{\theta}_K - \theta)^2 &\leq \left(\sum_{k=1}^{K} \frac{1}{\sigma^2 + \frac{W^2}{(2^{b_k}-1)^2}}\right)^{-1} \\ &= \frac{b\left(\sigma^2 + \frac{W^2}{(2^b-1)^2}\right)}{B}. \end{aligned} \tag{13}$$

It is noted that the numerator of the optimized target function in (13) is just the equivalent 1-bit MSE function $g(\sigma^2, b)$ defined in Section III. Hence, for homogeneous sensor networks, the optimal distributed estimation under the total bit rate constraint B can be treated in an alternative way, where there are B same equivalent 1-bit quantization sensors, thus minimizing

$$\begin{aligned} E(\bar{\theta}_K - \theta)^2 &\leq \Bigg(\underbrace{\frac{1}{f(\sigma_1^2, b_1)} + \cdots + \frac{1}{f(\sigma_K^2, b_K)}}_{K}\Bigg)^{-1} \\ &= \Bigg(\underbrace{\underbrace{\frac{1}{g(\sigma_1^2, b_1)} + \cdots + \frac{1}{g(\sigma_1^2, b_1)}}_{b_1} + \cdots + \underbrace{\frac{1}{g(\sigma_K^2, b_K)} + \cdots \frac{1}{g(\sigma_K^2, b_K)}}_{b_K}}_{b_1 + \cdots + b_K = B}\Bigg)^{-1}. \end{aligned} \tag{12}$$

the final estimation MSE bound becomes minimizing the equivalent 1-bit MSE function. The method based on the equivalent 1-bit MSE function is stated as follows.

1) For each sensor, the optimal quantization bit rate is identical and obtained by minimizing the corresponding equivalent 1-bit MSE function

$$\begin{aligned} b^{\mathrm{opt}} &= \arg\min_{b\in\mathbb{Z}^+} g(\sigma^2, b) \\ &= \arg\min_{b\in\mathbb{Z}^+} \left[b\left(\sigma^2 + \frac{W^2}{(2^b-1)^2}\right)\right] \end{aligned} \tag{14}$$

where the minimization involves just a simple one-dimensional numerical search.

2) The total number of active sensors (K^{opt}) under the total bit rate constraint B is

$$K^{\mathrm{opt}} = \left\lfloor \frac{B}{b^{\mathrm{opt}}} \right\rfloor. \tag{15}$$

It is obvious that the proposed method based on the equivalent 1-bit MSE function is optimal if $B/(b^{\mathrm{opt}})$ is an integer, i.e., $K^{\mathrm{opt}} = B/(b^{\mathrm{opt}})$. For the case where $B/(b^{\mathrm{opt}})$ is not an integer, there are $b_r = B - K^{\mathrm{opt}} \cdot b^{\mathrm{opt}} < b^{\mathrm{opt}}$ bits remaining after the two steps above, then we simply allocate these remaining b_r bits to one more sensor to quantize its observation. Obviously, it is quasi-optimal even though it is not necessarily optimal.

Remark 1: It is noted that the proposed method above is almost fully distributed. First, the optimal quantization bit rate b^{opt} of each sensor can be obtained locally by minimizing its corresponding equivalent 1-bit MSE function. With the given total bit rate constraint B and the optimal quantization bit rate of each sensor b^{opt}, the number of active sensors at each task period is $K^{\mathrm{opt}} = B/(b^{\mathrm{opt}})$ (we assume $B/(b^{\mathrm{opt}})$ is integer here), then each sensor will be in the active mode with a probability of $p = K^{\mathrm{opt}}/N$, where N is the total number of sensors. Assume each sensor has a unique index $i \in [0, N-1]$, we design a periodic scheduling for each sensor i as follows: sensor i is active when $t \in [kN+i, kN+i+K^{\mathrm{opt}}] (k \in \mathbb{Z})$; otherwise, it is in sleep mode. With the given scheduling scheme, there are K^{opt} active sensors at any task period t and each sensor will be active for K^{opt} task periods in any consecutive N task duration. Therefore, the energy cost at each sensor node is even, and the network lifetime is maximized, which is defined as the time for the first sensor node in the network to deplete.

V. Distributed Estimation in Heterogeneous Sensor Networks

In this section, we address the general distributed estimation under the total bit rate constraint for heterogeneous sensor networks. Assuming the observation noise variance for every sensor is $\sigma_k^2 (k = 1, \ldots, N)$, respectively. Without loss of generality, we assume $\sigma_1^2 \le \cdots \le \sigma_N^2$, so, if K sensors are needed, we just simply choose the first K sensors, which will minimize the estimation MSE. This scenario leads to the general problem stated in (9).

To find the optimal number of active sensors and the corresponding optimal quantization bit rate of each active sensor to minimize the estimation MSE bound at the fusion center, we adopt the Lagrange multiplier method to solve the following equivalent problem:

$$\begin{aligned} \max \quad & \sum_{k=1}^{K} \left(\sigma_k^2 + \frac{W^2}{(2^{b_k}-1)^2}\right)^{-1}, \\ \text{s.t.} \quad & \sum_{k=1}^{K} b_k \le B, \quad b_k \in \mathbb{Z}^+, \quad k = 1, \ldots, K. \end{aligned} \tag{16}$$

Its Lagrangian G is given as

$$\begin{aligned} G(b_k, \lambda) = \sum_{k=1}^{K} & \left(\sigma_k^2 + \frac{W^2}{(2^{b_k}-1)^2}\right)^{-1} \\ & + \lambda \left(\sum_{k=1}^{K} b_k - B\right) \end{aligned} \tag{17}$$

which leads to the following two optimization conditions:

$$\frac{\partial \left(\sigma_k^2 + \frac{W^2}{(2^{b_k}-1)^2}\right)^{-1}}{\partial b_k} + \lambda = 0, \forall k \in [1, K],$$

and

$$\sum_{k=1}^{K} b_k = B. \tag{18}$$

Unfortunately, we can verify that the optimal solution $b_k (k = 1, \ldots, K)$ cannot be found in a closed form from (18). Instead, we propose a quasi-optimal method to solve the problem, which is also based on the equivalent 1-bit MSE function. The procedure is stated as follows.

