Hindawi Publishing Corporation
Journal of Sensors
Volume 2014, Article ID 368643, 11 pages
http://dx.doi.org/10.1155/2014/368643

Research Article

Distributed Binary Quantization of a Noisy Source in Wireless Sensor Networks

Sahar Movaghati and Masoud Ardakani

Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G 2V4

Correspondence should be addressed to Sahar Movaghati; movaghat@ualberta.ca

Received 30 October 2013; Revised 21 May 2014; Accepted 18 July 2014; Published 12 August 2014

Academic Editor: Mike McShane

In distributed (decentralized) estimation in wireless sensor networks, an unknown parameter must be estimated from some noisy measurements collected at different sensors. Due to limited communication resources, these measurements are typically quantized before being sent to a fusion center, where an estimation of the unknown parameter is calculated. In the most stringent condition, each measurement is converted to a single bit. In this study, we propose a distributed quantization scheme which is based on single-bit quantized data from each sensor and achieves high estimation accuracy at the fusion centre. We do this by designing some local binary quantizers which define a multithreshold quantization rule for each sensor. These local binary quantizers are initially designed so that together they mimic the functionality of a multilevel quantizer. Later, their design is improved to include some error-correcting capability, which further improves the estimation accuracy from the sensors' binary data. The distributed quantization formed by such local binary quantizers along with the proper estimator proposed in this work achieves better performance, compared to the existing distributed binary quantization methods, specially when fewer sensors with low measurement noise are available.

1. Background and Introduction

The distributed quantization and estimation problem involves estimating an unknown parameter from several noisy measurements [1]. This problem appears in many distributed sensing systems in a wide range of applications, such as medical imaging, environmental monitoring, and target tracking. In many of these cases a central processing unit or fusion center (FC) has to determine an unknown signal from some noisy observations received from distributed sensors. In wireless sensor networks (WSNs), sensors encapsulate their measurement readings inside the communication packets, according to the standard protocols, such as IEEE 802.15.4 and ZigBee, and transmit these packets to the FC. These protocols take care of the lower layers' services (the term "layer" refers to the OSI model), such as routing and dealing with errors caused by channel noise. In this study, we assume that the WSN communication protocols provide error-free communication services for us. We focus our interest on the measurement data inside the packets and the advantage of optimum quantizer design at the application layer to improve estimation performance at the FC.

Normally, due to system constraints, such as shortage of energy and bandwidth resources, distributed observations need to be compressed before being sent to the FC [2]. Quantizing the observations reduces the transmission load in the network; however, it implies some information loss, reducing the estimation accuracy. Therefore, quantizer and estimator design is an important problem in distributed estimation, which has attracted considerable attention in the recent literature [3–15].

Suppose N noisy measurements of a random parameter X have the form $Z_n = X + w_n$, $1 \le n \le N$, where w_n are independent zero mean measurement noises. If all the measurements are available in the estimator without any distortion, an unbiased estimation with error variance as low as $(\sum 1/\sigma_n^2)^{-1}$ can be achieved, where σ_n is the variance of the measurement noise for the nth observation [3]. Because of communication constraints in a WSN, the measurements are quantized before transmission using a local quantization rule;

that is, $y_n = \mathcal{Q}_n(Z_n)$. The goal is then to design these local quantization rules, as well as the procedure to be used for estimating X from the quantized data.

Some existing distributed estimation algorithms are based on applying a uniform quantizer to quantize each analog measurement into a few bits [3–8, 16]. In all these algorithms, the number of quantization bits for each analog measurement is decided based on the signal-to-noise ratio (SNR). For example, in [3], the number of quantization bits for the nth measurement, $1 \leq n \leq N$, is $\log_2(1 + V/\sigma_n)$. In some applications such as WSNs, this means the sensors with higher SNR, that is, smaller σ_n, will have to send more bits; hence, they consume more power for data transmission. Consequently, the better sensors become more exhausted and die more quickly, which in turn reduces the long-term performance of the estimation task.

Another technique has been proposed by [9] which is based on adding some deterministic or random control input to the observation data, prior to the quantization. Therefore, the quantized value of the nth observation is $\tilde{y}_n = \mathcal{Q}_n(Z_n+v_n)$, where v_n is a deterministic or random signal. To find the optimal control input, a metric based on the Cramer-Rao lower bound (CRLB), or the Fisher information, is optimized [9, 17, 18].

Under severely stringent bandwidth or power conditions it is preferred to quantize each measurement into only one bit. Distributed estimation based on local binary quantizations has been studied in [4, 10–14, 19]. In [4], the local binary quantization is performed by comparing the analog measurement value to a fixed threshold in the middle of the analog data range. The estimation performance of a set of local binary quantizers is studied in [13] for asymptotic condition; that is, $N \to \infty$.

The local binary quantizers with a single fixed threshold are not very efficient, specially when the number of measurements N is small and the measurement noise is low. To improve the performance, [10, 11] suggest adaptive thresholds, which are sequentially adjusted according to the previously generated bits. In their approach, the threshold used for quantizing the nth measurement value is adjusted based on the previous $(n-1)$ bits. In [11], to find the optimum threshold for the nth measurement a maximum likelihood (ML) estimator must be derived every time based on the n previous generated bits and the n previous thresholds.

The estimation based on local binary-quantized observations is further studied in [12], where they consider different thresholds $\boldsymbol{\tau} = \{\tau_k, k \in \mathbb{Z}\}$ on the analog range to assure that there will always be a threshold close to the true parameter X. Out of the N observations, N_k measurements are quantized according to threshold τ_k, where $\sum_k N_k = N$. Through CRLB, they find the set of optimal threshold values τ_k^* and their associated frequencies N_k^*.

In the above distributed estimation methods based on binary data, the generation of each bit is performed using a single threshold. A multithreshold quantization scheme was suggested by [14]. Based on the distributed estimation method in [14], each measurement is used to estimate one of the bits in the binary representation of the unknown parameter. According to [14], if N measurements are available, $N/2$ of them estimate the first bit of the unknown parameter, $N/4$ estimate the second bit, and so on. So, the bth bit in the binary representation of the unknown parameter is estimated $N/2^b$ times. Hence, the value for that bit is determined by taking the average among the $N/2^b$ binary values, and, consequently, X is estimated by combining the final value of the individually estimated bits.

In this work we propose a new method for distributed binary quantization to improve the estimation performance at the FC. To do that, we design local quantizers to compress each local measurement to a single bit and suggest a centralized estimator to infer the unknown from those bits. Therefore, our goal in this work is to (i) first formulate the distributed quantization as a set of N local binary quantizers (local-Qs), (ii) jointly design these local-Qs to find the optimal set of N local-Qs, which can maximize the estimation accuracy of X, and (iii) find a centralized algorithm for the FC that combines these binary-quantized data to form an accurate estimate of X.

Our distributed quantization method benefits from multithreshold local quantizers to achieve high estimation accuracy. Compared to the binary quantization algorithms in [4, 11, 12, 14], our method achieves a better MSE performance at the FC, specially when limited number of sensors are available, with low measurement noise. This is while minimum computation and transmission load is imposed on the sensors.

The rest of this paper is organized as follows. In Section 2, the detailed setup of the problem and the required definitions and assumptions are provided. In Section 3, the design of a distributed quantizer based on different binary numeral systems is introduced. Section 4 proposes optimal local-Qs to improve the estimation performance. In Section 5, the appropriate decoder/estimator to be used in the FC is formulated. Finally, in Section 6, the simulation results are shown for performance evaluation.

