Distributed Adaptive Quantization and Estimation for Wireless Sensor Networks

Hongbin Li
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ 07030, USA
E-mail: Hongbin.Li@stevens.edu

***Abstract*— We consider distributed parameter estimation in a wireless sensor network, where due to bandwidth constraint, all sensor nodes have to quantize their observations and send quantized data to a fusion center. We consider the case where each sensor can send only one bit of information. In such a case, the achievable estimation performance is critically dependent on the choice of the one-bit quantizer used at the sensor nodes to perform quantization; it is also known that a fixed quantizer does not perform well, in particular when the quantization threshold is away from the unknown parameter to be estimated. In this paper, we propose a new distributed adaptive quantization scheme by which each individual sensor node dynamically adjusts the threshold of its quantizer based on earlier transmissions from other sensor nodes. We develop the maximum likelihood estimator (MLE) and derive the Cramér-Rao bound (CRB) associated with our distributed adaptive quantization scheme. Numerical results show that our approach does not suffer from the drawback of the fixed quantization approach and outperforms the latter.**

***Index Terms*— Wireless sensor networks, distributed estimation, adaptive quantization.**

I. INTRODUCTION

Consider a wireless sensor network that is composed of N spatially distributed sensor nodes. Each sensor node makes a noisy observation of an unknown parameter θ that is described by

$$x_n = \theta + w_n, \quad n = 1, 2, \ldots, N, \tag{1}$$

where N denotes the number of sensors and w_n the sensor noise which is assumed independent and identically distributed with respect to n. The problem of interest is to estimate θ from the observed signals $\{x_n\}$.

The traditional approach is based on *centralized* processing. In particular, the *unquantized* sensor data x_n are first transmitted to a fusion center (FC), and then a centralized estimation algorithm is run at the FC to find an estimate of θ, such as the sample mean estimator:

$$\hat{\theta} = \frac{1}{N} \sum_{n=1}^{N} x_n. \tag{2}$$

The above approach, however, is bandwidth inefficient. In a wireless sensor network, all sensor nodes have to share the communication bandwidth and, in addition, their battery life is limited. As such it may be impractical to send all unquantized data back to the FC.

This work was supported in part by the National Science Foundation under Grant CCF-0514938 and the US Army ARDEC.

An alternative approach is to utilize *distributed* processing, whereby sensor nodes first quantize x_n and send back quantized data which are next used at the FC to form an estimate of θ. A number of studies have considered such a distributed estimation approach, including stochastic methods that model θ as a random parameter and require knowledge of the joint distribution of θ and the observed signals (see, e.g., [1]), as well as deterministic methods that model θ as a deterministically unknown parameter. The latter can be further classified into methods that require knowledge of the conditional distribution of x_n (conditioned on θ), i.e., the likelihood function, (e.g., [2], [3]), and methods that do not (e.g., [4]). Details of these distributed estimation methods are found in [1], [2], [3], [4] and references therein.

This paper addresses the problem of distributed estimation for wireless sensor networks that have the stingiest bandwidth constraint: each sensor is allowed to transmit only 1 bit of information to the FC. The achievable estimation performance at the FC can be shown to critically depend on the choice of the 1-bit quantizer that is used to quantize the data at each sensor node [3]. One strategy is to choose a *fixed* quantizer for all sensor nodes with a fixed quantization threshold τ [2]. The optimum choice of τ, however, depends on θ which is unknown. It is found that if τ is set away from θ, the best achievable estimation performance at the FC has an exponentially increasing estimation error in $|\tau - \theta|$ [3]. An alternative strategy is to use a set of thresholds $\{\tau_k\}$, and each τ_k is used in a fraction ρ_k of the N sensor nodes [3], in the hope that some of the thresholds are close to the unknown θ. The problem is that finding the best set of $\{\tau_k, \rho_k\}$ is involved.

