Score-Function Quantization for Distributed Estimation

Parvathinathan Venkitasubramaniam and Lang Tong
School of Electrical and Computer Engineering
Cornell University
Ithaca, NY 14853
Email: {pv45, lt35}@cornell.edu

Ananthram Swami
Army Research Laboratory
Adelphi, Maryland
Email: a.swami@ieee.org

Abstract— We study the problem of quantization for distributed parameter estimation. We propose the design of score-function quantizers to optimize different metrics of estimation performance. Score-function quantizers are a class of quantizers known to maximize the Fisher Information for a fixed value of parameter θ. We show that for distributions that satisfy a monotonicity property, the class of score-function quantizers can be made independent of parameter θ. We then propose a generic algorithm to obtain the optimal Score-function quantizer that can be used to maximize three different metrics; Minimum Fisher Information, Bayesian Fisher Information and Minimum Asymptotic Relative Efficiency. Through numerical examples, we illustrate that these algorithms converge to the optimal quantizers obtained through known algorithms for maximin ARE and Bayesian Fisher Information.

I. INTRODUCTION

A. Motivation

The design of optimal quantizers for distributed estimation is a classical problem in statistical inference theory. The basic setup is represented in Figure 1, wherein nodes observe data and transmit quantized versions to a common fusion center which estimates the underlying parameter. Quantization of the observed data before transmission is an important problem, especially in the context of sensor networks. The limited battery energy at the nodes coupled with the necessity to share the wireless channel for transmission makes quantization of the data crucial. The goal of any optimal quantizer design is to minimize the estimation error at the fusion center.

Consider a system as shown in Figure 1. Each node measures a random observation X_i independent and identically distributed under some scalar parameter θ. The nodes then transmit quantized versions of the observations, $\{\gamma_i(X_i)\}$, to the fusion center. We assume that the quantizers at the nodes are identical, i.e., $\gamma_i(\cdot) = \gamma(\cdot)$. The motivation behind this assumption is two fold. Firstly, the distribution of observations is identical across nodes. Second, in large scale networks, the nodes deployed in a given area are typically identical.
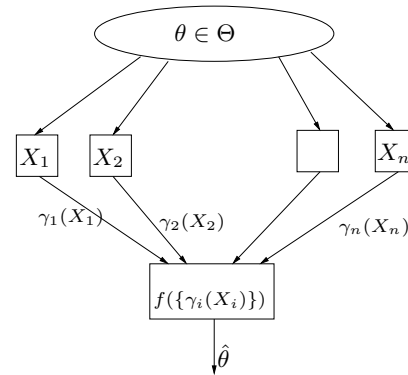
Fig. 1. Distributed Estimation System

The fusion center uses the received data $\{\gamma(X_i)\}$ to make an estimate $\hat{\theta}$ of the true parameter. The assumption of i.i.d observations at n different sensors is equivalent to that of n i.i.d observations at a single sensor. Depending on the complexity, a sensor may choose to perform a sub-optimal sample by sample quantization rather than a vector quantization on the set of n high resolution samples. Quantization is inevitable in digital processing, and trade-offs are between resolution, rate, power consumption of quantizer and the overall metric imposed by the problem.

The estimation performance is measured by the mean squared error of the estimate given by $\mathbb{E}(\hat{\theta} - \theta)^2$. Depending on the properties of the estimator, the mean squared error (MSE) can be expressed as a function of the distribution of quantized observations. It is well known that the MSE of any unbiased estimate is lower bounded by the inverse of Fisher Information (FI). Furthermore, if the quantized observations are distributed according to some probability mass function $\{\mathbf{q}_\theta\}$, and a Maximum-Likelihood (M-L) estimator based on $\{\mathbf{q}_\theta\}$ is implemented, then under certain regularity conditions, the MSE of the estimate asymptotically achieves this bound. The Fisher Information at a fixed value of parameter θ is maximized by a class of quantizers known as Score-function Quantizers (SFQ) [1]. Furthermore, it can be shown that this class of quantizers optimizes a general convex metric of the aposteriori distribution $\mathbf{q}_\theta$ at a given θ.