1) For each sensor, its optimal quantization bit rate $b_k^{\mathrm{opt}} (k = 1, \ldots, N)$ is obtained by minimizing its corresponding equivalent 1-bit MSE function

$$\begin{aligned} b_k^{\mathrm{opt}} &= \arg\min_{b_k\in\mathbb{Z}^+} g(\sigma_k^2, b_k) \\ &= \arg\min_{b_k\in\mathbb{Z}^+} \left[b_k\left(\sigma_k^2 + \frac{W^2}{(2^{b_k}-1)^2}\right)\right] \end{aligned} \tag{19}$$

where the minimization involves just a simple one-dimensional numerical search.

2) The total number of active sensors (K^{opt}) under the total bit rate constraint B is determined as

$$K^{\mathrm{opt}} = \max K, \quad \text{subject to} \sum_{k=1}^{K} b_k^{\mathrm{opt}} \le B. \tag{20}$$

In short, the whole solution is as follows: the K^{opt} sensors with the smallest observation noise variances are chosen to quantize and transmit their observations, and the quantization bit rate of each chosen sensor is $b_k^{\mathrm{opt}} (k = 1, \ldots, K^{\mathrm{opt}})$. Next, we will analyze the estimation MSE bound of the proposed method, which is stated in the following theorem. To simplify the statement, we assume $\sum_{k=1}^{K^{\mathrm{opt}}} b_k^{\mathrm{opt}} = B$ in the subsequent analysis.

Theorem 1: The estimation MSE of the proposed method based on the equivalent 1-bit MSE function under the total bit rate constraint B is

$$\left(\sum_{k=1}^{K^{\text{opt}}} \frac{1}{\sigma_k^2}\right)^{-1} < E(\bar{\theta}_P - \theta)^2 < 2.2872 \left(\sum_{k=1}^{K^{\text{opt}}} \frac{1}{\sigma_k^2}\right)^{-1} \quad (21)$$

where $\bar{\theta}_P$ denotes the estimation of the parameter θ by the proposed method, and K^{opt} is the optimal number of active sensors, obtained in (20).

Proof: The left part of the theorem is obvious since $(\sum_{k=1}^{K^{\text{opt}}} (1)/(\sigma_k^2))^{-1}$ is the lower bound of the estimation MSE of the BLUE estimator using the K^{opt} active sensors. To prove the right part of the theorem, we first introduce a lemma.

Lemma 1: $f(\sigma^2, b)$ is the estimation MSE bound function defined in (10), and $b^{\text{opt}}(\sigma^2)$ is the optimal quantization bit rate defined in (19), then

$$f(\sigma^2, b^{\text{opt}}(\sigma^2)) < 2.2872\sigma^2. \quad (22)$$

Lemma 1 is proved in Appendix I. By this lemma, then

$$\begin{aligned} E(\bar{\theta}_p - \theta)^2 &\leq \left(\sum_{k=1}^{K^{\text{opt}}} \frac{1}{f(\sigma_k^2, b^{\text{opt}}(\sigma_k^2))}\right)^{-1} \\ &< \left(\sum_{k=1}^{K^{\text{opt}}} \frac{1}{2.2872\sigma_k^2}\right)^{-1} \\ &= 2.2872 \left(\sum_{k=1}^{K^{\text{opt}}} \frac{1}{\sigma_k^2}\right)^{-1}. \end{aligned} \quad (23)$$

■

This theorem gives the estimation MSE bound of the proposed method. It is shown that the proposed method is quasi-optimal (up to a factor of 2.2872) when compared with the BLUE estimator using the same subset of active sensors.

As above, the performance bound of the proposed algorithm is analyzed. Nevertheless, the remaining question is what performance can be achieved if the total B bits are allocated to any number of sensors, say M sensors. More specifically, whether a lower bound of the estimation MSE less than $D_0 \equiv (\sum_{k=1}^{K^{\text{opt}}} (1)/(\sigma_k^2))^{-1}$ can be achieved using M sensors under the total bit rate constraint B? It is obvious that a lower bound less than D_0 cannot be achieved if $M < K^{\text{opt}}$ sensors are used, regardless of the quantization bit rate of each active sensor. It is also known that if $M > K^{\text{opt}}$ sensors are used, a lower bound $(\sum_{k=1}^{M}(1)/(\sigma_k^2))^{-1}$ (less than D_0) can be achieved if the quantization bit rate for each sensor is not limited. But under the total bit rate constraint B, whether a lower bound less than D_0 can be achieved if $M > K^{\text{opt}}$ sensors are used is a real question. To answer this question, we further analyze the lower bound of the estimation MSE by any quasi-BLUE estimation system with $M > K^{\text{opt}}$ active sensors under the total bit rate constraint B, which is stated in Theorem 2.

Theorem 2: For any quasi-BLUE estimation system under the total bit rate constraint B, where $M > K^{\text{opt}}$ sensors are used, the quantization bit rate for sensor k is $b_k (k = 1, \ldots, M)$ and $\sum_{k=1}^{M} b_k = B$, the lower bound of the estimation MSE is

$$E(\bar{\theta}_B - \theta)^2 > \left(\sum_{k=1}^{K^{\text{opt}}} \frac{1}{\sigma_k^2}\right)^{-1} \quad (24)$$

where $\bar{\theta}_B$ denotes the estimation of the parameter θ under the total bit rate constraint B, and K^{opt} is the optimal number of active sensors, obtained by our proposed algorithm as shown in (20) such that $\sum_{k=1}^{K^{\text{opt}}} b_k^{\text{opt}} = B$.

