Performance Bounds for the Rate-Constrained Universal Decentralized Estimators

Jin-Jun Xiao, *Student Member, IEEE*, Zhi-Quan Luo, *Senior Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—We consider decentralized estimation of a noise-corrupted deterministic parameter using a bandwidth-constrained sensor network with a fusion center (FC). Each sensor's noise is additive, zero mean, and independent across sensors. A decentralized estimator is said to be universal if the local sensor quantization rules and the final fusion rule at the FC are independent of sensor noise pdf. Assuming that information rate from each sensor to the FC is constrained to one bit per sample, we derive a Cramér–Rao lower bound (CRLB) on the mean-squared error (MSE) performance of a class of rate-constrained universal decentralized estimators. Our results show that if sensor observation noise has finite range in $[-U, U]$, then the minimum MSE performance of any one-bit rate-constrained universal decentralized estimator is at least $U^2/(4K)$, where K is the total number sensors. This bound implies that the recently proposed universal decentralized estimators are optimal up to a constant factor of 4.**

***Index Terms*—Cramér–Rao bound, distributed estimation, quantization, sensor network.**

I. Introduction

CONSIDER a wireless sensor network (WSN) with a fusion center (FC) in which the communication links from sensors to the FC are bandlimited and no intersensor communication is allowed. The distributed signal processing in such a WSN differs from the traditional signal processing framework in several important aspects. First, observation data in a WSN are collected at different nodes across the network, which necessitates communication from the sensors to the FC. Since local sensors typically have a low power budget that limits their resolution and communication capability, local data quantization/compression is needed so as to reduce the required expenditure of resources. Second, the sensor noise distribution can be difficult to characterize, especially for applications in dynamic sensing environment. To address these challenges, we are led to the design of rate-constrained universal decentralized estimation schemes (DES), which do not require the knowledge of the sensor noise probability density function (pdf) and have a low bandwidth requirement, to say, one bit per sample per sensor.

The problem of decentralized estimation has been studied first in the context of distributed control [1] and tracking [2], later in data fusion [3], [4], and most recently in wireless sensor networks [5]–[7]. Among these studies, the prevailing assumption has been that the joint distribution of the sensor observations is known. When the sensor noise distribution is unknown, [8] proposed to use a training sequence to aid the design of local data compression strategies. In some recent work [9], [10], we considered the universal decentralized estimation of a noise corrupted unknown parameter by a WSN with an FC, under a bandwidth constraint of one bit per sample per sensor. The proposed DES is universal in the sense that it does not require the knowledge of the noise pdf. It achieves a mean-squared error (MSE) of no more than W^2/K, where K is the total number sensors, and W denotes the dynamic range of sensor observations.

In this letter, we show that for any unknown parameter to be observed, the MSE performance of any one-bit rate-constrained universal DES is lower bounded by $U^2/4K$, where U is the range of sensor observation noise. This implies that if the unknown parameter has a small dynamic range, the recently proposed universal decentralized estimators [9], [10] are optimal up to a constant factor of 4. Our analysis consists of two steps: first we derive a Cramér–Rao lower bound (CRLB) for the one-bit rate-constrained decentralized estimators as a function of the noise pdf and local quantization rules, and then we maximize this CRLB with respect to the noise pdf and local quantization rules to obtain a lower bound on the MSE performance of any universal DES.

II. Decentralized Estimation

Consider a set of K distributed sensors, each making an observation related to an unknown parameter $\theta \in \mathbb{R}$. The observations are corrupted by additive noise and described by

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

where $\{n_k : k = 1, 2, \ldots, K\}$ are spatially independent zero-mean noise random variables with pdf $\{f_k(x) : k = 1, 2, \ldots, K\}$.

Suppose the information rate from each sensor to the FC is constrained to be one bit per sample. Let $m_k(x_k)$ denote the binary message function for sensor k (see Fig. 1). Under such an assumption, we can find $S_k \subseteq \mathbb{R}$ such that

$$m_k(x_k) = \begin{cases} 1, & x_k \in S_k \\ 0, & x_k \in S_k^c \end{cases} \tag{2}$$

where S_k^c is the complement of S_k in $\mathbb{R}$. The FC then combines the messages $m_k(x_k)$ to produce a final estimate of θ according to the fusion rule

$$\bar{\theta} = \Gamma(m_1(x_1), m_2(x_2), \ldots, m_K(x_K)).$$

Manuscript received January 22, 2006; revised June 9, 2006. This work was supported in part by the National Science Foundation under Grant DMS-0312416, in part by the USDOD ARMY under Grant W911NF-05-1-0567, and in part by Army Research Office under Grant 49453-MA. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Hongbin Li.