Although, SFQs are dependent on the parameter θ, for a

certain class of distributions, any SFQ can be expressed as a set of thresholds on the sufficient statistic of the distribution, independent of the parameter value. For this class of distributions, we propose the use of SFQs to optimize some estimation performance metrics. Since SFQs maximize Fisher Information for a given parameter value, we use this as a heuristic and design the best SFQ that optimizes known metrics of performance based on Fisher Information. The metrics that we consider are as follows. i) Maximin Fisher Information: Maximize the minimum Fisher Information over θ, ii) Maximin ARE: Maximize the minimum Asymptotic Relative Efficiency between quantized and unquantized M-L estimators over θ. iii) Maximum Bayesian Fisher Information: Maximize the average Fisher Information for a random parameter.

Optimal quantizers for maximin ARE and Bayesian Fisher Information have been proposed in literature [2], [3]. However, those algorithms do not specify a regular structure for the quantizer. Since SFQs can be expressed as thresholds on the sufficient statistic, they are easily represented and simple to implement. Furthermore, our numerical results indicate that for the class of distributions we consider, the SFQs are identical to the optimal quantizers for those two metrics.

B. Main Contributions

In this paper, we consider the design of quantizers within a sub-class of quantizers known as Score-Function Quantizers (SFQ). It has been shown that the maximum Fisher Information at a given parameter value is achieved by an SFQ. When the Score-function of the underlying distribution satisfies a monotonicity property, we show that the set of SFQs for a particular value of the parameter is the same for any value of the parameter. These quantizers can then be expressed as a set of thresholds on the sufficient statistic. For distributions that satisfy the monotonicity property, we present a generic algorithm to iteratively obtain the optimal SFQ for metrics based on Fisher Information. The metrics that we consider are Maximin Fisher Information, Maximin ARE between quantized and unquantized estimators and Bayesian Fisher Information. Through numerical simulations, we illustrate the performance of the algorithm for the three metrics. The optimal SFQs for Bayesian Fisher Information and Maximin ARE are shown to be identical to the quantizers obtained through known unrestricted quantizer design algorithms.

C. Related Work

There is extensive literature on data compression for distributed statistical inference. For a distributed detection setup, when sensors have i.i.d data, Tsitsiklis showed that quantizing the Likelihood Ratio optimizes detection performance metrics like Kullback-Leibler distance and Chernoff Information [4]. In the distributed estimation scenario, Lam and Reibman [3] developed an iterative quantizer to maximize the Bayesian Fisher Information of a random parameter. For certain restrictive types of estimators, Gubner [5] and Zhang and Li [6] provided optimal quantizers to minimize the mean squared

error for a random parameter. For a deterministic parameter in additive noise, Ribiero and Giannakis [7] showed that the quantizer that maximizes the Fisher Information at a particular θ is represented by a single threshold in the observation space. In a recent work [2], we presented an iterative algorithm to optimize the Maximin ARE metric without any assumptions on type of quantizer.

The paper is organized as follows. The basic system model and some fundamentals are discussed in Section II. Score-function quantizers and their optimality are explained in Section III. The estimation metrics and the iterative algorithm to optimize those metrics are presented in Section IV. Some numerical results and conclusions are given in Section 5 and 6 respectively.

II. SYSTEM MODEL

Consider a network of n nodes as shown in Figure 1. Each node observes i.i.d measurements X_i based on a parameter θ. We assume as a model, a family of distributions for the random measurement X_i, indexed by a parameter θ taking values in a parameter set Θ; we have the family $\{P_\theta; \theta \in \Theta\}$, where P_θ denotes a probability measure on the observation space $\mathscr{X}$ having σ-field G. Throughout this paper, we assume that the parameter set $\Theta \subseteq \mathbb{R}$.