Proof: Refer to Appendix II for the complete proof. ■

In conclusion, Theorem 1 shows that the estimation MSE bound of our proposed method is $(\sum_{k=1}^{K^{\text{opt}}} (1)/(\sigma_k^2))^{-1} < E(\bar{\theta}_p - \theta)^2 < 2.2872(\sum_{k=1}^{K^{\text{opt}}} (1)/(\sigma_k^2))^{-1}$, and Theorem 2 shows that $(\sum_{k=1}^{K^{\text{opt}}} (1)/(\sigma_k^2))^{-1}$ is the lower bound of the estimation MSE of any quasi-BLUE estimator under the total bit rate constraint B, regardless of the number of active sensors and the bit allocation among the active sensors. Therefore, the proposed algorithm gives a quasi-optimal tradeoff between the number of active sensors and the quantization bit rate of each sensor, and its estimation MSE is within a factor 2.2872 of the theoretical nonachievable lower bound.

VI. Simulation Results

In this section, we will present some simulation results for the proposed algorithms in Sections IV and V, respectively.

A. Homogeneous Sensor Networks

In this section, we simulate a homogeneous sensor network with $N = 500$ sensors, and the range of the observation signal is $[-1, 1]$, i.e., $W = 1$. Define the signal-to-noise ratio (SNR) as $\text{SNR} = 10\log_{10}(W^2/\sigma^2)$ and generate different SNR by changing the observation noise variance σ^2. Assuming the total bit rate constraint is $B = 500$ bits, Fig. 2 shows the estimation MSE with different quantization bit rates for the active sensors under different SNR. Notice that different quantization bit rate for each sensor implies different number of active sensors to perform the estimation task because of the total bit rate constraint B. For example, in the case of SNR = 20 dB, totally 125 active sensors with 4-bit quantized message per sensor will produce the minimum estimation MSE under the total bit rate constraint $B = 500$, which is better than all the other possible cases, such as 500 sensors with 1-bit quantized message per sensor, 250 sensors with 2-bit quantized message per sensor, 62.5 sensors with 8-bit quantized message per sensor, and so on. From the results shown in Fig. 2, we also can see that for low SNR case, such as 0 dB, 1-bit quantization per sensor will lead to the minimum estimation MSE, on the contrary, for the high-SNR case, a multiple-bit quantization per sensor will significantly decrease the estimation MSE compared to only 1-bit quantization per sensor under the same total bit rate constraint.

B. Heterogeneous Sensor Networks

In this section, we simulate a heterogeneous sensor network with $N = 500$ sensors, and the range of the observation signal

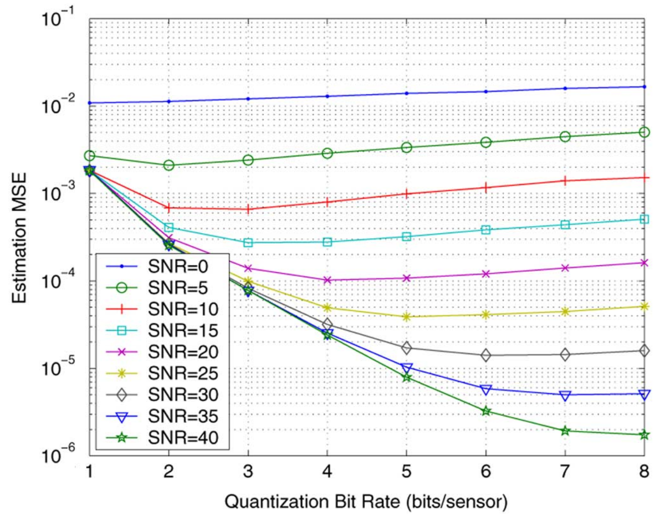

Fig. 2. Estimation MSE versus the quantization bit rate per sensor and different SNR under the total bit rate constraint $B = 500$ bits.

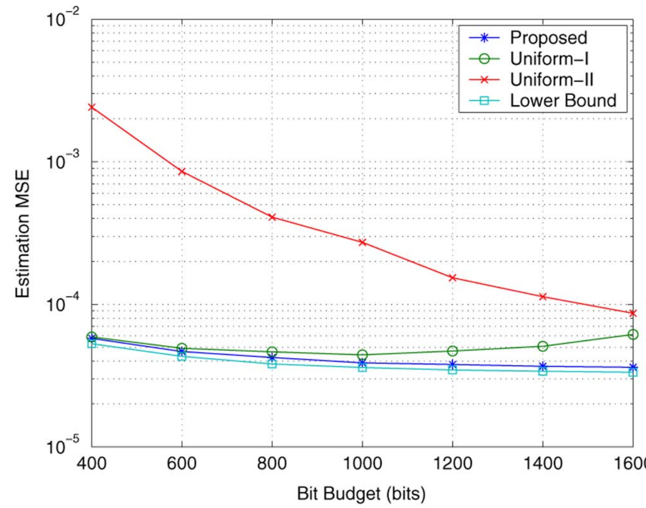

Fig. 3. Estimation MSE by the proposed method, Uniform-I method, Uniform-II method, and theoretical nonachievable lower bound of the estimation MSE.

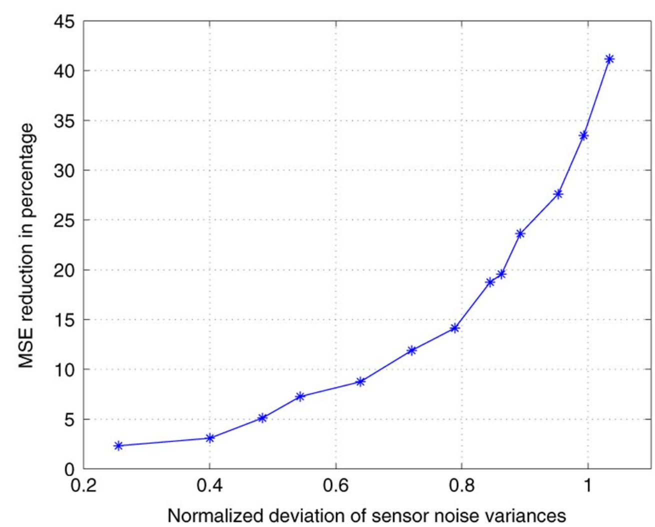

Fig. 4. Estimation MSE reduction in percentage of the proposed method compared with the Uniform-I method under the different normalized deviations of sensor noise variances.

is still $[-1, 1]$. We assume the observation noise variances to be Chi-squared distribution with 1 degree of freedom. In the simulation, for any given total bit rate constraint, our proposed estimation method is implemented to determine the number of active sensors and the quantization bit rate for each active sensor to minimize the estimation MSE.