The authors 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; georgios@ece.umn.edu).

Digital Object Identifier 10.1109/LSP.2006.881522

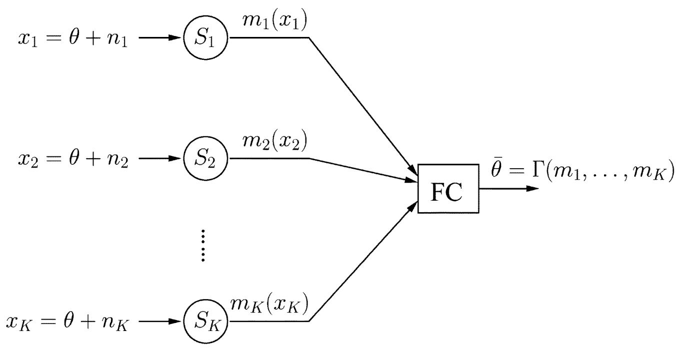

Fig. 1. Decentralized estimation scheme.

Therefore, the design of optimal DES consists of selecting the subsets $S_k,\ k = 1, 2, \ldots, K$ and a final fusion function Γ so that the resulting estimator is closest to the true θ in the statistical sense, while in this letter, the MSE criterion is adopted.

Notice that x_k has pdf $f_k(x - \theta)$, and m_k is a Bernoulli random variable with

$$\begin{aligned} P\{m_k = 1\} &= p_k(\theta) := \int_{S_k} f_k(x - \theta)\, dx \\ P\{m_k = 0\} &= 1 - p_k(\theta). \end{aligned} \tag{3}$$

We assume that the derivative $p_k'(\theta)$ exists. This is ensured if $f_k'(x)$ exists for all x and $f_k'(x)$ is integrable. Moreover

$$p_k'(\theta) = -\int_{S_k} f_k'(x - \theta)\, dx. \tag{4}$$

Let $J(\theta; m_k)$ denote the Fisher information about θ contained in the random data m_k. Based on (3), we can calculate $J(\theta; m_k)$ as [7], [10]

$$\begin{aligned} J(\theta; m_k) &= \frac{1}{p_k(\theta)} {p_k'}^2(\theta) + \frac{1}{1 - p_k(\theta)} {p_k'}^2(\theta) \\ &= \frac{{p_k'}^2(\theta)}{p_k(\theta)(1 - p_k(\theta))}. \end{aligned} \tag{5}$$

Letting $\mathbf{m} = (m_1, m_2, \ldots, m_K)$ and using the independence assumption, we obtain the Fisher information

$$J(\theta; \mathbf{m}) = \sum_{k=1}^{K} J(\theta; m_k) = \sum_{k=1}^{K} \frac{{p_k'}^2(\theta)}{p_k(\theta)(1 - p_k(\theta))}. \tag{6}$$

By the standard CRLB result [11], we have the following theorem for the lower bound for the MSE of any unbiased estimator of θ based on $\mathbf{m}$.

Proposition 1: Let the subsets $\{S_k : k = 1, 2, \ldots, K\}$ be fixed and suppose the derivative of the noise pdf f_k exists and is integrable. Then any unbiased, one-bit rate-constrained decentralized estimator $\bar{\theta} = \Gamma(\mathbf{m})$ has an MSE

$$E(|\bar{\theta} - \theta|^2) \geq (J(\theta; \mathbf{m}))^{-1} \tag{7}$$

where $J(\theta; \mathbf{m})$ is given by (6).

Remark 2: The assumption that $f_k'(x)$ exists and is integrable is needed to ensure that the following regularity condition

$$\frac{\partial P(m_k = 0)}{\partial \theta} + \frac{\partial P(m_k = 1)}{\partial \theta} = 0 \tag{8}$$

