Abstract—In this paper, we consider distributed estimation of a noise-corrupted deterministic parameter in energy-constrained wireless sensor networks from energy-distortion perspective. Given a total energy budget allowable to be used by all sensors, there exists a tradeoff between the subset of active sensors and the energy used by each active sensor in order to minimize the estimation MSE. To determine the optimal quantization bit rate and transmission energy of each sensor, a concept of equivalent unit-energy MSE function is introduced. Based on this concept, an optimal energy-constrained distributed estimation algorithm for homogeneous sensor networks and a quasi-optimal energy-constrained distributed estimation algorithm for heterogeneous sensor networks are proposed. Moreover, the theoretical energy-distortion performance bound for distributed estimation is addressed and it is shown that the proposed algorithm is quasi-optimal within a factor 2 of the theoretical lower bound. Simulation results also show that the proposed method can achieve a significant reduction in the estimation MSE when compared with other uniform schemes. Finally, the proposed algorithm is easy to implement in a distributed manner and it adapts well to the dynamic sensor environments.

Index Terms—Best linear unbiased estimator (BLUE), distributed estimation, energy-constrained wireless sensor networks, quadrature amplitude modulation (QAM).

I. INTRODUCTION

WIRELESS sensor networks (WSNs) is an emerging technology that has many current and future envisioned applications, such as environment monitoring, battlefield surveillance, health care, and home automation [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 due to various design considerations such as small size battery, bandwidth and cost, potentially powerful networks can be constructed to accomplish various high-level tasks via sensor cooperation [2], such as distributed estimation, distributed detection, and target localization and tracking.

Since sensors are equipped with small batteries, which are costly if not impossible to recharge or replace, sensor network operations must be energy efficient in order to maximize the network lifetime. Given the power constraints in sensors, one of the major objectives of the sensor network research is to design energy-efficient devices, protocols, and algorithms. In the context of energy-constrained wireless sensor networks, we study the optimal distributed parameter estimation by a set of distributed sensor nodes and a fusion center (FC) in this paper. Sensors collect real-valued data, perform a local data compression, and send the resulting messages to the fusion center that combines the received messages to produce a final estimation of the observed parameter.

Distributed estimation of unknown deterministic parameters by a set of distributed sensor nodes and a fusion center has become an important topic in signal processing research for wireless sensor networks [3]. Most of the early works [4]–[8] 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 bandwidth and energy cost.

Subject to severe bandwidth and energy constraints, each sensor in wireless sensor networks is allowed to transmit only a quantized version of its raw measurement to the fusion center. Recently, several bandwidth-constrained distributed estimation algorithms have been investigated [9]–[21]. The work of [9]–[11] 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 [12] and [13], a class of maximum likelihood estimators (MLE) was proposed to attain a variance that is close to the clairvoyant estimator when the observations are quantized to one bit. In [14] and [15], one-bit adaptive quantization schemes were proposed to asymptotically achieve an estimation mean square error (MSE) close to the clairvoyant estimator using unquantized data. The work of [16] addressed the maximum likelihood estimation over noisy channel for bandwidth-constrained sensor networks. Without the knowledge of noise distribution, the work of [17] and [18] proposed to use a training sequence to aid the design of local data quantization strategies, and the work of [19] and [20] proposed several universal (pdf-unaware) decentralized estimation systems based on best linear unbiased estimation (BLUE) rule for distributed parameter estimation in the presence of unknown, additive sensor noise. While most of the aforementioned work on bandwidth-constrained distributed estimation are posed for a given number of sensors (one observation per sensor) [9]–[20], the work of [21] proposed (quasi-)optimal distributed parameter estimation algorithms to minimize the...
estimation MSE for wireless sensor networks with a total rate constraint by optimally allocate the rate among all sensors.

To explicitly address the energy constraints in wireless sensor networks, energy-constrained distributed estimation algorithms have also been studied in [22]–[26]. In [22] and [23], the total sensor transmission energy is minimized by selecting the optimal quantization levels while meeting the target estimation MSE requirements. On the contrary, the work of [24] is to minimize the estimation MSE under the given energy constraints. The work of [25], [26] addressed the energy-constrained distributed estimation problem (under the BLUE fusion rule) by exploiting long-term noise variance statistics. Most of the aforementioned work on energy-constrained distributed estimation [22]–[26] is focused on specific estimation and transmission models with appropriate problem formulation and approximation.

In this paper, we study a generic framework for energy-constrained distributed estimation in wireless sensor networks from energy-distortion perspective, which can address different transmission and energy models. Here, the fundamental question is, *What is the optimal energy-distortion bound for distributed estimation and how to achieve the performance bound in a distributed manner?* More specifically, the problem we address is to minimize the estimation MSE under a given total energy budget by optimally scheduling the quantization bit rate and transmission energy for all sensors.

Based on the total energy constraint for all sensors, there exists an interesting tradeoff between the number of active sensors and the energy consumed at each active sensor. To solve this optimal tradeoff and design the optimal distributed estimation algorithm, we first define a novel concept of equivalent unit-energy MSE function, based on which, we further study how to i) select a subset of active sensors to observe the phenomenon, and ii) determine the quantizer and transmission energy for each active sensor to quantize its real-valued observation and transmit the quantized message to the fusion center, which performs the final estimation based on the quasi-BLUE fusion rule. Furthermore, the energy-distortion performance bound for distributed estimation is also addressed by utilizing the equivalent unit-energy MSE function concept. It is noted that the proposed concept and algorithm can be used to solve the energy-constrained distributed estimation with different transmission and energy models. It is also noted that the proposed algorithm is easy to implement in a distributed manner and it adapts well to the dynamic sensor environments, which both are desirable for wireless sensor network applications.

The rest of the paper is organized as follows. Section II states the distributed estimation problem under the total energy constraint. Section III introduces a concept of equivalent unit-energy 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 MSE of our proposed algorithm and a theoretical energy-distortion lower bound of distributed estimation are proved. Section VI presents some simulation and discussion that demonstrate the efficiency of the proposed algorithms. Finally, conclusions are given in Section VII.

II. PRELIMINARY AND PROBLEM STATEMENT

We consider a dense sensor network that includes $N$ distributed sensors, denoted as $\{1, \ldots, N\}$. Each sensor can observe, quantize and transmit its observation to the fusion center, which will estimate the unknown parameter $\theta$ based on the received messages. Since the total energy allowed to be used by all sensors is limited, there exists a tradeoff between the number of active sensors and the energy used by each active sensor, that is to say, only a subset of the sensors will be active at each task period. Assume there are $K$ active sensors and denote the subset of active sensors as $S_K = \{i_1, \ldots, i_K\}$ ($i_k \in [1, N]$ for $k = 1, \ldots, K$), the distributed estimation system can be described as follows (Fig. 1).

First, each active sensor $k \in S_K$ makes an observation on the unknown parameter $\theta$, which is corrupted by additive noise and is described by

$$x_k = \theta + n_k, k \in S_K.$$  

We assume that the observation noises of all sensors $n_k$ ($k = 1, \ldots, N$) are zero mean, spatially uncorrelated with variance $\sigma^2_{n_k}$, otherwise unknown. Second, each active sensor $k$ performs a local quantization $m_k = Q_k(x_k)$, where $Q_k(x_k)$ is a quantization function, and the quantization message $m_k$ is then transmitted to the fusion center where all the quantization messages are combined to produce a final estimation of $\theta$ using a fusion function. The quality of an estimation for $\theta$ is measured by the MSE criterion.

