Power Scheduling of Universal Decentralized Estimation in Sensor Networks

Jin-Jun Xiao, *Student Member, IEEE*, Shuguang Cui, *Student Member, IEEE*, Zhi-Quan Luo, *Senior Member, IEEE*, and Andrea J. Goldsmith, *Fellow, IEEE*

***Abstract*—We consider the optimal power scheduling problem for the decentralized estimation of a noise-corrupted deterministic signal in an inhomogeneous sensor network. Sensor observations are first quantized into discrete messages, then transmitted to a fusion center where a final estimate is generated. Supposing that the sensors use a universal decentralized quantization/estimation scheme and an uncoded quadrature amplitude modulated (QAM) transmission strategy, we determine the optimal quantization and transmit power levels at local sensors so as to minimize the total transmit power, while ensuring a given mean squared error (mse) performance. The proposed power scheduling scheme suggests that the sensors with bad channels or poor observation qualities should decrease their quantization resolutions or simply become inactive in order to save power. For the remaining active sensors, their optimal quantization and transmit power levels are determined jointly by individual channel path losses, local observation noise variance, and the targeted mse performance. Numerical examples show that in inhomogeneous sensing environment, significant energy savings is possible when compared to the uniform quantization strategy.**

***Index Terms*—Distributed estimation, inhomogeneous quantization, power scheduling, sensor networks.**

I. Introduction

RECENT technological advances in wireless sensor networks (WSN) have led to the emergence of small, inexpensive, and low-power sensor devices with limited on-board processing and communication capabilities. When suitably programmed and deployed in large scale, such networked sensors can cooperate to accomplish various high-level tasks. Sensor networks of this type are well-suited for situation awareness applications such as environmental monitoring (air, water, and soil), smart factory instrumentation, military surveillance, precision agriculture, intelligent transportation and space exploration [1]–[3], to name a few.

Manuscript received July 3, 2004; revised April 5, 2005. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada by Grant OPG0090391, by the Canada Research Chair Program, by the National Science Foundation by Grant DMS-0312416, by funds from National Semiconductor, and by the Alfred P. Sloan Foundation. This work was presented in part at the IEEE First Conference on Sensor and Ad Hoc Communications and Networks, Santa Clara, CA, October 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zidong Wang.

J.-J. Xiao and Z.-Q. Luo are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: xiao@ece.umn.edu; luozq@ece.umn.edu).

S. Cui and A. J. Goldsmith are with the Wireless System Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: shuguang@systems.stanford.edu; andrea@systems.stanford.edu).

Digital Object Identifier 10.1109/TSP.2005.861898

Since sensors have only small-size batteries whose replacement can be costly if not impossible, sensor network operations must be energy efficient in order to maximize network lifespan [4]. A main objective of current sensor network research is to design energy-efficient devices and algorithms to support all aspects of network operations. The μAMP's project [5] at MIT and the PicoRadio project [6] at Berkeley focus on energy-constrained radios and their impact on ultralow power sensor nodes and networking. Various energy-efficient algorithms have been proposed for network coverage [7], data gathering [8], and protocols of medium access control [9] and routing [10] (see also the survey paper [1] and the references therein). These references focus on collaborative strategies and cross-layer designs for distributed data collection, processing, and communication in an energy-efficient manner.

A common WSN architecture consists of a fusion center and a number of geographically distributed sensors. Such network architecture can be used to accomplish a joint signal processing task such as decentralized estimation and detection. In this paper, we consider decentralized estimation of an unknown by a set of distributed sensor nodes and a fusion center. The sensors collect real-valued data, perform a local data compression and send the resulting discrete messages to the fusion center, while the latter combines the received messages to produce a final estimate of the observed signal [11]. The problem of decentralized estimation has been extensively studied, first in the context of distributed control [12], and later in tracking [13] and data fusion [14]; see [15] for a more complete reference list. Most of these work assume that the joint distribution of sensor observations is known and that real-valued messages can be sent from sensors to the fusion center without distortion. These assumptions are unrealistic for practical sensor networks since the wireless links between the sensors and the fusion center invariably suffer from adverse channel effects such as attenuation and fading. Moreover, the characterization of probability distributions of sensor observations can be difficult for a large scale sensor network operating in a time-varying environment. In an effort to remove some of these restrictive assumptions, several recent work [11], [15]–[17] has proposed *universal* decentralized estimation schemes (DES) in which the local sensor messages and the final estimation formula are independent of the probability distribution of sensor observations. These universal DESs let each sensor send to the fusion center a short discrete message whose length is determined by the local signal-to-noise ratio (SNR), while guaranteeing a mean squared estimation error (mse) performance that is within a constant factor of that achieved by the centralized best linear unbiased estimator (BLUE). However, the work of

[11], [15]–[17] still assumes that the wireless channels between sensors and the fusion center are ideal, and all messages are received by the fusion center without any distortion.

In this paper, we model the wireless links between sensors and the fusion center as additive white Gaussian noise (AWGN) channels under suitable channel path loss. We adopt a universal decentralized quantization/estimation scheme and an uncoded quadrature amplitude modulated (QAM) transmission strategy for each sensor node. We derive an upper bound on the mse distortion of the universal DESs, and the target mse performance is then ensured by imposing the deduced upper bound to be within the desired distortion level. To minimize the total energy consumption under such a proposed distortion constraint, we optimally choose the number of quantization levels and transmit power for all sensors by taking into account both their local SNRs and individual channel path losses. Our approach is based on combining the recently proposed universal DESs [11], [15]–[17] with the energy models for the coded and uncoded M-QAM transmissions [18]–[20] so as to minimize the total sensor transmission power. We formulate this power scheduling problem as a convex program and derive its optimal solution analytically. The analytical form of the optimal power scheduling scheme suggests that the sensors with bad channels or poor observation qualities should decrease their quantization resolutions or simply become inactive in order to save energy. For the active sensors, their quantization and transmit power levels are determined jointly by individual channel path losses, local observation noise variance, and the targeted MSE performance. Computer simulations show that for an inhomogeneous sensing environment our power scheduling scheme can save substantial amount of sensor energy as compared to the simple uniform power scheduling strategy, thus dramatically increasing the sensor network lifespan.

Our paper is organized as follows. Section II describes the power scheduling problem based on a universal DES and the QAM transmission strategy. In Section III, we optimize the number of quantization bits and the local transmit power level for local sensors based on their local SNR's and channel path losses. Section IV shows some numerical results illustrating the energy savings when compared to the uniform power scheduling scheme, and Section V gives some concluding remarks.

II. Decentralized Estimation in WSNs

Consider a set of K distributed sensors, each making observations on a deterministic source signal θ. The observations are corrupted by additive noises and are described by

$$x_k = \theta + n_k, \quad k = 1, 2, \ldots, K. \tag{1}$$

We assume noises $\{n_k : k = 1, 2, \ldots, K\}$ are zero mean, spatially uncorrelated with variance σ_k^2, but otherwise unknown. By a suitable linear scaling, the above data model (1) is equivalent to the one where sensors observe θ with different attenuation, namely, $x_k = h_k\theta + n_k$. Indeed, if we let $x_k'(t) = x_k/h_k$ and $n_k' = n_k/h_k$, then $x_k' = \theta + n_k'$ which is identical to (1) with equivalent noise variances ${\sigma_k'}^2 = \sigma_k^2/h_k^2$.