We define a *deterministic quantizer* as a $G-$measurable mapping $\gamma : \mathscr{X} \mapsto \{1, ..., D\}$, where D is a constant. Let Γ denote the set of all deterministic quantizers. In our setup, the quantized observations $\gamma(X_i), i = 1, ..n$ are used to estimate the parameter θ. Let the probability mass function (p.m.f) of the quantized variable $\gamma(X_i)$ be specified by

$$\mathbf{q}_\theta^i = \{q_\theta^i(1), ..., q_\theta^i(D)\},$$
$$q_\theta^i(k) = P_\theta(\gamma(X_i) = k), \ k = 1, ..., D.$$

We assume that sensors use identical quantizers, hence the family of p.m.fs on the quantized observations is represented by $\{\mathbf{q}_\theta; \theta \in \Theta\}$. The fusion center receives the values of $\gamma(X_i), i = 1, \cdots, n$, and estimates θ. Suppose that $\hat{\theta}$ is an unbiased estimate of the parameter θ, and the family $\mathbf{q}_\theta$ satisfies certain regularity conditions ([8], pp. 169), then

$$\mathbb{E}_\theta[(\hat{\theta} - \theta)^2] \geq \frac{1}{nI_\theta},$$

where

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

is the Fisher Information in the variable $\gamma(X_i)$, and the bound is known as the *Cramer-Rao Bound*. In this work, we assume the fusion center implements a Maximum-Likelihood (M-L) estimator based on the distribution of quantized observations. It is known that the MSE of the M-L estimate asymptotically achieves the Cramer-Rao lower bound. Specifically, if $\hat{\theta}_{ML}$ represents the M-L estimate, then

$$\hat{\theta}_{ML} \sim \mathcal{N}\left(\theta, \frac{1}{nI_\theta}\right), \quad \text{as } n \to \infty.$$

Therefore, Fisher Information is an appropriate metric of performance for the M-L estimator. In the following section, we discuss a class of quantizers called Score-function Quantizers that have interesting properties with respect to Fisher Information.

III. SCORE FUNCTION QUANTIZERS

A. Definition and Optimality

In [4], Tsitsiklis considered quantization for a distributed detection setup and showed that Likelihood-Ratio quantizers (LRQ) have a number of optimal properties. In particular, they can be used to maximize detection performance metrics like Kullback-Leibler distance and Chernoff Information. The quantity corresponding to likelihood-ratio in an estimation setup is the Score-function which is defined as:

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

The definition is subject to regularity conditions mentioned in the previous section. The Score-function, by virtue of its dependence on the parameter θ, cannot be treated as a statistic for distributed estimation. Therefore, the Score-function quantizer (SFQ) is also a function of the parameter value, and does not directly lend itself for quantization in distributed estimation.

We define the *threshold set T* as the set of all vectors $t = (t_1, ..., t_{D-1}) \in \mathbb{R}^{D-1}$, satisfying $-\infty \leq t_1 \leq \cdots \leq t_{D-1} \leq \infty$. For any $t \in T$, the associated intervals $I_1, ..., I_D$ are defined by $I_1 = [-\infty, t_1], I_2 = [t_1, t_2], \cdots, I_D = [t_{D-1}, \infty]$.

Definition 1: A quantizer $\gamma \in \Gamma$ is a *monotone SFQ* with threshold vector $t \in T$, if $\gamma(x) = d \iff S_\theta(x) \in I_d, \forall x$. We say that a quantizer is a SFQ is there exists a permutation mapping $\pi : \{1, \cdots, D\} \mapsto \{1, \cdots, D\}$ such that $\pi \circ \gamma$ is a monotone SFQ ($\circ$ is the composition operator).
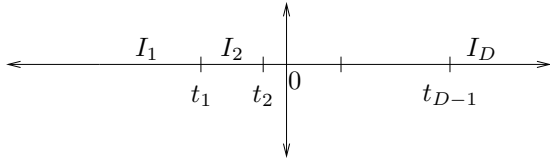
Fig. 2.   Monotone Score-function Quantizer : $I_i = (t_{i-1}, t_i)$ represent the partition i

In an SFQ, indices are assigned to the data by partitioning the score-function (see Figure 2) using $D - 1$ thresholds. This is analogous to applying a conventional, perhaps non-uniformly spaced, Analog-Digital Converter (ADC) on the Score-function. The thresholds on the score-function map onto corresponding thresholds on the data; data values corresponding to a given quantization index are, however, not necessarily contiguous (see Figure 3).