A. BLUE Estimation Rule

If the fusion center has the knowledge of the sensor noise variance $\sigma^2_{n_k}$ ($k \in S_K$) and the sensors can perfectly send their observations $x_k$ to the fusion center, the BLUE estimator [27] for $\theta$ is known to be

$$\bar{\theta} = \left( \sum_{k \in S_K} \frac{1}{\sigma^2_k} \right)^{-1} \sum_{k \in S_K} \frac{x_k}{\sigma^2_k}$$  

and the estimation MSE of the BLUE estimator is

$$E(\bar{\theta} - \theta)^2 = \left( \sum_{k \in S_K} \frac{1}{\sigma^2_k} \right)^{-1}.$$  

But the BLUE scheme is impractical for wireless sensor networks because of the high communication and thus high energy

![Fig. 1. Distributed estimation system under the total energy constraint.](image-url)
cost. Instead of sending the real-valued observations to the fusion center directly, quantization at the local sensors is essential to reduce the communication bandwidth and energy cost. In this paper, we adopt a probabilistic quantization scheme [22] as well as a quasi-BLUE estimation scheme, based on which the optimal tradeoff between the number of active sensors and the energy allocated at each active sensor is addressed.

Suppose the observation $x_k$ is bounded to $[-W, W]$, that is, $x_k = \theta + n_k \in [-W, W]$. The probabilistic quantization with $b_k$ bits is summarized as follows: uniformly divide $[-W, W]$ into intervals of length $\Delta = 2W/(2^{b_k} - 1)$, and round $x_k$ to the neighboring endpoints of these small intervals in a probabilistic manner. More specifically, suppose $-W + i\Delta \leq x_k \leq -W + (i + 1)\Delta$, where $0 \leq i \leq 2^{b_k} - 2$, then $x_k$ is quantized to $m_k(x_k, b_k)$ according to

$$P\{m_k(x_k, b_k) = -W + i\Delta\} = 1 - r$$

$$P\{m_k(x_k, b_k) = -W + (i + 1)\Delta\} = r$$ (4)

where $r = (x_k + W - i\Delta)/\Delta \in [0, 1]$. As shown in [22], the quantized message $m_k(x_k, b_k)$ is an unbiased estimator of $\theta$ with a variance

$$E[(m_k(x_k, b_k) - \theta)^2] \leq \sigma_k^2 + \frac{W^2}{(2^{b_k} - 1)^2} = \sigma_k^2 + \delta_k^2$$ (5)

where $\delta_k^2 = W^2/(2^{b_k} - 1)^2$ for $b_k > 0$ denotes the upper bound of the quantization noise variance.

Now suppose all the observations $x_k (k \in S_K)$ of the $K$ active sensors are quantized into $b_k$-bits discrete messages $m_k(x_k, b_k)$ respectively with the probabilistic quantization scheme. Treating all the quantized messages $m_k$ as the new observations for the fusion center, the quasi-BLUE estimator based on the quantized message has the following form:

$$\tilde{\theta} = \left(\sum_{k \in S_K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \sum_{k \in S_K} \frac{m_k}{\sigma_k^2 + \delta_k^2}.$$ (6)

Notice that $\tilde{\theta}$ is an unbiased estimator of $\theta$ since every $m_k$ is unbiased. Moreover, the estimation MSE of the quasi-BLUE estimator is

$$E(\tilde{\theta} - \theta)^2 \leq \left(\sum_{k \in S_K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1}.$$ (7)

B. Transmission and Energy Models

To transmit a $b$-bit message from a sensor to the fusion center, the transmission energy cost is generally a function of the transmission bit rate and the transmission distance. Assume that each sensor sends a message to the fusion center using a separate channel, which can be achieved by using a multiple access technique such as TDMA or FDMA; and the channel between the sensor $k$ and the fusion center experiences a path loss proportional to $d_k = \rho_k^{\alpha}$, where $d_k$ is the transmission distance and $\alpha$ is the path loss exponent.

To address the energy-constrained distributed estimation problem, we consider several different popular transmission models: 1) binary transmission model, 2) uncoded quadrature amplitude modulation (QAM) model, and 3) coded quadrature amplitude modulation model. To reliably transmit $b_k$-bit message from the sensor $k$ to the fusion center, the transmission energy cost for the binary transmission model, where each bit will be transmitted separately, is

$$P_{\text{BIN}}(b_k) = c_k^b \cdot a_k \cdot b_k$$ (8)

where $c_k^b$ is a system constant. To minimize the transmission bandwidth and transmission delay, the $b_k$ bits can be transmitted simultaneously using $M$-ary quadrature amplitude modulation (QAM) with constellation size $2^{b_k}$, then the transmission energy cost [28], [29] is given by

$$P_{\text{QAM}}(b_k) = c_k^b \cdot a_k \cdot (2^{b_k} - 1)$$ (9)

where $c_k^b$ is a system constant defined the same as in [28] and [29]. Furthermore, with embedded error correction codes, coded QAM can reduce the transmission energy cost by a constant factor $G_c$ [28], [29], i.e.,

$$P_{\text{CQAM}}(b_k) = c_k^b \cdot a_k \cdot (2^{b_k} - 1)$$ (10)

where $c_k^b = c_k^b/G_c$ is a system constant defined the same as in [28], [29]. Thereafter, we call the system constant $c$ the transceiver characteristic parameter. It is noted that, compared with the binary modulation and transmission scheme, the QAM schemes also minimize the transmission delay and the circuit energy consumption since it minimizes the number of transmissions by transmitting the whole $b_k$-bit message as a single symbol.

C. Distributed Estimation Under Energy Constraints

With the probabilistic quantization scheme and the quasi-BLUE fusion rule, our primary goal is to minimize the upper bound of the estimation MSE under the energy constraint, i.e.,

$$\min \left(\sum_{k \in S_K} \frac{1}{\sigma_k^2 + \frac{W^2}{2^{b_k} - 1}}\right)^{-1},$$

s.t. $\sum_{k \in S_K} P_k \leq P_e$, $P_k > 0, b_k > 0, k \in S_K$ (11)

where $S_K$ is the subset of $K$ active sensors, $b_k$ and $P_k$ are the quantization bit rate and transmission energy of active sensor $k \in S_K$, and $P_e$ is the total energy allowed to be used by all active sensors.

It is obvious that the solution to the energy-constrained distributed estimation problem stated in (11) depends on the energy model used. For a special energy model, where the energy cost $P$ is assumed to be a constant linear function of the transmission bit rate $b$, i.e., $P = c \cdot b$, the energy-constrained distributed problems is retrogressed to the rate-constrained distributed estimation problem addressed in our previous work [21]. So the rate-constrained distributed estimation can be treated as a special case of energy-constrained distributed estimation. In this work, we will consider the energy-constrained distributed estimation problem with the QAM-based models. It is worth noting that the similar concept and methodology proposed in this work
can be extended to solve the energy-constrained distributed estimation with different transmission and energy models.

Using the uncoded/coded QAM models, the original problem in (11) turns to the following problem:

$$\min \left( \sum_{k \in S_K} \frac{\sigma_k^2 + \frac{W^2}{(2^b - 1)^2}}{P_k} \right)^{-1},$$

s.t. $$\sum_{k \in S_K} P_k \leq P_c,$$

$$P_k = a_k b_k (2^b - 1), k \in S_K,$$

$$b_k > 0, k \in S_K$$  \hspace{1cm} (12)

where all the variables are defined as before. In practice, the quantization bit rate \( b_k \) must be integer, i.e., \( b_k \in \mathbb{Z} \). To facilitate the subsequent analysis, we will relax the integer condition \( b_k \in \mathbb{Z} \) to \( b_k \in \mathbb{R} \). Later, we will discuss how to constrain the quantization bit rate to integer numbers.