Suppose sensors and the fusion center wish to jointly estimate θ based on the sensor observations $\{x_k\}$. We will use mse to measure the quality of an estimator. If the fusion center has the knowledge of sensor noise variances and the sensors can send the observations $\{x_k : k = 1, 2 \ldots, K\}$ to the fusion center without distortion, then the fusion center can simply perform the linear combination of sensor observations to recover θ. This is the concept of the centralized BLUE [21]

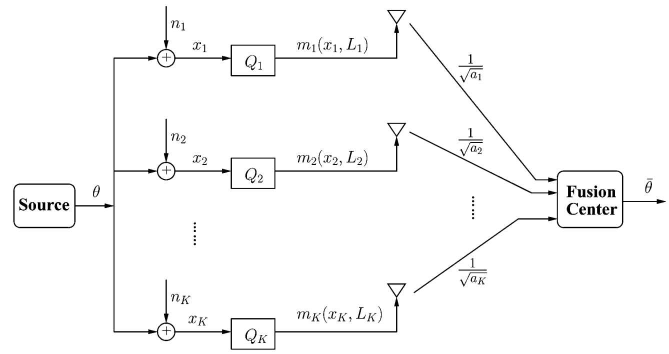


Fig. 1. A decentralized estimation scheme by a WSN with a fusion center.

$$\bar{\theta}_K = \bar{\theta}_K(x_1, x_2, \ldots, x_K) = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2}\right)^{-1} \sum_{k=1}^{K} \frac{x_k}{\sigma_k^2}. \tag{2}$$

A simple calculation shows that this estimator has an mse of

$$D_{\text{BLUE}} := \mathrm{E}\left(|\bar{\theta}_K - \theta|^2\right) = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2}\right)^{-1}. \tag{3}$$

The above scheme is only applicable in a centralized estimation situation where observations x_k's are either centrally located, or can be transmitted to a central location without distortion. Neither of these requirements is realistic in a WSN where the data are spatially distributed and the communication links between the fusion center and sensor nodes are power constrained and subject to the usual pathloss. As a result, we propose the following decentralized estimation scheme (see Fig. 1). First, each sensor performs a local quantization of x_k and generates a discrete message $m_k(x_k, L_k)$ of L_k bits, where the quantizer $Q_k : x_k \mapsto m_k(x_k, L_k)$ is to be designed. Each discrete message is then transmitted to the fusion center through a separate AWGN channel with a known pathloss coefficient, and the fusion center generates the final estimate $\bar{\theta}$ based on the received signals. The independent AWGN channels between sensors and the fusion center can be realized by any of the well-known multiaccess techniques such as TDMA, FDMA, or CDMA.

The main purpose of this paper is to investigate adaptive quantization of sensor observations and its impact on energy saving. More specifically, we will adopt certain specific quantization and transmission strategies for each sensor, and decide the optimal message lengths L_k and its corresponding transmit power levels. Due to the lack of noise pdf knowledge, we take $Q_k : x_k \mapsto m_k(x_k, L_k)$ to be a uniform randomized quantizer [15]. This quantization scheme works universally for all noise pdf, and it generates unbiased message functions. In addition, we adopt uncoded QAM transmission [19], [20] of the quantized bits. The estimator at the fusion center is a generalized version of the BLUE estimator (2) which weighs the message functions linearly with weights decided by both the observation noise and the quantization noise. We will optimally choose quantization and transmit power levels at local sensors so as to minimize the total transmit power, while ensuring a targeted mse performance.

A. Probabilistic Quantization of a Bounded Random Variable

Suppose $[-W, W]$ is the signal range that sensors can observe, that is, $x = \theta + n \in [-W, W]$, where W is a known parameter decided by the sensors' dynamic range, and θ is the unknown signal to be estimated. The noise n has zero mean and variance σ^2, but is otherwise unknown. Suppose we want to quantize x into L bits regardless of the probability distribution of x. This can be achieved by uniformly dividing $[-W, W]$ into intervals of length $\Delta = (2W)/(2^L - 1)$, and rounding x to the neighboring endpoints of these small intervals in a probabilistic manner (see Fig. 2). More specifically, suppose $i\Delta \le x < (i+1)\Delta$ where $-2^{L-1} \le i \le 2^{L-1}$, then x is quantized to $m(x, L)$ according to

$$\begin{aligned} &\mathrm{P}\{m(x,L) = i\Delta\} = 1 - r \\ &\mathrm{P}\{m(x,L) = (i+1)\Delta\} = r \end{aligned}$$

with $r \equiv (x - i\Delta)/(\Delta) \in [0, 1]$. Notice that r is chosen so that the quantization $m(x, L)$ is unbiased, namely, $\mathrm{E}_Q(x) = x$, where E_Q denotes expectation taken with respect to the probabilistic quantization noise. It is easy to see that m assumes 2^L discrete values $\{i\Delta : i = -(2^{L-1} - 1), \ldots, -1, 0, 1, \ldots, 2^{L-1} - 1\}$ which can be represented in L bits. The quantization noise $v(x, L) := m(x, L) - x$ can be viewed as a Bernoulli random variable taking values at $r\Delta$ and $(r-1)\Delta$, i.e.,

$$\begin{aligned} &\mathrm{P}\{v(x,L) = r\Delta\} = 1 - r \\ &\mathrm{P}\{v(x,L) = (r-1)\Delta\} = r. \end{aligned}$$

In terms of the quantization noise $v(x, L), m(x, L)$ can be written as

$$m(x, L) = \theta + n + v(x, L) \tag{4}$$

where n and $v(x, L)$ are independent. Next lemma, whose proof can be found in [15], shows that this message function is an unbiased estimator of θ with a variance approaching σ^2 at an exponential rate as L increases.

Lemma 1: Let $m(x, L)$ be an L-bit quantization of $x \in [-W, W]$ as defined. Then m is an unbiased estimator of θ and

$$\mathrm{E}\left(|m(x,L) - \theta|^2\right) \le \frac{W^2}{(2^L - 1)^2} + \sigma^2 \quad \text{for all } L \ge 1$$

where the expectation is taken with respect to both the sensor observation noise and quantization noise.

Now suppose the bit budget for sensor k is L_k for $1 \le k \le K$. With the strategy described above, we design local independent quantizers $Q_k : x_k \mapsto m_k(x_k, L_k)$, where $m_k(x_k, L_k)$ is a discrete message of L_k bits. According to (4), m_k can be represented as

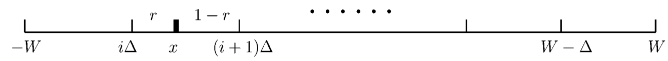


Fig. 2. A probabilistic uniform quantization scheme.

$$m_k(x_k, L_k) = \theta + n_k + v_k \tag{5}$$

where the quantization noise $v_k = v_k(x_k, L_k)$ across sensors are independent because quantizations are performed locally at each sensor without coordination. By Lemma 1, we have $\mathrm{E}(m_k) = \theta$ and

$$\mathrm{Var}(m_k) \le \frac{W^2}{(2^{L_k} - 1)^2} + \sigma_k^2 = \delta_k^2 + \sigma_k^2 \tag{6}$$

where

$$\delta_k^2 = \frac{W^2}{(2^{L_k} - 1)^2}$$

denotes an upper bound of the quantization noise variance. These message functions are then transmitted to the fusion center where they are combined to generate a final estimate of θ.

