

Quantization for Maximin ARE in Distributed Estimation

Parvathinathan Venkitasubramaniam, *Student Member, IEEE*, Lang Tong, *Fellow, IEEE*, and Ananthram Swami, *Senior Member, IEEE*

Abstract—We consider the design of optimal quantizers for the distributed estimation of a deterministic parameter. In particular, we design deterministic scalar quantizers to maximize the minimum asymptotic relative efficiency (ARE) between quantized and unquantized ML estimators. We first design identical quantizers using the class of score-function quantizers (SFQ). We show that the structure of SFQs generally depend on the parameter value, but can be expressed as thresholds on the sufficient statistic for a large class of distributions. We provide a convergent iterative algorithm to obtain the best SFQ that maximizes the minimum ARE for distributions of that class. We compare the performance of the optimal SFQ with a general quantizer designed without making any restrictions on the structure. This general quantizer is hard to implement due to lack of structure, but is optimal if the iterative design algorithm does not encounter local minima. Through numerical simulations, we illustrate that the two quantizers designed are identical. In other words, the optimal quantizer structure is that of an SFQ. For a distributed estimation setup, designing identical quantizers is shown to be suboptimal. We, therefore, propose a joint multiple quantizer design algorithm based on a person-by-person optimization technique employing the SFQ structure. Using numerical examples, we illustrate the gain in performance due to designing nonidentical quantizers.

Index Terms—Asymptotic relative efficiency (ARE), distributed estimation, quantization, score-function quantizer (SFQ).

I. INTRODUCTION

A. Motivation

DISTRIBUTED statistical inference is a classical problem in signal processing, where a fusion center receives data from several distributed nodes and infers the parameter of the underlying process. One of the key differences of this problem from classical point estimation is that the fusion center cannot have complete access to every observed measurement. The data

Manuscript received April 6, 2006; revised November 24, 2006. This work was supported in part by the National Science Foundation under Contract CNS-0435190 and prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. Results presented in part at ICASSP 2006 and CISS 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fulvio Gini.

P. Venkitasubramaniam and L. Tong are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: pv45@cornell.edu; lt35@cornell.edu).

A. Swami is with the Army Research Laboratory, Adelphi, MD 20783 USA (e-mail: a.swami@ieee.org).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2007.894279

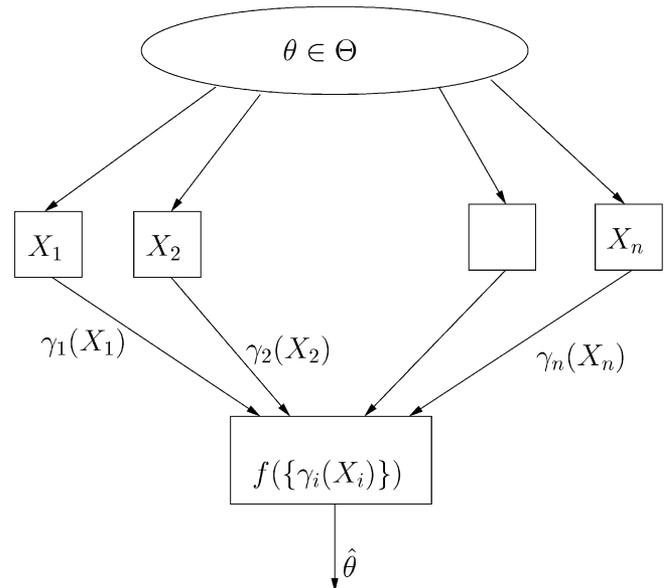


Fig. 1. Distributed estimation system. Observations : X_i . Quantizers : γ_i . Estimator : $\hat{\theta}$.

transmission from nodes to fusion center is restricted due to energy and bandwidth constraints. It is, therefore, imperative that the observations are quantized before transmission. This spawns two fundamental problems: design of quantizer at the node, and the estimator at the fusion center. The key metrics that need to be optimized are the quantizer size (number of quantization bits) and the estimation error at the fusion center.

The basic distributed estimation setup is shown in Fig. 1. The distribution of the observations X_1, \dots, X_n depends on an underlying parameter θ . Node i implements quantizer γ_i on the observation and transmits the quantized value $\gamma_i(X_i)$ to the fusion center, which then estimates the value of θ . When the quantized values are perfectly available at the fusion center, the estimation error of θ is only dependent on the quantizers $\{\gamma_i\}$ and the estimator $\hat{\theta}$. In this paper, we are interested in the optimal design of quantizers that minimize estimation error when $\hat{\theta}$ is a maximum-likelihood (ML) estimator.

The main challenge in designing quantizers for parameter estimation is the unavailability of information about the parameter. One class of approaches to this problem has been from an information-theoretic perspective, first proposed by Zhang and Berger [1]. In the information-theoretic setup, each node is assumed to observe a long sequence of measurements which is then compressed before transmission to the fusion center, and the key design metric is the asymptotic compression rate.

For that problem, Han and Amari [2] obtained the best known achievable rate region and also provided an ML estimator for their encoders. The lack of knowledge about the parameter was circumvented in their solution by using the *type* of the long sequence as an alternative indexing of distributions.

In networks where nodes observe a few measurements, the type provides little or no information about the underlying parameter. In such situations, if the parameter is random, the known statistics of the parameter can be utilized to design quantizers that minimize the average estimation error. Optimal scalar quantizers for random parameter estimation have been designed for different types of estimators. Lam and Reibman [3] provided an iterative algorithm to obtain a quantizer that maximizes the Bayesian Fisher Information (FI) of a random parameter. For some special classes of linear estimators, Gubner [4] and Zhang and Li [5] obtained the optimal MMSE quantizer.

For a deterministic parameter, the estimation performance varies with the parameter, and it is impossible to design a single scalar quantizer that achieves the optimal performance for all values of the parameter. It is, therefore, necessary to formulate an appropriate metric that represents the overall estimation performance. In some recent works, quantizers for deterministic parameter estimation have been designed under some restrictions on quantizer structure and underlying distributions. The idea of a maximin metric based on FI was considered by Fowler and Chen [6]. They proposed a transform coding based quantizer to minimize the maximum difference between FI of quantized and unquantized observations. When the observation is a parameter in additive noise, Ribiero and Giannakis [7] characterized the best possible FI. They also showed that when the range of the parameter is small, the performance of the optimal threshold quantizer is within a constant of the unquantized estimator. For non-i.i.d. additive noise observations, Luo [8] optimized the allocation of total bits among binary quantizers to maximize estimation performance when the noise is bounded. In an earlier work [9], we had proposed the idea of score-function quantization to maximize the FI for a given value of the parameter, which can be used as an upper bound for the performance of any quantizer.

In this paper, we consider the design of fixed size scalar quantizers for networks, where the number of nodes is large, but each node measures a few observations. The observations are identically distributed according to some deterministic parameter. The fusion center estimates the underlying parameter using an ML estimator. For a fixed distribution of quantized observations, it is known that the ML estimator is asymptotically efficient. Furthermore, the asymptotic performance can be characterized analytically using the FI of the distribution of quantized values.

The metric we propose is a measure of relative performance between quantized and unquantized estimators known as the asymptotic relative efficiency (ARE). The ARE represents the ratio of sample sizes required by the quantized and unquantized estimators to achieve the same asymptotic performance and it can be expressed as a ratio of FIs between quantized and unquantized distributions. Due to its dependence on the value of θ , the ARE cannot be maximized universally for all values of θ .