As discussed in Appendix A, the optimal solution for the energy-constrained distributed estimation problem in (12) cannot be found in a closed form. In the following sections, we will address this problem for homogeneous and heterogeneous sensor networks, respectively. To facilitate the solution, we first define an equivalent unit-energy MSE function in the next section.

### III. EQUIVALENT UNIT-ENERGY MSE FUNCTION

As shown in Section II, the \( b \)-bit quantization message from a sensor with observation noise variance \( \sigma^2 \) is an unbiased estimation of the parameter \( \theta \). We denote the estimation MSE bound as

$$f(\sigma^2, b) := \sigma^2 + \frac{W^2}{(2^b - 1)^2}. \hspace{1cm} (13)$$

**Definition 1 (Equivalent Unit-Energy MSE Function):** For a sensor with observation noise variance \( \sigma^2 \), quantization bit rate \( b \), transmission path loss \( a \), transceiver parameter \( c \), and transmission energy cost \( P(b, a, c) \), the equivalent unit-energy MSE function is defined as

$$g(\sigma^2, b, a, c) := P(b, a, c) \cdot f(\sigma^2, b). \hspace{1cm} (14)$$

With this definition, the estimation MSE of the quasi-BLUE estimator, shown in (7), can be rewritten as

$$E(\hat{\theta} - \theta)^2 \leq \left( \sum_{k \in S_K} \frac{1}{f(\sigma_k^2, b_k)} \right)^{-1}$$

$$= \left( \sum_{k \in S_K} \frac{P_k}{g(\sigma_k^2, b_k, a_k, c_k)} \right)^{-1}. \hspace{1cm} (15)$$

From (15), we can see that a sensor with transmission energy \( P_k \) and estimation MSE \( f(\sigma_k^2, b_k) \) achieves the same estimation MSE as \( P_k \) equivalent uni-energy sensors, each with the same estimation MSE \( g(\sigma_k^2, b_k, a_k, c_k) = P_k \cdot f(\sigma_k^2, b_k) \). That is why the function \( g(\sigma^2, b, a, c) \) is called equivalent unit-energy MSE function.

With the uncoded/coded QAM models, the equivalent unit-energy MSE function defined in (14) is

$$g(\sigma^2, b, a, c) = P(b, a, c) \cdot f(\sigma^2, b)$$

$$= ca(b^2 - 1) \left( \sigma^2 + \frac{W^2}{(2^b - 1)^2} \right). \hspace{1cm} (16)$$

As shown in Proposition 1, \( g(\sigma^2, b, a, c) \) is convex over \( b \). We further define the optimal unit-energy MSE function \( g_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \), and the corresponding optimal quantization bit rate \( b_{\text{opt}}(\sigma^2, a, c) \) and optimal transmission energy \( P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) for an original sensor with observation noise variance \( \sigma^2 \), transmission path loss \( a \), and transceiver parameter \( c \) as follows:

$$b_{\text{opt}}(\sigma^2, a, c) = \arg \min_{b \in \mathbb{R}^+} g(\sigma^2, b, a, c)$$

$$= \arg \min_{b \in \mathbb{R}^+} \left[ ca(2^b - 1) \left( \sigma^2 + \frac{W^2}{(2^b - 1)^2} \right) \right]$$

$$g_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) = \min_{b \in \mathbb{R}^+} g(\sigma^2, b, a, c)$$

$$= g(\sigma^2, b_{\text{opt}}(\sigma^2, a, c), a, c)$$

$$P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) = ca \left( 2^{b_{\text{opt}}(\sigma^2, a, c)} - 1 \right). \hspace{1cm} (17)$$

**Proposition 1:** The equivalent unit-energy MSE function \( g(\sigma^2, b, a, c) \), the optimal unit-energy MSE function \( g_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \), the optimal quantization bit rate function \( b_{\text{opt}}(\sigma^2, a, c) \), the optimal transmission energy function \( P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) and the MSE function \( f(\sigma^2, b) \) defined before have the following properties.

1. \( g(\sigma^2, b, a, c) \) is convex over \( b \in \mathbb{R}^+ \).
2. \( g_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) is achieved when the optimal quantization bit rate \( b_{\text{opt}}(\sigma^2, a, c) \) is used and the optimal transmission energy \( P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) is allocated, where

$$b_{\text{opt}}(\sigma^2, a, c) = \log_2 \left( 1 + \frac{W}{\sigma} \right)$$

$$g_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) = 2ca\sigma W$$

$$P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) = caW. \hspace{1cm} (18)$$

3. The estimation MSE \( f(\sigma^2, b) \) with the optimal quantization bit rate \( b_{\text{opt}}(\sigma^2, a, c) \) and transmission energy \( P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) is

$$f(\sigma^2, b_{\text{opt}}(\sigma^2, a, c)) = 2\sigma^2. \hspace{1cm} (19)$$

The Proposition 1 is easy to prove as follows: the convexity of \( g(\sigma^2, b, a, c) \) over \( b \) can be proved by checking \( \partial^2 g(\sigma^2, b, a, c)/\partial b^2 > 0 \) for any \( b < 0, 0 > b \); then \( b_{\text{opt}}(\sigma^2, a, c) \) can be obtained by solving \( \partial g(\sigma^2, b, a, c)/\partial b = 0 \), and

\( g_{\text{opt}}^{\text{eq}}(\sigma^2, a, c), P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) and \( f(\sigma^2, b_{\text{opt}}(\sigma^2, a, c)) \) can be obtained according to the definitions in (13), (16), and (17).

It is noted that the optimal transmission energy function \( P_{\text{opt}}^{\text{eq}}(\sigma^2, a, c) \) depends not only on the signal to noise ratio but
also on the transmission path loss and transceiver parameter, but the optimal quantization bit rate function $b^{\text{opt}}(\sigma^2, a, c)$ only depends on the signal to noise ratio as shown in (18).

IV. DISTRIBUTED ESTIMATION IN HOMOGENEOUS SENSOR NETWORKS

In homogeneous sensor networks, the noise variances for all sensors are identical, that is $\sigma_n^2 = \sigma^2 (k = 1, \ldots, N)$. We assume equal distances and same channel condition from all sensors to the fusion center, thus the transmission path loss is the same for all sensors too, i.e., $d_k = d$ and $a_k = a(k = 1, \ldots, N)$. Also, assume that the transceiver parameters are the same for all sensors, i.e., $c_k = c (k = 1, \ldots, N)$. Therefore, the equivalent unit-energy MSE function defined in (14) is the same for all sensors.

Assume there are $K$ active sensors and the transmission energy for each sensor is $P_k$ such that $\sum_{k=1}^{K} P_k = P_c$, then the estimation MSE shown in (15) can be reformulated as

$$E(\hat{\theta} - \theta)^2 = \left( \sum_{k=1}^{K} \frac{P_k}{g_k^{\text{opt}}(\sigma^2, a, c)} \right)^{-1} \left( \sum_{k=1}^{K} \frac{P_k}{g_k^{\text{opt}}(\sigma^2, a, c)} \right)^{-1} \left( \sum_{k=1}^{K} \frac{P_k}{g_k^{\text{opt}}(\sigma^2, a, c)} \right)^{-1} = \frac{P_c (\sigma^2, a, c)}{P_c}$$

where the equality is achieved when each active sensor adopts the optimal quantization bit rate $b^{\text{opt}}(\sigma^2, a, c)$ and optimal transmission energy $P^{\text{opt}}(\sigma^2, a, c)$ defined in (17) and (18). So the solution for energy-constrained distributed estimation in homogeneous sensor networks is stated as follows.

