MINIMUM ENERGY DECENTRALIZED ESTIMATION IN SENSOR NETWORK WITH CORRELATED SENSOR NOISE

Alexey Krasnopeev, Jin-Jun Xiao, and Zhi-Quan Luo

Department of Electrical and Computer Engineering,
University of Minnesota, Minneapolis, MN 55455 USA
E-mail: { krasnoa, xiao, luozq }@ece.umn.edu

ABSTRACT

We consider the problem of a single parameter estimation by a sensor network with a fusion center (FC). Sensor observations are corrupted by additive noises which can have arbitrary spatial correlation. Due to a bandwidth constraint each sensor is only able to transmit a finite number of bits. The fusion center combines messages from the sensors to produce a parameter estimator, which is required to have Mean Square Error (MSE) within a constant factor of that of the Best Linear Unbiased Estimator (BLUE). We show that total sensor transmitted power can be minimized while meeting target MSE requirement if quantization levels are determined jointly by the fusion center using the knowledge of noise covariance matrix. By numerical examples we show that energy saving up to 70% can be achieved when compared to uniform quantization strategy when each sensor generates the same number of bits.

1. INTRODUCTION

A typical wireless sensor network (WSN) consists of a fusion center and a large number of sensors, which are spatially distributed to monitor parameters of interest. Sensors are limited in their computation and communication capabilities due to limited power supply. Each sensor makes a measurement of the parameter, generates a local message, and sends it to the fusion center (FC), while the fusion center combines received messages to produce a final estimator of the parameter.

Power and bandwidth limit the length of messages and the data rates. Recently, several decentralized estimation schemes (DES) [1]-[4] have been proposed for parameter estimation in the presence of additive sensor noise. These DESs require each sensor to send only a few bits to the FC, and the message length is determined by the sensor's local SNR. Performance of the resulting estimator is shown to be within a constant factor of the Best Linear Unbiased Estimator (BLUE) performance.

Since sensors have only small size batteries, minimization of sensor power consumption is important to ensure the lifespan of a WSN. In [5] optimal coded and uncoded transmission strategies are proposed for sensor networks in order to minimize required energy per bit. In the recent work of [6], the authors consider the problem of optimal power allocation for decentralized estimation where sensor measurements are corrupted by additive noises, while communication links from sensors to the FC differ in quality. The authors of [6] consider an adaptive modulation scheme for message transmission which leads to the exponential dependence of energy on the message size. Then optimal power and quantization levels for sensors are decided jointly at the FC.

All the work mentioned has been done under the assumption that observation noises are spatially uncorrelated. However, sensors in the network may have correlated observations, for example, in case when they are densely deployed in the field. In our work, we consider parameter estimation in case when sensor observations are corrupted by additive noises which are spatially correlated with known correlation matrix. We assume that transmission of each bit requires a constant amount of energy, dependent on the quality of the channel between a sensor and the fusion center.

We use the following notations. Diagonal matrix with nonzero elements $a_1, \ldots, a_N$ is denoted by $\mathrm{diag}(a_1, \ldots, a_N)$. For any real number $x \in \mathbb{R}$, we denote $\lceil x \rceil$ to be the smallest integer greater or equal to x. For any random variable R, we use $\mathrm{E}\, R$ to denote expected value of R and $\mathrm{var}\, R$ to denote variance of R.

This research is supported in part by the National Science Foundation, Grant No. DMS-0312416.

2. PROBLEM FORMULATION

We consider a problem of estimating unknown parameter θ by a sensor network consisting of N sensors. Measurement of each sensor x_i is corrupted by additive noise n_i, so that

$$x_i = \theta + n_i, \quad i = 1, \ldots, N.$$

We assume that both θ and n_i have finite range, so that all x_i belong to a common finite interval $[-U, U]$, with $U > 0$ a known constant. The noises n_i are assumed to be zero mean and corre-

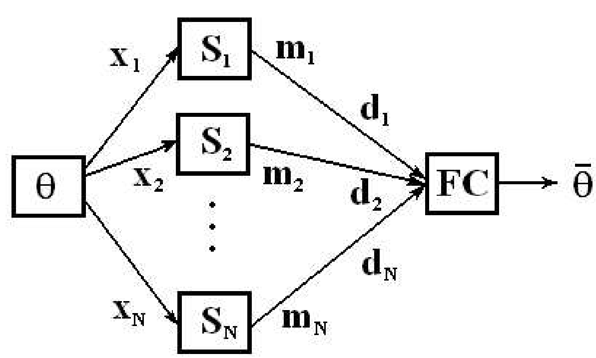

Fig. 1. Decentralized estimation scheme

lated across sensors with covariance matrix $\mathbf{C}$, but otherwise unknown. Measurements x_i are quantized to produce messages m_i to be passed on to the FC, which combines received messages in order to estimate θ.