2. Problem Setup

Suppose X is a random scalar distributed according to the pdf $P(X)$. A number of noisy measurements of X are observed as

$$Z_n = X + w_n \quad 1 \leq n \leq N, \tag{1}$$

where w_n for $1 \leq n \leq N$ are i.i.d. additive noise. In this work, we assume that $P(X)$ is a uniform distribution in the interval $[-V, V]$, and the measurement noise is Gaussian with zero mean and variance σ^2. It is straightforward to modify the proposed distributed estimation method to work with other signal and noise pdfs. An example for Gaussian X is discussed in Section 4.3.

Considering the most stringent scenario, only single-bit data transmission is allowed for sending each measurement to the FC. Therefore, each measurement Z_n is quantized to a bit y_n, according to a local binary quantization rule $\mathcal{Q}_n(\cdot)$.

Note that data transmission over nonideal channels can involve some errors due to channel noise, which can be dealt with separately using error detection and error correction

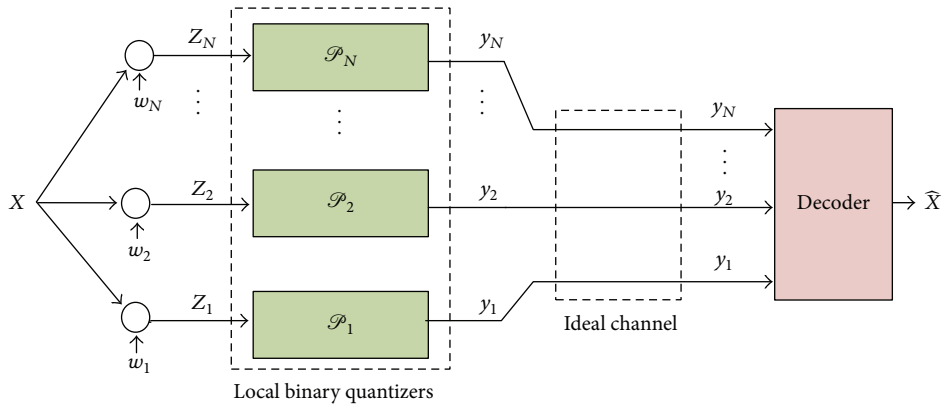

Figure 1: The overall distributed estimation method.

techniques implemented in the WSN communication protocols. In this study, we do not consider any noise or error added by the channel; that is, the communication channel is assumed to be error-free. Therefore, all errors discussed here are due to the measurement or quantization noise, not the channel noise. Thus, the goal is to design a set of local binary quantization rules $\mathscr{Q}_n(\cdot)$ to be used for quantizing the observations Z_n and also to design an estimation algorithm that combines the quantized binary data y_n to form an accurate estimate of X at the FC, that is, $\widehat{X}$ (see Figure 1).

Assume that the parameter range $[-V, V]$ is "partitioned" into L "divisions" δ^l; $1 \le l \le L$. Compressing the real-valued measurement Z_n to a bit y_n can be performed through introducing a local-Q $\mathscr{P}_n$, which is a function mapping each division δ^l, $1 \le l \le L$ to a binary value, that is, 0 or 1. Therefore, the quantization of each measurement can be described as

$$y_n = \mathscr{P}_n(l) \quad \text{iff } Z_n \in \delta^l. \tag{2}$$

Each local-Q is an ordered sequence of L binary values, that is, a binary vector of length L. Wherever the binary value alters between two successive divisions the threshold between those divisions is regarded as an edge in the local-Q; see Figure 2. Therefore, a local-Q denoted by $\mathscr{P}_n$ is associated with a set of edges, that is, $\{e_n^1, e_n^2, \ldots, e_n^{\epsilon_n}\}$, where ϵ_n is the number of edges in $\mathscr{P}_n$. These edges define a new partitioning of $[-V, V]$ into $\epsilon_n + 1$ "cells"; see Figure 2. Similar to multiresolution quantization [20, 21], different quantizers have different cell sizes (resolution); however, in our method all local quantizers have binary output.

It must be mentioned that, due to the additive Gaussian noise, the analog measurements Z_n are in the range $(-\infty, \infty)$. However, since the desired parameter which must be estimated in the FC is within the range $[-V, V]$, the local-Qs are defined over this range. Therefore, if Z_n is in $(-\infty, -V)$ or (V, ∞) it will be mapped to V or $-V$, respectively.

3. Distributed Quantization

In this section we describe how a set of local-Qs are designed. For now, assume that we have $N = B$ noiseless observations of X. In other words, each observation equals X, which

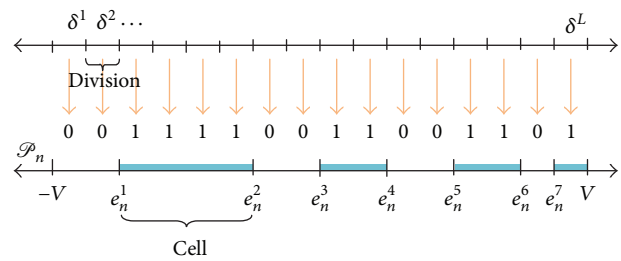

Figure 2: A sample local-Q.

must be quantized into one bit, y_b, $1 \le b \le B$. Since the combinations of B bits can specify maximum of $L = 2^B$ values, one can generate y_b; $1 \le b \le B$, so that together they identify the division of X among L distinct divisions within the range $[-V, V]$. In order to achieve that, B local-Qs must be appropriately designed.

Consider the analog range $[-V, V]$ to be partitioned into L equal-length divisions (considering equal-length divisions is intuitive when we have a uniform source. For nonuniform sources a nonequal partitioning must be considered; see Section 4.3), δ^l, $1 \le l \le L$. To identify to which division among the total L divisions X belongs, B bits are provided for the FC, using B local-Qs; that is, $\mathscr{P}_b$, $1 \le b \le B$. These local-Qs can be designed so that, for each division on the range $[-V, V]$, the B bits together make the B-digit binary label/word assigned to that division. Figure 3 illustrates $B = 4$ local-Qs that together assign 4-digit binary words to the 16 divisions. This assignment is conducted using the "natural" binary numeral system. In other words, reading vertically from left to right, the first division is assigned 0000, the second is assigned 0001, the third is assigned 0010, and so on. If a different binary numeral systems, referred to as the "labeling" scheme in this work, is used for labeling the divisions, a different set of B local-Qs would result. This is why the local-Qs in Figure 3 are identified as $\mathscr{P}_b^{\mathrm{Nr}}$, $1 \le b \le B$, where the superscript "Nr" stands for the natural labeling. If instead of the natural labeling, Gray labeling is used, a set of local-Qs, that is, $\mathscr{P}_b^{\mathrm{Gr}}$, will be produced; see Figure 4. Clearly, other labeling schemes can be considered.

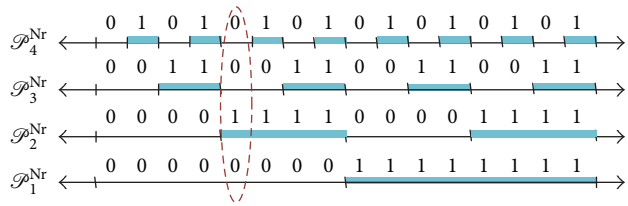

Figure 3: The natural basis local-Qs for $B = 4$.

