

QUANTIZATION FOR DISTRIBUTED ESTIMATION IN LARGE SCALE SENSOR NETWORKS

Parvathinathan Venkitasubramaniam, Gökhan Mergen, Lang Tong and Ananthram Swami

ABSTRACT

We study the problem of quantization for distributed parameter estimation in large scale sensor networks. Assuming a Maximum Likelihood estimator at the fusion center, we show that the Fisher Information is maximized by a score-function quantizer. This provides a tight bound on best possible MSE for any unbiased estimator. Furthermore, we show that for a general convex metric, the optimal quantizer belongs to the class of score function quantizers. We also discuss a few practical applications of our results in optimizing estimation performance in distributed and temporal estimation problems.

1. INTRODUCTION

The two main tasks in a sensor network are sensing of a phenomenon and communicating the sensed information to a fusion center (FC). Sensors have low battery energy and restricted communication capabilities whereas the fusion center is typically a centralized access point or a powerful cluster head capable of performing sophisticated data processing and communication.

A common operation at the fusion center is to use the data obtained from the sensors in estimating a parameter θ that is indicative of the phenomenon sensed. The communication from the sensors to the FC, however, is severely restricted due to the energy constraint of the sensors and due to the common wireless channel that is shared by several sensors. Therefore, it is imperative for the sensors to compress the information before transmission, although the measurement at each sensor may have a high resolution. We are interested in the problem of quantizing the observation at the sensor optimally to maximize the estimation performance.

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

Gökhan Mergen is with Qualcomm Inc., Campbell, CA. Email: gmergen@qualcomm.com

Ananthram Swami is with Army Research Laboratory, Adelphi, MD. Email: a.swami@ieee.org

This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564 and by the Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011.

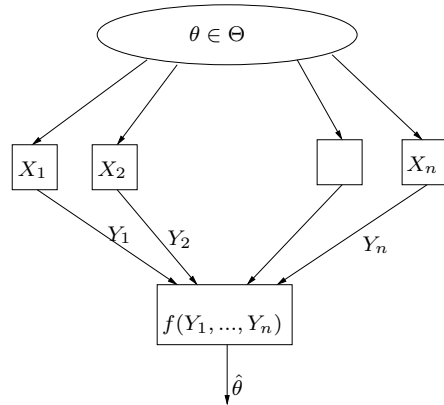


Fig. 1. Distributed Estimation

Consider a system as shown in Fig. 1. We assume that the data at each sensor X_i belongs to a set \mathcal{X} , and is distributed according to a family of distributions $p_\theta(x)$. $\{X_i\}$ are assumed i.i.d conditioned on the parameter θ . Each sensor uses an identical quantizer $\gamma : \gamma(X_i) \in \{1, \dots, D\}$ to compress the observation X_i . The use of identical quantizer is motivated by the identical nature of distributions and the possibility of using Type-Based Multiple Access [1] for efficient communication to the fusion center.

The quantized observations are used by the fusion center to make an estimate $\hat{\theta}$ of the true parameter. The criterion for performance evaluation is the mean squared error $\mathbb{E}_\theta(\theta - \hat{\theta})^2$. For a given distribution of observations, it is well known that the MSE of any unbiased estimate is lower bounded by the inverse of the Fisher Information (FI). Furthermore, when a Maximum Likelihood estimator is used at the fusion center, the MSE of the estimate asymptotically achieves this bound.

An ideal quantizer for an M-L fusion center should therefore maximize FI for every value of the parameter. However, it is not possible to design a single quantizer that delivers the maximum FI for every θ , as the FI is a function of θ . In this work, we assume that the fusion center implements an M-L estimator¹. We investigate the type of quantizer that

¹For a noisy multiple access channel with communication constraints, it was shown in [1] that using Type Based Multiple Access, it is possible to design a Maximum-Likelihood estimator that is asymptotically efficient.

maximizes the Fisher Information for a given θ and thereby obtain the maximum achievable FI for that value of θ . We show that such a quantizer falls into a category called *score function quantizers*. The score function quantizers depend on the value of θ , so, they are not practical. However, they can be used to evaluate the estimation performance of the best possible quantization. We show that the class of score-function quantizers can be used to maximize a general convex performance metric. We also propose some applications of using score-function quantization to optimize other estimation performance criteria.