We assume that sensors send messages using a TDMA scheme and the channel between each sensor and the FC is corrupted by Additive White Gaussian Noise (AWGN) with power spectral density $N_0/2$. The signal power received at the FC is assumed to be inversely proportional to d_i^κ where d_i is the distance between sensor i and the FC, and κ is the path loss exponent. Suppose that message m_i has length b_i bits. Let us assume that energy W_i required for transmission of m_i is proportional to the number of bits in the message. This is the case e.g., if sensors use MQAM or MPSK modulation to transmit messages. For example, if MQAM is used, W_i can be found as follows [5]:

$$W_i = \frac{2}{3} N_f N_0 G_0 d_i^\kappa (2^s - 1) \ln\left(\frac{4(1-2^{-s})}{sP_b}\right) \frac{b_i}{s},$$

where $s = \log M$ is the number of bits per symbol, N_f is the receiver noise figure, P_b is the required bit error probability, and G_0 is the system constant defined as in [5].

In case when $m_i = x_i$, BLUE estimator for θ is known to be [7]

$$\hat{\theta} = \frac{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{x}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}},$$

where $\mathbf{x} = (x_1, \ldots, x_N)^T$ and $\mathbf{1}$ is the vector of all ones. Estimation performance is characterized by variance of the estimator $\operatorname{var}\hat{\theta} = (\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1})^{-1}$. We assume that the FC has full knowledge of correlation between sensors. Upon receiving sensor messages m_i, the FC combines them into an estimator $\bar{\theta}$ as given by the expression below:

$$\bar{\theta} = \frac{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{m}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}, \tag{1}$$

where $\mathbf{m} = (m_1, \ldots, m_N)^T$. We can write the following sequence of equalities:

$$\begin{aligned}
\operatorname{MSE}(\bar{\theta}) &= \operatorname{E}(\bar{\theta}-\theta)^2 = \operatorname{E}(\bar{\theta}-\hat{\theta}+\hat{\theta}-\theta)^2 \\
&= \operatorname{E}(\bar{\theta}-\hat{\theta})^2 + \operatorname{E}(\hat{\theta}-\theta)^2 \\
&= \operatorname{E}\left(\frac{\mathbf{1}^T\mathbf{C}^{-1}(\mathbf{m}-\mathbf{x})}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}\right)^2 + \operatorname{var}\hat{\theta} \\
&= \left(\frac{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{Q}\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} + 1\right)\operatorname{var}\hat{\theta},
\end{aligned}$$

where $\mathbf{Q} = \operatorname{E}(\mathbf{m}-\mathbf{x})(\mathbf{m}-\mathbf{x})^T$ is the matrix of quantization noise. Here in the third step we have used the fact that cross term $\operatorname{E}(\bar{\theta}-\hat{\theta})(\hat{\theta}-\theta) = 0$, as $\hat{\theta}$ is independent of m_i and unbiased.

We wish to optimize the transmission energy while maintaining the estimation performance within a constant factor of BLUE performance, i.e. $\operatorname{MSE}(\bar{\theta}) \le (1+\alpha)\operatorname{var}\hat{\theta}$ for some constant $\alpha > 0$. Therefore, the following condition must hold

$$\frac{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{Q}\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} \le \alpha. \tag{2}$$

The total energy is equal to $W = \sum_{i=1}^N W_i = \sum_{i=1}^N w_i b_i$, where w_i is the energy required for transmission of a single bit from sensor i to the FC. Our goal is to determine the set of integer numbers $\{b_i\}$ such that W achieves its minimum provided MSE condition (2) is satisfied.