1) For all sensors, the optimal quantization bit rate $b^{\text{opt}}$ and transmission energy $P^{\text{opt}}$ are the same and obtained by minimizing the corresponding equivalent unit-energy MSE function, as shown in Proposition 1:

$$b^{\text{opt}} = \log_2 \left( 1 + \frac{W}{\sigma} \right)$$
$$P^{\text{opt}} = \frac{caW}{\sigma}.$$ (21)

2) The total number of active sensors $K^{\text{opt}}$ under the total energy constraint $P_c$ is

$$K^{\text{opt}} = \left\lceil \frac{P_c}{P^{\text{opt}}} \right\rceil.$$ (22)

It is obvious that the proposed method based on the equivalent unit-energy MSE function is optimal if $P_c/P^{\text{opt}}$ is an integer, otherwise, it is quasi-optimal.

Remark 1: It is noted that the proposed method based on equivalent unit-energy MSE function can be implemented in a fully distributed manner. First, the optimal quantization bit rate $b^{\text{opt}}$ and optimal transmission energy $P^{\text{opt}}$ of each sensor can be obtained locally by minimizing its corresponding equivalent unit-energy MSE function. Second, the subset of active sensors is chosen in a round-robin manner such that there are $K^{\text{opt}} = P_c/P^{\text{opt}}$ (we assume $P_c/P^{\text{opt}}$ is integer here) active sensors at any task period and each sensor will be active for $K^{\text{opt}}$ 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 heterogeneous sensor networks, the observation noise variance for sensor $k$ is $\sigma_n^2(k = 1, \ldots, N)$, respectively. Assume the distance from sensor $k$ to the fusion center is $d_k$, thus the transmission path loss is $a_k = d_k^2$, and assume the transceiver parameter of sensor $k$ is $c_k$. This scenario leads to the general problem stated in (12). The goal is to find the optimal number of active sensors and the corresponding optimal quantization bit rate and transmission energy allocation for each active sensor to minimize the estimation MSE bound at the fusion center.

Unfortunately, as discussed in Appendix A, the optimal solution for the general case in (12) cannot be found in a closed form. Instead, we propose a quasi-optimal method to solve this problem, which is also based on the equivalent unit-energy MSE function. The procedure is stated as follows.

1) For each sensor $k \in [1, N]$, determine its optimal quantization bit rate $b_k^{\text{opt}}$, optimal transmission energy $P_k^{\text{opt}}$ and optimal unit-energy MSE function $g_k^{\text{opt}}$ as shown in Proposition 1:

$$b_k^{\text{opt}} = \log_2 \left( 1 + \frac{W}{\sigma} \right)$$
$$g_k^{\text{opt}} = 2\sigma d_k a_k W$$
$$P_k^{\text{opt}} = \frac{ca_k a_k W}{\sigma}.$$ (23)

2) Sort all the sensors by their corresponding optimal unit-energy MSE function $g_k^{\text{opt}}(k \in [1, N])$ from the smallest to the largest, i.e., $g_1^{\text{opt}} \leq g_2^{\text{opt}} \leq \cdots \leq g_N^{\text{opt}}$. Let $S_k(k \in [1, N])$ denote the subset of all sensors consisting of the first $k$ sensors with the smallest optimal unit-energy MSE function, then

$$S_k = \{i_1, i_2, \ldots, i_k\}.$$ (24)

Let $S_k^c$ denote the complemental subset of $S_k$, then,

$$S_k^c = \{i_{k+1}, \ldots, i_N\}.$$ (25)

and

$$g_i^{\text{opt}} \leq g_j^{\text{opt}}, \text{if } i \in S_k \text{ and } j \in S_k^c.$$ (26)

Based on these definitions, the optimal number of active sensors $K^{\text{opt}}$ under the total energy constraint $P_c$ is determined by

$$K^{\text{opt}} = \max_k \text{ s.t., } \sum_{i \in S_k} P_i^{\text{opt}} \leq P_c$$ (27)

that is to say, the subset of active sensors is $S_{K^{\text{opt}}}$.

In short, the whole solution is: all sensors in the subset $S_{K^{\text{opt}}}$, i.e., the first $K^{\text{opt}}$ sensors with the smallest optimal unit-energy MSE function, are active to quantize their observations with
quantization bit rate $b_k^{\text{opt}}$ and transmit their quantized messages to the fusion center with transmission energy $P_k^{\text{opt}} (k \in S_{K^{\text{opt}}})$.

To implement the described algorithm above, each sensor needs to decide i) whether it should be active, i.e., whether it belongs to $S_{K^{\text{opt}}}$, and ii) its quantization bit rate and transmission energy if it will be active. Both tasks can be achieved in a distributed manner as follows.

- As shown in (24) and (27), the subset of active sensors $S_{K^{\text{opt}}}$ is determined at the fusion center based on the collected network information and the total energy constraint $P_c$. Denote the maximum optimal unit-energy MSE function of all the active sensors in the subset $S_{K^{\text{opt}}}$ as

$$g_k^{\text{opt}} = \arg \max_{k \in S_{K^{\text{opt}}}} g_k^{\text{opt}}.$$ (28)

Then the fusion center broadcasts the threshold $g_k^{\text{opt}}$ to all the local sensors. Upon receiving the threshold, each sensor compares the threshold with its own optimal unit-energy MSE function $g_k^{\text{opt}}$. If $g_k^{\text{opt}} \leq g_k^{\text{th}}$, then sensor $k$ is active; otherwise, it is inactive.

- As shown in (23), the optimal quantization bit rate $b_k^{\text{opt}}$ of sensor $k \in [1, N]$ depends only on its own signal to noise ratio, and the optimal transmission energy $P_k^{\text{opt}}$ and the optimal unit-energy MSE function $g_k^{\text{opt}}$ of sensor $k$ depend only on its own optimal quantization bit rate $b_k^{\text{opt}}$, transmission path loss $q_k$ and transceiver parameter $c_k$. Therefore, all of $b_k^{\text{opt}}$, $P_k^{\text{opt}}$, and $g_k^{\text{opt}}$ can be computed locally at each sensor without requiring information from other sensors.

Remark 2: As shown above, the total energy constraint $P_c$ is to determine the subset of active sensors according to (24), (27). It is interesting to see that if the total energy constraint $P_c$ is changed, we only need to wake up several more sleep sensors (energy constraint increased) or send several active sensors to sleep (energy constraint decreased), but do not need to change the quantization bit rate and transmission energy allocation of each active sensor. So the proposed method adapts well to the situations when the total energy constraints need to be changed frequently to achieve various estimation MSE performances, which is the case for dynamic sensor environments.

Next, we will analyze the estimation MSE bound of the proposed method, which is stated in the following theorem. To simplify the statements, we assume $\sum_{k \in S_{K^{\text{opt}}}} P_k^{\text{opt}} = P_c$ in the subsequent analysis.

**Theorem 1:** The estimation MSE of the proposed method based on the equivalent unit-energy MSE function under the total energy constraint $P_c$ is

$$\left( \sum_{k \in S_{K^{\text{opt}}}} \frac{1}{\sigma_k^2} \right)^{-1} < E(\tilde{\theta}_p - \theta)^2 < 2 \left( \sum_{k \in S_{K^{\text{opt}}}} \frac{1}{\sigma_k^2} \right)^{-1},$$ (29)

where $\tilde{\theta}_p$ denotes the estimation of the parameter $\theta$ by the proposed method, and $S_{K^{\text{opt}}}$ is the optimal subset of active sensors, obtained in (24) and (27).