To deal with the above difficulty, we present a new distributed adaptive quantization scheme by which each individual sensor node dynamically adjusts the threshold of its quantizer based on earlier transmissions from other sensor nodes. Our scheme is in essence a distributed Delta modulation technique, whereby each sensor node accumulates earlier transmissions from other sensor nodes, and uses the accumulated value s_{n-1} as the threshold for its 1-bit quantizer. The accumulated value s_{n-1} can be shown to converge (with respect to n) around the unknown θ. Based on our proposed adaptive quantization scheme, we develop the maximum likelihood estimator (MLE) that can be used at the FC to find the ML estimate of θ, and the Cramér-Rao bound (CRB) that tells about the best achievable estimation performance (among all unbiased estimators) for the proposed adaptive quantizer.

The rest of the paper is organized as follows. We first briefly review the fixed quantization scheme [2] and the associated MLE and CRB in Section II. Our distributed adaptive quantizer and the corresponding MLE and CRB are presented in Section III. Numerical results, comparisons, and discussions are contained in Section IV.

II. Fixed Quantization Approach

The fixed quantization approach is to apply a common threshold τ for all sensors and generate quantized data b_n as follows [2]:

$$b_n = \text{sgn}(x_n - \tau), \quad n = 1, 2, \ldots, N, \tag{3}$$

which are sent to the FC. The probability mass function (PMF) of the binary random variable b_n is given by

$$P(b_n; \theta) = [F_w(\tau - \theta)]^{(1+b_n)/2}[1 - F_w(\tau - \theta)]^{(1-b_n)/2}, \tag{4}$$

where $F_w(x)$ denotes the complementary cumulative density function (CDF) of w_n. Since $\{b_n\}$ are independent and identically distributed, the log-PMF or log-likelihood function is

$$\begin{aligned} L_{\text{FQ}}(\theta) &\triangleq \ln[P(b_1, \ldots, b_N; \theta)] \\ &= \sum_{n=1}^{N} \left\{ \left(\frac{1+b_n}{2}\right) \ln[F_w(\tau - \theta)] \right. \\ &\qquad \left. + \left(\frac{1-b_n}{2}\right) \ln[1 - F_w(\tau - \theta)] \right\}, \end{aligned} \tag{5}$$

where the subscript FQ is used to denote fixed quantization. The MLE is given by

$$\hat{\theta}_{\text{FQ}} = \arg\max_{\theta} L_{\text{FQ}}(\theta). \tag{6}$$

The CRB based on the above fixed quantization is [2] (also see [3]):

$$\text{CRB}_{\text{FQ}}(\theta) = \frac{F_w(\tau - \theta)[1 - F_w(\tau - \theta)]}{N p_w^2(\tau - \theta)}, \tag{7}$$

where $p_w(x)$ denotes the probability density function (PDF) of w_n. We see that $\text{CRB}_{\text{FQ}}(\theta)$ depends on the the threshold τ. Furthermore, it has been found that the CRB increases exponentially with $|\tau - \theta|$ [3].

III. Distributed Adaptive Quantization Approach

We first introduce our distributed adaptive quantization scheme, followed by our development of the MLE and CRB.

A. Adaptive Quantization

We assume that the sensors share the communication channel on a time-sharing basis (e.g., each sensor is polled by the FC), so that sensor 1 transmits first, followed by sensor 2, and so on and so forth. The 1-bit quantizer at sensor 1 uses a zero-threshold to generates b_1:

$$b_1 = \text{sgn}\{x_1\}. \tag{8}$$

Then, b_1 is sent (i.e. broadcast) to the FC as well as the other $N-1$ sensors. After receiving b_1, sensor 2 computes $s_1 = \Delta b_1$, where Δ is a parameter of user choice, and generates b_2:

$$b_2 = \text{sgn}\{x_2 - s_1\} \tag{9}$$

where Δ is a design parameter. In general, for sensor n, it first forms a cumulative sum:

$$s_{n-1} = \Delta \sum_{k=1}^{n-1} b_k, \tag{10}$$

and then it uses s_{n-1} as a threshold for quantization:

$$b_n = \text{sgn}\{x_n - s_{n-1}\}. \tag{11}$$

One can immediately recognize that the above process is reminiscent of the Delta modulation, but is implemented is a distributed fashion.