3. QUANTIZATION STRATEGY

Suppose that sensor observation x_i is bounded to a finite interval $[-U, U]$. Suppose further that we wish to quantize x_i in such a way that resulting message m_i has length b_i bits, where b_i is to be determined later. We therefore have $K_i = 2^{b_i}$ quantization points $\{a_j^{(i)} \in [-U, U],\ j = 1, \ldots, K_i\}$. These points are placed so that $a_1^{(i)} = -U < a_2^{(i)} < \ldots < a_{K_i}^{(i)} = U$, and $a_{k+1}^{(i)} - a_k^{(i)} = \Delta_i$ for every k. It is easy to see that $\Delta_i = 2U/(K_i - 1)$. Suppose that $x_i \in [a_k^{(i)}, a_{k+1}^{(i)})$. Then x_i is mapped into the point $a_{k+1}^{(i)}$ with probability p and into $a_k^{(i)}$ with probability $1-p$. Namely,

$$\begin{aligned}
\Pr(m_i = a_{k+1}^{(i)}) &= (x_i - a_k^{(i)})/\Delta_i &= p, \\
\Pr(m_i = a_k^{(i)}) &= (a_{k+1}^{(i)} - x_i)/\Delta_i &= 1-p.
\end{aligned}$$

This probabilistic mapping produces messages with $\operatorname{E} m_i = x_i$, where expected value is taken with respect to randomization. Therefore, $\bar{\theta}$ is an unbiased estimator for θ. Unbiasedness of the quantization strategy described together with the fact that random variables m_i and m_j are conditionally independent given x_i and x_j for all $i \neq j$ leads to the following important property.

Lemma 1. *The quantization noise matrix $\mathbf{Q}$ is diagonal.*

Further, the maximal variance of m_i is equal to $\Delta_i^2/4$. Let Q_i be the i-th diagonal element of $\mathbf{Q}$. We have $Q_i \le U^2/(2^{b_i}-1)^2$.

4. TOTAL ENERGY MINIMIZATION

Introducing notation $\mathbf{c} = \mathbf{C}^{-1}\mathbf{1}$and $\beta = \alpha/\operatorname{var}\hat{\theta}$, we can rewrite (2) as $\mathbf{c}^T\mathbf{Q}\mathbf{c} \le \beta$. Since the distribution of $\mathbf{x}$ is unknown in general, we enforce a stronger condition: $\max \mathbf{c}^T\mathbf{Q}\mathbf{c} \le \beta$. Recalling that $\mathbf{Q}$ is diagonal, we can write $\max \mathbf{c}^T\mathbf{Q}\mathbf{c} = \max \sum_{i=1}^N Q_i c_i^2 = \sum_{i=1}^N U^2 c_i^2/(2^{b_i}-1)^2$. Since b_i can only take integer values, the problem of energy minimization subject to MSE constraint is actually a non-convex integer programming problem. We relax b_i to take real positive values and formulate the following optimization problem:

$$\begin{aligned}
\text{minimize} \quad & \textstyle\sum_{i=1}^N w_i b_i \\
\text{subject to} \quad & \textstyle\sum_{i=1}^N \frac{c_i^2}{(2^{b_i}-1)^2} \le \frac{\beta}{U^2}, \\
& b_i > 0, \quad i = 1, \ldots, N.
\end{aligned} \tag{3}$$

Without loss of generality we consider case when $c_i > 0$ for all i. In case $c_i = 0$ for some sensors we can exclude corresponding m_i from consideration, as they do not contribute to $\bar{\theta}$. Solution to problem (3) can be efficiently found by the FC using the interior point method [8]. Then the FC can round the solution to the nearest greater integer and broadcast it to the sensors. However, here we will present a closed form approximately optimal solution to the problem. We drop the last constraints on b_i as they become inactive at the optimum point. Associating a dual variable λ with the MSE constraint we write the Lagrangian for the problem as follows:

$$L(b_i, \lambda) = \sum_{i=1}^N w_i b_i + \lambda\left(\sum_{i=1}^N \frac{c_i^2}{(2^{b_i}-1)^2} - \frac{\beta}{U^2}\right).$$