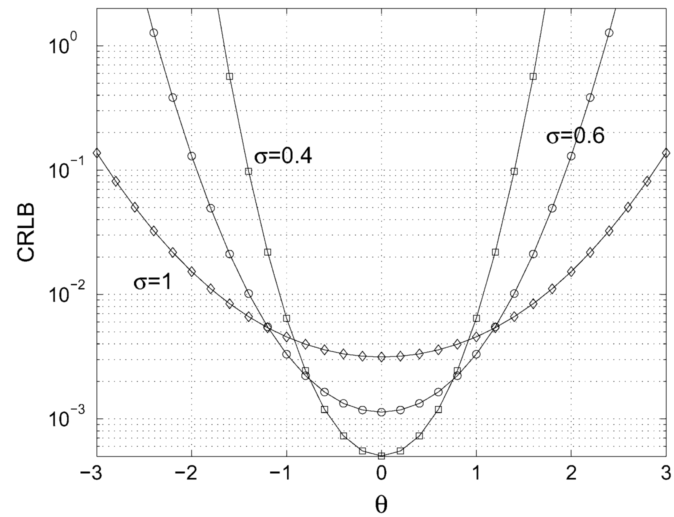

Fig. 2. CRLB versus θ for a DES with sign detectors under i.i.d. Gaussian noise.

holds true. This condition in turn guarantees the validity of the CRLB (7).

Example 3: Suppose that all sensors use a sign detector with $S_k = \mathbb{R}_+$. Then, $p_k(\theta) = \int_0^\infty f_k(x - \theta) dx = 1 - F_k(-\theta)$. It follows from (6) that for any unbiased estimator $\bar{\theta} = \Gamma(\mathbf{m})$

$$E(|\bar{\theta} - \theta|^2) \geq \left(\sum_{k=1}^{K} \frac{{f_k}^2(-\theta)}{F_k(-\theta)(1 - F_k(-\theta))} \right)^{-1}. \tag{9}$$

The CRLB given by (9) may blow up for some choices of $\{f_k : k = 1, 2, \ldots, K\}$ and θ. Fig. 2 plots the case when $K = 500, \theta \in [-3, 3]$ and noises are i.i.d. Gaussian with standard deviation σ. The curves clearly show that the CRLB increases exponentially when θ moves away from 0 (see also [10] and [12]).

In the next section, we apply the CRLB in Proposition 1 to derive performance bounds for a class of universal decentralized estimators in [9] and [10].

III. Performance Bounds of Universal Decentralized Estimation

Consider a universal DES whose message functions m_k and the FC estimator Γ do not exploit the knowledge of the noise pdf. We assume that sensor observation noise n_k has a finite range, i.e., $n_k \in [-U, U]$. The performance of a universal estimator is characterized by the worst-case MSE over all possible noise pdf $\{f_1, f_2, \ldots, f_K\}$ with support in $[-U, U]$. Thus, designing an optimal universal DES amounts to selecting $\{S_1, \ldots, S_K; \Gamma\}$ so that $\max_{f_k} E(|\bar{\theta} - \theta|^2)$ is minimized, where $\bar{\theta} = \Gamma(\mathbf{m})$ and m_k is given in (2). We define

$$D(K, U) := \min_{S_k, \Gamma} \max_{f_k} E(|\bar{\theta} - \theta|^2)$$

which characterizes the optimal (worst-case) performance of universal decentralized estimators.

In Example 3, the local quantization rules are taken to be the sign detectors, and the corresponding CRLB blows up if $f_k(-\theta) = 0$ for all $1 \leq k \leq K$. This implies that the sign detector is not a good universal DES. Recently, several universal decentralized estimators have been proposed in [9], [10], and [12]. For example, the DES proposed in [9] works as follows: $1/2$ of the sensors quantize their observations to the first most significant bit (MSB), $1/4$ of the sensors quantize their observations to the second MSB, and so on. The final estimator at the FC is a simple linear average of these received bits. It has been shown in [9] that if θ is in a finite range $[\theta_0 - V, \theta_0 + V]$ (with interval bounds known *a priori*), then this DES achieves an MSE of no more than W^2/K with $W = U + V$. This implies that

$$D(K,U) \leq \frac{W^2}{K} = \frac{(U+V)^2}{K}. \quad (10)$$

In what follows, we derive a lower bound on $D(K,W)$. From the CRLB definition, we obtain that

$$\begin{aligned} D(K,U) &= \min_{S_k,\Gamma} \max_{f_k} E(|\bar{\theta} - \theta|^2) \\ &\geq \min_{S_k} \max_{f_k} J^{-1}(\theta; \mathbf{m}) \end{aligned} \quad (11)$$

where $J(\theta; \mathbf{m})$ is given in (6). Next, we use (11) to prove the following lower bound for $D(K,U)$.