We, therefore, consider maximizing the minimum ARE across the parameter set as a criterion for quantizer design. The minimum ARE over θ corresponds to the least percentage of additional samples required by the quantizer to perform asymptotically, as well as the unquantized estimator. Further, it can be shown that for a certain class of distributions, the maximin ARE quantizer requires the least sample size amongst all deterministic quantizers that achieve the same asymptotic error variance.

B. Main Contributions

In this paper, we design quantizers that maximize the minimum ARE over the parameter set between quantized and unquantized observations. Assuming identical quantizers for the nodes, we first consider the quantizer structure that belongs to the class of score-function quantizers (SFQ). Although SFQs are dependent on parameter θ , we show that for a certain class of distributions that satisfy a monotonicity property, the SFQs are expressible as a set of thresholds on the sufficient statistic. For these distributions, we provide a convergent algorithm to obtain the best SFQ that maximizes the minimum ARE.

In order to compare the performance loss due to the restriction on structure, we propose an alternative algorithm to design the maximin ARE quantizer with no assumptions on quantizer structure or nature of distribution. This iterative algorithm, although optimal in the absence of local minima, does not provide any structural representation for the quantizer and is computationally intensive. However, through numerical simulations we illustrate that, for distributions that satisfy the monotonicity property, the two algorithms result in identical quantizers. In other words, the optimal quantizer structure is that of an SFQ.

Assuming identical quantizers for the nodes is restrictive and in general, suboptimal for distributed estimation. Hence, we utilize the optimality of SFQs and propose a person-by-person optimization (PBPO) technique to design multiple nonidentical quantizers that maximize the metric. Through numerical examples, we demonstrate a significant improvement in performance due to designing nonidentical quantizers especially at low noise levels.

The basic organization of this paper is as follows. In Section II, we describe the maximin ARE criterion for quantizer design and the relation to sample size of estimators. In Section III, we discuss the optimality of SFQs and also present an algorithm to obtain the best SFQ for the maximin criterion assuming identical quantizers. In Section IV, we discuss the alternate approach to maximin quantizer design that does not make assumptions on structure. The comparison of the two quantizers and the optimality through numerical simulations are given in Section V. The design of multiple nonidentical quantizers for maximin ARE and supplementary numerical results are presented in Section VI. Some concluding remarks and possible future extensions are mentioned in Section VII.

II. SYSTEM MODEL

Consider the distributed estimation setup as shown in Fig. 1. Let \mathcal{X} be some set endowed with a σ -field \mathcal{G} . The observations $X_i \in \mathcal{X}$ are independent and identically distributed (i.i.d.). The probability measure of X_i belongs to a family of probability measures $\{P_\theta; \theta \in \Theta\}$ on $(\mathcal{X}, \mathcal{G})$ indexed by parameter

θ lying in a set Θ . Throughout this paper, we assume that θ is real-valued and $\Theta \subset \mathbb{R}$ is a bounded set. Node i implements deterministic quantizer γ_i and transmits the quantized version $\gamma_i(X_i)$ to the fusion center. For the remainder of this section and two subsequent sections, we shall focus on designing identical quantizers for the nodes, or in other words, $\gamma_i = \gamma, \forall i$.

A *deterministic quantizer* can be formally defined as a \mathcal{G} -measurable mapping $\gamma : \mathcal{X} \mapsto \{1, \dots, D\}$, where D , a constant integer, is known as the *size* of the quantizer. A deterministic quantizer uniquely maps each observation to a positive integer $k \leq D$, which we refer to as the partition index of the observation. Since the observation space is uncountable, the structural representation of a quantizer can be very complex.

When nodes implement identical quantizers, the quantized variable $\gamma(X_i)$ at the nodes are i.i.d. according to a probability mass function (p.m.f) \mathbf{q}_θ indexed by θ , where

$$\begin{aligned} \mathbf{q}_\theta &= (q_\theta(1), \dots, q_\theta(D)) \\ q_\theta(i) &= P_\theta(\gamma(X) = i). \end{aligned}$$

The estimation is performed using only the quantized observations which are received perfectly at the fusion center. Therefore, for a fixed estimator $\hat{\theta}_{\text{ML}}$, the mean squared error can be expressed as a function of the distribution \mathbf{q}_θ . In this paper, we assume that an ML estimator $\hat{\theta}_{\text{ML}}$ based on \mathbf{q}_θ is implemented at the fusion center and our goal is to minimize the increase in MSE due to quantization. For ML estimators, MSE performance is characterized by the FI, and the relative performance with respect to an unquantized estimator can be expressed using the ARE. These concepts are discussed in the following section.

A. Cramér–Rao Bound and ARE

It is well known that the MSE of any unbiased estimator is lower bounded by the Cramér–Rao lower bound (CRLB). Specifically, under some regularity conditions (see [10, p. 169]) on $\{P_\theta; \theta \in \Theta\}$, the MSE satisfies

$$\mathbb{E}\{|\theta - \hat{\theta}|^2\} \geq \frac{1}{nI_\theta} \quad (1)$$

where

$$I_\theta = \sum_{i=1}^D q_\theta(i) \left(\frac{d}{d\theta} \log q_\theta(i) \right)^2 \quad (2)$$

is the FI of the quantized observation and n is the number of measurements. Although the CRLB may not be achievable for any finite n , under more regularity conditions (see [10, p. 183]), the MSE of an ML estimator converges to the bound asymptotically. Specifically, as $n \rightarrow \infty$

$$\sqrt{n}(\hat{\theta}_{\text{ML}} - \theta) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{I_\theta}\right).$$

As a result, when an ML estimator is implemented at the fusion center, the FI of the quantized variable $\gamma(X_i)$ measures the asymptotic performance.

A measure of relative performance between two estimators is ARE, which is defined as follows [11].

Definition: If two estimators W_n and V_n satisfy

$$\begin{aligned} \sqrt{n}[W_n - \theta] &\xrightarrow{d} \mathcal{N}(0, \sigma_W^2) \\ \sqrt{n}[V_n - \theta] &\xrightarrow{d} \mathcal{N}(0, \sigma_V^2) \end{aligned}$$

then the ARE of V_n with respect to W_n is

$$\text{ARE}(V_n, W_n) = \frac{\sigma_W^2}{\sigma_V^2}. \quad (3)$$

Since any ML estimator satisfies the convergence criterion as stated in Definition 1, the ARE can be used as a relative measure of performance between ML estimators based on quantized and unquantized observations. If p_θ is the pdf of unquantized observations, then the ARE is equivalently expressed as the ratio of FIs I_θ/J_θ where

$$J_\theta = \int_{\mathcal{X}} p_\theta(x) \left(\frac{d \log p_\theta(x)}{d\theta} \right)^2 dx \quad (4)$$

is the FI based on the unquantized observations. Since the FI J_θ is fixed for a given prior distribution, the ARE between quantized and unquantized ML estimators represents a measure of the asymptotic performance of the quantizer.