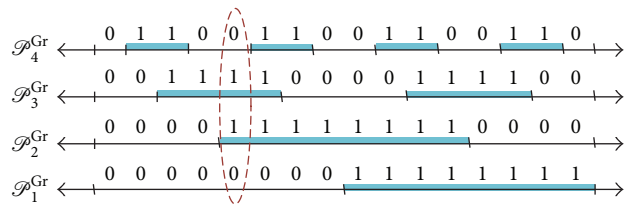

Figure 4: The Gray basis local-Qs for $B = 4$.

Similarly, for the noisy scenario, we can quantize each local noisy measurement Z_b, $1 \leq b \leq B$, using one of the local-Qs $\mathscr{P}_b$, $1 \leq b \leq B$. Depending on where Z_b falls, $\mathscr{P}_b$ decides the bth bit. Since there are $N = B$ measurements, the bth measurement provides the bth digit of the B-digit binary word, for $1 \leq b \leq B$. At the FC, the B bits are used to remake the B-digit binary word and thereby construct $\widehat{X}$.

The above described method can be viewed as a *distributed quantizer*, since a uniform L-level quantization is implemented using B separate binary quantizers. In a noiseless scenario, the performance of this distributed quantizer is the same as the centralized scalar quantizer, meaning that its MSE is $\Delta^2/12$, where $\Delta = 2V/L$. Moreover, when all observations are equal to X, using either $\mathscr{P}_b^{\mathrm{Nr}}$, $1 \leq b \leq B$, $\mathscr{P}_b^{\mathrm{Gr}}$, $1 \leq b \leq B$, or any other set of B local-Qs based on different labelings, results in the same MSE, that is, $\Delta^2/12$. However, in the presence of the measurement noise, any of the B bits of the distributed quantizer might be quantized wrong, leading to a larger estimation error in the FC. In this case, the overall MSE at the FC is affected totally differently by different local-Qs and also depends on the position of the unknown X. Thus, writing the MSE in closed form will not have a simple or insightful structure. In Section 5.3, we derive the CRLB to compare to the simulation results of our method.

One way to reduce the final MSE is to use better labeling schemes, which results in different local-Qs. In Section 4.1, the optimal local-Qs for achieving the best performance are discussed. For uniform X, using the Gray labeling to design the local-Qs achieves the optimal estimation performance. However, this is not the case for other distributions of X. For Gaussian X, the optimal local-Qs are discussed in Section 4.3.

It is possible to further enhance the estimation when more observations are available. In the same noise level, when $N > B$ measurements are used, it is possible to estimate X more accurately, using the same $L = 2^B$-level quantizer as the basic quantizer. In such cases, the extra measurements can be used to repeat some of the binary digits and reduce the bit error rate of those bits, that is, a similar approach as [14]. However, instead of simple bit repetition, one can better employ the extra measurements to get the optimal estimation performance. This is further discussed in Section 4.2.

In the following sections, we introduce an algebraic approach to further explain the distributed quantization. Through that, we find the optimal local-Qs to get the best estimation performance when $N \geq B$. In particular, for $N > B$, an analogy between error-correcting codes and the distributed quantization is used.

4. Optimal Distributed Quantization

Assume that there are N independent noisy measurements from the unknown X. To achieve the best estimation, one needs to obtain the optimal set of local-Qs to be assigned to the analog measurements. For a fixed B, if $N = B$, the solution to the above problem results in the optimal labeling scheme. If $N > B$, the result is an error-correcting distributed quantization. These two cases are studied separately in the following sections.

4.1. Best Labeling Scheme, $N = B$. In Section 3, we discussed two sets of local-Qs, resulting from applying two different labeling systems, that is, natural and Gray. In this section we discuss other sets of local-Qs. It is worth mentioning again that if no measurement noise is present all of these sets of local-Qs will have the same performance. However, in the presence of measurement noise, they have different performances in terms of the estimation MSE.

To locate a variable in L distinct divisions (in the interval $[-V, V]$), B bits are needed, where $B = \log_2 L$. To generate such B bits, an eligible set of B local-Qs can be chosen from a variety of choices. As mentioned in Section 2, each local-Q is basically a binary vector of length L. Choosing B "eligible local-Qs" that can identify L different divisions is analogous to choosing B linearly independent binary vectors (please note that the algebraic calculations are done in GF (2)) of size L. The B "eligible local-Qs" are called the "basis local-Qs," for brevity.

The set of natural local-Qs in Section 3 gives one example set of basis local-Qs, namely, the "natural basis local-Qs" $\mathscr{P}_b^{\mathrm{Nr}}$, $1 \leq b \leq B$, Figure 3. Taking the natural basis local-Qs as the reference vectors and linearly combining them to generate a new set of B linearly independent vectors, we obtain a different set of basis local-Qs. The linear combinations of the B reference local-Qs into a new set of B linearly independent local-Qs can be represented by a $B \times B$ binary matrix of rank B. For example, for $B = 4$ a special combination of natural basis local-Qs can be shown by

$$G_{BB} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

which indicates a combination of the local-Qs as

$$\begin{aligned}\mathscr{P}_1^{\mathrm{Gr}} &= \mathscr{P}_1^{\mathrm{Nr}},\\ \mathscr{P}_2^{\mathrm{Gr}} &= \mathscr{P}_1^{\mathrm{Nr}} \oplus \mathscr{P}_2^{\mathrm{Nr}},\\ \mathscr{P}_3^{\mathrm{Gr}} &= \mathscr{P}_2^{\mathrm{Nr}} \oplus \mathscr{P}_3^{\mathrm{Nr}},\\ \mathscr{P}_4^{\mathrm{Gr}} &= \mathscr{P}_3^{\mathrm{Nr}} \oplus \mathscr{P}_4^{\mathrm{Nr}},\end{aligned} \tag{4}$$

where $\oplus$ means modulo-2 summation of the L-valued binary vectors representing the local-Qs. The new set of basis local-Qs obtained by the G_{BB} example in (3) is the set of Gray local-Qs (Figure 4).

In this work, we focus on the natural basis local-Qs as the original local-Q set and their linear combinations, in the form of G_{BB}, for producing new local-Q sets. By searching all matrices of type G_{BB}, one can find the best basis local-Qs with the lowest MSE, for each SNR level. As mentioned above, G_{BB} must be full rank. Moreover, a permutation of the vectors of G_{BB} results in the same set of local-Qs. Hence, the size of the search space is

$$\Omega_1(B) = \frac{1}{B!}\prod_{i=1}^{B}\left(2^B - 2^{i-1}\right). \tag{5}$$

4.2. Error-Correcting Local-Qs, $N > B$. As mentioned earlier, having $N > B$ measurements available, the estimation accuracy can be improved by repeating some of the local-Qs. However, a more effective approach is proposed here, by choosing new local local-Qs for the redundant measurements. Taking the B basis local-Qs of Section 3 as starting point, N local-Qs are obtained so that together they generate N-bit words to identify $L = 2^B$ divisions. Obviously, using N-bit words instead of B-bit words (as in Section 3) is a redundant way of labeling an L-division quantization. However, this redundancy enables some error correction possibility in estimating the division of X. This property is analogous to error-correcting property in channel coding, where some redundant bits are added to the B-bit messages to make N-bit codewords [22]. This happens by adding extra vectors to the generator matrix.