The paper is organized as follows. Section 2 deals with the system model and general quantizer design. Section 3 defines *score function quantization* and explains the optimality of SFQ. Section 4 discusses some practical applications of score-function quantization in optimizing estimation performance. Section 5 provides some numerical results. Finally, conclusions and future extensions are given in Section 6.

1.1. 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 known performance metrics [2]. The optimality of Likelihood Ratio quantization for a non i.i.d setup was shown by Hoballah and Varshney in [3]. In the distributed estimation scenario, Lam and Reibman in [4] 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 as solutions of multi-dimensional equations. They minimized the mean squared error for a random parameter. More recently, for parameter in additive noise, Ribiero and Giannakis [7] obtained the optimal 1-bit quantizer that maximizes Fisher Information for a given value of the parameter.

The problem has also been posed in an information theoretic framework by Zhang and Berger in [8] where two nodes receive a sequence of observations, each pair jointly distributed under a parameter. They showed the existence of a "good" estimator for the problem. For the same setup, Han and Amari [9] proposed a universal coding scheme for rate constrained transmission. They also developed the maximum likelihood estimator based on their scheme and gave an expression for Fisher Information.

2. SYSTEM MODEL

Consider a network of n sensors in a field. Each sensor observes i.i.d measurements X_i based on a parameter θ . We assume as a model, a family of distributions for the ran-

dom measurement X_i , indexed by a deterministic 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 \mathcal{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 $\gamma: \mathcal{X} \mapsto \{1, \dots, D\}$, where D is a constant. Let Γ denote the set of all deterministic quantizers. In our setup, the quantized observations $\gamma(X_i), i = 1, \dots, n$ is used to estimate the parameter θ . Let the probability mass function (p.m.f) of the quantized variable $\gamma(X_i)$ be specified by

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

We assume that sensors use identical quantizers, hence the new family of densities on the quantized observations is represented by $\{\mathbf{q}_\theta; \theta \in \Theta\}$.

2.1. Cramer-Rao Lower Bound

The fusion center receives the values of $\gamma(X_i), i = 1, \dots, n$ and estimates θ . Suppose that $\hat{\theta}$ is an unbiased estimate of the parameter θ , and the family \mathbf{q}_θ satisfies certain regularity conditions ([10];pp 169), then

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

where

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

is the Fisher Information in the variable $\gamma(X_i)$ and the bound is the *Cramer-Rao* Bound.

Furthermore, if $\hat{\theta}$ is a maximum likelihood estimate based on observations $\gamma(X_i)$, then under the assumption of some more regularity conditions ([10];pp 183), we know that the estimator satisfies

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

In other words, the ML estimator asymptotically achieves the *Cramer-Rao* lower bound.

The Fisher Information of the quantized variable, I_θ is a good performance metric for the quantizer. A quantizer that maximizes I_θ is optimal with a M-L fusion center. Since I_θ is a function of the parameter θ , it is not possible to design a single quantizer that maximizes I_θ for every θ . However, for a given parameter θ , we wish to find out the maximum achievable Fisher Information I_θ^* and the type of quantizer that achieves this maximum value. Firstly, for a given family of distributions \mathbf{q}_θ , I_θ^* would provide a benchmark for

quantizer design. Secondly, this would give us an insight into the class of quantizers that optimize Fisher Information based metrics. Furthermore, in the case when sensors observes a long sequence of measurements, θ can be estimated locally and used to quantize the measurements optimally (this will be discussed in more detail in Section 4.1).