At the point of optimum we must have $\partial L/\partial b_i = 0$ for all i, which leads to the following set of conditions:

$$\frac{2^{b_i}}{(2^{b_i}-1)^3} = \frac{w_i\lambda'}{c_i^2}, \tag{4}$$

where $\lambda' = 1/2\lambda \ln 2$. The complementary slackness condition yields

$$\sum_{i=1}^{N} \frac{c_i^2}{(2^{b_i}-1)^2} = \frac{\beta}{U^2}. \tag{5}$$

As the reader can verify, optimal solution $\{b_i\}$ cannot be found in the closed form from the system (4)-(5). For this reason, we will find a feasible solution $\{b_i^*\}$, which will be used as an approximation to $\{b_i\}$. Namely, we impose

$$\sum_{i=1}^{N} \frac{c_i^2}{(2^{b_i^*}-1)^2} = \frac{\beta}{U^2} \tag{6}$$

as well as the following equality

$$\frac{2^{b_i^*}-1}{(2^{b_i^*}-1)^3} = \frac{1}{(2^{b_i^*}-1)^2} = \frac{w_i\lambda^*}{c_i^2} \tag{7}$$

for some λ^*. Note that the condition (7) is a modification of (4). Clearly, such a solution is unique, and parameter λ^* can be found from (6) and (7) as

$$\lambda^* = \frac{\beta}{U^2}\Big(\sum_{i=1}^{N} w_i\Big)^{-1}. \tag{8}$$

Approximation b_i^* can be found explicitly from (7) and (8) to be

$$b_i^* = \log\left(1 + \frac{|c_i|}{\sqrt{\lambda^* w_i}}\right). \tag{9}$$

Now suppose that optimal solution $\{b_i\}$ is such that $b_i \geq 1$ for all i. Then the following bound holds true.

Lemma 2. *Let $\{b_i\}$ be the optimal solution to the problem (3) such that $b_i \geq 1$ for all i, and let $\{b_i^*\}$ be its approximation defined by (9). Then*

$$b_i^* - \frac{1}{2} < b_i < b_i^* + \frac{1}{2}. \tag{10}$$

We conclude that $|b_i - b_i^*| < 1$. Thus, rounded optimal solution $\lceil b_i \rceil$ is at most one bit away from $\lceil b_i^* \rceil$. We can interpret this result as follows: in situation when b_i are sufficiently large, e.g. when high estimation precision is required, the optimal solution behaves approximately as $\log(1 + |c_i|/\sqrt{w_i})$.

5. NUMERICAL RESULTS

In this section, we present numerical simulation to compare the transmission energy of quantization using the closed form approximate solution (9) to that of uniform quantization when all sensors quantize their observations to the same number of bits to achieve the same MSE. Let us denote by b the number of bits used in case of uniform quantization. We can find the minimal W_{uniform} using the MSE constraint to be

$$W_{\text{uniform}} = \left\lceil \log\left(1 + \sqrt{\frac{U^2}{\beta}\sum_{i=1}^{N} c_i^2}\right)\right\rceil \sum_{i=1}^{N} w_i.$$

The optimal energy obtained by relaxing $\{b_i\}$ to take on real values is a lower bound on the actual optimal energy. If we round b_i up to the closest integer $\lceil b_i \rceil$, we can obtain an upper bound (denoted by $\overline{W}_{\text{opt}}$) on the actual energy. Even though we use $\lceil b_i^* \rceil$ to approximate the actual optimal solution, significant energy can be saved when compared with the uniform quantization strategy in which each sensor quantizes the signal into the same number of bits to achieve the same target distortion. The percentage of saving is defined as $(W_{\text{uniform}} - \overline{W}_{\text{opt}})/W_{\text{uniform}} \times 100$.

For a positive random variable R we define normalized deviation of R to be $\sqrt{\operatorname{var} R}/\operatorname{E} R$, which will be used as a measure of the absolute heterogeneity of R. The sensor noise variances $\{\sigma_i^2\}$ are taken to be $\sigma_i^2 = 1 + a^2 Z_i$, where Z_i are i.i.d. random variables with $Z_i \sim \chi_1^2(z)$. As can be easily verified, $\{\sigma_i^2\}$ are also i.i.d. with $\sigma_i \sim \chi_1^2((x-1)/a^2)$. We control heterogeneity of sensor noise variances by varying the parameter a. We suppose that only i-th and $i+1$-th sensors have nonzero noise correlation for all i, in other words, $\mathbf{C}$ is a tridiagonal matrix.