Among the N local-Qs used to quantize the N measurements, there must be at least B linearly independent local-Qs. Hence, the N local-Qs are produced by linear combination of a set of B basis local-Qs according to a $B \times N$ matrix of rank B. For example, consider the matrix

$$G_{BN} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1\\ 0 & 1 & 1 & 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1 & 1 & 0 & 1\\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}, \tag{6}$$

which is used to produce N local-Qs from the B natural basis local-Qs. In this example, the first part of G_{BN} is a 4×4 matrix similar to (3). Therefore, there will be the four Gray basis local-Qs among the 7 local-Qs produced by G_{BN}. Figure 5 shows the 3 extra local-Qs.

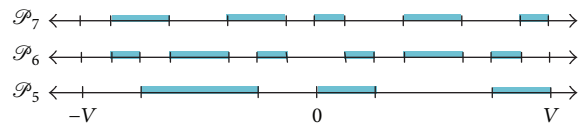

Figure 5: Local-Qs for the extra 3 bits.

Different G_{BN} matrices can be used to produce different sets of local-Qs, each having a different estimation performance. A repetition strategy, in which some of the basis local-Qs are repeated for quantizing the extra measurements, is a special case of the explained approach. However, by searching among different matrices one can find a set of N local-Qs that have better performance than a mere repetition scheme. The optimal set of local-Qs for each values of N and B depends on the SNR and can be obtained by exhaustive search. The size of the search space is given by

$$\Omega_2(B,N) = \Omega_1(B) \times \binom{2^B - 1 + (N - B - 1)}{N - B}. \tag{7}$$

In Section 6, some optimal results are shown and discussed. It is worth mentioning that the search for finding the optimal local-Qs is done in the FC and happens only once. After the optimal set of local-Qs is obtained, each local-Q is assigned to each sensor. As long as the measurement noise variance does not change a lot, the local-Qs do not need to be updated.

It is worthwhile pointing that there is a difference between the functionality of error-correcting codes when used against channel noise and when, in this case, used in the distributed estimation problem. In channel coding, the source of bit error is the additive channel noise which contaminates the coded bits after they are sent on the channel. On the contrary, in the case of distributed quantization, the bit error happens during the quantization process when the bits are being generated. The main difference this causes is that, unlike the channel coding, the a priori bit error probability (BEP) is not the same for all bits. Even in a homogeneous scenario with i.i.d. noise for all the observations, that is, the same σ, the probability of error for each bit depends on its local-Q and how small the *cell lengths* are in that local-Q. Moreover, the BEPs also depend on the value of X. Because of these fundamental differences, known optimal channel codes are no longer optimal for distributed estimation. Also, a separate study is required for the decoder/estimator of the proposed error-correcting quantization method.

4.3. Gaussian Source. Up to this end we had assumed that the unknown parameter X is uniformly distributed. Therefore, in Section 3, we had started our method by first assuming uniform partitioning of the range $[-V, V]$ into L equal divisions. If the parameter X is not uniformly distributed, the uniform partitioning of the range is not optimal. However, still the same methodology described in Sections 4.1 and 4.2 can be used to design the optimal local-Qs, but with different partitioning of the range of X. As an example, we discuss the Gaussian source. The results of applying our distributed

quantization method for a Gaussian parameter are presented in Section 6.

For a Gaussian source, we assume that the range of X, that is, $(-\infty, \infty)$, is partitioned according to the Lloyd-Max algorithm [23, 24] for centralized quantization. For example, for $X \sim \mathcal{N}(0,1)$, the optimal $B = 3$-bit centralized quantization has the range of X partitioned into $L = 8$ divisions as $(-\infty, -1.748, -1.05, -0.5006, 0, 0.5006, 1.05, 1.748, \infty)$. To derive B local-Qs for a Gaussian noise, we take the same steps as in Sections 4.1 and 4.2, except that the value of edges for every local-Q is a subset of the above edges.

4.4. Suboptimal Search Strategy. Due to the nature of the optimization problem, an exhaustive search must be performed to find the optimal local-Qs. In many WSN applications, B and N are small quantities, and the search is feasible. For those values of B and N where the search space size is too big, suboptimal strategies can be suggested to reduce the complexity. Some of those strategies are discussed here.

Strategy 1. When a B-bit precision is practiced with N sensors, there must be always B linearly independent local-Qs among the N local-Qs. Based on the simulation results, it can be investigated that, for uniform X and Gaussian noise, there are always the same B independent local-Qs for every SNR value. In this scenario, the B local-Qs derived based on Gray labeling, that is, $\mathscr{P}_1^{\mathrm{Gr}}, \ldots, \mathscr{P}_B^{\mathrm{Gr}}$, are always among the N optimal local-Qs. Therefore, for uniform X and Gaussian noise, to design the N local-Qs for any SNR, one can fix B of them to be the Gray local-Qs. This reduces the size of the search space from $\Omega_2(B,N)$ to $\Omega_2'(B,N)$, which is equal to the second term of the product in (7).

Strategy 2. Another simplifying strategy can be practiced for big values of N. During the simulations, it was observed that when N gets bigger while B is constant, some local-Qs are repeated among the optimal local-Qs. Therefore, even for very big N, one can limit the search to a smaller space, for example, $N_0 < N$, find the optimal local-Qs for that subspace, and repeat the same N_0 local-Qs for the rest of the observations.

Both of the above strategies can be combined to reduce the search space. For example, for a uniformly distributed unknown parameter with Gaussian measurement noises, a suboptimal search method is explained as follows. For $1 \le n \le B$, one uses $\mathscr{P}_1^{\mathrm{Gr}}, \ldots, \mathscr{P}_B^{\mathrm{Gr}}$. For $B+1 \le n \le N_0$, a search is performed on a space with size $\Omega_2'(B, N_0)$. For $N_0 + 1 \le n \le 2N_0$, $2N_0 + 1 \le n \le 3N_0, \ldots$, the same N_0 local-Qs are repeated. A third strategy which further reduces the complexity is presented in Section 5.3.

5. Decoding

In the FC, a decoding method must be used to estimate the unknown parameter from the received binary-quantized measurements y_n. Here, we explain two types of decoders, that is, the discrete and continuous decoders, which estimate X based on the discrete and continuous a posteriori distributions, respectively. For the sake of better understanding the overall behavior of the distributed estimation method, we first explain the discrete decoder. In both estimators the receiver only needs to do the computations once at the beginning of the algorithm to build a lookup table, which is used during the estimation procedure.

5.1. Discrete A Posteriori. As a discrete decoder, the estimator locates X in one of the discrete divisions δ^l, $1 \le l \le L$, from the received bits. In this section we propose an ML estimator that decodes the N received bits to estimate the *division* of X among the L divisions.

As discussed in Section 4, for each value of X a set of N local-Qs together generate an N-bit "codeword." Assuming no measurement noise, there are $L = 2^B$ different codewords; that is, $\mathbf{c}^l = [c_1^l c_2^l \cdots c_N^l]$, $1 \le l \le L$. There is a one-to-one correspondence between these valid codewords and the L divisions, that is, δ^l, $1 \le l \le L$. In the presence of measurement noise, each of the N bits, that is, y_n, $1 \le n \le N$, might be wrong, resulting in a "received word" (note that the notation "received" does not imply a communication channel or channel error) $\mathbf{y} = [y_1, y_2, \ldots, y_n]$, which might include some bit errors. The discrete decoder's function is to find the most likely valid codeword based on the received N-bit word $\mathbf{y}$, which in turn results in estimating the division of X. This estimator can be reduced to a lookup table that can be saved in the receiver.