B. Fusion Function: Quasi-BLUE

Our goal is to construct a linear estimator of θ from $\{m_1, m_2, \ldots, m_K\}$ such that the mse is minimized. Recall from the property of BLUE (2) that the optimal weight of m_k is proportional to $\mathrm{Var}(m_k)^{-1}$. Therefore, according to (6) we can set the weight for m_k as $1/(\sigma_k^2 + \delta_k^2)$, giving rise to the following estimator

$$\begin{aligned} \bar{\theta}_K &= \bar{\theta}_K(m_1, m_2, \ldots, m_K) \\ &= \left(\sum_{k=1}^K \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \sum_{k=1}^K \frac{m_k}{\sigma_k^2 + \delta_k^2}. \end{aligned} \tag{7}$$

Notice that $\bar{\theta}_K$ is an unbiased estimator of θ since every m_k is an unbiased quantization of x_k. Moreover, it has an mse

$$\begin{aligned} D &\equiv \mathrm{E}\left(|\bar{\theta}_K - \theta|^2\right) \\ &= \mathrm{E}\left[\left(\left(\sum_{k=1}^K \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \sum_{k=1}^K \frac{m_k - \theta}{\sigma_k^2 + \delta_k^2}\right)^2\right] \\ &= \left(\sum_{k=1}^K \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-2} \sum_{k=1}^K \frac{\mathrm{E}(|m_k - \theta|^2)}{(\sigma_k^2 + \delta_k^2)^2} \\ &\le \left(\sum_{k=1}^K \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \end{aligned}$$

where the third step follows from the fact that $m_k - \theta = n_k + v_k$ and $m_j - \theta = n_j + v_j$ [see (5)] are uncorrelated for $j \ne k$, and $\mathrm{E}(m_k) = \theta$ for all k as shown in Lemma 1, while the final inequality results from the application of (6) to each term in the sum.

So far we have assumed that the sensor messages $\{m_k : k = 1, 2, \ldots, K\}$ are perfectly received by the fusion center. When each m_k is transmitted to the fusion center through a nonperfect channel with finite power, bit error occurs, which further impacts on the estimation accuracy at the fusion center. We will model each sensor's channel to the fusion center as a memoryless binary symmetric channel, the following lemma (whose proof is given in Appendix A) analyzes the contribution of bit error rate (BER) to the mse performance.

Lemma 2 (MSE due to BER): Suppose the probability of bit error achieved by sensor k is p_b^k, and m'_k is the decoded version of m_k at the receiver. Let D' denote the mse achieved by the estimator (7) based on the received messages $\{m'_1, m'_2, \ldots, m'_K\}$. If $\{p_b^k\}$ satisfy (for some $p_0 > 0$)

$$p_0 \geq \frac{8W}{\sigma_k}\sqrt{\frac{Kp_b^k}{3}}, \quad \text{for all } k \tag{8}$$

then

$$D' \leq (1+p_0)^2 \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1}.$$

Lemma 2 shows that the actual achieved mse is at most a constant factor away from what is achievable with perfect sensor channels, provided that each sensor's BER is bounded above (8).

III. Power Scheduling

We assume that the channel between each sensor and the fusion center is corrupted with additive white Gaussian noise whose double-sided power spectrum density is given by $N_0/2$. In addition, the channel between sensor k and the fusion center experiences a pathloss proportional to $a_k = d_k^\alpha$, where d_k is the transmission distance and α is the passloss exponent. We further assume that sensors follow a time division multiple access scheme to send data to the fusion center. If sensor k sends L_k bits with quadrature amplitude modulation with constellation size 2^{L_k} at a bit error probability p_b^k, then the total amount of required transmission energy [19], [20] is given by

$$E_k = c_k a_k \left(\ln \frac{2}{p_b^k}\right)(2^{L_k} - 1) \tag{9}$$

with

$$c_k = 2N_f N_0 G_d \tag{10}$$

where N_f is the receiver noise figure, N_0 is the single-sided thermal noise spectral density, and G_d is a system constant defined in the same way as in [19], [20].

In addition, we suppose that sensor k first samples the observed signal at rate B_s, and then quantizes each sample to L_k bits. The transmission symbol rate is equal to the sampling rate B_s, and we take the QAM constellation size to be 2^{L_k} in order to minimize the delay. Neglecting circuit power, the average transmit power consumption of node k is $P_k = B_s E_k$. Our primary goal is to minimize the total power consumption while meeting a target overall mse performance. A secondary goal is to maintain fairness in the power scheduling among sensors. For any $q \geq 1$, the L^q-norm of the power vector $\mathbf{P} = (P_1, P_2, \ldots, P_K)$ is defined as

$$\|\mathbf{P}\|_q = \left(\sum_{k=1}^{K} P_k^q\right)^{1/q}.$$

Minimizing the total power consumption implies minimizing the L^1-norm of $\mathbf{P}$, while minimizing the maximum of the individual power values implies minimizing the L^∞-norm of $\mathbf{P}$. In our paper, we make a design compromise by choosing to minimize the L^2-norm of $\mathbf{P} : \|\mathbf{P}\|_2 = \sqrt{\sum_{k=1}^{K} P_k^2}$. In this way, we can penalize the large terms in the power vector while still keeping the total power consumption reasonably low. The L^2-norm also allows an easier analytical treatment of the power scheduling problem, and helps us to obtain a closed form solution.

With the goal of minimizing the L^2-norm of the sensor power vector, we obtain the following formulation of the power scheduling problem

$$\begin{aligned} &\min \|\mathbf{P}\|_2 \\ &\text{s.t. } D' \leq D_0 \end{aligned} \tag{11}$$

where D' and D_0 are the achieved and target mse performance respectively.

Assume that the constants $\{c_k\}$ [cf. (10)] and B_s are the same for all nodes, and that the same target BER is chosen for all sensors. To ensure that $D' \leq D_0$, we use the upper bound of D' deduced in Lemma 2. In addition, applying (9), we can reformulate (11) as

$$\begin{aligned} &\min_{L_k \in \mathbb{Z}} \sum_{k=1}^{K} a_k^2 (2^{L_k} - 1)^2 \\ &\text{s.t. } D' = (1+p_0)^2 \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \leq D_0 \\ &\delta_k^2 = \frac{W^2}{(2^{L_k} - 1)^2}, k = 1, \ldots, K \\ &L_k \geq 0, \quad k = 1, \ldots, K \end{aligned} \tag{12}$$

where δ_k^2 is the quantization noise variance, $2W$ is the range for the sensed signal, L_k is an integer signifying the number of quantized bits per sample at sensor k. To facilitate the subsequent analysis, we will relax the integer condition $L_k \in \mathbb{Z}$ to $L_k \in \mathbb{R}$. However, even with this relaxation, the above problem remains nonconvex in L_k. Fortunately, we show next that the problem (12) can be reformulated as a convex one by performing a change of variables.