It is easily seen that, for every value of θ , the performance of the quantized estimator is worse than that of the unquantized observations, or in other words, the ARE is less than 1 for all θ . The ARE is related to the ratio of sample sizes required by the two estimators for the same asymptotic performance. This relation can be derived using the notion of *asymptotic effective variance* [12]. The effective standard deviation of an estimator $\hat{\theta}^{(n)}$ can be defined as the solution $\tau(n, \epsilon)$ of the equation

$$\Pr\{|\hat{\theta}^{(n)} - \theta| \geq \epsilon\} = \Pr\{|N| \geq \epsilon/\tau\}$$

where N is the standard normal variable. It was shown in [12] that, for an ML estimator

$$\lim_{n \rightarrow \infty, \epsilon \rightarrow 0} n\tau^2(n, \epsilon) = \frac{1}{I_\theta}.$$

Therefore, if $\tau_\gamma(n_1, \epsilon)$ and $\tau_u(n_2, \epsilon)$ represent the effective standard deviations of the quantized and unquantized estimators, respectively, then the ARE between the two ML estimators is equivalent to

$$\text{ARE} = \frac{I_\theta}{J_\theta} = \lim_{n_1, n_2 \rightarrow \infty, \epsilon \rightarrow 0} \frac{n_2 \tau_u^2(n_2, \epsilon)}{n_1 \tau_\gamma^2(n_1, \epsilon)}.$$

If α is the ratio of sample sizes between the quantized and unquantized estimators such that their effective variances are asymptotically equal, then $I_\theta/J_\theta = \alpha$ implies

$$\lim_{n \rightarrow \infty, \epsilon \rightarrow 0} \frac{\tau_u^2(\alpha n, \epsilon)}{\tau_\gamma^2(n, \epsilon)} = 1. \quad (5)$$

In other words, the ARE is equivalent to the ratio of sample sizes between the unquantized and quantized ML estimators such that their effective variances are asymptotically equal.

B. Maximin ARE Criterion

At any fixed value of θ , the ARE between ML estimators for quantized and unquantized observations is a measure of de-

crease in performance due to quantization. Since a quantizer is designed without any knowledge of the parameter, it is not possible to design a single quantizer that has the best ARE at every value of θ . We, therefore, consider the design of a quantizer that maximizes the minimum ARE over Θ . In other words, if $\hat{\theta}_u$ and $\hat{\theta}_\gamma$ are the ML estimators that use unquantized and quantized (using γ) observations, respectively, then we wish to design quantizer γ^* such that

$$\gamma^* = \arg \max_{\gamma \in \Gamma_D} \min_{\theta} \text{ARE}(\hat{\theta}_\gamma, \hat{\theta}_u) = \arg \max_{\gamma \in \Gamma_D} \min_{\theta} \frac{I_\theta}{J_\theta}$$

where Γ_D represents the set of all deterministic quantizers of size D . If n_γ/n_u represents the ratio of sample sizes such that the ML estimators have asymptotically equal effective variances, then from (5), we know that

$$\gamma^* = \arg \min_{\gamma \in \Gamma_D} \max_{\theta} \frac{n_\gamma}{n_u}.$$

The optimal quantizer γ^* , thus, minimizes the maximum percentage of additional samples $(n_\gamma - n_u)/n_u$ required by the ML estimator due to quantization. Furthermore, if the observations can be expressed as $X = \theta + W$, where W is a random variable with infinite support, it is easily shown that J_θ is a constant for all θ . The optimal quantizer γ^* for such distributions requires the least sample size amongst all deterministic quantizers of size D .

In the following two sections, we provide techniques to design a single quantizer that maximizes minimum ARE and thereby infer the optimal structure for the quantizer. In Section VI, we use the proposed structure and design multiple nonidentical quantizers for maximin ARE estimation.

III. SCORE-FUNCTION QUANTIZERS

A deterministic quantizer is represented by a mapping from the observation space into a finite set. Unless some restrictions are imposed on the structure of the quantizer, this representation could be highly complex. To this extent, we propose the use of SFQs to provide the structure for the ARE quantizer.

SFQs, introduced in [9], are represented by $D - 1$ thresholds on the *score-function* of the observation, given by

$$S_\theta(x) = \frac{d}{d\theta} \log p_\theta(x)$$

where $p_\theta(x)$ is the density of the unquantized variable X . The existence of the score-function is subject to some regularity conditions (see [10, p. 183]). SFQs are formally defined as follows.

Definition 1: Let \mathcal{T} denote the set of all vectors $t = (t_1, \dots, t_{D-1}) \in \mathbb{R}^{D-1}$ satisfying $-\infty \leq t_1 \leq \dots \leq t_{D-1} \leq \infty$. For any $t \in \mathcal{T}$, the associated intervals I_1, \dots, I_D are defined by $I_1 = [-\infty, t_1], I_2 = [t_1, t_2], \dots, I_D = [t_{D-1}, \infty]$. We say that a quantizer γ_θ is a monotone SFQ at θ (see Fig. 2) with threshold vector $t \in \mathcal{T}$, if

$$\gamma_\theta(x) = d \iff S_\theta(x) \in I_d, \quad \forall d.$$

We say that a quantizer is an SFQ at θ if there exists a permutation mapping $\pi : \{1, \dots, D\} \mapsto \{1, \dots, D\}$ such that $\pi \circ \gamma_\theta$ is a monotone SFQ at θ .

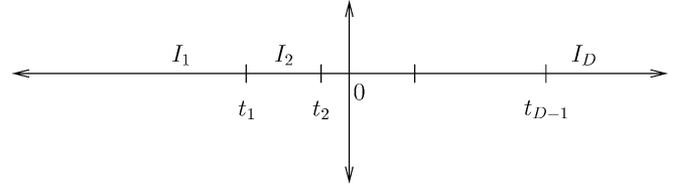


Fig. 2. Monotone SFQ : $I_i = (t_{i-1}, t_i)$ represent the partition i .

The motivation to use SFQs to design the maximin ARE quantizer comes from [9] where it was shown that SFQs, when coupled with ML estimators, exhibit some optimal properties. In particular, our approach is based on the following result.

Theorem 1: (from [9]) At a given θ , the FI I_θ is maximized by an SFQ at that value of θ .

In other words, by evaluating the optimal SFQ at any value of θ , the maximum possible FI for that θ can be achieved. Since J_θ is fixed for a given distribution, this quantizer also achieves the maximum ARE at that θ . Therefore, by evaluating the best SFQ at every value of θ , one can obtain an upper bound (genie bound) on quantizer performance for a given distribution of observations. The optimality of SFQs in estimation is analogous to the optimal properties of likelihood-ratio quantizers (LRQ) in detection [13].

Using this optimality as a heuristic, we restrict our criterion to the design of quantizer γ^* that satisfies

$$\gamma^* = \arg \max_{\gamma \in \mathcal{S}_D} \min_{\theta} \text{ARE}(\hat{\theta}_\gamma, \hat{\theta}_u)$$

where \mathcal{S}_D represents the set of all SFQs of size D .

SFQs, owing to their dependence on the value of θ , are not directly applicable to the quantizer design problem. However, under some conditions on the underlying family of distributions $\{p_\theta\}$, the class of SFQs at all parameter values can be shown to be identical. Furthermore, under those conditions, the SFQs have a simple structural representation in the observation space. This is evident from the following Lemma.

Lemma 1: Let \mathcal{S}_θ represent the class of SFQs at the parameter value θ , i.e.,

$$\mathcal{S}_\theta = \{\gamma_\theta : \gamma_\theta \text{ is an SFQ at } \theta\}.$$

Let T be a sufficient statistic for p_θ . If the score function can be expressed as $S_\theta(x) = f_\theta[T(x)]$, such that f_θ is monotone increasing for all θ , as follows.

- 1) The class of SFQs \mathcal{S}_θ at parameter θ is identical for all θ .
- 2) Every SFQ γ_θ can be equivalently expressed as $D - 1$ thresholds on the sufficient statistic. In other words, there exists $(t_1, \dots, t_{D-1}) \in \mathbb{R}^{D-1}$ such that $\forall k$

$$\gamma_\theta(x) = k \iff T(x) \in [t_{k-1}, t_k].$$

Proof: Since the score-function is monotonic in the sufficient statistic, any sequence of thresholds on the score-function will retain the same order in the sufficient statistic domain as well. Hence, the class of SFQs are independent of θ . \square

Therefore, when the family of distributions satisfy the monotonicity property, any SFQ corresponds to a sequence of $D - 1$ thresholds on the sufficient statistic $T(x)$. An example of a

class of distributions that satisfy the monotonicity property is the exponential family of distributions. Distributions such as the asymmetric Gaussian mixture do not belong to this class.