**Proof:** The left part of the theorem is obvious since $(\sum_{k \in S_{K^{\text{opt}}}} 1/\sigma_k^2)^{-1}$ is the lower bound of the estimation MSE of the BLUE estimator using the subset $S_{K^{\text{opt}}}$ of sensors without energy constraint. To prove the right part of the theorem, by Proposition 1, we have

$$E(\tilde{\theta}_p - \theta)^2 < \left( \sum_{k \in S_{K^{\text{opt}}}} \frac{1}{\sigma_k^2} \right)^{-1} = 2 \left( \sum_{k \in S_{K^{\text{opt}}}} \frac{1}{\sigma_k^2} \right)^{-1}.$$ (30)

This theorem gives the lower and upper bounds of the estimation MSE of the proposed method. It is shown that the proposed method is quasi-optimal (up to a factor of 2) when compared with the BLUE estimator using the same subset of active sensors without energy constraint.

As shown above, the performance bound of the proposed algorithm is analyzed. Nevertheless, the remaining question is what is the optimal energy-distortion bound for distributed estimation, i.e., what is the minimal estimation MSE that be achieved if the total energy $P_c$ is allocated to any subset of the sensors. To answer this question, Theorem 2 states the lower bound of the estimation MSE by any quasi-BLUE estimation system with any subset of sensors under the total energy constraint $P_c$. Surprisingly, under the same total energy constraint $P_c$, the lower bound of the estimation MSE by any quasi-BLUE estimation system with any subset $S$ of sensors is same as the lower bound of the estimation MSE of the BLUE estimator using the subset $S_{K^{\text{opt}}}$ of sensors obtained by the proposed algorithm in (24) and (27).

**Theorem 2:** Assume any subset of sensors $S = \{i_1, \ldots, i_q\}$ is used, where $i_k \in [1, N]$ and $|S|$ denotes the cardinality of the set $S$, i.e., the total number of sensors in the set $S$. The energy allocated to each active sensor $k \in S$ is $P_k$, such that $\sum_{k \in S} P_k = P_c$. Then the lower bound of the estimation MSE is

$$E(\tilde{\theta}_c - \theta)^2 > \left( \sum_{k \in S_{K^{\text{opt}}}} \frac{1}{\sigma_k^2} \right)^{-1},$$ (31)

where $\tilde{\theta}_c$ denotes the estimation of the parameter $\theta$ by the subset of active sensors $S$ under the given total energy constraint $P_c$, and $S_{K^{\text{opt}}}$ is the optimal subset of active sensors, obtained by our proposed algorithm as shown in (24) and (27) such that $\sum_{k \in S_{K^{\text{opt}}}} P_k^{\text{opt}} = P_c$.

**Proof:** Refer to Appendix B for the complete proof.

In conclusion, Theorem 1 shows that the bound of estimation MSE of our proposed method is $(\sum_{k \in S_{K^{\text{opt}}}} 1/\sigma_k^2)^{-1} < E(\tilde{\theta}_p - \theta)^2 < 2(\sum_{k \in S_{K^{\text{opt}}}} 1/\sigma_k^2)^{-1}$, and Theorem 2 shows that $(\sum_{k \in S_{K^{\text{opt}}}} 1/\sigma_k^2)^{-1}$ is the lower bound of the estimation MSE of any quasi-BLUE estimator under the total energy constraint $P_c$, regardless of the subset of active sensors and the energy allocation among the active sensors. Therefore, the proposed algorithm gives a quasi-optimal tradeoff between the number of active sensors and the energy allocation at each active sensor, and its estimation MSE is within a factor of 2 of the theoretical non-achievable lower bound.
Remark 2: As we mentioned before, in all the prior analysis, we assume the quantization bit rate can be real-valued number. But in practice, the quantization bit rate must be integer. Denote the optimal integer quantization bit rate as $b_{\text{opt}}^r(\sigma^2, a, c) \in \mathbb{Z}^+$, the corresponding optimal transmission energy as $P_{\text{opt}}^r(\sigma^2, a, c)$ and optimal equivalent unit-energy MSE function as $\tilde{\gamma}^r(\sigma^2, a, c)$ for a sensor with observation noise variance $\sigma^2$, transmission path loss $a$, and transceiver parameter $c$, thus,

$$
\begin{align*}
\tilde{\gamma}^r(\sigma^2, a, c) &= \arg\min_{b \in \mathbb{Z}^+} g(\sigma^2, b, a, c) \\
&= \arg\min_{b \in \mathbb{Z}^+} \left[ b(2^b - 1) \left( \sigma^2 + \frac{W^2}{(2^b - 1)^2} \right) \right], \\
\tilde{\gamma}^r(\sigma^2, a, c) &= g(\sigma^2, b_{\text{opt}}^r(\sigma^2, a, c), a, c), \\
P_{\text{opt}}^r(\sigma^2, a, c) &= c \sigma^2 \left( 2^{b_{\text{opt}}^r(\sigma^2, a, c)} - 1 \right). \\
\end{align*}
$$

(32)

Different from $b^r(\sigma^2, a, c), \tilde{\gamma}^r(\sigma^2, a, c)$ cannot be written in a closed form, however, it can be easily solved since the minimization in (32) involves just a simple one-dimensional numerical search. So, in practice, the proposed distributed estimation algorithms above can be easily implemented by using $b^r(\sigma^2, a, c), P_{\text{opt}}^r(\sigma^2, a, c)$ and $\tilde{\gamma}^r(\sigma^2, a, c)$ instead of $b^r(\sigma^2, a, c), P_{\text{opt}}(\sigma^2, a, c)$ and $g^r(\sigma^2, a, c)$.

Since $g(\sigma^2, b, a, c)$ is convex over $b$ as shown in Proposition 1, $b^r(\sigma^2, a, c) \leq b^r(\sigma^2, a, c)$ or $[b^r(\sigma^2, a, c)]$, where $[b^r(\sigma^2, a, c)]$ denotes the maximum integer no more than $b^r(\sigma^2, a, c)$, and $[b^r(\sigma^2, a, c)]$ denotes the minimum integer no less than $b^r(\sigma^2, a, c)$. Fig. 2 shows the optimal real-valued quantization bit rate $b^r(\sigma^2, a, c)$ and the optimal integer quantization bit rate $b_{\text{opt}}^r(\sigma^2, a, c)$ versus different signal to noise ratios (SNR) defined as $\text{SNR} = 10 \log_{10}(W^2/\sigma^2)$. Further, Fig. 3 shows the ratio of the estimation MSE $f(\sigma^2, b)$ to $\sigma^2$ using the optimal real-valued quantization bit rate $b = b^r(\sigma^2, a, c)$ with transmission energy $P = P_{\text{opt}}^r(\sigma^2, a, c)$, or integer quantization bit rate $b = b_{\text{opt}}^r(\sigma^2, a, c)$ with transmission energy $P = P_{\text{opt}}^r(\sigma^2, a, c)$. From Fig. 3, we can see that the upper bound of the estimation MSE is twice of the observation noise variance when the optimal real-valued quantization bit rate is used, as it is proved in Proposition 1, and the upper bound of the estimation MSE is within a small factor (up to $4$) of the observation noise variance when the integer constraint is imposed on the optimal quantization bit rate. It is also worth to note that, when the optimal integer quantization bit rate is used, the theoretical lower bound shown in Theorem 2 is still valid, and it can be proved in the same way.