Let us define $r_k = 1/(\sigma_k^2 + \delta_k^2), k = 1, \ldots, K$. We see that L_k and δ_k^2 can both be replaced by functions of r_k as

$$\delta_k^2 = \frac{1}{r_k} - \sigma_k^2, \quad (2^{L_k} - 1)^2 = \frac{W^2}{\delta_k^2} = \frac{W^2}{\frac{1}{r_k} - \sigma_k^2} = \frac{W^2 r_k}{1 - r_k \sigma_k^2}.$$

Therefore, the problem (12) can be transformed into a problem with variables $\mathbf{r} = (r_1, r_2, \ldots r_K)$ shown as follows:

$$\min_{r_k \in \mathbb{R}} \sum_{k=1}^{K} a_k^2 \left(\frac{W^2 r_k}{1 - r_k \sigma_k^2} \right)$$
$$\text{s.t.} \sum_{k=1}^{K} r_k \geq \frac{1}{D_0'}$$
$$0 \leq r_k < \frac{1}{\sigma_k^2}, \quad k = 1, \ldots, K \tag{13}$$

where $D_0' = D_0/(1+p_0)^2$. The upper limit on r_k is given by the fact that $r_k = 1/(\sigma_k^2 + \delta_k^2)$ and $\delta_k^2 \geq 0$. It is easy to check that the optimization problem (13) is convex with respect to the new variables $\mathbf{r} = (r_1, r_2, \ldots r_K)$.

Though the convex problem (13) is efficiently solvable by numerical methods, we show below that it can actually be solved analytically. In particular, we can write the Lagrangian function G as

$$G(\mathbf{r}, \lambda_0, \mu) = \sum_{k=1}^{K} a_k^2 \left(\frac{W^2 r_k}{1 - r_k \sigma_k^2} \right) + \lambda_0 \left(\frac{1}{D_0'} - \sum_{k=1}^{K} r_k \right) - \sum_{k=1}^{K} \mu_k r_k$$

which gives the following Karush-Kuhn-Tucker (KKT) conditions [22]

$$\frac{W^2 a_k^2}{(1 - r_k \sigma_k^2)^2} - \lambda_0 - \mu_k = 0, \quad k = 1, 2, \ldots K \tag{14}$$
$$\lambda_0 \left(\sum_{k=1}^{K} r_k - \frac{1}{D_0'} \right) = 0$$
$$\sum_{k=1}^{K} r_k - \frac{1}{D_0'} \geq 0 \tag{15}$$
$$\lambda_0 \geq 0$$
$$\mu_k r_k = 0, \quad k = 1, 2, \ldots K \tag{16}$$
$$\mu_k \geq 0, \quad k = 1, 2, \ldots K$$
$$0 \leq r_k < \frac{1}{\sigma_k^2}, \quad k = 1, 2, \ldots K.$$

We notice that, if $\lambda_0 = 0$, then (14) would imply $u_k > 0$ for all k; Combining this with (16) shows $r_k = 0$ for all k, which violates the condition (15). Thus, we must have $\lambda_0 > 0$, which further implies $\sum_{k=1}^{K} r_k = \frac{1}{D_0'}$.

We proceed to solve the above KKT system. Without loss of generality, we assume that $a_1 \leq a_2 \leq \ldots \leq a_K$, and define

$$f(M) = a_M \left(\sum_{k=1}^{M} \frac{a_k}{\sigma_k^2} \right)^{-1} \left(\sum_{k=1}^{M} \frac{1}{\sigma_k^2} - \frac{1}{D_0'} \right) \quad \text{for } 1 \leq M \leq K. \tag{17}$$

Let K_1 be such that

$$f(K_1) < 1 \quad \text{and} \quad f(K_1 + 1) \geq 1. \tag{18}$$

We show in Appendix B that such K_1 is unique unless $f(M) < 1$ for all $1 \leq M \leq K$, in which case we take $K_1 = K$. Simple manipulations of the KKT system lead to

$$\lambda_0 = \left(\frac{W \sum_{k=1}^{K_1} \frac{a_k}{\sigma_k^2}}{\sum_{k=1}^{K_1} \frac{1}{\sigma_k^2} - \frac{1}{D_0'}} \right)^2$$

and

$$r_k^{\text{opt}} = \frac{1}{\sigma_k^2} \left(1 - \frac{a_k W}{\sqrt{\lambda_0}} \right)^+, \quad k = 1, 2, \ldots, K \tag{19}$$

where $(x)^+$ equals 0 when $x < 0$, and otherwise is equal to x.

By definition, we have

$$r_k = \frac{1}{\sigma_k^2 + \frac{W^2}{(2^{L_k} - 1)^2}}.$$

Thus, (19) implies that the optimal value of L_k is

$$L_k^{\text{opt}} = \log_2 \left(1 + \sqrt{\frac{W^2}{\frac{1}{r_k^{\text{opt}}} - \sigma_k^2}} \right) = \begin{cases} 0 & \text{for } k \geq K_1 + 1 \\ \log \left(1 + \frac{W}{\sigma_k} \sqrt{\frac{\eta_0}{a_k} - 1} \right) & \text{for } k \leq K_1 \end{cases} \tag{20}$$

where

$$\eta_0 = \frac{\sqrt{\lambda_0}}{W} = \left(\sum_{k=1}^{K_1} \frac{1}{\sigma_k^2} - \frac{1}{D_0'} \right)^{-1} \sum_{k=1}^{K_1} \frac{a_k}{\sigma_k^2}.$$

The transmission power for node k is given as

$$P_k = c_k B_s \left(\ln \frac{2}{p_b} \right) \frac{W a_k}{\sigma_k} \sqrt{\left(\frac{\eta_0}{a_k} - 1 \right)^+}.$$

Notice that when $\frac{\eta_0}{a_k} \leq 1$, or $a_k \geq \eta_0$, we have $L_k = 0$, and therefore $P_k = 0$. Since a_k is the channel loss factor (inverse of the channel gain), this implies that when the channel quality for sensor k is worse than the threshold η_0, we should discard its observation in order to save energy. In the simulations, we can see that in some cases, a large number of sensors with bad channel qualities or poor observations shut off (see Figs. 3 and 4). The message length formula in (20) is intuitively appealing as it indicates that the message length should be proportional to the logarithm of local SNR scaled by channel path gain. This is in the same spirit as the message length formula when the channel is distortionless; see [17]. For example, if the channel quality is very poor (indicated by $a_k > \eta_0$), the message length formula (20) may still assign $L_k = 0$, even if the sensor has good quality observation (i.e., small σ_k^2). This again conforms with intuition.

To implement the described power scheduling scheme, the central node (fusion center) needs to broadcast the threshold η_0 whose value is based on the collected network information. Local sensors then decide the quantization message length L_k according to its own local information (σ_k and a_k) and η_0 [cf. (20)]. In particular, sensors with $a_k \geq \eta_0$ should be inactive.

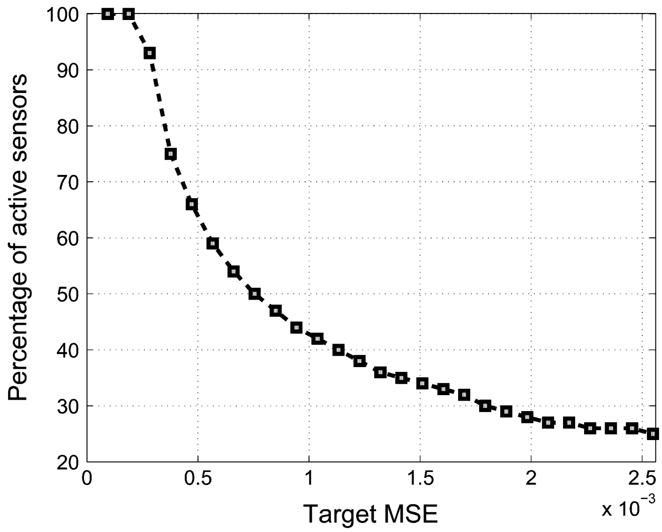


Fig. 3. Number of active sensors decreases as the target mse increases.