In order to demonstrate the efficiency of the proposed method, we compare the proposed method with other two kinds of uniform schemes.

i) *Uniform-I*: For the given total bit rate constraint, the same subset of active sensors as that used by our proposed method is used, but the quantization bit rate is uniform among all the active sensors.

ii) *Uniform-II*: All the sensors in the simulated heterogeneous sensor network are used and the quantization bit rate is uniform among all the sensors.

Fig. 3 shows the estimation MSE by our proposed method, the Uniform-I method and the Uniform-II method, and the theoretical lower bound of the estimation MSE presented in Theorem 2 under the total bit rate constraint. From Fig. 3, we can see that the proposed method outperforms the other two uniform schemes. Further, it also can be seen that the estimation MSE of our proposed method is very close to the theoretical nonachievable lower bound (about 1.1 times).

Noted that both our proposed method and the *Uniform-I* method are based on the same subset of active sensors, and the only difference is that the optimal bit allocation is performed in our proposed method, while uniform bit allocation is performed in the *Uniform-I* method. Because of the heterogeneity of the network, a better estimation performance is obtained by our proposed method.

Next, we further show how the heterogeneity of the sensor networks will influence the estimation performance. We define the normalized deviation of sensor noise variances as

$$\alpha = \frac{\sqrt{\mathrm{Var}(\sigma^2)}}{E(\sigma^2)} \tag{25}$$

which will be used as a measure of the heterogeneity of the sensor network. Then, we define the reduction in the estimation MSE achieved by our proposed method in comparison with the Uniform-I method as

$$\beta = \frac{D_u - D_p}{D_u} \tag{26}$$

where D_u denotes the estimation MSE by the Uniform-I method, and D_p denotes the estimation MSE by our proposed method. Fig. 4 plots the estimation MSE reduction of our proposed method compared with the Uniform-I method versus the normalized deviations of sensor noise variances. From Fig. 4, we conclude that when compared with the Uniform-I method, the amount of estimation MSE reduction of our proposed method becomes more significant when the local sensor noise variances become more heterogeneous.

VII. Conclusion

We considered the distributed estimation of a noise-corrupted deterministic parameter under the total bit rate constraint in wireless sensor networks. Because of the total bit rate constraint, a tradeoff between the number of active sensors and the quantization bit rate of each active sensor is addressed to minimize the estimation MSE. To determine the optimal quantization bit rate of each sensor, a concept of the equivalent 1-bit MSE function is introduced, based on which an optimal rate-constrained distributed estimation algorithm for homogeneous sensor networks and a quasi-optimal rate-constrained distributed estimation algorithm for heterogeneous sensor networks are proposed. Moreover, a theoretical nonachievable lower bound of estimation MSE under the total bit rate constraint is proposed and it is shown that our proposed algorithm is quasi-optimal and within a factor 2.2872 of the theoretical lower bound. Simulation results also show that our proposed method can achieve a significant amount of the estimation MSE reduction when compared with the several uniform schemes in which each sensor quantizes its observation with the same number of bits.

To facilitate the problem, we have assumed in this paper that the observation noises among different sensors are uncorrelated and the channels from the local sensors to the fusion center are error free. As the future work, we plan to relax the above assumptions and study the general distributed parameter estimation under the total bit rate constraint. In this general case, it is likely that the quantization scheme and the fusion rule need to be revised by taking into account the sensor correlation and the channel fading, based on which the sensor scheduling and the bit allocation also need to be jointly optimized to minimize the estimation MSE.

Appendix I
Proof of Lemma 1

Based on the definition of the equivalent 1-bit MSE function $g(\sigma^2, b)$ in (11) and the optimal quantization bit rate $b^{\rm opt}(\sigma^2)$ in (14) and (19), we first present the following proposition.

Proposition 1: The functions $g(\sigma^2, b)$ and $b^{\rm opt}(\sigma^2)$ have the following properties:

1) $g(\sigma^2, b)$ increases over σ^2, i.e.,

$$g(\sigma_i^2, b) > g(\sigma_j^2, b), \quad \text{if } \sigma_i^2 > \sigma_j^2. \tag{27}$$

2) $g(\sigma^2, b)$ is convex over $b (b > 0)$, i.e., [see (28), shown at the bottom of the page].

Proof:

1) From the definition of $g(\sigma^2, b)$, it is obvious that it increases over σ^2.
2) Its convexity over b can be easily proved by checking $(\partial^2 g(\sigma^2, b))/(\partial b^2) > 0$. ∎

Based on the definition of equivalent 1-bit MSE function $g(\sigma^2, b)$, we further define the optimal equivalent 1-bit MSE function $g^{\rm opt}(\sigma^2)$ and the corresponding optimal number of equivalent 1-bit sensors $l^{\rm opt}(\sigma^2)$ for each original sensor with observation noise variance σ^2 as

$$\begin{aligned} l^{\rm opt}(\sigma^2) &= \arg\min_{b\in\mathbb{R}^+} g(\sigma^2, b) = \arg\min_{b\in\mathbb{R}^+} \left[b\left(\sigma^2 + \frac{W^2}{(2^b-1)^2}\right)\right] \\ g^{\rm opt}(\sigma^2) &= \min_{b\in\mathbb{R}^+} g(\sigma^2, b) = g(\sigma^2, l^{\rm opt}(\sigma^2)). \end{aligned} \tag{29}$$

Here, $l^{\rm opt}(\sigma^2) \in \mathbb{R}^+$, while $b^{\rm opt}(\sigma^2) \in \mathbb{Z}^+$ defined in (14) and (19). It is obvious that $b^{\rm opt}(\sigma^2) = \lfloor l^{\rm opt}(\sigma^2)\rfloor$ or $\lceil l^{\rm opt}(\sigma^2)\rceil$ since $g(\sigma^2, b)$ is convex over b as stated in Proposition 1, where $\lfloor l^{\rm opt}(\sigma^2)\rfloor$ denotes the maximum integer no more than $l^{\rm opt}(\sigma^2)$, and $\lceil l^{\rm opt}(\sigma^2)\rceil$ denotes the minimum integer no less than $l^{\rm opt}(\sigma^2)$.