B. MLE

Different from the fixed quantization approach, the binary data bits $b_1, b_2, \ldots, b_N$ generated by our distributed adaptive quantization are no longer independent and identically distributed. The conditional PMF of b_n is given by

$$\begin{aligned} P(b_n|b_1, \ldots, b_{n-1}; \theta) &= [F_w(s_{n-1} - \theta)]^{(1+b_n)/2} \\ &\quad \times [1 - F_w(s_{n-1} - \theta)]^{(1-b_n)/2}. \end{aligned} \tag{12}$$

Using conditional probabilities, we can write the joint PMF of $b_1, b_2, \ldots, b_N$ as

$$\begin{aligned} &P(b_1, \ldots, b_N; \theta) \\ &= P(b_1; \theta) P(b_2|b_1; \theta) \ldots P(b_N|b_1, \ldots, b_{N-1}; \theta) \\ &= \prod_{n=1}^{N} P(b_n|b_1, \ldots, b_{n-1}; \theta) \\ &= \prod_{n=1}^{N} P(b_n|s_{n-1}; \theta) \end{aligned} \tag{13}$$

It follows from (12) and (13) that the log-likelihood function is given by

$$\begin{aligned} L_{\text{AQ}}(\theta) = \sum_{n=1}^{N} &\left\{ \left(\frac{1+b_n}{2}\right) \ln[F_w(s_{n-1} - \theta)] \right. \\ &\left. + \left(\frac{1-b_n}{2}\right) \ln[1 - F_w(s_{n-1} - \theta)] \right\}, \end{aligned} \tag{14}$$

where the subscript AQ is used to denote our adaptive quantization. As such, the MLE is given by

$$\hat{\theta}_{\text{AQ}} = \arg\max_{\theta} L_{\text{AQ}}(\theta). \tag{15}$$

C. CRB

Noting that $F'_w(x) \triangleq \frac{\partial F_w(x)}{\partial x} = -p_w(x)$, we can quickly verify that the first- and second-order derivatives of $L_{\text{AQ}}(\theta)$ are

$$\begin{aligned} \frac{\partial L_{\text{AQ}}(\theta)}{\partial \theta} = \sum_{n=1}^{N} &\left\{ \left(\frac{1+b_n}{2}\right) \frac{p_w(s_{n-1} - \theta)}{F_w(s_{n-1} - \theta)} \right. \\ &\left. - \left(\frac{1-b_n}{2}\right) \frac{p_w(s_{n-1} - \theta)}{1 - F_w(s_{n-1} - \theta)} \right\}, \end{aligned} \tag{16}$$

$$
\begin{aligned}
&\frac{\partial^2 L_{\text{AQ}}(\theta)}{\partial \theta^2} \\
=& \sum_{n=1}^{N}\Bigg\{\left(\frac{1+b_n}{2}\right)\left(\frac{p_w'(s_{n-1}-\theta)}{F_w(s_{n-1}-\theta)}-\frac{p_w^2(s_{n-1}-\theta)}{F_w^2(s_{n-1}-\theta)}\right) \\
&-\left(\frac{1-b_n}{2}\right)\Bigg(\frac{p_w'(s_{n-1}-\theta)}{\big[1-F_w(s_{n-1}-\theta)\big]} \\
&+\frac{p_w^2(s_{n-1}-\theta)}{\big[1-F_w(s_{n-1}-\theta)\big]^2}\Bigg)\Bigg\} \\
\triangleq& \sum_{n=1}^{N} A(b_n, s_{n-1}, \theta),
\end{aligned}
\tag{17}
$$

where $p_w'(x) \triangleq \frac{\partial p_w(x)}{\partial x}$. The Fisher information for the estimation problem is given by (e.g., [5])