Lemma 4: For any $\theta \in \mathbb{R}$, it holds that $D(K,U) \geq (U^2)/(4K)$.

Proof: It is easy to see that

$$\begin{aligned} &\min_{S_k} \max_{f_k} J^{-1}(\theta; \mathbf{m}) \\ &\quad \geq \max_{f_k} \min_{S_k} J^{-1}(\theta; \mathbf{m}) = \left(\min_{f_k} \max_{S_k} J(\theta; \mathbf{m}) \right)^{-1} \\ &\quad \overset{(a)}{=} \left(\min_{f_k} \max_{S_k} \sum_{k=1}^{K} J(\theta; m_k) \right)^{-1} \\ &\quad \overset{(b)}{=} \left(\sum_{k=1}^{K} \min_{f_k} \max_{S_k} J(\theta; m_k) \right)^{-1} \\ &\quad \overset{(c)}{=} \frac{1}{K} \left(\min_{f} \max_{S} J(\theta; m) \right)^{-1} \end{aligned} \quad (12)$$

where (a) is due to the independence of $\{m_k\}$ given θ, (b) holds since both the objective function $J(\theta; m_k)$ and variables f_k, S_k are nicely decoupled for different k, and (c) follows from the fact that the min-max is achieved at a common S and f. Moreover, by appealing to (5), (11), and (12), we have

$$\begin{aligned} D(K,U) &\geq \frac{1}{K} (\min_{f} \max_{S} J(\theta; m))^{-1} \\ &= \frac{1}{K} \left(\min_{f} \max_{S} \frac{{p'_S}^2(\theta)}{p_S(\theta)(1-p_S(\theta))} \right)^{-1} \\ &= \frac{1}{K} \max_{f} \min_{S} \frac{p_S(\theta)(1-p_S(\theta))}{{p'_S}^2(\theta)} \end{aligned} \quad (13)$$

where $p_S(\theta; f) = \int_S f(x-\theta)\, dx; p'_S(\theta; f) = -\int_S f'(x-\theta)\, dx$.

Next, we choose a specific choice of f to lower bound the right-hand side of (13). Let us define (see the left plot in Fig. 3)

$$f_0(x) := \begin{cases} \frac{U-|x|}{U^2}, & |x| \leq U \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

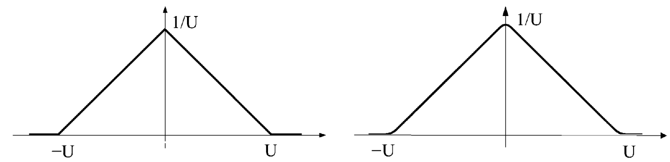

Fig. 3. Plot of $f_0(x)$ and $f_\epsilon(x)$.

We wish to choose $f = f_0$ in (13). Unfortunately, $f_0(x)$ is nondifferentiable at three points $x = \{-U, 0, U\}$, which prevents us from directly substituting $f = f_0$ in (13). We can circumvent this technical difficulty by smoothing $f_0(x)$ and taking limit. Specifically, for any $\epsilon > 0$, let $\phi_\epsilon(x) \geq 0$ denote a smooth (i.e., infinitely differentiable) kernel function,[1] with support in $[-\epsilon, \epsilon]$, symmetric (i.e., $\phi_\epsilon(x) = \phi_\epsilon(-x)$), and $\int_{-\infty}^{\infty} \phi_\epsilon(x) dx = 1$. We define a smoothed version of $f_0(x)$ as (see the right plot in Fig. 3)

$$f_\epsilon(x) = \int_{-\infty}^{\infty} f_0(y) \phi_\epsilon(x-y)\, dy.$$

Since $f_0(x)$ is continuous over $\mathbb{R}$, it follows that $f_\epsilon(x), f'_\epsilon(x) \to f_0(x), f'_0(x)$ uniformly for all $x \in \mathbb{R} - \{\pm U, 0\}$. Moreover, both $f_\epsilon(x), f'_\epsilon(x)$ are bounded by $1/U$ and $1/U^2$, respectively, in $[-U-\epsilon, U+\epsilon]$. By invoking the Lebesgue dominance theorem [13, p. 176], we obtain that when $\epsilon \to 0$

$$\begin{aligned} \int_S f_\epsilon(x)\, dx \to \int_S f_0(x)\, dx, \ \int_S f'_\epsilon(x)\, dx \\ \to \int_S f'_0(x)\, dx \end{aligned} \quad (15)$$

for every measurable set $S \subseteq \mathbb{R}$.