In order to design the optimal SFQ that maximizes the minimum ARE, it suffices to obtain the optimal thresholds on $T(x)$. In general, an exhaustive search for optimal thresholds after finely discretizing the observation space would yield the required quantizer within a small margin of error. Such techniques would, however, be computationally intensive. We, therefore, propose an iterative algorithm that evaluates the optimal quantizer efficiently.

A. Iterative Algorithm

The algorithm is initialized with arbitrary thresholds $-\infty < t_1 < \dots < t_{D-1} < \infty$ on the sufficient statistic. At every iteration, the lowest threshold is chosen to maximize the metric while keeping the other thresholds fixed. This process is carried out for each subsequent threshold and the entire iteration is repeated until the change in thresholds is negligible. Since the minimum ARE cannot decrease at any iteration, the algorithm converges.

Let $M(\gamma)$ denote the minimum ARE of a quantizer γ , i.e.,

$$M(\gamma) = \min_{\theta} \frac{I_{\theta}(\gamma)}{J_{\theta}}. \quad (6)$$

- 1) *Initialization:* Let $\mathbf{t} = \{t_1, \dots, t_{D-1}\} \in \mathcal{T}^{D-1}$ represent the threshold set of the quantizer. Set $t_0 = -\infty, t_D = \infty$ and initialize \mathbf{t} arbitrarily such that $t_i < t_{i+1}$. Evaluate the probability mass function of the quantized variable as

$$q_{\theta}(i) = P_{\theta}\{T(x) \in [t_{i-1}, t_i]\}.$$

For this q_{θ} , evaluate the FI I_{θ} [in (2)] and the metric $M(\gamma)$ [in (6)].

- 2) *Iteration $n+1$:* Let t_j^n represent the value of threshold j in iteration n . Each iteration is divided into $D-1$ steps. At step j , the new values for thresholds t_1, \dots, t_{j-1} would have been obtained. Threshold t_j is then picked as that value between the new t_{j-1} and the old t_{j+1} which maximizes the metric. The j th step, therefore, corresponds to obtaining the optimal value for t_j given the other thresholds. More formally, for every $t \in [t_{j-1}^{n+1}, t_{j+1}^n]$, let γ_j^t represent the quantizer

$$\gamma_j^t(x) = \begin{cases} i & T(x) \in [t_{i-1}^{n+1}, t_i^{n+1}], \quad i < j \\ i & T(x) \in [t_{i-1}^n, t_i^n], \quad i > j + 1 \\ j & T(x) \in [t_{j-1}^{n+1}, t] \\ j + 1 & T(x) \in [t, t_{j+1}^n]. \end{cases}$$

The optimal threshold is chosen as

$$t_j^{n+1} = \arg \max_t M(\gamma_j^t).$$

Let the optimal quantizer at the end of $D-1$ steps be denoted as γ^{n+1} and the corresponding value of the metric be $M(\gamma^{n+1})$.

- 3) *Termination:* Choose ϵ as close to zero as required. At iteration n , evaluate $M(\gamma^n)$. If $|M(\gamma^n) - M(\gamma^{n-1})| < \epsilon$,

then the algorithm terminates. In other words, when the change in performance is very close to zero, we terminate the algorithm.

Theorem 2: At the n th iteration, let $R^{(n)} = \max_{\theta} J_{\theta}/I_{\theta}^{(n)}$. Then $R^{(n+1)} \leq R^{(n)}, \forall n$. For any $\epsilon > 0$, the algorithm converges in finite steps.

Proof: See Appendix.

At every step of the algorithm, a new set of thresholds on the sufficient statistic are evaluated. The quantizer at the end of every iteration is, therefore, an SFQ. From Theorem 2, we know that the algorithm converges in finite iterations. Furthermore, in the absence of local minima, the algorithm converges to the best SFQ that optimizes the maximin ARE

$$\gamma^* = \arg \max_{\gamma \in \mathcal{S}_D} \min \text{ARE}(\hat{\theta}_{\gamma}, \hat{\theta}_u).$$

Note that the algorithm is valid only for those distributions that satisfy the monotonicity property of Lemma 1. The iterative algorithm can also be used to optimize other metrics within the class of SFQs by appropriately modifying the quantity $M(\gamma)$ [14].

Since the use of SFQ to design the quantizer was based on a heuristic, it is not known if it is the optimal quantizer structure for the maximin ARE criterion. In the following section, we, therefore, consider an alternate approach to design the optimal quantizer, when no restrictions are placed on quantizer structure or nature of distribution.

IV. GENERAL QUANTIZER DESIGN

In this section, we present an iterative algorithm to obtain the maximin ARE quantizer without any assumptions on partition structure in the observation space or nature of distribution. Such an algorithm is generally computationally intensive and may not be useful for practical implementation. However, we present this approach to estimate the loss in performance due to the SFQ assumption.

The basic idea for this technique comes from the fact that the FI at any fixed θ is the variance of the score-function at that θ . Mathematically, if $\{S_{\theta}^u(x), x \in \mathcal{X}\}$ and $\{S_{\theta}^q(i), i = 1 \dots D\}$ represent the score-functions of the unquantized and quantized observations, respectively, then

$$S_{\theta}^q(i) = \frac{d}{d\theta} \log q_{\theta}(i), \quad I_{\theta} = \sum_{i=1}^D q_{\theta}(i) (S_{\theta}^q(i))^2$$

$$S_{\theta}^u(x) = \frac{d}{d\theta} \log p_{\theta}(x), \quad J_{\theta} = \int_{\mathcal{X}} p_{\theta}(x) (S_{\theta}^u(x))^2 dx.$$

Furthermore, it is easily shown that the difference in FIs I_{θ} and J_{θ} can be expressed as MSE of the quantizer score-function $S_{\theta}^q(i)$ with respect to $S_{\theta}^u(x)$, i.e.,

$$J_{\theta} - I_{\theta} = \sum_{i=1}^D \int_{\{x: \gamma(x)=i\}} p_{\theta}(x) (S_{\theta}^u(x) - S_{\theta}^q(i))^2 dx. \quad (7)$$

We use (7) as the basis for iteration in the quantizer algorithm. The iterative algorithm is, in principle, similar to the SFQ algorithm in Section III-A. Instead of evaluating thresholds in

each iteration, we compute the optimal partition index $\gamma(X)$ for each observation¹ X , keeping the others fixed. The algorithm converges because the minimum ARE cannot decrease at any iteration.

Let the partitions in \mathcal{X} in iteration n be represented by $(L^{(n)}(1), \dots, L^{(n)}(D))$. If the quantizer at the iteration n is represented by $\gamma^{(n)}$, then

$$L^{(n)}(i) = \{x : \gamma^{(n)}(x) = i\}.$$

The formal statement of the algorithm is as follows.