2.2. Class of Quantizers

Since we assume identical quantizers and i.i.d measurements, the performance of a quantizer with respect to any metric is dependent on the set of posterior p.m.fs $\{\mathbf{q}_\theta, \forall \theta\}$. From the definition of I_θ in (2), we know that the Fisher Information for any given θ is a function of \mathbf{q}_θ and $\frac{d}{d\theta}\mathbf{q}_\theta$. Therefore, the problem of maximizing I_θ for a fixed θ can be divided into two steps:

- i) Find the optimal pair of posterior vectors $\mathbf{r}_\theta = (\mathbf{q}_\theta, \frac{d}{d\theta}\mathbf{q}_\theta)$ that maximizes I_θ .
- ii) Obtain the quantizer γ that generates the optimal \mathbf{r}_θ .

For a given θ , we now define the set of posterior vectors:

$$\begin{aligned} Q_\theta &= \{\mathbf{q}_\theta(\gamma) | \gamma \in \Gamma\} \\ R_\theta &= \{(\mathbf{q}_\theta(\gamma), \frac{d}{d\theta}\mathbf{q}_\theta(\gamma)) | \gamma \in \Gamma\}, \end{aligned}$$

where the p.m.fs are generated by deterministic quantizers. It is possible that the set of deterministic quantizers alone is not sufficient to obtain maximum possible Fisher Information. We therefore enlarge the set to include a certain class of randomized quantizers.

The definition of a deterministic quantizer was given in Section 2. We now define a randomized quantizer as follows [2]. Let K be an arbitrary positive integer and $\gamma_1, \dots, \gamma_K$ be some deterministic quantizers. Let $\mathbf{p} = (p_1, \dots, p_K)$ be a point on the standard K simplex. Consider a random variable W whose p.m.f is given by \mathbf{p} . Specifically, $\Pr(W = k) = p_k$. A *randomized quantizer* can be defined as a function $\gamma : \mathcal{X} \times \{1, \dots, K\} \mapsto \{1, \dots, D\}$ such that $\gamma(Y, W) = \gamma_W(Y)$. It is easy to see that

$$P_\theta(\gamma(Y, W) = d) = \sum_{k=1}^K p_k P_\theta(\gamma_k(Y) = d), \forall d \quad (3)$$

Let $\bar{\Gamma}$ represent the set of all randomized quantizers. We now define the new set of posterior vectors as

$$\bar{R}_\theta = \{\mathbf{r}_\theta(\gamma) | \gamma \in \bar{\Gamma}\},$$

Clearly, \bar{R}_θ is the convex hull of R_θ .

3. OPTIMALITY OF SCORE FUNCTION QUANTIZATION

In order to obtain the optimal class of quantizers, we adopt a technique similar to the one used in [2] to optimize per-

formance of distributed detection. In [2], the class of Likelihood ratio quantizers was shown to optimize performance metrics like Chernoff Information (Bayesian detection) and Kullback-Leibler distance (Neyman-Pearson Detection). The equivalent of the likelihood ratio in parameter estimation is the score function, which is defined as follows.

A1: Let P_θ satisfy the following conditions:

- i) The distributions P_θ have common support, so that without loss of generality the set $A = \{x : p_\theta(x) > 0\}$ is independent of θ .
- ii) For any x in A and $\theta \in \Theta$, the derivative $p'_\theta(x) = \frac{\partial p_\theta(x)}{\partial \theta}$ exists and is finite.
- iii) p'_θ is absolutely continuous with respect to P_θ on the set A .

The score function is then defined as:

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

The score function measures the sensitivity of the likelihood function w.r.t θ .