Let us fix a small $\epsilon > 0$ and choose $f = f_\epsilon$ in (13). This gives

$$D(K,U) \geq \frac{D_\epsilon}{K} \quad (16)$$

where $D_\epsilon := \min_S (p_S(\theta; f_\epsilon)(1 - p_S(\theta; f_\epsilon)))/({p'_S}^2(\theta; f_\epsilon))$. By (15), we have

$$\begin{aligned} \lim_{\epsilon \to 0} p_S(\theta; f_\epsilon) &= \lim_{\epsilon \to 0} \int_S f_\epsilon(x-\theta)\, dx \\ &= \int_S f_0(x-\theta)\, dx = p_S(\theta; f_0) \end{aligned}$$

and similarly, $\lim_{\epsilon \to 0} p'_S(\theta; f_\epsilon) = p'_S(\theta; f_0)$. Therefore, we obtain from the above two equalities that

$$\lim_{\epsilon \to 0} D_\epsilon = \min_S \frac{p_S(\theta; f_0)(1 - p_S(\theta; f_0))}{{p'_S}^2(\theta; f_0)} := D_0. \quad (17)$$

Thus, once we show

$$D_0 \geq U^2/4 \quad (18)$$

then Lemma 4 follows from (16)–(18).

[1]Such $\phi_\epsilon(x)$ obviously exists. An example is $\phi_\epsilon(x) = (\int_{-\epsilon}^{\epsilon} \psi_\epsilon(x) dx)^{-1} \psi_\epsilon(x)$, where

$$\psi_\epsilon(x) = \begin{cases} \exp(\frac{1}{x^2 - \epsilon^2}), & \text{if } |x| \leq \epsilon \\ 0, & \text{otherwise.} \end{cases}$$

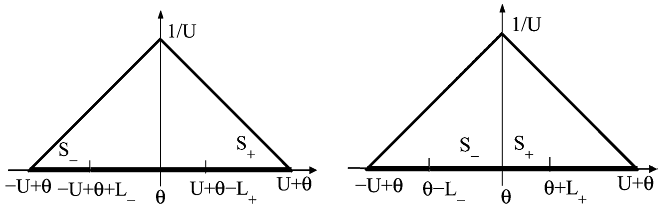

Fig. 4. Illustration of two cases of S_+ and S_- that minimize $p_S(\theta; f_0)(1 - p_S(\theta; f_0))$ when L_+, L_- are fixed.

It remains to prove (18). It is easy to see that $x_k \in [-U + \theta, U + \theta]$ for all k. Consider any measurable set $S \subseteq \mathbb{R}$. Since $f_0(x - \theta)$ has support in $[-U + \theta, U + \theta]$, we only need to consider the part of S that is in this interval. Let $S_+ = S \cap [\theta, U + \theta]$ and $S_- = S \cap [-U + \theta, \theta]$. Also, let L_+ and L_- denote the Lebesque measures of S_+ and S_-, respectively. It follows that

$$p_S(\theta; f_0) = \int_{S_+} f_0(x - \theta)\, dx + \int_{S_-} f_0(x - \theta)\, dx \quad (19)$$

and

$$p'_S(\theta; f_0) = -\int_S f'_0(x) dx = \frac{1}{U^2}(L_+ - L_-) \quad (20)$$

since according to (14), $f'_0(x) = 1/U^2$ if $x \in S_-$, and $f'_0(x) = -1/U^2$ if $x \in S_+$.

When L_+ and L_- are fixed, $p'_S(\theta; f_0)$ is also fixed [see (20)]. Thus, minimizing (over S) the middle term in (17) to obtain D_0 is equivalent to minimizing $p_S(\theta; f_0)(1 - p_S(\theta; f_0))$. Since the sizes of S_- and S_+ are fixed, the minimum of $p_S(\theta; f_0)(1 - p_S(\theta; f_0)$ is reached only when $p_S(\theta; f_0)$ is either a minimum or a maximum. Since $f_0(x)$ is monotonically increasing on $[-U, 0]$ and monotonically decreasing over $[0, U]$, it follows that $p_S(\theta; f_0)(1 - p_S(\theta; f_0))$ reaches the minimum only if (see Fig. 4)

1) $S_- = [-U + \theta, -U + \theta + L_-]$, $S_+ = [U + \theta - L_+, U + \theta]$ (left plot), or
2) $S_- = [\theta - L_-, \theta]$, $S_+ = [\theta, \theta + L_+]$(right plot).