To solve $l^{\rm opt}(\sigma^2)$ from (29), we need to solve $(\partial g(\sigma^2, b))/(\partial b) = 0$, which leads to the following equation:

$$\left(2^{l^{\rm opt}(\sigma^2)} - 1\right)^3 - \frac{W^2}{\sigma^2}\left[2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1\right] = 0 \tag{30}$$

so

$$W^2 = \sigma^2 \cdot \frac{\left(2^{l^{\rm opt}(\sigma^2)} - 1\right)^3}{2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1}. \tag{31}$$

By solving (19) with (31), we get the following relationship between $b^{\rm opt}(\sigma^2)$ and $l^{\rm opt}(\sigma^2)$:

$$b^{\rm opt}(\sigma^2) = \begin{cases} 1, & \text{if } 0 < l^{\rm opt}(\sigma^2) < 1.41 \\ 2, & \text{if } 1.41 \le l^{\rm opt}(\sigma^2) < 2.44 \\ 3, & \text{if } 2.44 \le l^{\rm opt}(\sigma^2) < 3.45 \\ \text{others.} & \end{cases} \tag{32}$$

Based on the above results, now we turn to $f(\sigma^2, b^{\rm opt}(\sigma^2))$

$$\begin{aligned} f(\sigma^2, b^{\rm opt}(\sigma^2)) &= \sigma^2 + \frac{W^2}{\left(2^{b^{\rm opt}(\sigma^2)} - 1\right)^2} \\ &= \sigma^2 \left[1 + \frac{\left(2^{l^{\rm opt}(\sigma^2)} - 1\right)^3}{2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1} \cdot \frac{1}{\left(2^{b^{\rm opt}(\sigma^2)} - 1\right)^2}\right] \\ &= \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2)) \cdot \frac{1}{\left(2^{b^{\rm opt}(\sigma^2)} - 1\right)^2}\right] \end{aligned} \tag{33}$$

where

$$y(l^{\rm opt}(\sigma^2)) \equiv \frac{\left(2^{l^{\rm opt}(\sigma^2)} - 1\right)^3}{2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1}$$

and it is easy to verify that $y(l^{\rm opt}(\sigma^2))$ increases over $l^{\rm opt}(\sigma^2) > 0$. Next, we discuss four cases.

1) $0 < l^{\rm opt}(\sigma^2) < 1.41$

$$\begin{cases} g(\sigma^2, b_1) > g(\sigma^2, b_2) \ge g(\sigma^2, b^{\rm opt}(\sigma^2)), & \text{if } b_1 < b_2 \le b^{\rm opt}(\sigma^2), \\ g(\sigma^2, b^{\rm opt}(\sigma^2)) \le g(\sigma^2, b_3) < g(\sigma^2, b_4), & \text{if } b^{\rm opt}(\sigma^2) \le b_3 < b_4. \end{cases} \tag{28}$$

In this case, $b^{\rm opt}(\sigma^2) = 1$ as shown in (32), so

$$f(\sigma^2, b^{\rm opt}(\sigma^2)) = \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2)) \cdot \frac{1}{(2^{b^{\rm opt}(\sigma^2)} - 1)^2}\right] \le \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2))\right] \overset{(*)}{<} 2.2872\sigma^2 \quad (34)$$

where the step $(*)$ holds because $y(l^{\rm opt}(\sigma^2))$ increases over $l^{\rm opt}(\sigma^2)$ and $y(1.41) < 1.2872$.

2) $1.41 \le l^{\rm opt}(\sigma^2) < 2.44$
In this case, $b^{\rm opt}(\sigma^2) = 2$ as shown in (32), so

$$f(\sigma^2, b^{\rm opt}(\sigma^2)) = \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2)) \cdot \frac{1}{(2^{b^{\rm opt}(\sigma^2)} - 1)^2}\right] \le \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2)) \cdot \frac{1}{9}\right] \overset{(*)}{<} 1.6918\sigma^2 \quad (35)$$

where the step $(*)$ holds because $y(l^{\rm opt}(\sigma^2))$ increases over $l^{\rm opt}(\sigma^2)$ and $y(2.44) < 6.2265$.

3) $2.44 \le l^{\rm opt}(\sigma^2) < 3.45$
In this case, $b^{\rm opt}(\sigma^2) = 3$ as shown in (32), so

$$f(\sigma^2, b^{\rm opt}(\sigma^2)) = \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2)) \cdot \frac{1}{(2^{b^{\rm opt}(\sigma^2)} - 1)^2}\right] \le \sigma^2 \left[1 + y(l^{\rm opt}(\sigma^2)) \cdot \frac{1}{49}\right] \overset{(*)}{<} 1.4717\sigma^2 \quad (36)$$

where the step $(*)$ holds because $y(l^{\rm opt}(\sigma^2))$ increases over $l^{\rm opt}(\sigma^2)$ and $y(3.45) < 23.115$.