VI. SIMULATION AND DISCUSSION

In this section, we first present some simulation results for the proposed algorithms in Sections IV and V, respectively. In all the simulations, we assume the transceiver parameters are the same for all sensors, i.e., $c_k = c$, and the quantization bit rates to be integer number as we mentioned in Remark 3. All the final results are obtained by repeating the experiments for 10000 times and averaging the corresponding results. Then, we give a brief discussion and comparison between the proposed algorithm in this paper and the algorithm in [22].

A. Homogeneous Sensor Networks

In this section, we simulate a homogeneous sensor network with $N = 500$ sensors, where the noise variances of all sensors are the same and the distances from all sensors to the fusion center are also the same. Without loss of generality, we assume the range of the observation signal is $[-1, 1]$, i.e., $W = 1$, and the distance from each sensor to the fusion center is $d = 1$. 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 $\sigma^2$. Define the normalized energy as $P' = P/c = a(2^b - 1)$, where $c$ is the transceiver parameter, $a$ is the transmission path loss, and $b$ is the quantization bit rate.

Assuming the normalized total energy constraint is $P' = 500$, Fig. 4 shows the estimation MSE with different quantization bit rates for the active sensors under different SNR, where different quantization bit rates, amounting to different energy allocation, imply different total number of active sensors to perform the
estimation task because of the total energy constraint. Explicitly, for the given total normalized energy constraint $P^* = 500$, we can have 500 active sensors with 1-bit quantization message for each sensor, or 167 active sensors with 2-bit quantization message for each sensor, or 71 active sensors with 3-bit quantization message for each sensor, or 33 active sensors with 4-bit quantization message for each sensor and so on. For example, for the case of $\text{SNR} = 20 \text{ dB}$, totally 71 active sensors out of all 500 sensors with 3-bit quantization message per sensor will produce the minimum estimation MSE among all the possible energy allocation strategies as shown in Fig. 4. From Fig. 4, we also can see that there exists an optimal quantization bit rate for any given SNR under total energy constraint, and that too small or too big quantization bit rate will sacrifice the estimation MSE performance significantly. More specifically, 1-bit quantization per sensor will lead to the minimum estimation MSE for low SNR cases, such as 0 dB, while for high SNR cases, multiple-bit quantization per sensor will significantly decrease the estimation MSE compared to only 1-bit quantization per sensor under the same total energy constraint.

B. Heterogeneous Sensor Networks With Equal Distances

In this section, we simulate a heterogeneous sensor network with $N = 500 \text{ sensors}$, where the noise variance of each sensor is different, which is assumed to be Chi-squared distribution with one degree of freedom, while the distance from each sensor to the fusion center is the same. Same as before, we assume the range of the observation signal is $[-1, 1]$ and the distance from each sensor to the fusion center is $d = 1$.

For any given total energy constraint, our proposed estimation method in Section V is implemented to determine the subset of active sensors and the energy allocation at 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 uniform schemes.

1) Uniform-I: For the given total energy constraint, the same subset of active sensors as that used by our proposed method is used, but all energy is uniformly allocated among all the active sensors.

2) Uniform-II: all sensors in the simulated heterogeneous sensor network are used and all energy is uniformly allocated among all sensors.

Fig. 5 shows the estimation MSE by our proposed method, the Uniform-I method, the Uniform-II method, and the theoretical lower bound of the estimation MSE presented in Theorem 2 under the total energy constraint. From Fig. 5, 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 within a factor 2 of the theoretical non-achievable lower bound.

Note 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 energy allocation is performed in our proposed method, while uniform energy allocation is performed in the Uniform-I method. Because of the heterogeneity of the network, a better estimation performance is obtained using our proposed method. Define the normalized deviation of sensor noise variances as

$$\alpha = \frac{\sqrt{\text{Var}(\sigma^2)}}{E(\sigma^2)}$$

(33)

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

$$\beta = \frac{D_m - D_p}{D_m}$$

(34)

where $D_m$ denotes the estimation MSE by the Uniform-I method, and $D_p$ denotes the estimation MSE by our proposed method. Fig. 6 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. 6, 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.

C. Heterogeneous Sensor Networks With Random Distances

In this part of the simulation, we relax the assumption in Section VI-B that the distance from each sensor to the fusion center is the same. We assume the distance $d_k$ from the sensor $k$ to the fusion center is independently and uniformly distributed from 1 to 10, i.e., $d_k \sim U[1,10](k = 1,\ldots,N)$. Same with in Section VI-B, we simulate a heterogeneous sensor network with $N = 500$ sensors, where the noise variances of all sensors are different and are assumed to be Chi-squared distribution with 1 degree of freedom.

The proposed method in Section V is implemented and compared with two uniform schemes: Uniform-I and Uniform-II method defined as before. Fig. 7 shows the estimation MSE by our proposed method, the Uniform-I method, the Uniform-II method, and the theoretical lower bound of the estimation MSE under the total energy constraint. From Fig. 7, we can see that the proposed method outperforms the other two uniform schemes, and that the estimation MSE of our proposed method is within a factor 2 of the theoretical non-achievable lower bound. Comparing the results in Figs. 7 and 5, it can be seen that the proposed method obtains more gain for heterogeneous networks with random distances than for heterogeneous networks with equal distances, especially when the total energy constraint is more stringent, since there exists more randomness in the networks.

D. Discussion

In this section, we highlight the distinction between this work and the work in [22], which also addresses the energy-constrained distributed estimation with a similar system model.

In [22], the energy-constrained distributed estimation is addressed by minimizing the $L^2$-norm of the power consumption while meeting the target MSE performance. On the contrary, in this study, we address the energy-constrained distributed estimation from energy-distortion perspective, where the goal is to minimize the estimation MSE under a given total energy budget in $L^1$-norm. Furthermore, this work offers an energy-distortion performance analysis for distributed estimation, analogous to the classical rate-distortion analysis for conventional source coding in information theory.

Next, we compare the energy consumption by the methods in this work and in [22] while achieving the same target MSE performance under the same system setup.

1) Homogeneous Sensor Networks: First, we compare the energy consumption by the two methods for a special case—homogeneous sensor networks, where each sensor has the same noise variance, i.e., $\sigma_k^2 = \sigma^2(k = 1,\ldots,N)$, the same transmission pathloss, i.e., $d_k = a$, and the same transceiver parameters, i.e., $c_k = c$. Assume the total number of sensors as $N$, then the estimation MSE of the centralized BLUE is $D_0 = (\sum_{k=1}^{N} 1/\sigma_k^2)^{-1} = \sigma^2/N$. Denote the target MSE in each estimation cycle as $D = \gamma D_0$, where $\gamma > 1$ is called the normalized target MSE. Note that the distortion bound $D_0$ is non-achievable since it requires perfect transmission from all sensors to the fusion center. Also, as shown in Theorem 1, $D = 2D_0$ is the performance bound of the proposed method when all sensors are active with optimal quantization bit rate and transmission energy. So in a dense network, the normalized target MSE $\gamma$ in each estimation cycle is generally much greater than 2, i.e., $\gamma > 2$, otherwise, each sensor will consume significant amount of its limited energy every time and will deplete quickly.

For homogeneous sensor network case, it is easy to show that the ratio of the energy consumption by the method in [22] to that by the proposed method in this work is given as

$$r = \frac{\gamma}{2 \sqrt{\frac{1}{\gamma - 1}}}$$

(35)
energy saving becomes more significant when the normalized target MSE becomes bigger according to (35).