Definition 1 a) We define the *threshold set* T as 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 $t \in T$, the associated intervals I_1, \dots, I_D are defined by $I_1 = [0, t_1]$, $I_2 = [t_1, t_2], \dots, I_D = [t_{D-1}, \infty]$.

b) Let $t \in T$. We say that a quantizer $\gamma \in \Gamma$ is a monotone Score-function quantizer (SFQ) with threshold vector t , if $\gamma(x) = d \iff S_\theta(x) \in I_d \forall d$ (see Fig. 3). We say that a quantizer is an SFQ if there exists a permutation mapping $\pi : \{1, \dots, D\} \mapsto \{1, \dots, D\}$ such that $\pi \circ \gamma$ is a monotone SFQ (\circ is the composition operator).

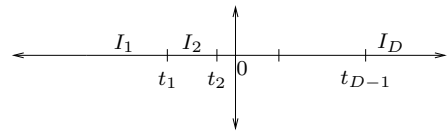


Fig. 2. Monotone Score Function Quantizer

It is to be noted that a score function quantizer is dependent on the value of θ . Since, currently our focus is to maximize I_θ for a given θ , we consider this class of quantizers. Under certain conditions on the underlying family of distributions (mentioned in Section 4), the class of SFQs can be made independent of θ .

3.1. Optimality

Theorem 1 Let P_θ satisfy the conditions given in A1. Furthermore, let $p'_\theta(y)$ be absolutely integrable w.r.t Lebesgue measure. Then, the following statements are true:

- i) Let $f : R_\theta \rightarrow \mathbb{R}$ be a convex function of \mathbf{q}_θ' for a given \mathbf{q}_θ and A be any closed subset of Q_θ . Then, there exists a SFQ that maximizes f over all \bar{R}_θ subject to $\mathbf{q}_\theta \in A$.
- ii) There exists $\mathbf{r}_\theta(\gamma) \in \bar{R}_\theta$ that maximizes I_θ over \bar{R}_θ , such that γ is a deterministic SFQ.

Proof: See [11].

From the above theorem, it can be seen that, in order to maximize I_θ for a given θ , it is sufficient to search within the set of score function quantizers at that θ . Furthermore, part i) of the theorem is a general statement and states that to maximize any metric which is a function of q_θ and $\frac{d}{d\theta}q_\theta$, the search can be restricted to the set of SFQs.

3.2. Finding the Optimal SFQ

The set of SFQs for any given θ is uncountable, and it is not possible to conduct an exhaustive search for the optimal set of thresholds on the score function that maximizes Fisher Information. Therefore, an efficient algorithm needs to be formulated to obtain the SFQ that maximizes I_θ .

The algorithm for finding the optimal SFQ is based on the same principle as in the Lloyd-Max algorithm used in quantization for minimum distortion. Let the score function of the quantized observation be denoted as $S_\theta(i)$, $1 \leq i \leq D$. It is easy to see that for any SFQ, $S_\theta(i)$ is the centroid of the score functions $S_\theta(x)$ in bin i .

$$S_\theta(i) = \frac{d}{d\theta} \log q_\theta(i) = \frac{\int_{t_{i-1}}^{t_i} p_\theta(x) S_\theta(x) dx}{\int_{t_{i-1}}^{t_i} p_\theta(x) dx},$$

where $\{t_i\}$ are the thresholds of the SFQ. The Fisher Information can therefore be expressed as a sum of mean distortions of score functions.

$$I_\theta = J_\theta - \sum_{i=1}^D \int_{t_{i-1}}^{t_i} p_\theta(x) (S_\theta(x) - S_\theta(i))^2 dx$$

where J_θ is the Fisher Information of the unquantized X . Since J_θ is a constant for any given θ , the problem of maximizing I_θ is equivalent to minimizing the distortion in score functions. Therefore, the Lloyd-Max algorithm can be used to successively increase I_θ . Since every iteration yields a score function quantizer, the algorithm converges to the SFQ yielding maximum possible Fisher Information at the given θ . In the following section, we discuss the applications of the previous results.