Score-function quantizers, like their counterparts in detection, the LRQs, have some optimal properties, albeit with respect to a fixed parameter value. By adopting the technique used in [4], it can be shown that any extreme point in the set of aposteriori probability mass functions at a particular θ is
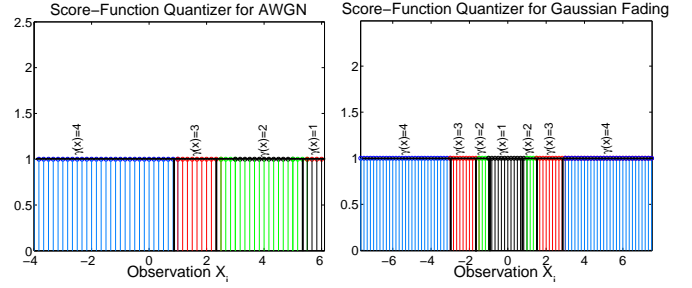
Fig. 3.   Partitions in $\mathscr{X}$ domain for two different distributions

achieved by an SFQ. In particular, the Fisher Information at a given parameter value θ is maximized by an SFQ at that θ [1]. The maximal Fisher Information which is achieved at every θ using a different SFQ, serves as a useful benchmark for quantizer design. We shall use this bound to compare the performances of different quantizers in Section V.

The SFQ, as iterated earlier, is dependent on the parameter value, rendering it not directly applicable to practical quantizer design. It can be shown, however, that for distributions which satisfy a monotonicity property, the set of score-function quantizers at a particular value of parameter is independent of the parameter itself.

B. Class of Score Function Quantizers

The class of Score-Function Quantizers $\mathscr{S}_\theta$ at a particular θ can be defined as

$$\mathscr{S}_\theta = \{\gamma_\theta : \gamma_\theta \text{ is a SFQ at } \theta\}.$$

Let $\{P_\theta; \theta \in \Theta\}$ correspond to the family of densities $\{p_\theta; \theta \in \Theta\}$. We know that a statistic T is sufficient for θ if and only if there are functions $g(\theta)$ and h such that

$$p_\theta(x) = g_\theta[T(x)]h(x), \quad \forall x, \theta$$

Under certain conditions on the prior distribution of observations, $p_\theta(x)$, $\mathscr{S}_\theta$ can be made independent of θ. This is clear from the following lemma.

Lemma 1: Suppose that probability measure for observation X, $\{P_\theta; \theta \in \Theta\}$ has a corresponding family of densities $\{p_\theta; \theta \in \Theta\}$. Let T be a sufficient statistic for p_θ. If the score function can be expressed as $S_\theta(x) = f_\theta[T(x)]$, where f_θ is monotone increasing for all θ, then

i) The *class of SFQs* at parameter θ is identical for all θ. i.e., $\mathscr{S}_\theta = \mathscr{S}_{\theta'}, \forall \theta, \theta' \in \Theta$.

ii) Every SFQ γ_θ can be equivalently expressed as $D - 1$ thresholds on the sufficient statistic $T(x)$. In other words, there exists $-\infty = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_{D-1}) \leq t_D = \infty$ 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. □

In order to design an optimal Score-Function Quantizer for distributions that satisfy the monotonicity property, it is therefore sufficient to obtain the corresponding thresholds on the sufficient statistic. The decoupling of the set of SFQs and the parameter values, together with the known optimal properties of SFQs make it a good candidate for practical quantizer design. In the following section, we discuss some known metrics of performance which are suited to the properties of SFQ, and also present an algorithm to obtain the corresponding SFQs for those metrics.