2) Heterogenous Sensor Networks: In this section, we compare the two methods for general heterogeneous sensor networks. Assume a heterogeneous sensor network with \( N = 1000 \) sensors, and the noise variance \( \sigma_k^2 \) is generated according to the distribution \( 0.1 + \chi^2(1) \), where \( \chi^2(1) \) is the Chi-square distribution with one degree of freedom. Fig. 8(a) and (b) shows the ratio of energy cost of the method in [22] to that of the proposed method in this work under different normalized target MSE. In Fig. 8(a), the distance \( d_k \) from each sensor \( k \) to the fusion center is assumed to be the same, while in Fig. 8(b), the distance from each sensor to the fusion center is assumed to be uniformly distributed from 1 to 5, i.e., \( d_k \sim U[1, 10] \). The pathloss exponent is assumed to be \( \alpha = 2 \) in both cases. From Fig. 8(a) and (b), the similar conclusions as in homogeneous sensor network case can be drawn.

VII. CONCLUSION

In this paper, we considered the distributed parameter estimation in energy-constrained wireless sensor networks from energy-distortion perspective. For a given constraint on the allowable total energy to be used by all sensors at each estimation cycle, we studied the optimal tradeoff between the subset of active sensors and the energy used by each active sensor to minimize the estimation MSE. To facilitate the solution, a concept of equivalent unit-energy MSE function was introduced. Then, an optimal distributed estimation algorithm for homogeneous sensor networks and a quasi-optimal distributed estimation algorithm for heterogeneous sensor networks, which are both based on the equivalent unit-energy MSE function, were proposed. Furthermore, the lower and upper bounds of the estimation MSE of the proposed algorithm were discussed and a theoretical energy-distortion bound for distributed estimation was proved. It is shown that our proposed algorithm is quasi-optimal within a factor 2 of the theoretical lower bound. Simulation results also show that a significant reduction in estimation MSE is achieved by our proposed algorithm when compared with other uniform methods. It is worth noting that the proposed framework is generic to address rate-constrained distributed estimation and energy-constrained distributed estimation with different transmission and energy models.

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 problems under the total energy constraint. For the general cases, the quantization scheme and the fusion rule need to take into account the sensor correlation and channel fading, based on which the sensor scheduling and the quantization bit rate and transmission energy allocation also need to be jointly optimized to minimize the estimation MSE.

APPENDIX A

DISCUSSION ON CLOSED-FORM SOLUTION

To solve the energy-constrained distributed estimation problem stated in (12), we adopt the Lagrange multiplier method to solve the following equivalent problem:

\[
\begin{align*}
\max & \quad \sum_{k \in S_K} \left( \sigma_k^2 + \frac{W^2}{B_k} \right)^{-1} \\
\text{s.t.} & \quad \sum_{k \in S_K} P_k \leq P_c \\
& \quad P_k = c_k a_k B_k, k \in S_K
\end{align*}
\]

where \( B_k = 2^{b_k} - 1 \). Its Lagrangian \( G \) is given as

\[
G(B_k, \lambda) = \sum_{k \in S_K} \left( \sigma_k^2 + \frac{W^2}{B_k} \right)^{-1} - \lambda \left( \sum_{k \in S_K} c_k a_k B_k - P_c \right)
\]

which leads to the following optimization conditions:

\[
2W^2 \chi^2(1) + \chi^2(1) \sigma_k^2 + \frac{W^2}{B_k} = 0, \forall k \in S_K
\]

and

\[
\sum_{k \in S_K} c_k a_k B_k = P_c
\]
From these equations, the optimal solution for λ and B_k (k ∈ S_k) cannot be obtained in a closed-form.

**APPENDIX B**

**PROOF OF THEOREM 2**

Assume a subset of sensors S_k = \{i_1, \ldots, i_k, \ldots, i_{|S_d|}\} (i_k \in [1, N]) are used, and the quantization bit rate of each active sensor k ∈ S_k is \(b_k^q\) and the corresponding transmission energy allocated is \(E_k\), such that \(\sum_{k \in S_k} E_k = P_c\). Denote this estimation system as C_k, the estimation of θ as \(\hat{\theta}\), and its estimation MSE as \(D_k\), so the objective is to show that \(D_k = E(\hat{\theta} - \theta)^2 > (\sum_{k \in S_k} 1/\sigma_k^2)^{-1}\), where S_k is the optimal subset of active sensors, obtained by our proposed algorithm as shown in (24), (27) such that \(\sum_{k \in S_k} P_k^\text{opt} = P_c\). The basic idea to prove this statement is to construct another quasi-BLUE estimation system, denoted as C_\eta with estimation MSE D_\eta such that \(D_\eta \geq D_k \geq (\sum_{k \in S_k} 1/\sigma_k^2)^{-1}\). The estimation system C_\eta is constructed as follows: only the sensors in S_k are used, i.e., the subset of active sensors is S_\eta = S_k, the quantization bit rate of each active sensor k ∈ S_k is \(b_k^\eta\), and the corresponding transmission energy allocated is \(P_k^\eta\). More specifically

\[
b_k^\eta = \begin{cases} \max \left( b_k^q, b_k^\text{opt} \right), & \text{if } k \in S_{k}\text{opt} \cap S_k \\ b_k^q, & \text{if } k \in S_{k}\text{opt} \setminus S_k \\ 0, & \text{otherwise} \end{cases} \tag{40}
\]

thus

\[
P_k^\eta = \begin{cases} \max \left( P_k^\text{opt}, P_k^\text{opt} \right), & \text{if } k \in S_{k}\text{opt} \cap S_k \\ P_k^\text{opt}, & \text{if } k \in S_{k}\text{opt} \setminus S_k \\ 0, & \text{otherwise} \end{cases} \tag{41}
\]

where \(k \in S_{k}\text{opt} \setminus S_k\) means that \(k \in S_{k}\text{opt}\) but \(k \notin S_k\). It is noted that \((S_{k}\text{opt} \setminus S_k) \cup (S_k \setminus S_{k}\text{opt}) = S_k\).  

1) Show that \(D_\eta > (\sum_{k \in S_{k}\text{opt}} 1/\sigma_k^2)^{-1}\).  
Since in the constructed estimation system C_\eta, only the sensors in the subset S_{k}\text{opt} are active and limited quantization bit rate \(b_k^\eta\) and limited transmission energy \(P_k^\eta\) are used for each sensor k ∈ S_{k}\text{opt}, and \(D_0 = (\sum_{k \in S_{k}\text{opt}} 1/\sigma_k^2)^{-1}\) is the lower bound of the estimation MSE of BLUE estimator using the subset of sensors S_{k}\text{opt} without quantization bit rate and transmission energy constraints, so \(D_\eta > D_0 = (\sum_{k \in S_{k}\text{opt}} 1/\sigma_k^2)^{-1}\).