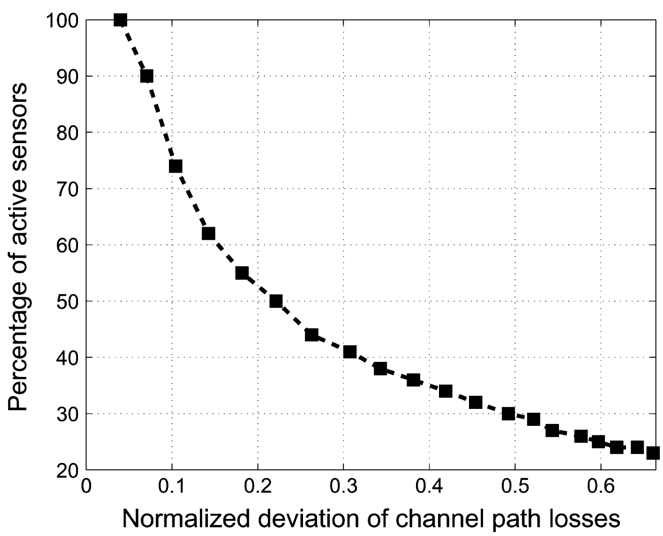


Fig. 4. Number of active sensors decreases as the channel path losses become more heterogeneous.

IV. Numerical Results

According to (18), the number of active sensors K_1 depends on the target mse D_0, the distribution of channel path losses as well as sensor noise variances. This dependence will be illustrated by numerical results in this section. In all simulations, the total number of sensors $K = 1000$. We generate sensor noise variances σ_k^2 according to the distribution $0.1 + a^2\chi^2(1)$, where $\chi^2(1)$ is the Chi-square distribution of degree 1. Also, the channel path loss coefficients $a_k = d_k^\alpha$ are generated by a uniformly distributed $d_k \in [1, 10]$ (the distance between sensor k and the fusion center), and some passloss exponent α. With different choices of a and α, we can generate $\{\sigma_k^2, k = 1, 2, \ldots, K\}$ and $\{a_k, k = 1, 2, \ldots, K\}$ to model different sensing environments.

For a positive random variable R, we define

$$\text{normalized deviation of } R = \frac{\sqrt{\text{Var}(R)}}{\text{E}(R)}$$

which will be used as a measure of the absolute heterogeneity of R. The purpose of our simulations is to observe how the percentage of active sensors, and amount of energy saving vary over the heterogeneity of sensor noise variances or channel path losses.

According to (17), when target D_0 increases or when channel path losses become more diverse, more sensors will become inactive. Such inactive sensors neither perform quantization nor transmit any messages to the fusion center in order to conserve energy. Fig. 3 plots the percentage of active sensors versus some target mse $D_0 \geq D_{\text{BLUE}}$ when the distribution of channel path losses and sensor noise variance are kept fixed by choosing $\alpha = 3.5, a = 1$. In Fig. 4, the percentage of active sensors versus the normalized deviation of channel path losses is plotted by keeping distribution of local sensor noise variances fixed choosing $a = 1$, and the target mse $D_0 = 2D_{\text{BLUE}}$ where D_{BLUE} is the mse of the centralized BLUE defined in (3). Similar curve can be obtained if we plot the percentage of active sensors versus the normalized deviation of sensor noise variances by keeping the target D_0 and channel path losses fixed.

Since L_k can only take integer values, the original problem (12) is actually a nonconvex integer programming problem. We have relaxed L_k to take real values to make the problem convex. Therefore, the optimal power consumption obtained by allowing L_k to take on real values is a lower bound (denoted as P_-) of the actual power consumption. If we round the L_k^{opt} up to the closest integer $\bar{L}_k^{\text{opt}}$ that is larger than L_k^{opt}, we can obtain an upper bound (denoted as P_+) of the actual power consumption.

However, even the approximate solution $\bar{L}_k^{\text{opt}}$ can achieve significant energy saving compared with the following two strategies:

- *Uniform quantization*: each sensor quantizes the observation into the minimal same number of bits to achieve the target mse distortion D_0;
- *Uniform power scheduling*: each sensor schedules the minimal identical amount of power to achieve the target mse distortion D_0.

Notice that the uniform power scheduling can also be obtained by solving (11) with objective function replaced by $\|\mathbf{P}\|_\infty = \max_k\{P_k\}$. The comparison is shown in Figs. 5 and 6. In Fig. 5, we suppose there is no fading and all channel path losses $\{a_k\}$ take the same value 1, and observe how the percentage of energy savings varies over the standard deviation of observation noise variances $\{\sigma_k^2\}$. In such a case, the uniform quantization strategy coincides with the uniform power scheduling strategy due to the same path losses. In Fig. 6, all $\{\sigma_k^2\}$ take the same value 0.1 as the normalized deviation of $\{a_k\}$ increases. We conclude that when compared to either uniform power scheduling or uniform quantization scheme, the amount of energy savings of our proposed strategy becomes more significant when either the local noise variances or channel path losses become more heterogeneous.

V. Concluding Remarks

In this paper we have derived a power scheduling strategy for decentralized estimation in wireless sensor networks, whereby sensor nodes are assumed to adopt a uniform randomized

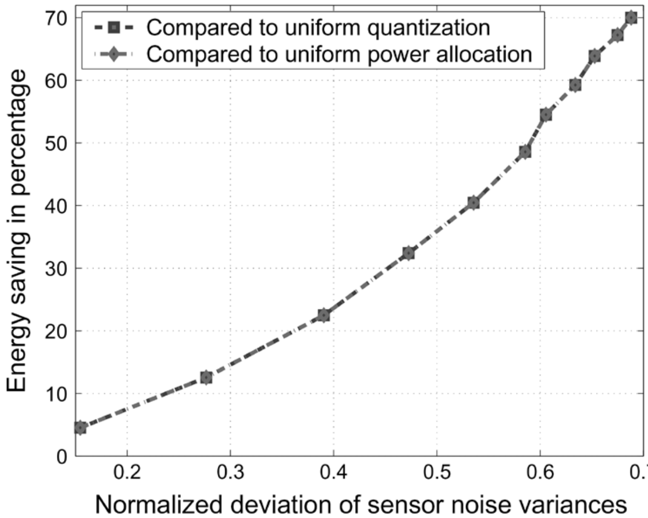


Fig. 5. Percentage of energy saving increases when sensor noise variances become more heterogeneous.