- **Initialization:** Divide \mathcal{X} into D arbitrary nonoverlapping partitions represented by $\{L^{(0)}(i)\}$. Evaluate the score-function and FI as

$$S_{\theta}^{(0)}(i) = \frac{d}{d\theta} \log q_{\theta}^{(0)}(i) = \frac{d}{d\theta} \log P_{\theta}\{x \in L^{(0)}(i)\}$$

$$I_{\theta}^{(0)} = \sum_{i=1}^D \frac{1}{q_{\theta}^{(0)}(i)} \left(\frac{dq_{\theta}^{(0)}(i)}{d\theta} \right)^2.$$

- **Iteration $n+1$:** Let $\theta^* = \arg \max_{\theta} J_{\theta}/I_{\theta}^{(n)}$. For every $x \in L^{(n)}(j)$, let

$$k^* = \arg \min_k \{M_n^k(x), \quad k = 1, \dots, D\}$$

$$M_n^k(x) = \begin{cases} \max_{\theta} \frac{p_{\theta}(x)}{J_{\theta}} (S_{\theta}(x) - S_{\theta}^{(n)}(k))^2 & k \neq j \\ \frac{p_{\theta^*}(x)}{J_{\theta^*}} (S_{\theta^*}(x) - S_{\theta^*}^{(n)}(k))^2 & k = j \end{cases}$$

$$x \in L^{(n+1)}(k^*).$$

- **Termination:** Choose ϵ close to zero. At each iteration, evaluate $I_{\theta}^{(n)}$. If $\max_{\theta} |I_{\theta}^{(n)} - I_{\theta}^{(n-1)}| < \epsilon$, then the algorithm terminates. In other words, when the change in FI is very close to zero, we terminate the algorithm.

Theorem 3: At the n^{th} iteration, let $R^{(n)} = \max_{\theta} J_{\theta}/I_{\theta}^{(n)}$. Then $R^{(n+1)} \leq R^{(n)}, \forall n$. For any $\epsilon > 0$, the algorithm converges in finite steps.

Proof: Refer to Appendix.

This algorithm, although computationally intensive, can be used for any underlying distribution (subject to regularity conditions) and, in the absence of local minima, provides the optimal quantizer

$$\gamma^* = \arg \max_{\gamma \in \Gamma_D} \min \text{ARE}(\hat{\theta}_{\gamma}, \hat{\theta}_u).$$

As is evident from the description of the algorithm, the final quantizer is represented as an index for each observation and, therefore, does not specify a regular structure on the observation space. On the other hand, the quantizer resulting from the SFQ algorithm in Section III-A has a simple structure, but the optimization is over a subset of deterministic quantizers. It remains to be seen if there is a loss in performance due to the structural constraints. In the following section, we present some numerical results and illustrate that, for the distributions where the SFQ algorithm is valid, the two quantizers are identical.

¹The observation space is uniformly partitioned into fine bins such that the set of observations is large but finite.

V. NUMERICAL RESULTS

As mentioned in Section III, it is known that at any particular θ , the FI can be maximized [9] by an SFQ (optimized at that θ). The maximal FI, thus, obtained serves as a good upper bound to the performance of any quantizer.

In this section, we consider the performance of the optimal quantizer for two different distributions : parameter in additive and multiplicative Gaussian noise. Since these distributions satisfy the monotonicity property of Lemma 1, the optimal SFQ was also evaluated through the algorithm in Section III. In both situations, the two algorithms resulted in identical quantizers. Although, in the examples considered in this section, the algorithms discussed did not result in a local minima, we do not have an analytical guarantee for their absence.

A. Multiplicative Gaussian Noise

The observation X_i is a faded version of the parameter θ , where the fading is assumed Gaussian

$$X = H\theta, \quad H \sim \mathcal{N}(0, \sigma_h^2), \Theta \subset \mathbb{R}^+.$$

The FI in this case is a decreasing function of θ .

Fig. 3 plots the ARE and FI of the maximin ARE quantizer. The performance is compared to the genie bound obtained by evaluating the best SFQ at every value of θ (Theorem 1). The figures also plot the performance of a quantizer which guesses the value of θ to be the average value $((\theta_{\min} + \theta_{\max})/2)$ in the parameter set. It can be seen that the shape of the maximin ARE quantizer curve follows that of the unquantized FI and the genie bound. Furthermore, the performance of the maximin ARE quantizer is significantly better than that obtained by just using the average value of the parameter.

B. Parameter in AWGN

The observation is of the form

$$X = \theta + N, \quad N \sim \mathcal{N}(0, \sigma^2).$$

The FI of the unquantized variable in this scenario is a constant for all values of θ

$$J_{\theta} = \frac{1}{\sigma^2}.$$

Therefore, maximizing the minimum ARE is equivalent to maximizing the minimum FI across the parameter. Fig. 4 plots the FI of the optimal quantizer for different values of D .

It is interesting to observe that for $D = 2$, the maximin ARE quantizer and the optimal SFQ at $\theta = (\theta_{\min} + \theta_{\max})/2$ are identical. In other words, the quantizer designed by guessing the parameter value to be $(\theta_{\min} + \theta_{\max})/2$ is optimal. However, for $D = 4$, designing a quantizer by making a similar guess results in poor minimum ARE, although the quantizer equals the genie bound at the guessed value of θ .

As number of quantization bins D increases, we see that the performance of the maximin quantizer tracks the genie bound more closely. It is also interesting to see that the maximin ARE

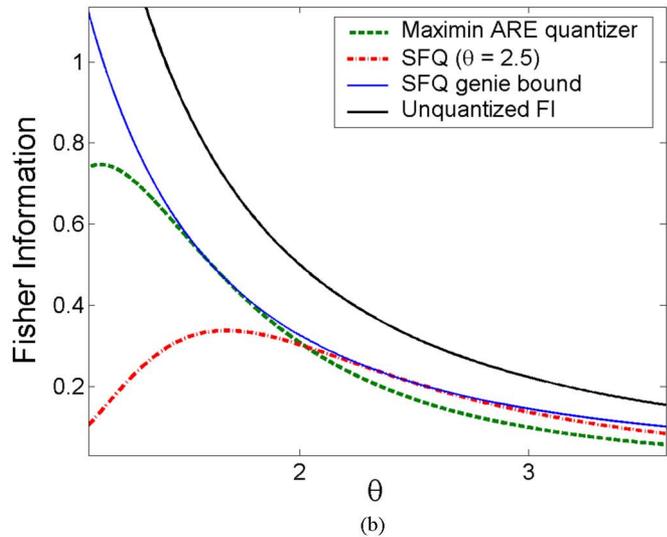
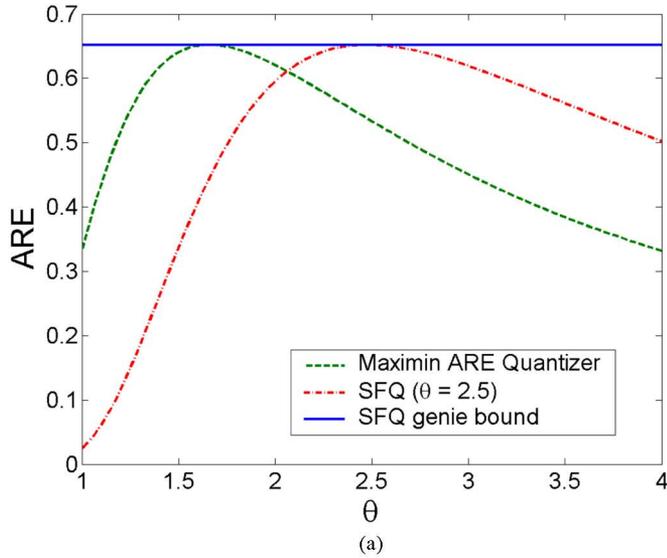


Fig. 3. Maximin ARE quantizer for multiplicative noise: $\theta \in [1, 4], \sigma_h^2 = 1, D = 2$. (a) Comparison of quantizer ARE. (b) Comparison of quantizer FI.

quantizer for both distributions result in ARE curves where the extreme values of the parameter have nearly equal performance.