$$
\begin{aligned}
J_{\text{AQ}}(\theta) =& E\left\{\frac{\partial^2 L_{\text{AQ}}(\theta)}{\partial \theta^2}\right\} \\
=& \sum_{n=1}^{N} E_{b_n, s_{n-1}}\{A(b_n, s_{n-1}, \theta)\},
\end{aligned}
\tag{18}
$$

where $E_{b_n, s_{n-1}}$ denotes the expectation with respect to the joint distribution of b_n and s_{n-1}. Since

$$
P(b_n, s_{n-1}; \theta) = P(s_{n-1}; \theta) P(b_n | s_{n-1}; \theta), \tag{19}
$$

we can write

$$
J_{\text{AQ}}(\theta) = \sum_{n=1}^{N} E_{s_{n-1}}\Big\{E_{b_n|s_{n-1}}\{A(b_n, s_{n-1}, \theta)\}\Big\}, \tag{20}
$$

where $E_{b_n|s_{n-1}}$ denotes the expectation with respect to the conditional distribution $P(b_n|s_{n-1};\theta)$. From (12), we have

$$
\begin{aligned}
& E_{b_n|s_{n-1}}\{A(b_n, s_{n-1}, \theta)\} \\
=& F_w(s_{n-1}-\theta) A(1, s_{n-1}, \theta) \\
&+ [1-F_w(s_{n-1}-\theta)] A(-1, s_{n-1}, \theta).
\end{aligned}
\tag{21}
$$

To carry out the outer expectation in (20), we need find out the distribution of s_{n-1}, which is addressed in the next section. Once we have the Fisher information $J_{\text{AQ}}(\theta)$, the CRB is given by

$$
\text{CRB}_{\text{AQ}}(\theta) = -\frac{1}{J_{\text{AQ}}(\theta)}. \tag{22}
$$

D. *PMF of s_n*

Note that s_n is a random walk process with time-varying probabilities of the increment of Δ and $-\Delta$. Also note that $s_1 \in \{\pm\Delta\}$, $s_2 \in \{-2\Delta, 0, 2\Delta\}$, $s_3 \in \{-3\Delta, -\Delta, \Delta, 3\Delta\}$, $s_4 \in \{-4\Delta, -2\Delta, 0, 2\Delta, 4\Delta\}$, and so on and so forth. In general, we have

$$
\begin{aligned}
& s_{2k-1} \in \{\pm\Delta, \ldots, \pm(2k-1)\Delta\}, \quad k = 1, 2, \ldots \\
& s_{2k} \in \{0, \pm 2\Delta, \ldots, \pm 2k\Delta\}, \quad k = 1, 2, \ldots
\end{aligned}
\tag{23}
$$

Although finding a closed-form expression for the PMF of s_n seems cumbersome, it can be computed rather straightforwardly by recursive calculation. To facilitate presentation, we introduce the following notation

$$
P_{i,j} \triangleq \Pr(s_i = j\Delta). \tag{24}
$$

To initialize the recursion, we first compute the distribution of s_n for $n = 2k-1 = 1$ (i.e., $k = 1$):

$$
\begin{aligned}
P_{1,1} &= \Pr(x_1 > 0) = F_w(-\theta) \\
P_{1,-1} &= \Pr(x_1 < 0) = 1 - F_w(-\theta)
\end{aligned}
\tag{25}
$$

Then, the recursion for even n, i.e., $n = 2k$, $k = 1, 2, \ldots$, is given by

$$
\begin{aligned}
P_{2k,2l} =& P_{2k-1,2l+1} \Pr(x_{2k} < (2l+1)\Delta) \\
&+ P_{2k-1,2l-1} \Pr(x_{2k} > (2l-1)\Delta) \\
=& \big[1 - F_w((2l+1)\Delta - \theta)\big] P_{2k-1,2l+1} \\
&+ F_w((2l-1)\Delta - \theta) P_{2k-1,2l-1} \\
& \qquad l = -k, -k+1, \ldots, k-1, k,
\end{aligned}
\tag{26}
$$

where the boundary probabilities are zero: $P_{2k-1,2k+1} = P_{2k-1,-2k-1} = 0$. In essence, the above equation says that if the accumulation s_{2k} takes a value of $2l\Delta$, it is either because the previous accumulation s_{2k-1} takes a value of $(2l+1)\Delta$ and in conjunction the observation x_{2k} is less than $(2l+1)\Delta$, or s_{2k-1} takes a value of $(2l-1)\Delta$ and in conjunction x_{2k} is greater than $(2l-1)\Delta$.