IV. SCORE-FUNCTION QUANTIZER DESIGN

A. Metrics of Performance

Since the known optimality of SFQs is with respect to Fisher Information, their utility is most beneficial when the fusion center implements an M-L estimator. An ideal quantizer for a M-L estimator would maximize the Fisher Information for every value of the parameter. However, it is impossible to design a single quantizer that is optimal for all θ because the Fisher Information is a function of θ. Hence, we consider the following metrics which are dependent on Fisher Information but serve as measures of the performance over the parameter set.

i) Maximin Fisher Information: We wish to design a quantizer γ^* that achieves:

$$\gamma^* = \max_\gamma \min_\theta I_\theta.$$

For a given asymptotic error variance, we know that Fisher Information is inversely related to sample size. Therefore, a quantizer that satisfies the above metric would also require the least sample size amongst all deterministic quantizers. Mathematically, for given asymptotic error variance ϵ,

$$\gamma^* = \min_\gamma \max_\theta n_q,$$

where n_q is the number of samples required by a quantizer to achieve that asymptotic error variance.

ii) Maximin Asymptotic Relative Efficiency (ARE) : The quantizer that optimizes the maximin ARE is stated as:

$$\gamma^* = \max_\gamma \min_\theta \frac{I_\theta}{J_\theta},$$

where J_θ is the Fisher Information of the unquantized variable. It can be shown that the maximin ARE quantizer is equivalently expressed as:

$$\gamma^* = \min_\gamma \max_\theta \frac{n_q}{n_{uq}}.$$

In other words, the optimal quantizer tries to reduce the sample size as close to that required by the unquantized estimator. For certain distributions, J_θ is independent of θ, in which case, the metric is equivalent to Maximin Fisher Information.

iii) Bayesian Fisher Information : For a random parameter, the Bayesian Fisher Information is a known metric of

performance. If we assume that the parameter is distributed according to some distribution $f(\theta)$, the quantizer that maximizes this metric can be specified as:

$$\gamma^* = \max_\gamma \mathbb{E}\{I_\theta\} = \max_\gamma \int_\Theta f(\theta) I(\theta) d\theta.$$

The Bayesian Fisher Information serves as a good lower bound on the mean-squared error in random parameter estimation.

B. Iterative Algorithm

In this section, we propose an algorithm for iteratively obtaining the optimal SFQs for the metrics discussed previously. The basic idea behind the algorithm is this: We start with an initial guess of thresholds on the sufficient statistic. We then iteratively improve the quantizer by successively optimizing each threshold keeping the other thresholds fixed. The algorithm converges because the performance metric improves with every iteration.

Denote the metric under consideration as $M(\gamma) = f_\theta(I_\theta)$. From the definition of the metrics in the Section IV-A, we know that $M(\gamma)$ is independent of θ. More specifically, $M(\gamma)$ for the three metrics are given by:

i) Maximin Fisher Information : $M(\gamma) = \min_\theta I_\theta$.
ii) Maximin ARE : $M(\gamma) = \min_\theta \frac{I_\theta}{J_\theta}$.
iii) Bayesian Fisher Information : $M(\gamma) = \mathbb{E}(I_\theta)$.

Since we are only interested in distributions that satisfy the monotonicity property, any SFQ corresponds to a set of thresholds on the sufficient statistic. Let $T(x)$ denote the sufficient statistic and $\mathscr{T}$ represent its domain. We assume that $\mathscr{T} \subset \mathbb{R}$. The formal statement of the algorithm is as follows:

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

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

For this q_θ, evaluate the Fisher Information I_θ and the metric $M(\gamma)$.

ii) *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, we would have already computed the new threshold for t_{j-1}. We consider all possible values of threshold t_j that lie between the updated value of t_{j-1} and the old value of t_{j+1} and we pick the value that optimizes the metric. More formally:

For every $t \in [t_{j-1}^n, t_{j+1}^n]$, let γ_j^t represent the quantizer:

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

The optimal threshold is chosen as:

$$t_j^{n+1} = \arg \max_{t \in [t_{j-1}^{n+1}, t_{j+1}^n]} M(\gamma_j^t)$$

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

iii) *Termination:* Choose ϵ close to zero. At each iteration, 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.

It is easily shown that the metric cannot decrease at any iteration. Since these metrics are strictly bounded from above, the algorithm has to terminate. Furthermore, since we restrict ourselves to score-function quantizers, the algorithm obtains the best possible score-function quantizer in the absence of local minima in the path of iteration.

We currently do not have an analytical proof that the globally optimal quantizer for the metrics is an SFQ. However, our numerical results along with the known optimal properties of SFQ seem to suggest that it is indeed so.