For the multiplicative noise distribution, the score-function $S_\theta(x) = x^2/\theta^3 - 1/\theta$ which is monotone in x^2 , whereas the score-function for the AWGN distribution is given by $S_\theta(x) = x - \theta$ which is monotone in x . The SFQs for the multiplicative and additive noise observations are, therefore, expressible as thresholds on the x^2 and x , respectively. Fig. 5 illustrates the partition information of the optimal quantizers for the two distributions when $D = 4$. The use of SFQs significantly simplifies the design of quantizer, especially when the observations are vectors. For example, if the observation at each node is comprised of multiple elements, each of which are independent and identically distributed as the parameter in AWGN, then the SFQ corresponds to $D - 1$ thresholds on the sum of the elements of the vector observation.

In both of the examples considered, the unrestricted quantizer algorithm of Section IV resulted in SFQs. Furthermore, for $D = 2$, the quantizer ARE is equal to the maximum possible value

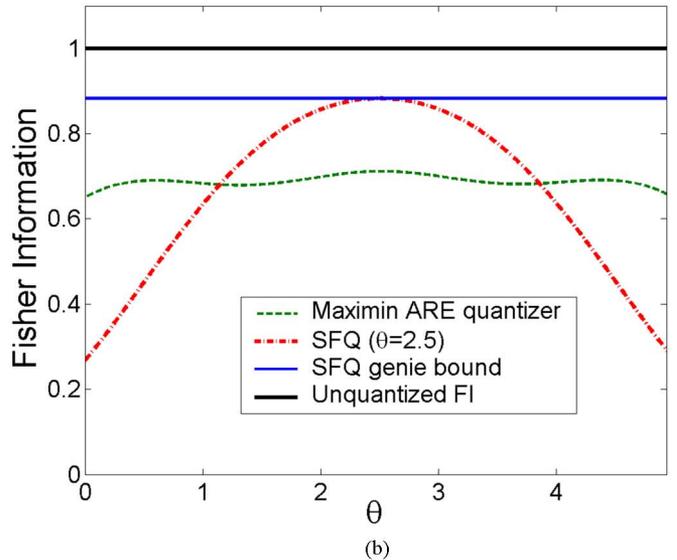
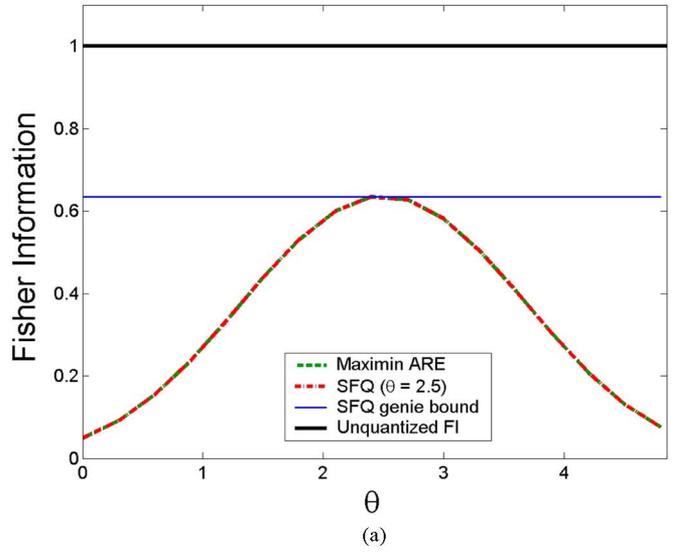


Fig. 4. FI comparison for parameter in AWGN: $\theta \in [0, 5], \sigma^2 = 1$. (a) $D = 2$. (b) $D = 4$.

(genie bound) at some θ . These examples and some others (such as Gaussian mixture) which we simulated seem to indicate that the SFQ provides the optimal structure. We conjecture that for distributions that satisfy the monotonicity property, the maximin ARE quantizer is an SFQ.

Designing identical quantizers is generally suboptimal, as will be seen in the next section. However, it provides some practical advantages especially in large-scale networks, such as sensor networks, where it is preferable for nodes to operate identically. Moreover, when quantizers are distinct, it would be necessary for the fusion center to obtain the quantizer information of each node as well, which could lead to communication overhead.

VI. MULTIPLE QUANTIZER DESIGN

The material discussed so far dealt with optimal quantization assuming all sensors use identical quantizers. For a distributed detection setup with binary hypotheses, it has been shown [15]

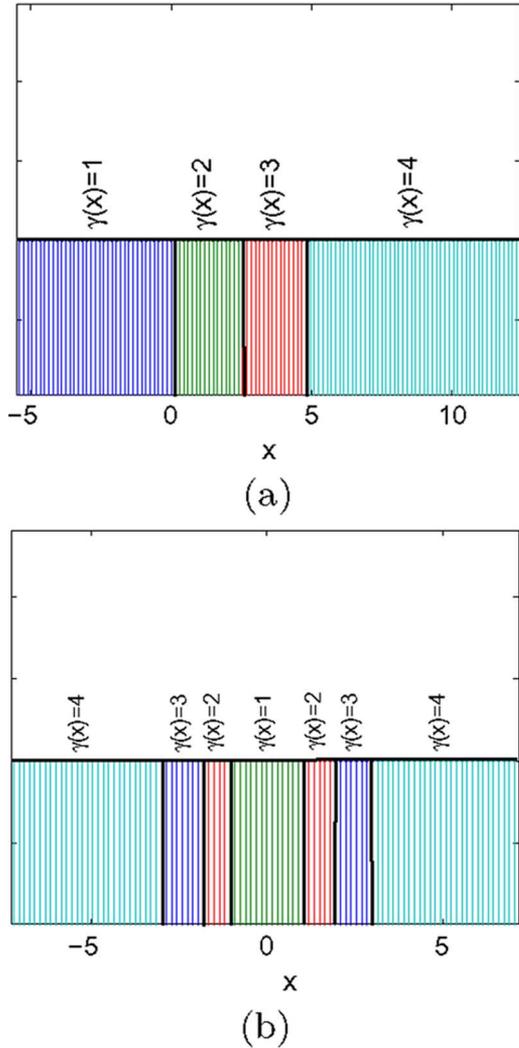


Fig. 5. Partitions in \mathcal{X} domain for a 2-bit quantizer. (a) Additive noise with $\sigma^2 = 1$: thresholds on x . (b) Multiplicative noise with $\sigma^2 = 1$: thresholds on x^2 .

that use of identical quantizers is optimal. However, for a general M -ary hypothesis testing, the maximum number of quantizers required is given by $M(M - 1)/2$. We, therefore, expect that for distributed estimation, using identical quantizers would not be optimal. This can be seen from the following example.

Consider the estimation of a parameter $\theta \in [\theta_1, \theta_2]$ in the presence of additive white noise. If we restrict ourselves to SFQs, the optimal 1-bit quantizer can be shown to have a single threshold at $(\theta_1 + \theta_2)/2$. However, if we were to design two 1-bit quantizers with different thresholds and use the average ARE, it is easily seen from Fig. 6 that the estimation performance improves significantly. Therefore, within the class of SFQs, identical quantizers are suboptimal.