In all simulations, the total number of sensors $N = 200$. The correlation coefficient for each pair of sensors is taken to be 0.2. Since all coefficients w_i are scaled by a common factor, $\{w_i\}$ are taken to be channel path losses $w_i = d_i^\kappa$.

Assume that the target estimation performance is fixed. From Fig. 2 we can see that the amount of energy saving becomes significant when the local noise variances become more and more heterogeneous, supposing that sensors have identical w_i. In Fig. 3, we plot the percentage of energy savings versus the heterogeneity of channel gains, supposing that sensors have same observation noise variances. Here we suppose that all sensors are uniformly distributed inside a unitary disk whose center is at the FC. It is easy to show that in this case normalized deviation of w_i depends only on κ. In our simulation, we choose $1 \leq \kappa \leq 8$. We observe that percentage of saving depends more on the heterogeneity of sensor noise variances than that of channel gains. This can be understood regarding expression (9) for b_i^*, where in the logarithm, the quantity depends on the distribution of c_i, but only on the distribution of $1/\sqrt{w_i}$.

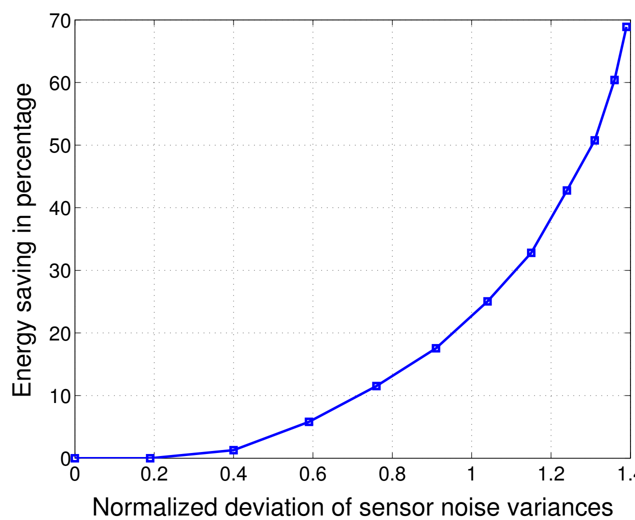

Fig. 2. Percentage of energy saving increases when sensor noise variances become more heterogeneous.

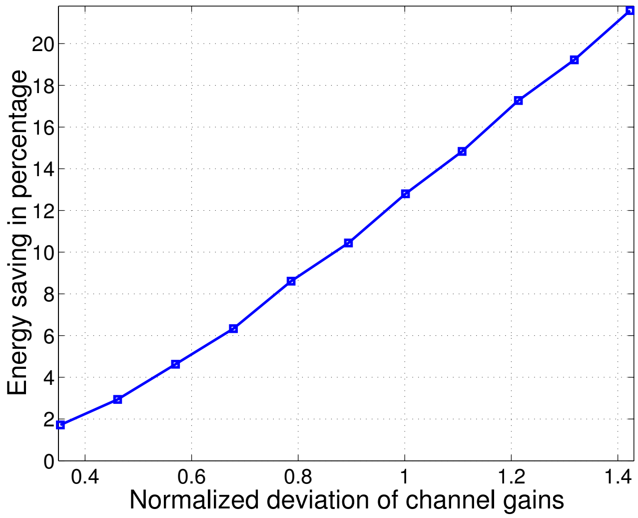

Fig. 3. Percentage of energy saving increases when channel gains become more heterogeneous.