To build the lookup table for the discrete decoder, the likelihoods, or equivalently the a posteriori probabilities, of all the L valid codewords conditioned on the received word $\mathbf{y}$ are calculated. The codeword with the maximum likelihood is chosen as the decoder's decision corresponding to each $\mathbf{y}$. The center of mass [23] for that division is recognized as the estimated value $\widehat{X}$. This decoder is therefore referred to as the discrete ML estimator. The a posteriori likelihood of each codeword can be written as

$$P\left(\mathbf{c}^l \mid \mathbf{y}\right) \propto P\left(c^l\right) \prod_{n=1}^{N} P\left(y_n \mid \mathbf{c}^l\right), \tag{8}$$

where $P(c^l)$ is the a priori probability of codeword c^l, which is equal to $P(\Gamma^{l-1} \le X \le \Gamma^l)$, where Γ^{l-1} and Γ^l indicate the left and right edges of the lth quantization division δ^l, corresponding to $\mathbf{c}^l$. For uniform X, $\Gamma^{l-1} = -V + (l-1)\Delta$, and $\Gamma^l = -V + l\Delta$.

Since each valid codeword is associated with one of the L divisions δ^l, $1 \le l \le L$, each term of the product in (8) can be described as

$$\begin{aligned} P\left(y_n \mid \mathbf{c}^l\right) &= P\left(y_n \mid \Gamma^{l-1} \le X \le \Gamma^l\right) \\ &= \frac{\int_{\Gamma^{l-1}}^{\Gamma^l} P\left(y_n \mid X\right) P\left(X\right) dX}{\int_{\Gamma^{l-1}}^{\Gamma^l} P\left(X\right) dX}. \end{aligned} \tag{9}$$

The denominator in the second line is by definition $P(c^l)$, the a priori probability of the codeword c^l, which for uniform X

is equal for every c^l. Remembering that y_n can be either 0 or 1, (9) can be written as

$$P\left(y_n \mid \mathbf{c}^l\right) = \int_{\Gamma^{l-1}}^{\Gamma^l} P1_n(X)^{(y_n)}\left(1 - P1_n(X)\right)^{(1-y_n)} P(X)\, dX, \tag{10}$$

where $P1_n(X)$ is the probability that the nth bit is 1 when the parameter value is X, that is, $P(y_n = 1 \mid X)$. The formula to calculate $P1_n(X)$ is given in Appendix B. Using (10) and (8) one can find the a posteriori probability of the codewords.

At the beginning of the algorithm a decoding table is formed by calculating the a posteriori probabilities for different values of the received word, that is, $\mathbf{y}^k$, $1 \le k \le 2^N$, and different valid codewords $\mathbf{c}^l$, $1 \le l \le L$. This calculation is done once in the FC and does not need to repeat every time. The decoding table has 2^N entries, one for every possible value of $\mathbf{y}$, associating it with a codeword which has the highest likelihood $P(\mathbf{c}^l \mid \mathbf{y})$. For every estimation instance, based on the received word, the decoder chooses the corresponding codeword from the table as the most likely codeword and announces the center of mass of the associated division as $\widehat{X}$. It is easy to see that the lookup table depends on the local-Qs set, as well as σ.

The discrete ML decoder is based on finding the a posteriori probability of the valid codewords, which have one-to-one correspondence with the division of X. Therefore, they can only locate X with a limited precision. To better estimate the analog parameter X, a continuous decoder can be used instead of the discrete decoder, resulting in better estimation performance.

5.2. Continuous A Posteriori. The continuous decoder is designed based on the continuous a posteriori pdf of X. Using the a posteriori pdf, a MAP or MMSE estimator can be designed. Again, similar to the discrete decoder, the receiver only needs to do the computations once at the beginning of the algorithm to build a lookup table, which is used during the estimation procedure.

The continuous a posteriori distribution of X, that is, $P(X \mid \mathbf{y})$, can be written as

$$\begin{aligned} P(X \mid \mathbf{y}) &\propto P(X) \prod_{n=1}^{N} P(y_n \mid X) \\ &\propto P(X) \prod_{n=1}^{N} P1_n(X)^{(y_n)}\left(1 - P1_n(X)\right)^{(1-y_n)}. \end{aligned} \tag{11}$$

The factorization in (11) is valid since the measurement noises as well as the quantization procedures are independent for different measurements. Once the above a posteriori conditional distribution is derived, one can find $\widehat{X} = \arg\max_X P(X \mid \mathbf{y})$ or $\widehat{X} = E\{X \mid \mathbf{y}\}$ as a MAP or MMSE estimator of X, respectively.

In the FC, the following calculations are done once to build a decoding table. For each possible received word $\mathbf{y}^k$, $1 \le k \le 2^N$, the MAP or MMSE estimation $\widehat{X}^k$ is calculated to build an entry in the lookup table, mapping $\mathbf{y}^k$ to $\widehat{X}^k$. Every time an N-bit word is received at the decoder this lookup table is used to find $\widehat{X}$.

It is easy to investigate that the MSE performance of the continuous decoder is better than the discrete one, while the computational complexity of the continuous decoder is only slightly higher than the discrete decoder. The performance of different decoders is studied through simulations in Section 6 and compared to the CRLB for the optimum unbiased estimator.

5.3. Performance Bounds. To evaluate the performance of our estimator, we have calculated the CRLB for a set of local-Qs. In our distributed estimation method, for the case of Gaussian or uniform X and Gaussian noise, the CRLB can be the indicator of the best estimation performance if the estimator is unbiased. The unbiasedness of our MMSE estimator is proven in Appendix A. The CRLB of the estimation method is given by

$$\left(\sum_{n=1}^{N} \int_X \left\{ \frac{\left(\partial P1_n(X)/\partial X\right)^2}{P1_n(X)} + \frac{\left(\partial P1_n(X)/\partial X\right)^2}{1 - P1_n(X)} \right\} \times P(X)\, dX + J_2 \right)^{-1}, \tag{12}$$

where, for a uniform X, $J_2 = 0$, and, for a Gaussian X with zero mean and variance σ_x^2, $J_2 = 1/\sigma_x^2$. See Appendix B for derivation of (12).

Strategy 3. Based on the CRLB, a strategy can be devised to reduce the computational complexity of the search method described in Section 4. As explained there, the optimal local-Qs are found by minimizing the MSE. That is to say, for every candidate set of N local-Qs, the MSE of estimation is found through simulation, and finally the set of local-Qs with the lowest MSE is selected. To reduce the complexity, we suggest using an MSE bound, that is, the CRLB, instead of finding the actual MSE. So, for every candidate set of local-Qs, the CRLB is calculated based on (12), and the set with the lowest CRLB value is selected. Using this strategy along with the other two strategies in Section 4.4, the complexity of finding the local-Qs will be reduced greatly. The MSE results for the suboptimal local-Qs have been compared with the optimal local-Qs in Section 6.