4) $l^{\rm opt}(\sigma^2) \ge 3.45$
From the definition of $l^{\rm opt}(\sigma^2)$ and $b^{\rm opt}(\sigma^2)$, it is obvious that $b^{\rm opt}(\sigma^2) > l^{\rm opt}(\sigma^2) - 1$, so

$$\begin{aligned} f(\sigma^2, b^{\rm opt}(\sigma^2)) &= \sigma^2 \left[1 + \frac{(2^{l^{\rm opt}(\sigma^2)} - 1)^3}{2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1} \cdot \frac{1}{(2^{b^{\rm opt}(\sigma^2)} - 1)^2}\right] \\ &< \sigma^2 \left[1 + \frac{(2^{l^{\rm opt}(\sigma^2)} - 1)^3}{2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1} \cdot \frac{1}{(2^{l^{\rm opt}(\sigma^2)-1} - 1)^2}\right] \\ &\overset{(*)}{\le} 2.1599\sigma^2 \end{aligned} \quad (37)$$

where the step $(*)$ holds because

$$z(l^{\rm opt}(\sigma^2)) \equiv \frac{(2^{l^{\rm opt}(\sigma^2)} - 1)^3}{2\ln 2 \cdot l^{\rm opt}(\sigma^2) \cdot 2^{l^{\rm opt}(\sigma^2)} - 2^{l^{\rm opt}(\sigma^2)} + 1} \cdot \frac{1}{(2^{l^{\rm opt}(\sigma^2)-1} - 1)^2}$$

decreases over $l^{\rm opt}(\sigma^2)$ and $z(3.45) < 1.1599$.

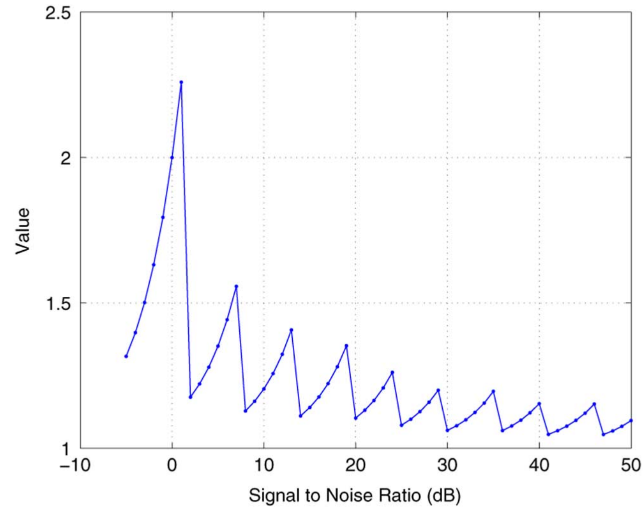

Fig. 5. Value of $(f(\sigma^2, b^{\rm opt}(\sigma^2)))/(\sigma^2)$ versus SNR defined as SNR $= 10\log_{10}(W^2/\sigma^2)$.

From above, for any given σ^2, we get

$$f(\sigma^2, b^{\rm opt}(\sigma^2)) < 2.2872\sigma^2. \quad (38)$$

Remark 2: The numerical relationship between $f(\sigma^2, b^{\rm opt}(\sigma^2))$ and σ^2 is shown in Fig. 5, from which we can see that the maximum value of $(f(\sigma^2, b^{\rm opt}(\sigma^2))/\sigma^2)$ is less than 2.2872 and it is very close to 1 when the SNR is high.

APPENDIX II
PROOF OF THEOREM 2

We begin by introducing two definitions and two corresponding lemmas required by the proof of this theorem.

Definition 2 (Pairwise Bit Rate Exchange): Assuming in a sensor network with M sensors to estimate an unknown parameter, there are two sensors i and j with the observation noise variances $\sigma_i^2 \le \sigma_j^2$ and the quantization bit rates $b_i < b_j$, then exchange the quantization bit rates of the two sensors, i.e., sensor i quantizes its observation using $b'_i = b_j$ bits and sensor j quantizes its observation using $b'_j = b_i$ bits. We call this operation as pairwise bit rate exchange.

Lemma 2: Let D and $D_{\rm ex}$ denote the estimation MSE bound before and after a pairwise bit rate exchange, respectively, then $D_{\rm ex} \le D$.

Proof: Let $D = (1/D')$, and $D_{\rm ex} = (1/D'_{\rm ex})$, then

$$\begin{aligned} D' &= \sum_{\substack{k=1 \\ k\ne i,j}}^{M} \frac{1}{\sigma_k^2 + \delta_k^2} + \frac{1}{\sigma_i^2 + \frac{W^2}{(2^{b_i}-1)^2}} + \frac{1}{\sigma_j^2 + \frac{W^2}{(2^{b_j}-1)^2}}, \\ D'_{\rm ex} &= \sum_{\substack{k=1 \\ k\ne i,j}}^{M} \frac{1}{\sigma_k^2 + \delta_k^2} + \frac{1}{\sigma_i^2 + \frac{W^2}{(2^{b_j}-1)^2}} + \frac{1}{\sigma_j^2 + \frac{W^2}{(2^{b_i}-1)^2}}. \end{aligned} \quad (39)$$

Let $u = \sigma_i^2, v = \sigma_j^2, t = (W^2)/((2^{b_i} - 1)^2)$ and $s = (W^2)/((2^{b_j} - 1)^2)$; then, it can be easily verified that $D' - D'_{\rm ex} \le 0$, and thus, $D_{\rm ex} \le D$, by the following algebra fact: For fixed positive numbers s, t, u, v with $u \leq v$, then

$$\frac{1}{u+s} + \frac{1}{v+t} \geq \frac{1}{u+t} + \frac{1}{v+s}, \quad \text{if } t > s.$$

■

Definition 3 (Equivalent 1-Bit Quantization Sensor Replacement): In a sensor network with M sensors to estimate an unknown parameter, if there are two sensors $i \in [1, \ldots, K^{\text{opt}}]$ and $j \in [K^{\text{opt}} + 1, \ldots, N]$ with the observation noise variances $\sigma_i^2 \leq \sigma_j^2$ and the quantization bit rates $b_i \geq b_j$, and $b_i < b_i^{\text{opt}}$ (that is, $b_i + 1 \leq b_i^{\text{opt}}$), then we replace an equivalent 1-bit quantization sensor corresponding to sensor j by increasing the quantization bit rate of sensor i by 1, that is, sensor i quantizes its observation using $b_i' = b_i + 1$ bits. We call this operation as equivalent 1-bit quantization sensor replacement.