6. LOCAL SENSOR COOPERATION

We have seen that our suboptimal solution works well in terms of energy saving. Let us now consider a special case when the use of $\{b_i^*\}$ is especially appealing. Suppose that covariance matrix $\mathbf{C}$ has a block-diagonal structure: $\mathbf{C} = \mathrm{diag}(\mathbf{C}_1, \ldots, \mathbf{C}_n)$. This situation may occur when sensors in the network are partitioned into several groups in such a way that sensors within each group are placed relatively close to each other and far from the rest of the sensors. Thus, only the observations of sensors belonging to one and the same group have nonzero correlation. In this case matrix $\mathbf{C}^{-1}$ is also block-diagonal: $\mathbf{C}^{-1} = \mathrm{diag}(\mathbf{C}_1^{-1}, \ldots, \mathbf{C}_n^{-1})$. We assume further that sensors within each group can cooperate to learn the corresponding covariance sub-matrix $\mathbf{C}_j$. Then quantization levels b_i can be found in a distributive manner as follows. Value of λ^* can be computed by the FC and broadcasted back to the sensors. Then each sensor can easily compute $c_i = [\mathbf{C}_j^{-1}\mathbf{1}]_i$ and independently find its own quantization level b_i^*. The advantage of this method is that the FC needs to broadcast only one universal message for all sensors as opposed to the case of general $\mathbf{C}$ when all quantization levels b_i need to be broadcasted.

7. MINIMAX FORMULATION

Minimizing total transmission energy results in sensors having different lifetime. An alternative approach is to minimize maximal energy W_i which leads to maximum network lifetime. Relaxing $\{b_i\}$ as in (3), we can state the problem as follows:

$$\begin{array}{ll} \text{minimize} & \max\limits_i w_i b_i \\ \text{subject to} & \sum_{i=1}^N \dfrac{c_i^2}{(2^{b_i}-1)^2} \leq \dfrac{\beta}{U^2}, \\ & b_i > 0, \quad i = 1, \ldots, N, \end{array} \tag{11}$$

or alternatively

$$\begin{array}{ll} \text{minimize} & \max t \\ \text{subject to} & w_i b_i \leq t \\ & \sum_{i=1}^N \dfrac{c_i^2}{(2^{b_i}-1)^2} \leq \dfrac{\beta}{U^2}, \\ & b_i > 0, \quad i = 1, \ldots, N. \end{array} \tag{12}$$

It can be shown by analyzing the KKT conditions for (12) that at the optimum point equality $w_i b_i = t_{\text{opt}}$ must hold for all i, where t_{opt} can be found as a solution to the following equation

$$\sum_{i=1}^N \frac{c_i^2}{(2^{\frac{t}{w_i}}-1)^2} = \frac{\beta}{U^2}. \tag{13}$$

The solution t_{opt} is unique due to the monotonicity of the left hand side function in (13). The FC can solve (13) and broadcast t_{opt} to the sensors, which in turn can determine their quantization levels locally. In this case sensor lifetime is not affected by transmitted power.

8. CONCLUSION

We have shown that total transmission energy consumption in a WSN can be minimized if quantization levels for sensors are determined jointly by the FC using information about correlation of sensor observations. We have also presented an approximate suboptimal solution to the energy minimization problem achieving the same target estimation performance as the optimal solution. It is shown by numerical simulations that energy saving up to 70% can be achieved compared to uniform quantization when each sensor sends the same number of bits.

9. REFERENCES

[1] Z.-Q. Luo, "Universal Decentralized Estimation in a Bandwidth Constrained Sensor Network," submitted to *IEEE Trans. on Inform. Theory*.

[2] Z.-Q. Luo, "An Isotropic Universal Decentralized Estimation for a Bandwidth Constrained Ad Hoc Sensor Network," submitted to *IEEE Journal on Selected Areas in Comm.*, 2004.

[3] Z.-Q. Luo, and J.-J. Xiao, "Decentralized Estimation in Inhomogenious Sensing Environment," submitted to *IEEE Trans. on Inform. Theory*.

[4] Z.-Q. Luo, and J.-J. Xiao, "Decentralized Estimation in Inhomogenious Environment," *Proc. IEEE International Syposium on Inform. Theory*, pp. 520, Chicago, June 2004.

[5] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained Modulation Optimization," to appear in *IEEE Trans. on Wireless Comm.*, 2004.

[6] J.-J. Xiao, S. Cui, A. J. Goldsmith, and Z.-Q. Luo, "Joint Estimation in Sensor Networks under Energy Constraints," to be presented at SECON, Santa Clara, 2004.

[7] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall, Upper Saddle River, New Jersey, 1993.

[8] S. Boyd, L. Vandenberghe, *Convex Optimization.* Cambridge University Press, Cambridge, U. K., 2003.