6. Numerical Results

In this section, different simulation results are discussed in order to study the performance of the proposed distributed estimation method. Both scenarios in Section 4.1, where $N = B$, and Section 4.2, where $N > B$, are considered. The results are shown for two distributions of the unknown parameter, that is, uniform and Gaussian. For the case of uniform unknown parameter, X is uniformly distributed in the interval $[-1, 1]$. For Gaussian X, it is distributed according to $\mathcal{N}(0, 0.1)$. Comparison with other methods, such as [4, 11, 12, 14], is discussed. The performance is evaluated in terms of the MSE. In all simulations, the measurement noise is i.i.d.

Table 1: Optimal local-Qs for uniform X, $B = 3$, $N = 8$, and SNR = 18.2 dB.

$\mathscr{P}_1$	0	0	0	0	1	1	1	1
$\mathscr{P}_2$	0	0	1	1	1	1	0	0
$\mathscr{P}_3$	0	1	1	0	0	1	1	0
$\mathscr{P}_4$	0	1	1	0	0	1	1	0
$\mathscr{P}_5$	0	1	1	0	0	1	1	0
$\mathscr{P}_6$	0	0	1	1	0	0	1	1
$\mathscr{P}_7$	0	1	0	1	0	1	0	1
$\mathscr{P}_8$	0	1	0	1	1	0	1	0

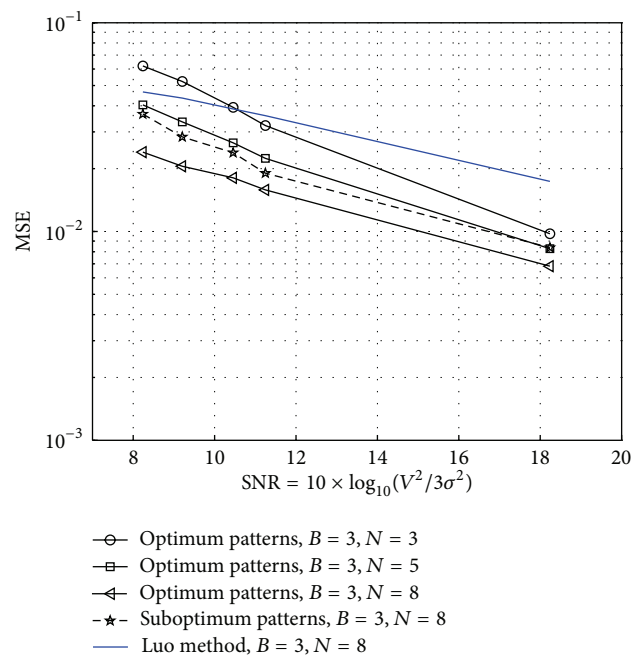

Figure 6: Effect of error-correcting local-Qs in estimation performance.

zero mean Gaussian with variance σ^2. Note that this SNR is only related to the measurement noise, while the channel is assumed to be error-free.

Figure 6 shows the performance of the distributed estimation method for a uniform source. Since X is uniform, $E\{X^2\} = V^2/3$; therefore, the average SNR is $10\log_{10}(V^2/3\sigma^2)$. At each SNR, the optimal set of local-Qs was found by searching all matrices G_{BN} for $B = 3$ and $N = 3$, 5, and 8. For the case of $N = 8$, the results of the suboptimal search method based on both Strategy 1 and Strategy 3 are also shown in the figure. Note that, for all cases, the continuous MMSE estimator is used in the decoder. The results show that when $N = B$ the optimal local-Q set for every SNR is the Gray basis local-Qs. When $N > B$, for each SNR a different set of optimal local-Qs are found, reducing the estimation error. For example, for the case of $B = 3$ and $N = 8$ the optimal local-Qs for SNR = 18.2 dB are presented in Table 1. From the table, we can see that the local-Qs $\mathscr{P}_1$, $\mathscr{P}_2$, and $\mathscr{P}_3$ are the 3 Gray basis local-Qs, and the rest are linear combinations of those.

Table 2: Optimal local-Qs for Gaussian X, $N = B = 3$, SNR = 5.2 dB.

$\mathscr{P}_1$	0	1	1	0	0	1	1	0
$\mathscr{P}_2$	1	0	0	0	1	1	1	0
$\mathscr{P}_3$	0	1	0	1	1	0	1	0

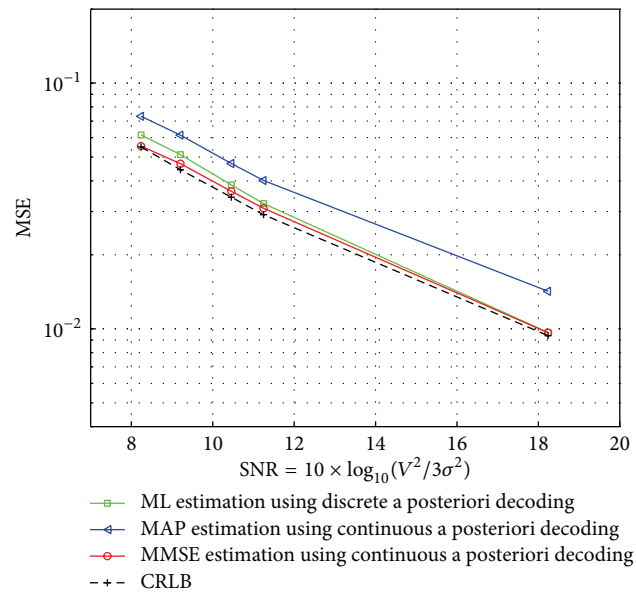

Figure 7: Performance of different decoding methods, for Gaussian X.

For the case of $N = 8$ the performance of our algorithm is compared with the distributed estimation proposed by Luo [14]. Based on Luo's algorithm, $N/2 = 4$ measurements are quantized to the first bit in a natural binary system, $N/4 = 2$ measurements are used to estimate the second bit, and the remaining 2 measurements are used for the third bit. As can be seen in Figure 6, the optimal sets of 8 local-Qs for different SNRs outperform the method of [14].

The three different decoding methods based on the discrete and continuous a posteriori functions in Section 5 are compared in Figure 7. The first decoding method is the ML estimation based on the discrete a posteriori likelihood of the codewords discussed in Section 5.1. The other two methods are the MAP and MMSE estimations based on the continuous a posteriori density of X in Section 5.2. Also, the CRLB for uniform X has been shown for comparison.

For a Gaussian parameter with zero mean and variance $\sigma_x^2 = 0.1$, the results for $N = B = 3$ are shown in Figure 8. The average SNR in this case is $10\log_{10}(\sigma_x^2/\sigma^2)$. Please note that, as explained in Section 4.3, the values for division edges of the centralized quantizer, that is, Γ^l, $1 \le l \le L$, are set to $\sigma_x \times \{-1.748, -1.05, -0.5006, 0, 0.5006, 1.05, 1.748\}$ [23]. As an example, the optimal local-Qs for SNR = 5.2 dB are presented in Table 2. Note that the local-Qs are lineally independent, but, unlike the case of uniform X, the optimal local-Qs are not the Gray basis local-Qs.