V. NUMERICAL EXAMPLES

In this section, we compare the performance of the quantizers for different metrics under two different conditional distributions. These distributions satisfy the monotonicity property stated in Lemma 1 of Section III. The parameter is assumed to be in a bounded set in $\mathbb{R}$.

A. Parameter in AWGN

In the first example, the observation X is Gaussian with mean θ and variance σ^2. Here, the Fisher Information of the unquantized random variable J_θ is a constant, and hence the metrics Maximin Fisher Information and Maximin ARE are equivalent. Figure 4 plots the Fisher Information of the optimal Score-function quantizers for the Bayesian and Maximin metrics. The prior distribution of parameter is assumed uniform for the Bayesian metric.
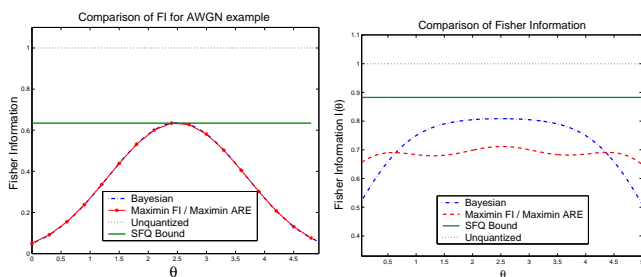
Fig. 4. FI Comparison for Parameter in AWGN. Parameter Set : $\theta \in [0, 5]$. Noise Variance $\sigma^2 = 1$. No. of bin indices : $D = 2, 4$

The maximum possible Fisher Information obtained through evaluating SFQs at each value of parameter is also shown in the figure. As can be seen from the figure, the Maximin quantizer has the highest lower bound on Fisher Information, indicating that asymptotically, it requires fewer samples to achieve the same error variance. Both these quantizers are expressible as $D-1$ thresholds on the sufficient statistic, which in this case, is the observation X. It is interesting to see that

the optimal quantizers for all the metrics are identical when $D = 2$.

B. Gaussian Faded Parameter

The observation is a faded value of the parameter and the fading coefficient is assumed to be Gaussian. In other words,

$$X = H\theta, \quad H \sim \mathcal{N}(0, 1).$$

Since the Fisher Information is no longer independent of θ, the Maximin Fisher Information metric is different from the Maximin ARE. Figures 5 and 6 plot the Fisher Information and the Asymptotic Relative Efficiency versus the parameter respectively for the different quantizers.
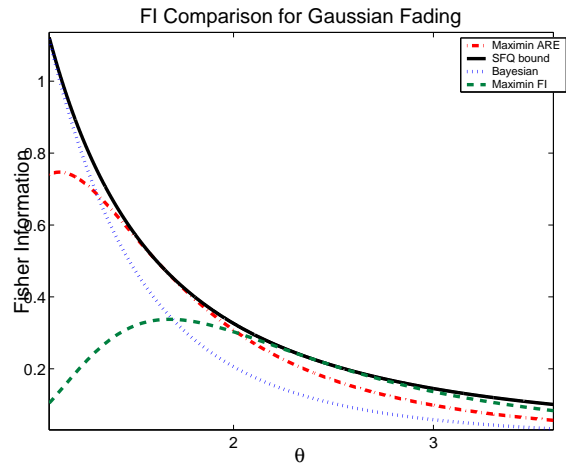
Fig. 5. FI Comparison for Gaussian Fading. Parameter Set : $\theta \in [1, 4]$. Noise Variance : $\sigma^2 = 1$, No. of indices : $D = 2$
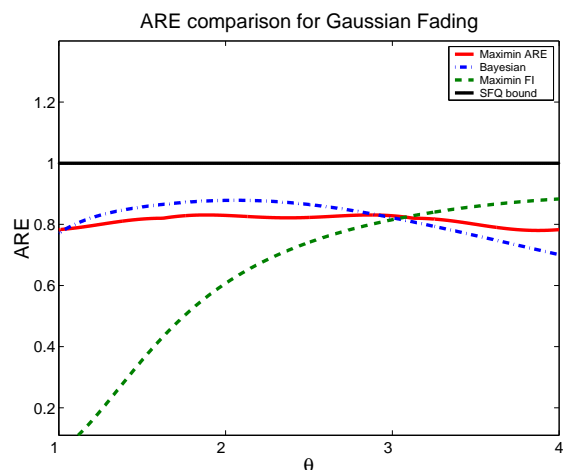
Fig. 6. ARE Comparison for Gaussian Fading. Parameter Set : $\theta \in [1, 4]$. Noise Variance : $\sigma^2 = 1$, No. of indices : $D = 4$