Likewise, the recursion for odd n, i.e., $n = 2k-1$, $k = 2, 3, \ldots$, is given by

$$
\begin{aligned}
P_{2k-1,2l-1} =& P_{2k-2,2l} \Pr(x_{2k-1} < 2l\Delta) \\
&+ P_{2k-2,2l-2} \Pr(x_{2k-1} > (2l-2)\Delta) \\
=& \big[1 - F_w(2l\Delta - \theta)\big] P_{2k-2,2l} \\
&+ F_w((2l-2)\Delta - \theta) P_{2k-2,2l-2} \\
& \qquad l = -k+1, -k+2, \ldots, k-1, k,
\end{aligned}
\tag{27}
$$

where the boundary probabilities are zero: $P_{2k-2,2k} = P_{2k-2,-2k} = 0$.

With the probabilities $P_{i,j}$ obtained recursively as above, the PMF of s_n for odd and even values of n can be expressed as

$$
\begin{aligned}
P(s_{2k-1}; \theta) &= \sum_{l=-k+1}^{k} P_{2k-1,2l-1} I(s_{2k-1} - (2l-1)\Delta), \\
P(s_{2k}; \theta) &= \sum_{l=-k}^{k} P_{2k,2l} I(s_{2k} - 2l\Delta),
\end{aligned}
\tag{28}
$$

where $I(x)$ denotes the indicator funciton:

$$
I(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases} \tag{29}
$$

IV. Numerical Results and Discussions

To illustrate the performance of the proposed distributed adaptive quantization and estimation scheme, we consider the case where the sensor noise $\{w_n\}$ are independent and identically distributed Gaussian random variables with zero mean and variance $\sigma^2 = 1$. We compare the fixed quantization approach described in Section II, our adaptive quantization approach, and the clairvoyant approach that uses unquantized