4. APPLICATION OF SCORE FUNCTION QUANTIZATION

The score function quantizer is a function of θ and hence cannot be used as a practical quantizer for parameter estimation. However, for any given value of θ , it is possible to obtain the maximum achievable Fisher Information at that θ . As mentioned earlier, it is not possible to design a single quantizer that achieves this bound for every value of θ . This bound would therefore serve as a benchmark for practical quantizer design. The performance of a quantizer can be quantified by evaluating the distance of the Fisher Information function from the bound obtained using SFQs.

4.1. Class of Score Function Quantizers

Suppose that $\{P_\theta; \theta \in \Theta\}$ has a corresponding 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

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

Definition 2: If T is a sufficient statistic for p_θ and the score function can be expressed as $S_\theta(x) = f_\theta[T(x)]$, where f_θ is monotone increasing for all θ , then the set of all threshold quantizers on $T(x)$ is defined as the *class of score-function quantizers*.

Therefore, in order to optimize any Fisher Information type metric, it is sufficient to characterize thresholds on the *sufficient statistic* rather than a score function which is dependent on θ . By Fisher Information type metric, we mean metrics which involve \mathbf{r}_θ such that Theorem 1 would apply [12].

4.2. Minimax Quantizer

As mentioned earlier, it is not possible to design a quantizer that maximizes I_θ for every θ . We know that $I_\theta < J_\theta, \forall \theta$, where J_θ is the Fisher information of the unquantized variable X . Therefore, a good metric for practical quantizer design would be to minimize the maximum distance between the two quantities over the parameter. Mathematically, we wish to design γ^* such that

$$\gamma^* = \arg \min_{\gamma} \max_{\theta} J_\theta - I_\theta$$

From theorem 1, we know that for any given θ , a strictly convex function of \mathbf{r}_θ is maximized by a score function quantizer. Although we cannot explicitly show that the minimax criterion corresponds to optimizing a convex metric, Theorem 1 can be used as a heuristic and the search for the optimal minimax quantizer can be restricted to the *class of score function quantizers*. Therefore, when the underlying family of distributions satisfy the conditions in definition 2,

we need to find optimal thresholds on the sufficient statistic T that optimize the criterion.

The minimax quantization problem has been dealt with in detail for a generic framework in [12] and we shall only present a simple example in this document to explain the use of score function quantization. Consider the simple AWGN case when $D = 2$.

$$X = \theta + N, \quad N \sim \mathcal{N}(0, \sigma^2), \theta \in [\theta_{\min}, \theta_{\max}]$$

A deterministic 1-bit quantizer corresponds is represented by the set $\Lambda_1 = \{x : \gamma(x) = 1\}$. We therefore wish to find the optimal Λ_1^* that optimizes the minimax criterion. It can be shown that

$$\begin{aligned} \Lambda_1^* &= \arg \min_{\Lambda_1} \max_{\theta} J_{\theta} - I_{\theta}, \\ &= \arg \min_{\Lambda_1} \max_{\theta} \text{Var}_{\theta}(x|x \in \Lambda_1) + \text{Var}_{\theta}(x|x \in \Lambda_1^c) \end{aligned} \quad (4)$$

The criterion is analytically not tractable beyond this point. However, it is clear that $p_{\theta}(x)$ in this case satisfies the conditions of Definition 2 and the sufficient statistic is x . Therefore the class of 1-bit score-function quantizers ($D = 2$) would be of the form:

$$\gamma(x) = \begin{cases} 1, & x \leq \tau \\ 2, & x > \tau \end{cases} \quad (5)$$

We call the quantizer of the form in (5) a *measurement quantizer*. The expression in (4) can now be expressed as

$$\tau^* = \arg \min_{\tau} \max_{\theta} \text{Var}_{\theta}(x|x \leq \tau) + \text{Var}_{\theta}(x|x > \tau).$$

The optimal value for $\tau^* = \frac{\theta_{\min} + \theta_{\max}}{2}$. The required minimax quantizer is therefore the quantizer of the form (5) with threshold τ^* . The performance of the minimax quantizer for some distributions are also illustrated in Section 5.