In case 1), it follows from (19) that

$$p_S(\theta; f_0) = 1 - \frac{L_+^2 + L_-^2}{2U^2}. \quad (21)$$

Substituting (20) and (21) into the definition of D_0 in (17), we obtain that

$$D_0 = \min_{L_+, L_-} \left[g(L_+, L_-) := \frac{\left(1 - \frac{L_+^2 + L_-^2}{2U^2}\right) \frac{L_+^2 + L_-^2}{2U^2}}{\frac{1}{U^4}(L_+ - L_-)^2} \right]. \quad (22)$$

Manipulating the expression in (22) gives

$$\begin{aligned} g(L_+, L_-) &= \frac{U^2}{4} \frac{\left[2U^2 - (L_+^2 + L_-^2)\right](L_+^2 + L_-^2)}{U^2(L_+ - L_-)^2} \\ &\geq U^2/4 \end{aligned} \quad (23)$$

for any $0 \leq L_+, L_- \leq U$ since

$$\begin{aligned} &[2U^2 - (L_+^2 + L_-^2)](L_+^2 + L_-^2) - U^2(L_+ - L_-)^2 \\ &= (U^2 - L_+^2)L_+^2 + (U^2 - L_-^2)L_-^2 + 2L_+L_-(U^2 - L_+L_-) \geq 0. \end{aligned}$$

Combining (22) and (23) gives (18).

In case 2), a similar calculation shows that

$$\begin{aligned} p_S(\theta; f_0) &= \frac{L_+ + L_-}{U} - \frac{L_+^2 + L_-^2}{2U^2} \\ &= 1 - \frac{(U - L_+)^2 + (U - L_-)^2}{2U^2} \\ &= 1 - \frac{\alpha^2 + \beta^2}{2U^2} \end{aligned}$$

where $\alpha = U - L_+, \beta = U - L_-$. Plugging the above equation and (20) into (17), we obtain

$$\begin{aligned} D_0 &= \min_{\alpha, \beta} \frac{\left(1 - \frac{\alpha^2 + \beta^2}{2U^2}\right) \frac{\alpha^2 + \beta^2}{2U^2}}{\frac{1}{U^4}(\alpha - \beta)^2} \\ &\overset{(a)}{=} \min_{\alpha, \beta} g(\alpha, \beta) \overset{(b)}{\geq} \frac{U^2}{4} \end{aligned}$$

where (a) follows from the definition of g in (22), and (b) is due to (23) since $0 \leq \alpha, \beta \leq U$. This completes the proof of (18). ■

The upper bound (10) and Lemma 4 imply that the universal DES proposed in [9] and [10] achieves an MSE that is optimal within a constant factor of 4 if the range of the parameter's uncertainty is much smaller than the dynamical range of the noise, i.e., $V \ll U$.

References

[1] D. A. Castanon and D. Teneketzis, "Distributed estimation algorithms for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 30, no. 5, pp. 418–425, May 1985.
[2] 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. Autum. Control*, vol. 27, no. 4, pp. 799–813, Aug. 1982.
[3] Z. Chair and P. K. Varshney, "Distributed Bayesian hypothesis testing with distributed data fusion," *IEEE Trans. Syst., Man, Cybern.*, vol. 18, no. 5, pp. 695–699, Sep.–Oct. 1988.
[4] Z.-Q. Luo and J. N. Tsitsiklis, "Data fusion with minimal communication," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1551–1563, Sep. 1994.
[5] K. Liu, H. El-Gamal, and A. Sayeed, "On optimal parametric field estimation in sensor networks," in *Proc. IEEE/Signal Process. Soc. 13th Workshop Statistical Signal Processing*, Jul. 2005, pp. 1170–1175.
[6] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.
[7] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, Part I: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, Mar. 2006.
[8] V. Megalooikonomou and Y. Yesha, "Quantizer design for distributed estimation with communication constraints and unknown observation statistics," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 181–184, Feb. 2000.
[9] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2210–2219, Jun. 2005.
[10] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, Part II: Unknown pdf," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2784–2796, Jul. 2006.
[11] S. M. Kay, *Fundamentals of Statistical Signal Process.: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[12] 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.
[13] S. K. Berberian, *Fundamentals of Real Analysis*. New York: Springer-Verlag, 1998.