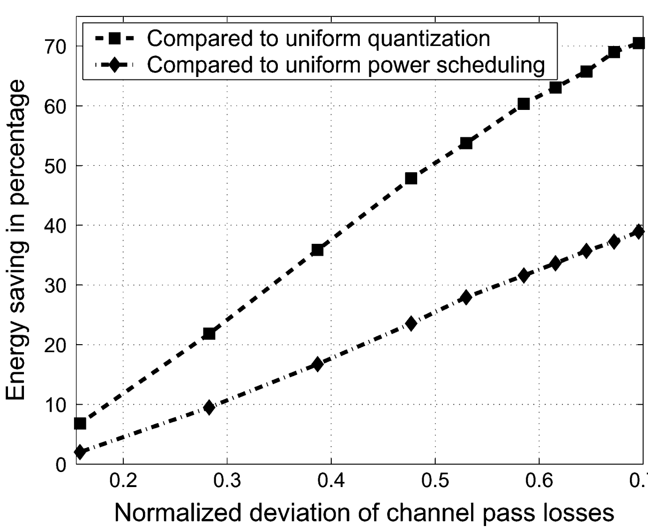


Fig. 6. Percentage of energy saving increases when channel path losses become more heterogeneous.

quantization scheme as well as an uncoded QAM transmission scheme. Our design minimizes the total energy consumption subject to the constraint that the worst distortion is within a given level. We show that the optimal quantization level and transmission power for each sensor can be determined jointly in terms of the channel path losses and the local observation noise levels. When the channel quality is below a (computable) threshold, the corresponding sensor will be completely shut off to save energy. In contrast, when the channel quality is good and the observation noise is low, the corresponding sensor will be active: it will first quantize its observation to a specific (computable) number of bits and then transmit them to the fusion center using an appropriate amount of transmission power. Numerical examples show that in an inhomogeneous sensing environment, our design can achieve a significant amount of energy saving when compared to the uniform quantization strategy in which each sensor generates the same number of bits regardless of its channel quality.

To obtain the desired quantization and transmit power levels, we have assumed in this paper that the fusion center knows $\{(\sigma_k^2, a_k) : k = 1, 2, \ldots, K\}$. This assumption is reasonable in cases where the network condition and the signal being estimated change slowly in a quasistatic manner. Thus, once $\{(\sigma_k^2, a_k) : k = 1, 2, \ldots, K\}$ are acquired by the fusion center, they can be used for a reasonably long period of time. Also, our approach can be generalized to the estimation of a memoryless discrete-time random process $\theta(t)$. Due to the temporal memoryless property of the source and sensor observations, we can impose sample-by-sample estimation without significant estimation performance loss, but obtain important features such as easy implementation and no coding and estimation delay.

For future work, we wish to extend the current work to the vector signal model: $\mathbf{x}_k = \mathbf{H}_k\boldsymbol{\theta} + \mathbf{n}_k$, for all k, where $\boldsymbol{\theta}$ is the vector of unknown parameters, $\mathbf{x}_k$ is the vector of sensor observations. Our initial investigation shows the corresponding energy minimization problem becomes nonconvex, which makes the optimal power scheduling difficult to compute. We also plan to explore tighter universal source coding bounds and other energy-efficient coded transmission schemes for decentralized estimation in wireless sensor networks. It is likely that joint source and channel coding approaches can achieve higher energy efficiency than the strategy considered in this paper. Moreover, designing a completely distributed algorithm for optimal power scheduling which does not require local sensor information $\{(\sigma_k, a_k) : k = 1, 2, \ldots, K\}$ at the fusion center is also part of our future work.

Appendix

A. Proof of Lemma 2

The quantized message m_k has L_k bits and can be written as

$$m_k = \left(\sum_{i=1}^{L_k} b_i 2^{L_k - i} - 2^{L_k - 1}\right)\Delta_k$$

where $\Delta_k = (2W)/(2^{L_k} - 1)$, b_1 is the first most significant bit (MSB) of m_k, and b_2 is the second MSB, etc.

Suppose the BER of sensor k is p_b^k, and m'_k are the decoded version of m_k at the fusion center. Let $\delta_k^2 = \Delta_k^2/4$, then the estimator based on the received messages $\{m'_1, m'_2, \ldots, m'_K\}$ is

$$\bar{\theta}'_K = \bar{\theta}'_K(m'_1, m'_2, \ldots, m'_K) = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \sum_{k=1}^{K} \frac{m'_k}{\sigma_k^2 + \delta_k^2}.$$

Notice that we may have $L_k = 0$ for some k. In this case, $\delta_k^2 = (W^2)/((2^{L_k} - 1)^2) = +\infty$, indicating that the corresponding weight of such message is $(1)/(\sigma_k^2 + \delta_k^2) = 0$. Therefore, such sensors do not participate the estimation in order to save energy, and will have no contribution in the final mse.

We now analyze the mse of $\bar{\theta}'_K$ by taking into account the bit error caused by the channel. To transmit m_k, the binary bits $\{b_1, b_2, \ldots, b_{L_k}\}$ must be sent. Suppose the decoded bits at the fusion center are $\{b'_1, b'_2, \ldots, b'_{L_k}\}$. Let A_ℓ denote the event that the first bit decoded incorrectly is b_ℓ, i.e., $b'_\ell \neq b_\ell$, but $b'_i = b_i$

for all $1 \leq i \leq \ell - 1$, and B denote the event that all bits are decoded correctly. Then for any $1 \leq \ell \leq L_k$, we have

$$\begin{aligned} \mathrm{P}(A_\ell) &= \mathrm{P}\{b'_1 = b_1, \ldots, b'_{\ell-1} = b_{\ell-1}; b'_\ell \neq b_\ell\} \\ &= \left(1 - p_b^k\right)^{\ell-1} p_b^k \\ \mathrm{P}(B) &= \mathrm{P}\{b'_1 = b_1, \ldots, b'_{L_k-1} = b_{L_k-1}, b'_{L_k} = b_{L_k}\} \\ &= \left(1 - p_b^k\right)^{L_k}. \end{aligned} \tag{21}$$

When the event A_ℓ happens, it holds that

$$\begin{aligned} |m_k - m'_k| &= \left| \sum_{i=1}^{L_k} b_i 2^{L_k - i} - \sum_{i=1}^{L_k} b'_i 2^{L_k - i} \right| \Delta_k \\ &\leq \sum_{i=\ell}^{L_k} |b_i - b'_i| 2^{L_k - i} \Delta_k \\ &\leq \sum_{i=\ell}^{L_k} 2^{L_k - i} \Delta_k = 2^{L_k - \ell} \Delta_k \sum_{i=0}^{L_k - \ell} 2^{-i} \\ &\leq 2^{L_k - \ell + 1} \Delta_k \end{aligned} \tag{22}$$

where the second and the third steps follow from $|b_i - b'_i| = 0$ for $1 \leq i \leq \ell - 1$, and $|b_i - b'_i| \leq 1$ for $\ell \leq i \leq L_k$. Thus, we can obtain from (21) and (22) that

$$\begin{aligned} \mathrm{E}\left(|m_k - m'_k| \,|\, m_k\right) &= \sum_{\ell=1}^{L_k} \mathrm{P}(A_\ell) \mathrm{E}(|m_k - m'_k| \,|\, A_\ell) \\ &\leq \sum_{\ell=1}^{L_k} \left(1 - p_b^k\right)^{\ell-1} p_b^k 2^{L_k - \ell + 1} \Delta_k \\ &= 2^{L_k} p_b^k \Delta_k \sum_{\ell=1}^{L_k} \left(\frac{1 - p_b^k}{2} \right)^{\ell-1} \\ &< 2^{L_k} p_b^k \frac{1}{1 - \frac{1 - p_b^k}{2}} \Delta_k \\ &< 2^{L_k + 1} p_b^k \Delta_k. \end{aligned}$$