4.3. Quantized Local Estimates

The nature of problems discussed so far have been focused on scalar quantization, wherein each node quantizes a single measurement and transmits the quantized value to the fusion center. The idea however can be extended to the case of vector quantization when the nodes receive a sequence of measurements. An information theoretic approach to this problem has been considered in [8], [9].

Consider two nodes, receiving sequences of measurements, $\mathbf{x} = \{x_i, i = 1 \dots n\}$ and $\mathbf{y} = \{y_i, i = 1 \dots n\}$ respectively. Each pair of measurements (x_i, y_i) is *i.i.d* with joint probability distribution $p_{\theta}(x, y)$. The observations are quantized subject to a rate constraint for each node. The quantized vectors \mathbf{u}, \mathbf{v} are transmitted to a fusion center, which estimates the underlying parameter θ based on the

received vectors. In [9], the quantizers at the individual nodes were abstracted by conditional probability distributions $w(u|x; p_{\theta}(x))$ and $\sigma(v|y; p_{\theta}(y))$. Given these distributions, the authors established the existence of a universal coding scheme to compress the data and designed a maximum-likelihood estimator that asymptotically achieves the Cramer-Rao bound given the conditional distributions.

Given a long sequence of observations at each node, it is possible to make an estimate of θ at each node based on the marginals $p_{\theta}(x), p_{\theta}(y)$. Therefore, using the estimated value of θ , the observation can be quantized using the SFQ evaluated at that estimate. The fusion center uses the received quantized information from both the nodes and makes an accurate estimate of the parameter. Since we have already established that the score-function quantizer maximizes Fisher Information at each node based on the marginals, we believe that the theory would extend to the joint estimation as well. This is a direction which we have started exploring recently, so we do not have concrete results yet.

5. NUMERICAL EXAMPLES

In this section we shall discuss a couple of numerical examples to illustrate the usefulness of the score-function quantization as a tight bound on Fisher Information. We shall also plot the performance of the minimax quantizer for these examples and compare it with the SFQ bound.

5.1. Parameter in AWGN

We consider the example as given in Section 4.2. The observation X is Gaussian with mean θ and variance σ^2 . In Fig. 3, we plot the maximum quantized Fisher Information for different values of D and compare it to the unquantized Fisher Information. The figure also shows the performance of two specific score function quantizers.

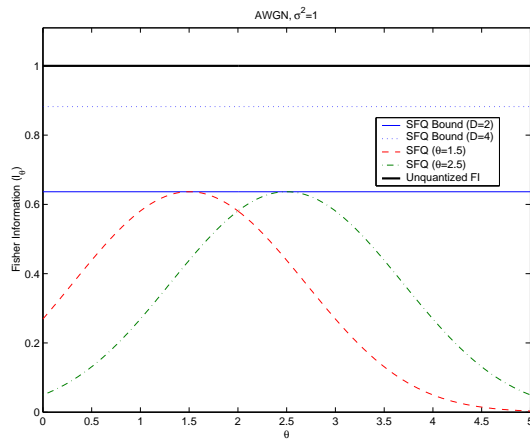


Fig. 3. Parameter in AWGN

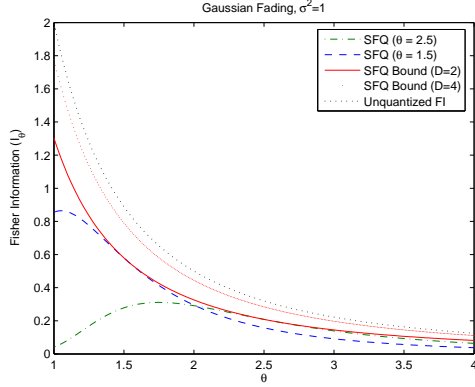


Fig. 4. Gaussian Fading Example

5.2. Gaussian Faded Parameter

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

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

Figure 4 plots the maximum quantized and unquantized Fisher Information for the Gaussian fading model. The plot also shows performance of two specific SFQs.