In order to design multiple quantizers, the algorithm presented in Section III can be coupled with a PBPO [16] across quantizers. Since the observations at multiple nodes are independently distributed, the FI of the total received data at the fusion center is the sum of FIs of the individual quantizers. In other words, let $\gamma_1, \dots, \gamma_k$ represent the quantizers for k nodes

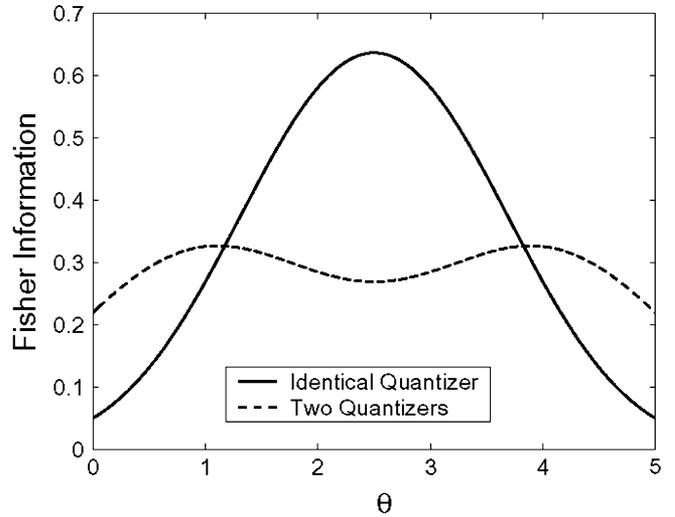


Fig. 6. AWGN parameter with $\sigma^2 = 1$: Comparison between 1) FI of optimal SFQ and 1) Average FI of two 1-bit SFQs with thresholds $t = 1$ and $t = 4$.

and $I_\theta(\gamma_i)$ be the FI for quantizer γ_i . Then, the ARE of the joint observation at the fusion center is given by

$$\text{ARE}(\gamma_1, \dots, \gamma_k) = \frac{\sum_{i=1}^k I_\theta(\gamma_i)}{kJ_\theta}$$

The algorithm in Section III is modified as follows. The metric $M(\gamma)$ is rewritten as

$$M(\gamma_1, \dots, \gamma_k) = \min_{\theta} \frac{\sum_i I_\theta(\gamma_i)}{J_\theta}$$

Each iteration now involves optimizing $k(D - 1)$ thresholds. Each threshold of the first quantizer is optimized assuming all other thresholds and the remaining $k - 1$ quantizers are fixed. This process is subsequently repeated for each quantizer in succession. The entire iteration is repeated until the change in $M(\cdot)$ is negligible. Since the basic iteration is identical to the algorithm in Section III, it is easily seen that the metric does not increase at any iteration. Furthermore, since the ARE is bounded, the algorithm converges in finite steps.

For a large-scale network, the algorithm would be computationally intensive. However, as will be seen from numerical results, it may not be necessary to design as many quantizers as the number of nodes in the network.

A. Numerical Results

We consider the example of multiplicative Gaussian noise and implement the algorithm for multiple quantizer design. The ARE of nonidentical quantizers is plotted in Fig. 7. It can be seen that, by increasing the number of quantizers, the minimum ARE at the fusion center can improve significantly.

The design of multiple nonidentical quantizers becomes imperative for low noise observations. Fig. 8 plots the improvement in minimum ARE with design of additional quantizers for different noise levels of observations. As can be seen from the figure, as the noise level reduces, the improvement due to making quantizers nonidentical increases.

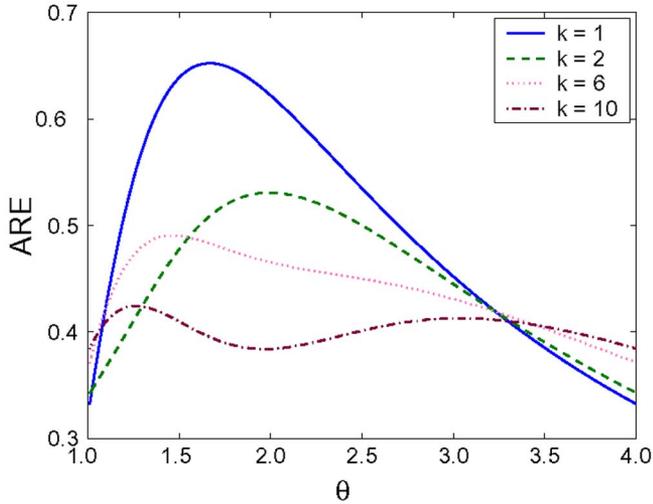


Fig. 7. ARE for nonidentical quantizers in multiplicative noise : $\theta \in [1, 4], \sigma^2 = 1$.

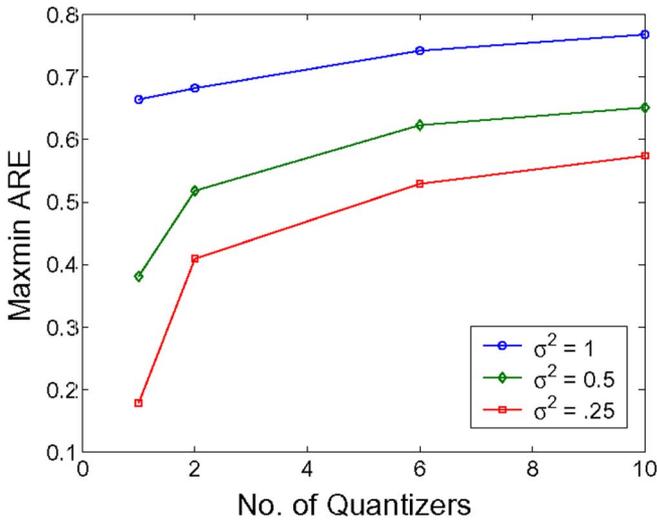


Fig. 8. Minimum ARE for nonidentical quantizers at different multiplicative noise levels : $\theta \in [1, 4]$.

It is interesting to note that the improvement in minimum ARE saturates with the increase in number of quantizers designed. For a large-scale network, it would, therefore, be sufficient to design much fewer quantizers than the number of nodes in the network. The nodes can be divided into groups based on the number of quantizers designed, and each group implements one of the designed quantizers. This greatly simplifies node design and reduces the communication overhead in large-scale networks.

VII. CONCLUSION AND FUTURE EXTENSIONS

In this paper, we considered the performance criterion, *maximin* ARE, to design quantizers for distributed estimation of a deterministic parameter. We proposed the use of SFQs to design the optimal quantizer for a certain class of distributions that satisfy a monotonicity property. Through numerical simulations, we illustrated that under those conditions, the SFQ provides the optimal structure. Furthermore, SFQs can be expressed as a set of thresholds on the sufficient statistic which significantly re-

duces the complexity of implementation. SFQs also have some generic optimal properties [9] that strongly suggest their utility value in ML estimation.

The application of this metric is not restricted to distributed estimation. In practical situations of classical point estimation, the estimator may not be capable of handling several high resolution measurements to estimate the parameter, and it is necessary to compress the data into few bits to reduce complexity. Currently, the setup is simple and assumes identical distributions and quantizers. The extension of this idea to non-i.i.d. distributions and noisy channels is an interesting direction to pursue. Improving the metric to handle vector parameters is also a nontrivial future extension.

APPENDIX

Proof of Theorem 2: Let the quantizer have D partitions. Since the quantizer is a monotone SFQ, we know that this corresponds to $D - 1$ thresholds on the real line. Therefore, the FI can be expressed as the sum of expectations at each partition, i.e.,

$$I_\theta = \sum_i I_\theta^i$$

where

$$I_\theta^i = q_\theta(i) \left(\frac{d}{d\theta} \log q_\theta(i) \right)^2.$$

We know that, at every step, threshold t_i affects only the terms I_θ^i and I_θ^{i+1} . Therefore, at iteration n , the new threshold $t_i^{(n)}$ is chosen such that

$$\min_\theta \frac{I_\theta^{i,(n)}}{J_\theta} + \frac{I_\theta^{i+1,(n)}}{J_\theta} \geq \min_\theta \frac{I_\theta^{i,(n-1)}}{J_\theta} + \frac{I_\theta^{i+1,(n-1)}}{J_\theta}.$$

Therefore, the new minimum ARE increases at every step, i.e.,

$$\min_\theta \frac{I_\theta^{(n)}}{J_\theta} \geq \min_\theta \frac{I_\theta^{(n-1)}}{J_\theta}.$$

Since the ARE is upper bounded by 1, the algorithm converges in finite steps.