In this example, it can be clearly seen that the quantizer for Maximin Fisher Information has the highest lower bound for Fisher Information amongst other quantizers. Similarly, Figure 6 shows that the Maximin ARE quantizer has the highest lower bound on ARE across the parameter set. When $D = 2$ (Figure

6), each of the quantizers meets the SFQ bound curve at some value of the parameter. Since the performance at any parameter cannot be better than the SFQ bound curve, this strongly suggests that the Score-function quantizer is the optimal type of quantizer for these metrics.

In [3], the authors had developed an algorithm to obtain the quantizer that maximizes the Bayesian Fisher Information of a random parameter iteratively. Their algorithm did not pose any restrictions on quantizer type or underlying distribution. Similarly, in [2], we provided an iterative algorithm to design a quantizer for the maximin ARE metric without making assumptions on the nature of the quantizer. Both these quantizers, although optimal in the absence of local minima, do not possess a regular structure in the observation domain, and hence are hard to implement. Our numerical simulations, however, indicate that for the distributions that satisfy the monotonicity property, the optimal quantizers obtained through these algorithms and the optimal Score-function quantizer are identical. Although, we do not have an analytical proof for this, we believe that with respect to these metrics, Score-Function quantizers are optimal.

VI. Conclusions

In this paper, we considered the class of score-function quantizers and showed that if the distribution of the observations satisfies a monotonicity property, it suffices to obtain optimal thresholds on sufficient statistic to maximize useful estimation performance metrics. One of the key advantages of score-function quantizers over known algorithms for Bayesian Fisher Information and Maximin ARE is the simplicity in structure and representation. Furthermore, some optimality properties of SFQs with regard to Fisher Information make it a likely candidate for quantization when coupled with M-L estimation. The application of this idea is not restricted to the distributed estimation alone. In situations, where data is received in a sequence, it may not be practical for a single processor to hold large number of observations in full resolution. The use of such quantizers is essential to reduce the complexity of processing while ensuring as little loss of information as possible.

References

[1] P. Venkitasubramaniam, G. Mergen, L. Tong, and A. Swami, "Quantization for Distributed Estimation in Large Scale Sensor Networks," in *Proc. 2005 ICISIP*, Bangalore, India, Dec. 2005.

[2] P. Venkitasubramaniam, L. Tong, and A. Swami, "Minimax Quantization for Distributed Maximum Likelihood Estimation," accepted to *IEEE ICASSP 2006*, May 2006.

[3] W. M. Lam and A. R. Reibman, "Design of Quantizers for Decentralized Estimation Systems," *IEEE Trasactions Comm.*, vol. 41, no. 11, pp. 1602–1605, Nov 1993.

[4] John N. Tsitsiklis, "Extremal Properties of Likelihood Ratio Quantizers," *IEEE Trans. Comm.*, vol. 41, no. 4, pp. 550–558, April 1993.

[5] J. A. Gubner, "Distributed Estimation and Quantization," *IEEE Trans. Info. Thy.*, vol. 39, no. 4, pp. 1456–1459, July 1993.

[6] K. Zhang and X. R. Li, "Optimal Sensor R=Data Quantization for Best Linear Unbiased Estimation Fusion," in *Proc. 43rd IEEE Conf. on Decision and Control*, Dec 2004, vol. 3, pp. 2656–2661.

[7] A. Ribiero and G. B. Giannakis, "Non-Parametric Distributed Quantization-Estimation Using Wireless Sensor Networks," in *Proc. 2005 ICASSP Conf.*, Philadelphia, PA, March 2005.

[8] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1994.