Figure 9 compares the performance of our proposed method with other binary quantization methods, for $N =$

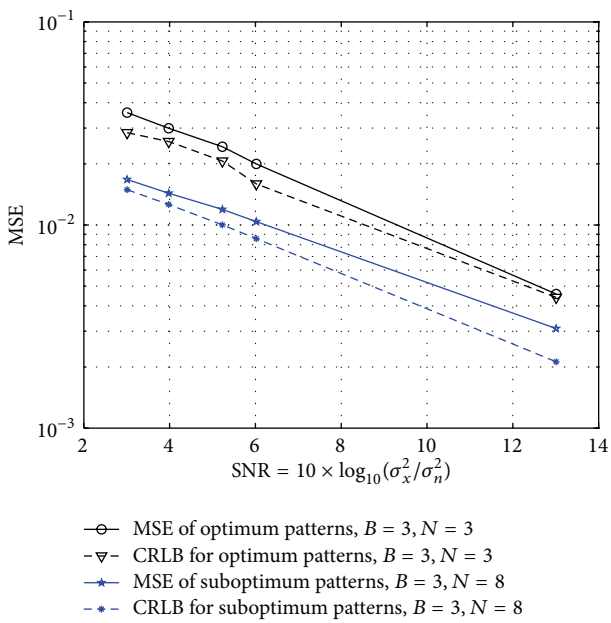

Figure 8: Performance of our distributed estimation method for Gaussian parameter with $\sigma_x^2 = 0.1$, $B = 3$.

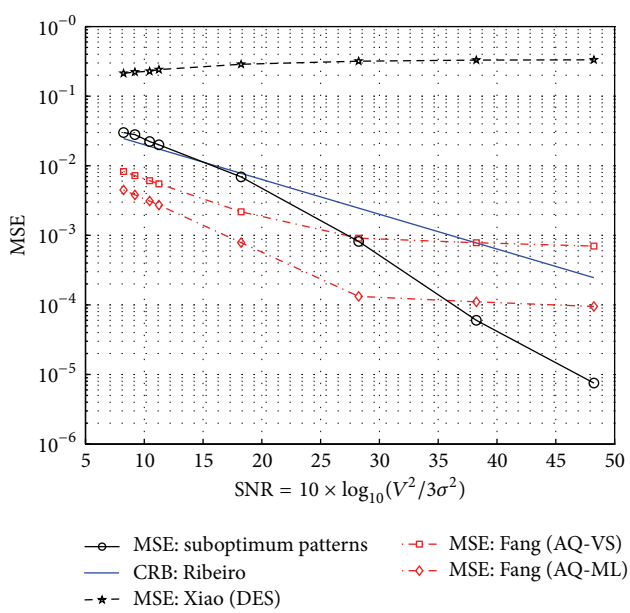

Figure 9: Performance comparison for different methods for $N = 10$.

10. The circle points in Figure 9 show the MSE for the suboptimal local-Qs found using the three suboptimal search strategies. For the method proposed by Ribeiro and Giannakis [12], the CRLB for the optimal solution is indicated. Two adaptive quantization methods proposed by Fang and Li [11], that is, AQ-VS and AQ-ML, are also shown in the figure. For the AQ-ML, before quantizing each measurement a ML estimation must be calculated. It can be seen from Figure 9 that for SNR > 15 dB our method has lower MSE compared to [11, 12], while its complexity for "each estimation job" (as long as the SNR does not change a lot) is less than those algorithms.

7. Conclusion

In this study we proposed an optimum distributed quantization method to estimate unknown parameter form sensors' noisy measurements, each compressed to one bit. The estimation method is based on designing multithreshold local-Qs to generate a single bit from each analog measurement and an appropriate decoder to be used at the fusion center for estimating the unknown. The results show that, for Gaussian X, when B measurements are used to achieve an estimation with B-bit precision, the optimal B local-Qs are those based on the Gray labeling for all SNRs. For scenarios where $N > B$ measurements are available, extra local-Qs are found to achieve an error-correcting capability compared to the case when only B measurements are available. The results have been compared with other distributed estimation methods with binary quantization and show better performance when a limited number of sensors with lower measurement noise are available.

Appendices

A. Unbiasedness of the Estimator

The amount of bias for an estimator is defined as

$$E\left\{X - \widehat{X}\right\} = E\{X\} - E\left\{\widehat{X}\right\}, \tag{A.1}$$

where X is the unknown random parameter and $\widehat{X}$ is its estimated value. In our method, X is going to be estimated from N binary data using a MMSE estimator in the FC; that is, $\widehat{X} = E\{X \mid y_1, \ldots, y_N\}$. Therefore,

$$\begin{aligned} E\left\{\widehat{X}\right\} &= E\{E\{X \mid y_1, \ldots, y_N\}\} \\ &= \sum_{y_1,\ldots,y_N} E\{X \mid y_1, \ldots, y_N\}\, p(y_1, \ldots, y_N) \\ &= \sum_{y_1,\ldots,y_N} \left(\int_X X p(X \mid y_1, \ldots, y_N)\, dX\right) p(y_1, \ldots, y_N) \\ &= \int_X X \sum_{y_1,\ldots,y_N} p(X, y_1, \ldots, y_N)\, dX \\ &= \int_X X p(X)\, dX = E\{X\}. \end{aligned} \tag{A.2}$$

The last equality ensures the unbiasedness of the method, when an MMSE estimator is used in the FC.

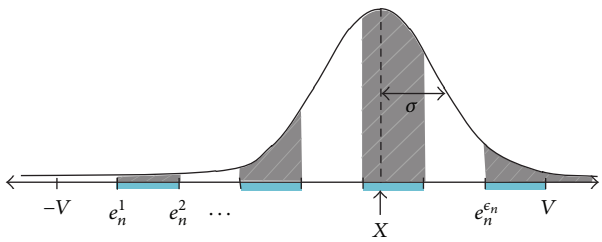

Figure 10: The probability $P(y_n = 1 \mid X)$.

B. CRLB

The CRLB for estimating an unknown random variable X from a noisy measurement Y is the inverse of the Fisher information metric, which can be obtained as

$$J = E\left\{\left[\frac{\partial \ln P(Y,X)}{\partial X}\right]^2\right\}, \tag{B.1}$$

where $P(Y,X)$ is the joint probability distribution of X and Y and E indicates the expected value with respect to both X and Y. It can be proved that the above definition is identical to [25]:

$$J = -E\left\{\left[\frac{\partial^2 \ln P(Y,X)}{\partial X^2}\right]\right\}. \tag{B.2}$$

Having $P(Y,X) = P(Y \mid X)P(X)$, (B.2) can be written as

$$J = \underbrace{-E\left\{\left[\frac{\partial^2 \ln P(Y \mid X)}{\partial X^2}\right]\right\}}_{J_1}\underbrace{-E\left\{\left[\frac{\partial^2 \ln P(X)}{\partial X^2}\right]\right\}}_{J_2} = J_1 + J_2, \tag{B.3}$$

where the first term J_1 is related to the likelihood of measurements and the second term J_2 depends only on the distribution of X.