Similarly, we have

$$\begin{aligned} \mathrm{E}(|m_k - m'_k|^2 \,|\, m_k) &= \sum_{\ell=1}^{L_k} \mathrm{P}(A_\ell) \mathrm{E}(|m_k - m'_k|^2 \,|\, A_\ell) \\ &\leq \sum_{\ell=1}^{L_k} \left(1 - p_b^k\right)^{\ell-1} p_b^k 2^{2L_k - 2\ell + 2} \Delta_k^2 \\ &= 2^{2L_k} p_b^k \Delta_k^2 \sum_{\ell=1}^{L_k} \left(\frac{1 - p_b^k}{4} \right)^{i-1} \\ &< 2^{2L_k} p_b^k \frac{1}{1 - \frac{1 - p_b^k}{4}} \Delta_k^2 \\ &< \frac{2^{2L_k + 2} p_b^k}{3} \Delta_k^2. \end{aligned}$$

Since the above bounds are independent of m_k, we obtain the following estimates for the unconditioned means

$$\begin{aligned} \mathrm{E}(|m_k - m'_k|) &< 2^{L_k + 1} p_b^k \Delta_k \\ \mathrm{E}(|m_k - m'_k|^2) &< \frac{2^{2L_k + 2} p_b^k}{3} \Delta_k^2. \end{aligned} \tag{23}$$

Let D be the mse of the centralized BLUE $\bar{\theta}_K$ [defined in (7)]

$$D = \mathrm{E}(|\bar{\theta}_K - \theta|^2) = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1}. \tag{24}$$

Now we calculate the part of mse due to the channel distortion. We ignore the terms in the estimators $\bar{\theta}_K$ and $\bar{\theta}'_K$ when $L_k = 0$ because the corresponding terms vanish in both estimators

$$\begin{aligned} &\mathrm{E}(|\bar{\theta}'_K - \bar{\theta}_K|^2) \\ &= \mathrm{E}\left(\left(\sum_{k=1, L_k \geq 1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1} \sum_{k=1, L_k \geq 1}^{K} \frac{m_k - m'_k}{\sigma_k^2 + \delta_k^2} \right)^2 \\ &\leq \left(\sum_{k=1, L_k \geq 1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-2} \sum_{k=1, L_k \geq 1}^{K} \frac{K \mathrm{E}(|m_k - m'_k|^2)}{\left(\sigma_k^2 + \delta_k^2\right)^2} \\ &\leq \left(\sum_{k=1, L_k \geq 1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-2} \sum_{k=1, L_k \geq 1}^{K} \frac{K \frac{2^{2L_k+2} p_b^k}{3} \Delta_k^2}{\left(\sigma_k^2 + \delta_k^2\right)^2} \\ &= \left(\sum_{k=1, L_k \geq 1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-2} \sum_{k=1, L_k \geq 1}^{K} \frac{\frac{K}{3} 2^{2L_k+2} p_b^k \frac{4W^2}{(2^{L_k} - 1)^2}}{\left(\sigma_k^2 + \delta_k^2\right)^2} \\ &\leq \left(\sum_{k=1, L_k \geq 1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-2} \sum_{k=1, L_k \geq 1}^{K} \frac{\frac{64K}{3} p_b^k W^2}{\left(\sigma_k^2 + \delta_k^2\right)^2}. \end{aligned}$$

In the above derivation, we have used (23) in step 3. A factor of K is introduced in step 2 since error term $m_k - m'_k$ does not have zero mean in general.

Choose $p_0 > 0$ such that for any k

$$\frac{64K}{3} p_b^k W^2 \leq p_0^2 \left(\sigma_k^2 + \delta_k^2\right) \tag{25}$$

then

$$\begin{aligned} \mathrm{E}(|\bar{\theta}'_K - \bar{\theta}_K|^2) &\leq p_0^2 \left(\sum_{k=1, L_k \geq 0}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1} \\ &= p_0^2 \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1}. \end{aligned}$$

Let D' be the mse of $\bar{\theta}'_K$. Then we can combine the above bound with (24) to obtain

$$\begin{aligned} D' &= \mathrm{E}(|\bar{\theta}'_K - \theta|^2) = \mathrm{E}(|\bar{\theta}'_K - \bar{\theta}_K + \bar{\theta}_K - \theta|^2) \\ &\leq \mathrm{E}(|\bar{\theta}'_K - \bar{\theta}_K|^2) + \mathrm{E}(|\bar{\theta}_K - \theta|^2) \\ &\quad + 2\sqrt{\mathrm{E}(|\bar{\theta}'_K - \bar{\theta}_K|^2) \mathrm{E}(|\bar{\theta}_K - \theta|^2)} \\ &\leq (p_0^2 + 2p_0 + 1) D, \end{aligned}$$

where the third term in the second line does not vanish because $\mathrm{E}((\bar{\theta}'_K - \bar{\theta}_K)(\bar{\theta}_K - \theta)) \neq 0$ due to the possible bias caused by BER. We bound this term by $\mathrm{E}(|\bar{\theta}'_K - \bar{\theta}_K|^2) \cdot \mathrm{E}(|\bar{\theta}_K - \theta|^2)$ using the Cauchy-Schwartz inequality. According to (25), the constant p_0 can be chosen as

$$p_0 = \max_{1 \leq k \leq K} \frac{8W}{\sigma_k} \sqrt{\frac{K p_b^k}{3}}.$$

This completes the Proof of Lemma 2. ■

B. Uniqueness of K_1

Lemma 3: For $f(M)$ defined in (17), i.e.,

$$f(M) = \frac{a_M\left(\sum_{k=1}^{M}\frac{1}{\sigma_k^2} - \frac{1}{D_0'}\right)}{\sum_{k=1}^{M}\frac{a_k}{\sigma_k^2}}, \quad \text{for } 1 \le M \le K$$

where $a_1 \le a_2 \le \ldots \le a_K$. Suppose $f(M) \ge 1$ for some $1 \le M \le K$, then there exists a unique K_1 such that $f(K_1) < 1$ and $f(K_1+1) \ge 1$, where $1 \le K_1 \le K$.

Proof: It is easy to see that

$$f(1) = \frac{\frac{a_1}{\sigma_1^2} - \frac{a_1}{D_0'}}{\frac{a_1}{\sigma_1^2}} < 1.$$

We find the smallest $K' \le K$ such that $f(K') \ge 1$ (the existence of such K' follows from the assumption). We claim that $f(M) \ge 1$ for any $M \ge K'$. This can be proved by showing that if $f(M) \ge 1$, then $f(M+1) \ge 1$. Specifically, suppose $f(M) \ge 1$ for some $K' \le M \le K$, then we have

$$\begin{aligned} f(M+1) &= \frac{a_{M+1}\left(\sum_{k=1}^{M+1}\frac{1}{\sigma_k^2} - \frac{1}{D_0'}\right)}{\sum_{k=1}^{M+1}\frac{a_k}{\sigma_k^2}} \\ &= \frac{a_M\left(\sum_{k=1}^{M}\frac{1}{\sigma_k^2} - \frac{1}{D_0'}\right) + \frac{a_{M+1}}{\sigma_{M+1}^2}}{\sum_{k=1}^{M}\frac{a_k}{\sigma_k^2} + \frac{a_{M+1}}{\sigma_{M+1}^2}} \\ &\quad + \frac{(a_{M+1} - a_M)\left(\sum_{k=1}^{M}\frac{1}{\sigma_k^2} - \frac{1}{D_0'}\right)}{\sum_{k=1}^{M}\frac{a_k}{\sigma_k^2} + \frac{a_{M+1}}{\sigma_{M+1}^2}} \\ &\ge \frac{a_M\left(\sum_{k=1}^{M}\frac{1}{\sigma_k^2} - \frac{1}{D_0'}\right) + \frac{a_{M+1}}{\sigma_{M+1}^2}}{\sum_{k=1}^{M}\frac{a_k}{\sigma_k^2} + \frac{a_{M+1}}{\sigma_{M+1}^2}} \\ &\ge 1 \end{aligned}$$

where the last inequality is due to the fact that

$$\frac{a+b}{c+b} \ge 1, \quad \text{if } \{a \ge c, a > 0, b > 0, c > 0\}.$$

Specifically, we can make the following identification

$$a = a_M\left(\sum_{k=1}^{M}\frac{1}{\sigma_k^2} - \frac{1}{D_0'}\right), \quad b = \frac{a_{M+1}}{\sigma_{M+1}^2}, \quad c = \sum_{k=1}^{M}\frac{a_k}{\sigma_k^2}$$

and use the fact that $f(M) = (a/c) \ge 1$.