Lemma 3: Let D and D_{re} denote the estimation MSE bound before and after an equivalent 1-bit quantization sensor replacement, respectively, then $D_{\text{re}} < D$.

Proof: Let $D = (1/D')$, and $D_{\text{re}} = (1/D_{\text{re}}')$, then

$$D' = \sum_{\substack{k=1 \\ k \neq i,j}}^{M} \frac{b_k}{g(\sigma_k^2, b_k)} + \frac{b_i}{g(\sigma_i^2, b_i)} + \frac{b_j}{g(\sigma_j^2, b_j)},$$

$$D_{\text{re}}' = \sum_{\substack{k=1 \\ k \neq i,j}}^{M} \frac{b_k}{g(\sigma_k^2, b_k)} + \frac{b_i + 1}{g(\sigma_i^2, b_i + 1)} + \frac{b_j - 1}{g(\sigma_j^2, b_j)} \tag{40}$$

so

$$\begin{aligned} D' - D_{\text{re}}' &= \frac{b_i}{g(\sigma_i^2, b_i)} - \frac{b_i + 1}{g(\sigma_i^2, b_i + 1)} + \frac{1}{g(\sigma_j^2, b_j)} \\ &= \left(\frac{b_i}{g(\sigma_i^2, b_i)} - \frac{b_i}{g(\sigma_i^2, b_i + 1)} \right) \\ &\quad + \left(\frac{1}{g(\sigma_j^2, b_j)} - \frac{1}{g(\sigma_i^2, b_i + 1)} \right) \\ &\overset{(*)}{<} 0 + 0 = 0 \end{aligned} \tag{41}$$

where the step $(*)$ holds for the following reasons:

1) $b_i < b_i + 1 \leq b_i^{\text{opt}}$, so from Proposition 1, we get

$$g(\sigma_i^2, b_i) > g(\sigma_i^2, b_i + 1); \tag{42}$$

2) $b_j \leq b_i < b_i + 1 \leq b_i^{\text{opt}}$, and $\sigma_i^2 \leq \sigma_j^2$, so from Proposition 1, we get

$$g(\sigma_j^2, b_j) \geq g(\sigma_i^2, b_j) \geq g(\sigma_i^2, b_i) > g(\sigma_i^2, b_i + 1). \tag{43}$$

So, $D' < D_{\text{re}}'$, and $D_{\text{re}} < D$. ■

Now, we begin to prove Theorem 2. Assuming M sensors $i_1, \ldots, i_{K^{\text{opt}}}, \ldots, i_M (i_1 < \cdots < i_{K^{\text{opt}}} < \cdots < i_M)$ are used, and the corresponding observation noise variance are $\sigma_{i_1}^2 < \cdots < \sigma_{i_{K^{\text{opt}}}}^2 < \cdots < \sigma_{i_M}^2$, respectively, it is obvious that $i_k \geq k$ and $\sigma_{i_k}^2 \geq \sigma_k^2$. The quantization bit rates are $b_1, \ldots, b_{K^{\text{opt}}}, \ldots, b_M$, respectively, and $\sum_{k=1}^{M} b_k = B$. Let D_1 denote the estimation MSE bound under this condition (denoted as C_1).

Step 1: Considering replace the active sensor $i_k (k = 1, \ldots, M)$ in the condition C_1 by the sensor $k (k = 1, \ldots, M)$, while the quantization bit rate doesn't change. That is to say, the first M sensors are active to observe and quantize their observations using $b_1, \ldots, b_M$ bits, respectively. Let D_2 denote the estimation MSE bound under this condition (denoted as C_2). Obviously, $D_2 \leq D_1$, because

$$D_2 = \left(\sum_{k=1}^{M} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1} \leq \left(\sum_{k=1}^{M} \frac{1}{\sigma_{i_k}^2 + \delta_k^2} \right)^{-1} = D_1. \tag{44}$$

Step 2: Construct another sequence $\{b_k'\} (k = 1, \ldots, M)$ by exchanging the order of the sequence $\{b_k\} (k = 1, \ldots, M)$ in the condition C_2 to make that $b_i' \geq b_j'$ if $0 < i < j \leq M$, and let the sensor $k (k = 1, \ldots, M)$ quantizes its observation with b_k' bits instead of b_k bits. Let D_3 denote the estimation MSE bound under this condition (denoted as C_3). It is obvious that the condition C_3 can be implemented from condition C_2 by a serial of pairwise bit rate exchange operations defined in Definition 2. Since each pairwise bit rate exchange operation will not increase the estimation MSE bound as shown in Lemma 2, then $D_3 \leq D_2$. After the two steps above, we constructed a new scenario where the first M sensors $k (k = 1, \ldots, M)$ with the smallest observation noise variances $\sigma_1^2 \leq \cdots \leq \sigma_{K^{\text{opt}}}^2 \cdots \leq \sigma_M^2$ are used, and the quantization bit rates are $b_1' \geq \cdots \geq b_{K^{\text{opt}}}' \cdots \geq b_M'$. To simplify the notation, in the following we denote the quantization bit rates as $b_k (k = 1, \ldots, M)$ and $b_1 \geq \cdots \geq b_{K^{\text{opt}}} \cdots \geq b_M$.