Proof of Theorem 3: Consider the minimax criterion

$$\begin{aligned} \arg \min_\gamma \max_\theta \frac{J_\theta}{I_\theta} &= \arg \max_\gamma \min_\theta \frac{I_\theta}{J_\theta} \\ &= \arg \min_\gamma \max_\theta \left(1 - \frac{I_\theta}{J_\theta} \right) \\ &= \arg \min_\gamma \max_\theta \frac{J_\theta - I_\theta}{J_\theta}. \end{aligned}$$

We know from (7) that

$$\frac{J_\theta - I_\theta}{J_\theta} = \sum_i \int_{T_i} \frac{p\theta(x)}{J_\theta} [S_\theta(x) - S_\theta(i)]^2 dx$$

where $T_i = \{x : \gamma(x) = i\}$. Using the steps of iteration, we get

$$\begin{aligned} \max_\theta \frac{J_\theta - I_\theta^{(n)}}{J_\theta} &= \max_\theta \sum_i \int_{T_i^{(n)}} \frac{p\theta(x)}{J_\theta} [S_\theta(x) - S_\theta^{(n)}(i)]^2 dx \end{aligned}$$

$$\begin{aligned}
&\leq \max_{\theta} \sum_i \int_{T_i^{(n)}} \frac{p_{\theta}(x)}{J_{\theta}} \left[S_{\theta}(x) - S_{\theta}^{(n-1)}(i) \right]^2 dx \\
&\leq \sum_i \int_{T_i^{(n-1)}} \frac{p_{\theta^*}(x)}{J_{\theta^*}} \left[S_{\theta^*}(x) - S_{\theta^*}^{(n-1)}(i) \right]^2 dx \\
&= \max_{\theta} \frac{J_{\theta} - J_{\theta}^{(n-1)}}{J_{\theta}}.
\end{aligned}$$

We know that $I_{\theta} \leq J_{\theta}$ for every value of θ . Therefore, the metric is lower bounded by 0. Since the algorithm cannot increase the metric at every iteration, in the absence of local minima, the inequalities are strict, and, hence, the algorithm converges to the optimal quantizer in finite steps.

ACKNOWLEDGMENT

The authors would like to thank G. Mergen for his contribution in formulating the concept of score-function quantization. They would also like to thank Prof. Veeravalli for his valuable suggestions regarding the multiple quantizer design.

REFERENCES

- [1] Z. Zhang and T. Berger, "Estimation via compressed information," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 198–211, Mar. 1988.
- [2] T. S. Han and S. Amari, "Parameter estimation with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1802–1833, Nov. 1995.
- [3] W. M. Lam and A. R. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. Commun.*, vol. 41, no. 11, pp. 1602–1605, Nov. 1993.
- [4] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.
- [5] K. Zhang and X. R. Li, "Optimal sensor data quantization for best linear unbiased estimation fusion," in *Proc. 43rd IEEE Conf. Decision Control*, Dec. 2004, vol. 3, pp. 2656–2661.
- [6] M. L. Fowler and M. Chen, "Fisher-information-based data compression for estimation using two sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 3, pp. 1131–1137, Jul. 2005.
- [7] A. Ribiero and G. B. Giannakis, "Non-parametric distributed quantization-estimation using wireless sensor networks," presented at the ICASSP Conf., Philadelphia, PA, Mar. 2005.
- [8] 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.
- [9] P. Venkatasubramaniam, G. Mergen, L. Tong, and A. Swami, "Quantization for distributed estimation in large scale sensor networks," in *Proc. ICISIP*, Bangalore, India, Dec. 2005, pp. 121–126.
- [10] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [11] G. Casella and R. Berger, *Statistical Inference*. Pacific Grove, CA: Brooks/Cole, 1990.
- [12] R. R. Bahadur, "On the asymptotic efficiency of tests and estimates," *SANKHYA: Ind. J. Statist.*, vol. 22, no. 3, pp. 229–252, 1960.
- [13] J. N. Tsitsiklis, "Extremal properties of likelihood ratio quantizers," *IEEE Trans. Commun.*, vol. 41, no. 4, pp. 550–558, Apr. 1993.
- [14] P. Venkatasubramaniam, L. Tong, and A. Swami, "Score-function quantization for distributed estimation," presented at the CISS, Princeton, NJ, Mar. 2007.
- [15] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control Signals Syst.*, pp. 167–182, 1988, no. 1.
- [16] R. Radner, "Team decision problems," *Ann. Math. Statist.*, vol. 33, pp. 857–881, 1962.



Parvathinathan Venkatasubramaniam (S'03) received the B.Tech. degree from the Department of Electrical Engineering, Indian Institute of Technology, Madras, in 2002. He is currently pursuing the Ph.D. degree in electrical engineering at Cornell University, Ithaca, NY.

His research interests include network security, signal processing in wireless networks, and multiple access protocols.

Mr. Venkatasubramaniam received the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with Dr. S. Adireddy and Prof. L. Tong) and a Best Student Paper Award at the 2006 IEEE International Conference of Acoustics, Speech, and Signal Processing.



Lang Tong (F'05) received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1991.

He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University, Stanford, CA, in 1991. He joined Cornell University, Ithaca, NY, in 1998, where he is now the Irwin and Joan Jacobs Professor in Engineering. Prior to joining Cornell University, he was on faculty at West Virginia University, Morgantown, and the University of Connecticut, Storrs. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology, Delft, The Netherlands.

Dr. Tong received the Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 best paper award (with M. Dong) from the IEEE Signal Processing Society, and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with P. Venkatasubramaniam and S. Adireddy). He also received the Young Investigator Award from the Office of Naval Research and is the coauthor of five student paper awards.

Dr. Tong received the Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 best paper award (with M. Dong) from the IEEE Signal Processing Society, and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with P. Venkatasubramaniam and S. Adireddy). He also received the Young Investigator Award from the Office of Naval Research and is the coauthor of five student paper awards.



Ananthram Swami (SM'96) received the B.Tech. degree from the Indian Institute of Technology, Bombay, the M.S. degree from Rice University, Houston, TX, and the Ph.D. degree from the University of Southern California, Los Angeles, all in electrical engineering.

He has held positions with Unocal Corporation, USC, CS-3, and Malgudi Systems. He was a Statistical Consultant to the California Lottery, developed a Matlab-based toolbox for non-Gaussian signal processing, and has held visiting faculty positions at

INP, Toulouse, France. He is currently with the U.S. Army Research Laboratory, Adelphi, MD, where his work is in the broad area of signal processing, wireless communications, and networking, including both sensor networks and MANETs.

Dr. Swami is the Chair of the IEEE Signal Processing Society's TC on Signal Processing for Communications, and an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II, EXPRESS BRIEFS and the *IEEE Signal Processing Magazine*. He was Co-Organizer and Co-Chair of the 1993 IEEE-SPS HOS Workshop, the 1996 IEEE-SPS SSAP Workshop, and the 1999 ASA-IMA Workshop on Heavy-Tailed Phenomena. He was Co-Guest Editor of a 2004 special issue of the *IEEE Signal Processing Magazine* on Signal Processing for Networking, a 2006 special issue of the *IEEE Signal Processing Magazine* on Distributed Signal Processing in Sensor Networks, a 2006 special issue of the *EURASIP JASP* on Reliable Communications over Rapidly Time-Varying Channels, and a 2006 special issue of the *EURASIP JWCN* on Wireless Mobile Ad Hoc Networks.