Although it is not possible to design a quantizer that achieves the Fisher Information bound for every value of θ , from the above plots, it is evident that practical quantizers can achieve the bound for specific values of θ . The bound is therefore tight from a point perspective, and hence can be used as a comparison for the performance of any practical quantizer.

In the AWGN case, as we have established earlier, the minimax quantizer is a score-function quantizer evaluated at $\theta = \frac{\theta_{\min} + \theta_{\max}}{2}$. It is easily verified that the Gaussian fading distribution also satisfies the conditions of Definition 2 and the sufficient statistic is given by x^2 . Hence the 1-bit minimax quantizer in this case would correspond to the optimal threshold on x^2 . The performance of this quantizer is shown in Figure 5 and is compared to the bound for $D = 2$. The figure also plots the performance of a quantizer of the form (5) optimized for every value of θ . As is evident, the score-function based minimax quantizer has a better performance than a measurement quantizer that knows θ .

6. CONCLUSIONS

In this work, we proposed the concept of *score function quantization* and showed the optimality in maximizing Fisher Information for a given parameter. Apart from providing a benchmark for the best possible estimation performance, we

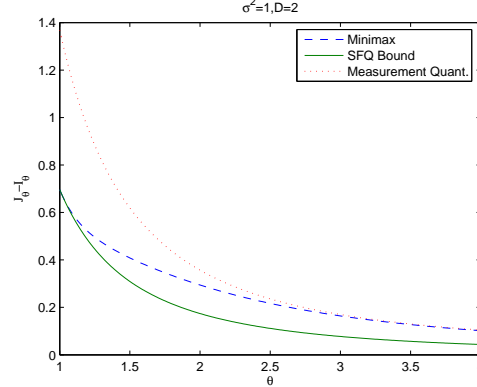


Fig. 5. Minimax Quantizer

also showed that this can be applied for quantization in practical scenarios such as the temporal estimation and minimax quantization. It is clear from our results that the use of score function in quantization is crucial for optimizing estimation performance. Our results are currently focused on a scalar parameter. The extension of our results to the vector case would be non-trivial. It would also be interesting to explore the nature of quantization when observations are not i.i.d.

7. REFERENCES

- [1] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *to appear in IEEE Trans. SP*, 2005.
- [2] John N. Tsitsiklis, "Extremal Properties of Likelihood Ratio Quantizers," *IEEE Trans. Comm.*, vol. 41, no. 4, pp. 550–558, April 1993.
- [3] I. Y. Hoballah and P. K. Varshney, "Distributed Bayesian Signal Detection," *IEEE Trans. Info. Thy.*, vol. 35, no. 5, pp. 995–1000, Sep. 1989.
- [4] W. M. Lam and A. R. Reibman, "Design of Quantizers for Decentralized Estimation Systems," *IEEE Transactions Comm.*, vol. 41, no. 11, pp. 1602–1605, Nov 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] Z. Zhang and T. Berger, "Estimation via Compressed Information," *IEEE Trans. Info. Thy.*, vol. 34, no. 2, pp. 198–211, March 1988.
- [9] Te Sun Han and S. Amari, "Parameter Estimation with Multiterminal Data Compression," *IEEE Trans. Info. Theory*, vol. 41, no. 6, pp. 1802–1833, Nov. 1995.
- [10] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1994.
- [11] P. Venkitasubramaniam, G. Mergen, and L. Tong, "Quantization for distributed estimation," Tech. Rep. ACSP-TR-05-07-01, Cornell University, July 2005.
- [12] P. Venkitasubramaniam, L. Tong, and A. Swami, "Minimax Quantization for Distributed Maximum Likelihood Estimation," submitted to *IEEE ICASSP 2006*, Oct. 2006.