Now, suppose that there are N binary measurements available from X. In other words, $Y = [y_1, \ldots, y_N]$. The likelihood function $P(Y \mid X)$ can be written as [16]

$$\begin{aligned} P(y_1, \ldots, y_N \mid X) &= \prod_{n=1}^{N} P(y_n \mid X) \\ &= \prod_{n=1}^{N} P1_n(X)^{(y_n)} P0_n(X)^{(1-y_n)}, \end{aligned} \tag{B.4}$$

where $P1_n(X)$ is the probability that the nth bit is 1 when the parameter value is X, that is, $P(y_n = 1 \mid X)$. Similarly, $P0_n(X) = P(y_n = 0 \mid X)$. If the nth local-Q is defined by the set of cell edges $\{e_n^1, e_n^2, \ldots, e_n^{\epsilon_n}\}$ mapping the first cell to 0, for Gaussian measurement noise with variance σ, $P1_n(X)$ can be written as (see Figure 10)

$$P1_n(X) = \sum_{i=1}^{\epsilon_n} (-1)^{i+1} Q\left(\frac{e_n^i - X}{\sigma}\right). \tag{B.5}$$

Now, $\ln P(Y \mid X)$ can be written as

$$\ln P(Y \mid X) = \sum_{n=1}^{N} \left((y_n) \ln P1_n(X) + (1 - y_n) \ln P0_n(X)\right). \tag{B.6}$$

Therefore, its first and second derivative can be calculated as

$$\frac{\partial \ln P(Y \mid X)}{\partial X} = \sum_{n=1}^{N} \left\{(y_n) \frac{\partial P1_n(X)/\partial X}{P1_n(X)} + (1 - y_n) \frac{\partial P0_n(X)/\partial X}{P0_n(X)}\right\}, \tag{B.7}$$

$$\begin{aligned} \frac{\partial^2 \ln P(Y \mid X)}{\partial X^2} &= \sum_{n=1}^{N} \left\{(y_n) \left\{\frac{\partial^2 P1_n(X)/\partial X^2}{P1_n(X)} - \left(\frac{\partial P1_n(X)/\partial X}{P1_n(X)}\right)^2\right\}\right. \\ &\quad \left. + (1 - y_n) \left\{\frac{\partial^2 P0_n(X)/\partial X^2}{P0_n(X)} - \left(\frac{\partial P0_n(X)/\partial X}{P0_n(X)}\right)^2\right\}\right\}. \end{aligned} \tag{B.8}$$

To find the first term, J_1, in (B.3), the expectation of the second derivative of $P(Y \mid X)$, that is, (B.8), is calculated as

$$\begin{aligned} J_1 &= -E\left\{\left[\frac{\partial^2 \ln P(Y \mid X)}{\partial X^2}\right]\right\} \\ &= -\int_X \sum_Y \left[\frac{\partial^2 \ln P(Y \mid X)}{\partial X^2}\right] P(Y \mid X) P(X)\, dX. \end{aligned} \tag{B.9}$$

Inserting (B.8) in (B.9) and calculating the integral over Y, we have

$$\begin{aligned} J_1 &= -\sum_{n=1}^{N} \int_X \left(\frac{\partial^2 P1_n(X)}{\partial X^2} - \frac{(\partial P1_n(X)/\partial X)^2}{P1_n(X)}\right) P(X)\, dX \\ &\quad - \sum_{n=1}^{N} \int_X \left(\frac{\partial^2 P0_n(X)}{\partial X^2} - \frac{(\partial P0_n(X)/\partial X)^2}{P0_n(X)}\right) P(X)\, dX. \end{aligned} \tag{B.10}$$

To further simplify the above, note that, for any binary local-Qs,

$$P0_n(X) = 1 - P1_n(X); \tag{B.11}$$

hence,

$$\frac{\partial P0_n(X)}{\partial X} = -\frac{\partial P1_n(X)}{\partial X}, \qquad \frac{\partial^2 P0_n(X)}{\partial X^2} = -\frac{\partial^2 P1_n(X)}{\partial X^2}. \tag{B.12}$$

Therefore, (B.10) is reduced to

$$J_1 = \sum_{n=1}^{N} \int_X \left\{\frac{(\partial P1_n(X)/\partial X)^2}{P1_n(X)} + \frac{(\partial P1_n(X)/\partial X)^2}{1 - P1_n(X)}\right\} P(X)\, dX. \tag{B.13}$$

For a random variable with uniform distribution, the second term, J_2, in (B.3) is zero; therefore, the CRLB reduces to $1/J_1$. For a Gaussian random variable, J_2 can be easily calculated as

$$J_2 = -E\left\{\left[\frac{\partial^2 \ln \mathcal{N}\left(x, \sigma_x^2\right)}{\partial X^2}\right]\right\} = \frac{1}{\sigma_x^2}, \tag{B.14}$$

and the CRLB will be $1/(J_1 + J_2)$.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. Çetin, L. Chen, J. W. Fisher III et al., "Distributed fusion in sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 42–55, 2006.

[2] M. Gastpar, M. Vetterli, and P. L. Dragotti, "Sensing reality and communicating bits: a dangerous liaison," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 70–83, 2006.

[3] J. Li and G. AlRegib, "Distributed estimation in energy-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3746–3758, 2009.

[4] J.-J. Xiao and Z.-Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, 2005.

[5] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27–41, 2006.

[6] Z.-Q. Luo and J.-J. Xiao, "Universal decentralized estimation in a bandwidth constrained sensor network," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 829–832, March 2005.

[7] J. Xiao, S. Cui, Z. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, 2006.

[8] A. Krasnopeev, J. Xiao, and Z. Luo, "Minimum energy decentralized estimation in a wireless sensor network with correlated sensor noises," *Eurasip Journal on Wireless Communications and Networking*, vol. 2005, no. 4, pp. 473–482, 2005.

[9] H. C. Papadopoulos, G. W. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 978–1002, 2001.

[10] H. Li and J. Fang, "Distributed adaptive quantization and estimation for wireless sensor networks," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 669–672, 2007.

[11] J. Fang and H. Li, "Distributed adaptive quantization for wireless sensor networks: from delta modulation to maximum likelihood," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, part 2, pp. 5246–5257, 2008.

[12] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, 2006.

[13] R. Gray, S. Boyd, and T. Lookabaugh, "Low rate distributed quantization of noisy observations," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, pp. 354–358, 1985.

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

[15] I. D. Schizas, G. B. Giannakis, and N. Jindal, "Distortion-rate bounds for distributed estimation using wireless sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 748605, 2008.

[16] S. Movaghati and M. Ardakani, "Particle-based message passing algorithm for inference problems in wireless sensor networks," *IEEE Sensors Journal*, vol. 11, no. 3, pp. 745–754, 2011.

[17] F. Chapeau-Blondeau and D. Rousseau, "Noise-enhanced performance for an optimal Bayesian estimator," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1327–1334, 2004.

[18] G. Balkan and S. Gezici, "CRLB based optimal noise enhanced parameter estimation using quantized observations," *IEEE Signal Processing Letter*, vol. 17, no. 5, pp. 477–480, 2010.

[19] S. Movaghati and M. Ardakani, "Energy-efficient quantization for parameter estimation in inhomogeneous WSNs," in *Proceedings of the Annual IEEE Global Telecommunications Conference (GLOBECOM '11)*, pp. 1–5, Houston, Tex, USA, December 2011.

[20] J. Chen, S. Dumitrescu, Y. Zhang, and J. Wang, "Robust multiresolution coding," *IEEE Transactions on Communications*, vol. 58, no. 11, pp. 3186–3195, 2010.

[21] S. Dumitrescu, "Fast encoder optimization for multi-resolution scalar quantizer design," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1520–1529, 2011.

[22] S. Lin and D. J. Costello, *Error Control Coding*, Prentice-Hall, New York, NY, USA, 2nd edition, 2004.

[23] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[24] J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, 1960.

[25] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*, John Wiley & Sons, New York, NY, USA, 2001.