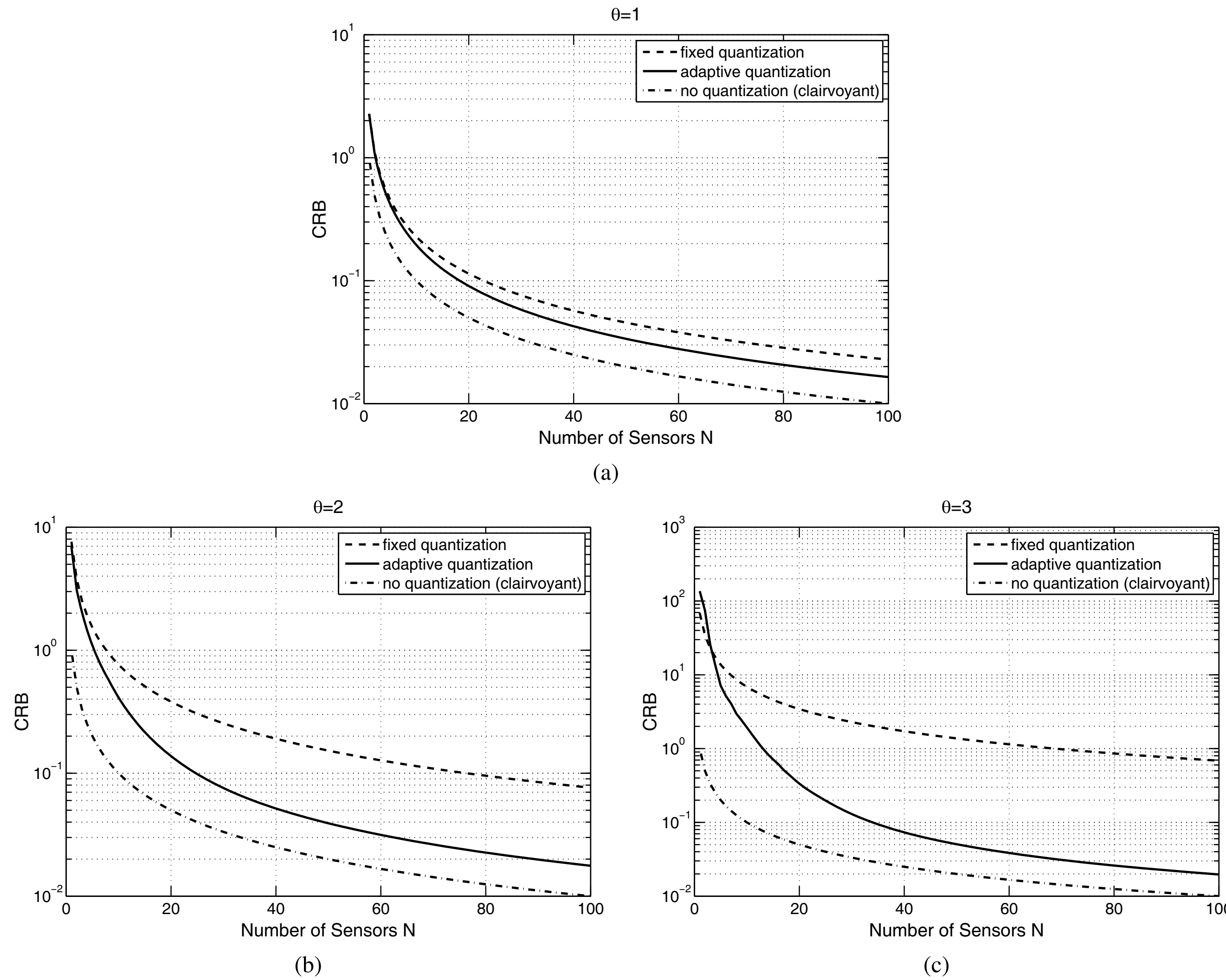

Fig. 1. CRBs of distributed estimation with fixed, adaptive, and no quantization versus N, the number of sensor nodes N, when $\sigma^2 = 1$, and $\Delta = 0.1$. (a) $\theta = 1$. (b) $\theta = 2$. (c) $\theta = 3$.

data (cf. the sample mean estimator (1)). Although not desirable due to its bandwidth inefficiency, the clairvoyant approach provides a benchmark (lower bound) on the achievable performance with quantized data. The CRB for the clairvoyant approach is well known, which is given by

$$\mathrm{CRB}_{\mathrm{NQ}}(\theta) = \frac{\sigma^2}{N}, \tag{30}$$

where the subscript NQ denotes that no quantization is used.

Figures 1(a) to 1(c) compare the CRB of the above three approach when $\theta = 1, 2$, and 3, respectively. For the fixed quantization approach, we set the threshold $\tau = 1$, and for our adaptive quantization approach, we choose $\Delta = 0.1$. As we can see, the fixed quantization approach is sensitive to the value of θ or, equivalently, the value of τ; as the two become more apart from each other (even not too far apart), the performance of the fixed quantization approach degrades significantly.

On the other hand, our adaptive quantization scheme does not have the above problem. In addition, in all three cases considered, it outperforms the fixed quantization scheme and is closer to the clairvoyant approach.

REFERENCES

[1] J. Gubner, "Distributed estimation and quantization," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1456–1459, July 1993.

[2] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 978–1002, March 2001.

[3] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, Part I Gaussian PDF," *IEEE transactions on signal processing*, vol. 54, no. 3, pp. 1131–1143, March 2006.

[4] Z. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2210–2219, June 2005.

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