Step 3: Expressing the estimation MSE bound D_3 with the concept of the equivalent 1-bit MSE function $g(\sigma^2, b)$ as

$$\begin{aligned} D_3 &= \left(\sum_{k=1}^{M} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1} \\ &= \left(\sum_{k=1}^{K^{\text{opt}}} \frac{b_k}{g(\sigma_k^2, b_k)} + \sum_{m=K^{\text{opt}}+1}^{M} \frac{b_m}{g(\sigma_m^2, b_m)} \right)^{-1}. \end{aligned} \tag{45}$$

From the total bit rate constraint, we get

$$\begin{aligned} &\sum_{k=1}^{K^{\text{opt}}} b_k^{\text{opt}} = B, \text{ and} \\ &\sum_{k=1}^{M} b_k = \sum_{k=1}^{K^{\text{opt}}} b_k + \sum_{m=K^{\text{opt}}+1}^{M} b_m = B \end{aligned} \tag{46}$$

so there must exist some $b_k < b_k^{\text{opt}} (k = 1, \ldots, K^{\text{opt}})$ and

$$\sum_{k=1}^{K^{\text{opt}}} I_0 \left(b_k^{\text{opt}} - b_k \right) \geq \sum_{m=K^{\text{opt}}+1}^{M} b_m \tag{47}$$

where $I_0 : R \rightarrow R$ is an indicator function defined as follows:

$$I_0(u) = \begin{cases} 0, & u \leq 0 \\ u, & u > 0 \end{cases} \quad (48).$$

From (46) and (47), we notice that there are $\sum_{m=K^{\rm opt}+1}^{M} b_m$ equivalent 1-bit quantization sensor corresponding to the sensors $m(m = K^{\rm opt}+1,\ldots,M)$, and they all can be replaced by a serial of the equivalent 1-bit quantization sensor replacement operations defined in Definition 3. After finishing the replacement operations, we get a new condition where only sensors $k(k=1,\ldots,K^{\rm opt})$ are used, and the quantization bit rates are changed to $\bar{b}_k$ ($\bar{b}_k$ is not necessarily equal to $b_k^{\rm opt}$), and the total bit rate constraint is still satisfied, i.e., $\sum_{k=1}^{K^{\rm opt}} \bar{b}_k = B$. Let D_4 denote the estimation MSE bound of this condition (denoted as C_4), then $D_4 < D_3$ since every equivalent 1-bit quantization sensor replacement operation will not increase the estimation MSE bound according to Lemma 3. On the other hand, in the condition C_4, only sensors $k(k = 1,\ldots,K^{\rm opt})$ are used to quantize their observations with limited bit rates $\bar{b}_k(k = 1,\ldots,K^{\rm opt})$ and $D_0 = (\sum_{k=1}^{K^{\rm opt}} (1)/(\sigma_k^2))^{-1}$ is the lower bound of the estimation MSE of BLUE estimator using $K^{\rm opt}$ sensors with observation noise variances $\sigma_1^2,\ldots,\sigma_{K^{\rm opt}}^2$, so $D_4 > D_0 = (\sum_{k=1}^{K^{\rm opt}} (1)/(\sigma_k^2))^{-1}$.

From all the steps above, we get

$$D_1 \geq D_2 \geq D_3 > D_4 > \left(\sum_{k=1}^{K^{\rm opt}} \frac{1}{\sigma_k^2} \right)^{-1} \quad (49)$$

thus the theorem is proved.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarsubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Netw.*, vol. 38, pp. 393–422, Mar. 2002.

[2] S. Kumar, F. Zhao, and D. Shepherd, "Special issue on collaborative information processing," *IEEE Signal Process. Mag.*, vol. 19, no. 2, p. 13–14, Mar. 2002.

[3] D. A. Castanon and D. Teneketzis, "Distributed estimation algorithms for nonlinear systems," *IEEE Trans. Autom. Control*, vol. AC-30, pp. 418–425, 1985.

[4] A. S. Willsky, M. Bello, D. A. Castanon, B. C. Levy, and G. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Trans. Autom. Control*, vol. AC-27, pp. 799–813, 1982.

[5] Z. Chair and P. K. Varshney, "Distributed Bayesian hypothesis testing with distributed data fusion," *IEEE Trans. Syst., Man, Cybern.*, vol. 18, pp. 695–699, 1988.

[6] J. L. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem," *IEEE Trans. Autom. Control*, vol. AC-24, pp. 266–269, 1979.

[7] C. Y. Chong, "Hierarchical estimation," presented at the 2nd MIT/ONR Workshop on C3, Monterey, CA, Jul. 1979.

[8] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.

[9] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[10] W. M. Lam and A. R. Reibman, "Design of quantizers for decentralized systems with communication constraints," *IEEE Trans. Commun.*, vol. 41, pp. 1602–1605, Aug. 1993.

[11] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, Part I: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, Mar. 2006.

[12] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, Part II: Unknown probability density function," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2784–2796, Jul. 2006.

[13] V. Megalooikonomou and Y. Yesha, "Quantizer design for distributed estimation with communications constraints and unknown observation statistics," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 181–184, Feb. 2000.

[14] V. Megalooikonomou and Y. Yesha, "Space efficient quantization for distributed estimation by a multi-sensor fusion system," *Inf. Fusion*, no. 5, pp. 299–308, 2004.

[15] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2210–2219, Jun. 2005.

[16] Z.-Q. Luo and J.-J. Xiao, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.

[17] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2006.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. 1, Estimation Theory.

Junlin Li (S'05) received the B.S. degree and M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2001 and 2004, respectively. He is currently working towards the Ph.D. degree at the School of Electrical and Computer Engineering in the Georgia Institute of Technology.

His research interests include multimedia signal processing, distributed signal processing, and wireless sensor networks.

Ghassan AlRegib (S'01–M'03) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology in 2003.

He joined the faculty at the Georgia Institute of Technology in August 2003, where he is currently an Assistant Professor in the School of Electrical and Computer Engineering. His research group is currently working on projects related to multimedia communications, video and 3-D graphics streaming, multimodal processing of 3-D objects, distributed processing, collaborative systems, and wireless sensor networks.

Dr. AlRegib was the recipient of the Outstanding Graduate Teaching Award in 2000–2001 at the School of Electrical and Computer Engineering at Georgia Tech, the Center for Signal and Image Processing Research Award in spring 2003 at Georgia Tech, and the Center for Signal and Image Processing Service Award in spring 2003 at Georgia Tech. He is the Co-Chair of the First International Conference on Immersive Telecommunications to be held in Italy in May 2007. He is also the Chair of the Special Sessions Program at the IEEE International Conference on Image Processing 2006 in Atlanta, GA. He is also serving as Associate Editor of the *IEEE Signal Processing Magazine* 2006–2007.