2) Show that \(D_\eta \leq D_k\).  
Divide \(S_k\) into three disjoint subset \(S_{\xi1}, S_{\xi2}\) and \(S_{\xi3}\) as follows:

\[
S_{\xi1} = \left\{ k : b_k^q \geq b_k^\text{opt}, \text{and } k \in S_{k}\text{opt} \cap S_k \right\} \\
S_{\xi2} = \left\{ k : b_k^q < b_k^\text{opt}, \text{and } k \in S_{k}\text{opt} \cap S_k \right\} \\
S_{\xi3} = S_k \setminus S_{k}\text{opt}. \tag{42}
\]

Similarly, divide \(S_\eta\) into three disjoint subset \(S_{\eta1}, S_{\eta2}\) and \(S_{\eta3}\) as follows:

\[
S_{\eta1} = \left\{ k : b_k^\eta \geq b_k^\text{opt}, \text{and } k \in S_{k}\text{opt} \cap S_k \right\} \\
S_{\eta2} = \left\{ k : b_k^\eta < b_k^\text{opt}, \text{and } k \in S_{k}\text{opt} \cap S_k \right\} \\
S_{\eta3} = S_\eta \setminus S_{k}\text{opt}. \tag{43}
\]

**Proposition 2:** According to the definitions of \(S_{k}\text{opt}, S_k, S_\eta, b_k^q, b_k^\eta, b_k^\text{opt}\), and \(P_k^\eta\) in (27), (40), (41), (42), and (43), it is easy to see that:

1) \(S_{\xi1} \cup S_{\xi2} \cup S_{\xi3} = S_k\) and \(S_{\eta1} \cup S_{\eta2} \cup S_{\eta3} = S_\eta\) and \(S_{\eta1} \subseteq S_{k}\text{opt}\); 
2) \(S_{\eta1} = S_{\xi1}\), \(b_k^\eta = b_k^\text{opt}\) and \(P_k^\eta = P_k^\text{opt}\) for any \(k \in S_{\xi1}\); 
3) \(S_{\eta2} = S_{\xi2}\), \(b_k^\eta = b_k^\text{opt}\) and \(P_k^\eta = P_k^\text{opt}\) for any \(k \in S_{\xi2}\); 
4) \(b_k^\eta = b_k^\text{opt}\) and \(P_k^\eta = P_k^\text{opt}\) for any \(k \in S_{\xi3}\); 
5) \(S_{\eta2} \subseteq S_{k}\text{opt}, S_{\eta3} \subseteq S_{k}\text{opt},\) and \(S_{\xi2} \subseteq S_{k}\text{opt}\); thus for any \(i \in S_{\xi2} \cup S_{\xi3}\) and \(j \in S_{\xi3}, g(\sigma_k^2, b_k^q, a_k, c_k) = g(\sigma_k^2, b_k^\eta, a_k, c_k) = g(\sigma_k^2, \sigma_k^2, a_k, c_k)\) according to (24), (25), and (26). Let \(g_1 = \max_{k \in S_{\xi2} \cup S_{\xi3}} g(\sigma_k^2, b_k^q, a_k, c_k)\) and \(g_2 = \min_{j \in S_{\xi3}} g(\sigma_k^2, b_k^\eta, a_k, c_k)\), then \(g_1 \geq g_2\). 

Let \(D_\xi = 1/D_\eta\), and \(D_\eta = 1/D_\xi\). Expressing \(D_\xi\) and \(D_\eta\) with the concept of the equivalent unit-energy MSE functions as follows:

\[
D_\xi = \sum_{k \in S_{\xi1} \cup S_{\xi2} \cup S_{\xi3}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} \\
D_\eta = \sum_{k \in S_{\eta1} \cup S_{\eta2} \cup S_{\eta3}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} \tag{44}
\]

According to Proposition 2, then

\[
D_\eta - D_\xi = \left( \sum_{k \in S_{\xi1}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} - \sum_{k \in S_{\xi1}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} \right) \]

\[
+ \left( \sum_{k \in S_{\xi2}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} - \sum_{k \in S_{\xi2}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} \right) \\
+ \left( \sum_{k \in S_{\xi3}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} - \sum_{k \in S_{\xi3}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} \right) \\
= \left( \sum_{k \in S_{\xi1}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} - \sum_{k \in S_{\xi1}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} \right) \\
+ \left( \sum_{k \in S_{\xi2}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} - \sum_{k \in S_{\xi2}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} \right) \\
+ \left( \sum_{k \in S_{\xi3}} \frac{P_k^\xi}{g(\sigma_k^2, b_k^q, a_k, c_k)} - \sum_{k \in S_{\xi3}} \frac{P_k^\eta}{g(\sigma_k^2, b_k^\eta, a_k, c_k)} \right) \]
thus the theorem is proved.

From (1) and (2) above, we get

\[
\sum_{k \in S_{\eta}} P_{k}^{\text{opt}} - P_{k}^{\text{c}} \geq \sum_{k \in S_{\eta}} g \left( \frac{1}{\sigma_{k}^{2}} \right) \left( P_{k}^{\text{opt}} - P_{k}^{\text{c}} \right)
\]

From the total energy constraint, we have

\[
\sum_{k \in S_{\eta}} P_{k}^{\text{opt}} = P_{\eta}
\]

Since \( S_{\eta} = S_{\bar{c}} \) and \( P_{k}^{\text{opt}} \geq P_{k}^{\text{c}} \) for any \( k \in S_{\eta} \) as shown in Proposition 2, then

\[
\sum_{k \in S_{\eta}} P_{k}^{\text{opt}} \leq \sum_{k \in S_{\eta}} P_{k}^{\text{c}}
\]

thus,

\[
D_{\eta} - D_{k} \geq \left( \sum_{k \in S_{\eta}} P_{k}^{\text{opt}} - \sum_{k \in S_{\eta}} P_{k}^{\text{c}} \right) \frac{1}{g_{1}} \geq 0
\]

therefore,

\[
D_{k} \geq D_{\eta}
\]

From (1) and (2), we get

\[
D_{k} \geq D_{\eta} > \left( \sum_{k \in S_{\eta}} \frac{1}{\sigma_{k}^{2}} \right)^{-1}
\]

thus the theorem is proved.

REFERENCES


Junlin Li (M’09) received the B.S. degree and M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2001 and 2004, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 2008. He spent summer 2006 as a Research Intern in the Microsoft Research (MSR), Seattle, WA, working on multiparty audio conferencing. Since 2008, he has been with Cisco Systems, Inc., San Jose, CA, working on video coding and processing. His research interests include distributed and statistical signal processing, wireless sensor networks, as well as multimedia processing and communications (including video coding, audio–visual processing and conferencing, 3-D graphic compression and transmission, etc.).

Dr. Li was the recipient of the Outstanding Research Award of the Center for Signal and Image Processing (CSIP) at Georgia Tech.

Ghassan AlRegib (M’07) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 2003. He joined the Faculty of Georgia Tech in 2003 and is currently an Associate Professor in the School of Electrical and Computer Engineering (ECE). His research group is working on projects related to multimedia processing and communications, distributed processing, collaborative systems, immersive communications, and wireless sensor networks.

Dr. AlRegib received the ECE Outstanding Graduate Teaching Award in spring 2001, the Center for Signal and Image Processing Research Award in spring 2003, and the Center for Signal and Image Processing Service Award in spring 2003, all from Georgia Tech. In April 2008, he received the ECE Outstanding Junior Faculty Member Award. He was the General Co-Chair and Co-Founder of the First International Conference on Immersive Telecommunication (IMMERSCOM) held in Italy in November 2007. He is the Steering Committee Co-Chair for the Second International Conference on Immersive Telecommunications to be held in California in May 2009. He was the Chair of the Special Sessions Program at the IEEE International Conference on Image Processing (ICIP) in 2006. He also served as an Associate Editor of the *IEEE Signal Processing Magazine* from 2006 to 2008. He serves as the Area Editor of the *IEEE Signal Processing Magazine* for columns and forums from 2009 to 2011. He is the Editor-in-Chief of the *ICST Transactions on Immersive Telecommunications* and a member of the Editorial Board of the *Wireless Networks (WINET) Journal.*