Since $f(M) \ge 1$ for any $M > K'$, it follows that there is a unique K_1 satisfying $f(K_1) < 1$ and $f(K_1+1) \ge 1$, and $K_1 = K' - 1$. The proof is complete. ■

Acknowledgment

The authors are indebted to one anonymous reviewer whose incisive and detailed comments helped to improve several aspects of this paper.

References

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

[2] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *ACM/IEEE Trans. Netw.*, vol. 11, pp. 2–16, Feb. 2002.

[3] W. Tsujita, S. Kaneko, T. Ueda, H. Ishida, and T. Moriizumi, "Sensor-based air-pollution measurement system for environmental monitoring network," in *Proc. 12th IEEE Int. Conf. Solid-State Sensors, Actuators and Microsystems*, vol. 1, Jun. 2003, pp. 544–547.

[4] M. Bhardwaj, T. Garnett, and A. P. Chandrakasan, "Upper bounds on the lifetime of sensor networks," in *Proc. IEEE Int. Conf. Commun.*, vol. 3, Jun. 2001, pp. 785–790.

[5] A. P. Chandrakasan, R. Min, M. Bhardwaj, S. Cho, and A. Wang, "Power aware wireless microsensor systems," in *Keynote Paper ESSCIRC*, Florence, Italy, Sep. 2002.

[6] J. M. Rabaey, M. J. Ammer, J. L. da Silva, D. Patel, and S. Roundy, "PicoRadio supports ad hoc ultra-low power wireless networking," *IEEE Comput.*, vol. 33, pp. 42–48, Jul. 2000.

[7] M. Cardei and J. Wu, "Energy-efficient coverage problems in wireless ad hoc sensor networks," *J. Computer Commun. Sensor Networks*, to be published.

[8] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Trans. Parallel Distrib. Syst.*, vol. 13, pp. 924–935, Sep. 2002.

[9] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proc. 21st Annual Joint Conf. IEEE Computer Commun. Societies*, vol. 3, New York, Jun. 2002, pp. 1567–1576.

[10] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," *IEEE Wireless Commun. Mag.*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

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

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

[13] 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, Aug. 1982.

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

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

[16] Z.-Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 735–744, Apr. 2005.

[17] Z.-Q. Luo and J.-J. Xiao, "Decentralized estimation in an inhomogeneous environment," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, p. 517.

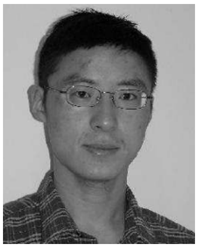
[18] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun. Mag.*, pp. 8–27, Aug. 2002.

[19] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimizatoin," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[20] ——, "Joint modulation and multiple access optimization under energy constraints," in *Proc. IEEE Global Telecomm. Conf.*, Dallas, Texas, Dec. 2004, pp. 151–155.

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

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.



Jin-Jun Xiao (S'04) received the B.Sc. degree in applied mathematics from Jilin University, Changchun, China, in 1997, and the M.Sc. degree in mathematics from the University of Minnesota, Twin Cities, in 2003.

He is currently pursuing the Ph.D. degree in electrical engineering at the University of Minnesota. During the summer of 2002, he worked on barcode signal processing at Symbol Technologies, Holtsville, NY. His current research interests are in distributed signal processing, multiuser information theory and their application to wireless sensor networks.

Shuguang Cui (S'99) received the B.Eng. degree in radio engineering with the highest distinction from Beijing University of Posts and Telecommunications, Beijing, China, in 1997, and the M.Eng. degree in electrical engineering from McMaster University, Hamilton, Canada, in 2000. He is currently working toward the Ph.D. degree in electrical engineering at Stanford University, CA.

From 1997 to 1998, he was with Hewlett-Packard, Beijing, P.R. China, as a system engineer. In the summer of 2003, he was with National Semiconductor, Santa Clara, CA, working on the ZigBee project. His current research interests include cross-layer energy minimization for low-power sensor networks, hardware and system synergies for high-performance wireless radios, and general communication theories.

Dr. Cui was a recipient of the NSERC graduate fellowship from the National Science and Engineering Research Council of Canada and the Canadian Wireless Telecommunications Association (CWTA) graduate scholarship.

Zhi-Quan (Tom) Luo (SM'03) received the B.Sc. degree in mathematics from Peking University, China, in 1984. He received the Ph.D. degree in operations research in 1989 from the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge.

During 1984–1985, he was with Nankai Institute of Mathematics, Tianjin, China. In 1989, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, where he became a Professor in 1998 and has held the Canada Research Chair in Information Processing since 2001. Since April 2003, he has been a Professor with the Department of Electrical and Computer Engineering, University of Minnesota, and holds an ADC Chair in Digital Technology. His research interests lie in the union of large-scale optimization, signal processing, data communications and information theory.

Prof. Luo is a member of SIAM and MPS. He is presently serving as an Associate Editor for several international journals including *SIAM Journal on Optimization*, *Mathematical Programming*, *Mathematics of Computation*, and *Mathematics of Operations Research*.

Andrea J. Goldsmith (S'90–M'93–SM'99–F'05) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1986, 1991, and 1994, respectively.

She was an Assistant Professor with the Department of Electrical Engineering, California Institute of Technology (Caltech) from 1994 to 1999. In 1999, she joined the Electrical Engineering Department, Stanford University, Stanford, CA, where she is currently an Associate Professor. Her industry experience includes affiliation with Maxim Technologies, Santa Clara, CA, from 1986 to 1990, where she worked on packet radio and satellite communication systems, and with AT&T Bell Laboratories, Holmdel, NJ, from 1991 to 1992, where she worked on microcell modeling and channel estimation. Her research includes work in capacity of wireless channels and networks, wireless information and communication theory, multiantenna systems, joint source and channel coding, cross-layer wireless network design, communications for distributed control, and adaptive resource allocation for cellular systems, *ad hoc* wireless networks, and sensor networks.

Dr. Goldsmith is the Bredt Faculty Development Scholar at Stanford University and a recipient of the Alfred P. Sloan Fellowship, the National Academy of Engineering Gilbreth Lectureship, a National Science Foundation CAREER Development Award, the Office of Naval Research Young Investigator Award, a National Semiconductor Faculty Development Award, an Okawa Foundation Award, Stanford's Terman Faculty Fellowship, and the David Griep Memorial Prize from the University of California, Berkeley. She was an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1995 to 2001, and has been an Editor for the IEEE *Wireless Communications Magazine* since 1995. She is also an elected member of Stanford's Faculty Senate and the IEEE Board of Governors.