

# OPTIMAL QUANTIZERS FOR DISTRIBUTED BAYESIAN ESTIMATION

Aditya Vempaty      Biao Chen      Pramod K. Varshney

Department of EECS, Syracuse University, NY, 13244 USA  
email: {avempaty, bichen, varshney}@syr.edu

## ABSTRACT

In this paper, we consider the problem of quantizer design for distributed estimation under the Bayesian criterion. We derive general optimality conditions under the assumption of conditionally independent observations at the local sensors and show that for a conditionally unbiased and efficient estimator at the Fusion Center, identical quantizers are optimal when local observations have identical distributions. This results in an  $N$ -fold reduction in complexity where  $N$  is the number of sensors. We illustrate our approach by applying it to the location parameter estimation problem.

**Index Terms**— Distributed Estimation, Quantizer Design, Posterior Cramér Rao Lower Bound (PCRLB)

## 1. INTRODUCTION

Distributed parameter estimation from quantized data has been an active area of research [1–5]. Identical quantizers at the sensors have traditionally been used by researchers as it simplifies the design problem [3] [6]. However, relatively little research has focused on the optimality of these identical quantizers. For decentralized detection, Tsitsiklis [7] showed the asymptotic optimality of identical quantizers with conditionally independent and identically distributed sensor observations. In [1], the authors considered the design of optimal quantizers for distributed estimation under different distortion criteria. Using the minimax criterion, optimal quantizers have been found in [5] and [4]. The maximum likelihood estimator has been used at the Fusion Center (FC) in [8] for which the optimal quantizers have been shown to be the score functions. A discussion on the design of quantizers with design goals of bandwidth efficiency, scalability and robustness to network changes can be found in [9]. In [10], an algorithm was developed for the design of a non-linear multiple-sensor distributed estimation system by partitioning the real line for quantization.

In this paper, we design optimal quantizers for a distributed estimation problem under the Bayesian criterion for an arbitrary cost function. We establish that when the FC uses a conditionally unbiased and efficient estimator, identical quantizers are optimal. We illustrate this result by considering the location parameter estimation problem and obtain conditions under which threshold quantizers are optimal.

## 2. PROBLEM FORMULATION

Consider a distributed estimation problem where the problem is to estimate a random scalar parameter  $\theta$  at the fusion center (FC). The

parameter  $\theta$  has a prior probability density function (pdf)  $p(\theta)$  where  $\theta \in \Theta$ . There are a total of  $N+1$  sensors  $S_0, S_1, \dots, S_N$  in the network and sensor  $S_0$  plays the role of FC whereas the other  $N$  sensors are peripheral sensors. Each sensor  $S_i$ , for  $i = 0, 1, \dots, N$  receives a local observation  $Y_i$  which is a noisy realization of the parameter  $\theta$  and takes values in a set  $\mathcal{Y}_i$ . We assume throughout this paper that  $Y_i$ 's are conditionally independent and identically distributed, hence the overall likelihood function is  $p(\mathbf{y}|\theta) = \prod_{i=0}^N p(y_i|\theta)$ . This likelihood function is known at the FC.

Each sensor  $S_i$ ,  $i \neq 0$ , quantizes its observation  $y_i$ , which is a realization of the random variable  $Y_i$ , using a local quantizer  $\gamma_i(\cdot)$ . The quantizer output  $u_i = \gamma_i(y_i) \in 1, \dots, D$  is transmitted to the FC error free. The FC uses  $u_1, \dots, u_N$  along with its own observation  $y_0$  (realization of  $Y_0$ ) and estimates the random parameter  $\theta$  as  $\hat{\theta} = \gamma_0(y_0, u_1, \dots, u_N) \in \Theta$ . Here  $\gamma_0 : \mathcal{Y}_0 \times \{1, \dots, D\}^N \rightarrow \Theta$  is a function that will be referred to as the estimator. For  $i = 1, 2, \dots, N$ , we use  $\Gamma_i$  to denote the set of all possible quantizers of sensor  $S_i$ . The collection  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N)$  of quantizers will be referred to as a strategy. The estimator is assumed to be given and, therefore, the strategy only involves local quantizers. We let  $\Gamma = \Gamma_1 \times \Gamma_2 \times \dots \times \Gamma_N$ , which is the set of all strategies. For  $i \neq 0$ , once a quantizer  $\gamma_i \in \Gamma_i$  is fixed, the quantizer output  $u_i$  at sensor  $S_i$  can be viewed as a realization of a random variable  $U_i$  defined by  $U_i = \gamma_i(Y_i)$ . Clearly, the probability distribution of  $U_i$  depends on the distribution of  $Y_i$  and on the choice of the quantizer  $\gamma_i$ . Similarly, once the estimator and the strategy are fixed, the global estimate  $\hat{\theta}$  becomes a random variable defined by  $\hat{\theta} = \gamma_0(Y_0, U_1, \dots, U_N)$ .

In the most general Bayesian formulation, we define a cost function  $C : \Theta \times \{1, \dots, D\}^N \times \Theta \rightarrow \mathcal{R}$ , with  $C(\hat{\theta}, u_1, \dots, u_N, \theta)$  representing the cost associated with an FC estimate  $\hat{\theta}$  and quantizer outputs  $u_1, \dots, u_N$ , when the true parameter is  $\theta$ . For any given strategy  $\gamma \in \Gamma$ , its Bayesian cost (or risk)  $J(\gamma)$  is defined as  $J(\gamma) = E[C(\hat{\theta}, U_1, \dots, U_N, \theta)]$ , where the arguments of  $C(\cdot)$  are all random variables. An equivalent expression, in which the dependence on  $\gamma$  is more explicit is

$$J(\gamma) = \int_{\Theta} p(\theta) E[C(\gamma_0(Y_0, \gamma_1(Y_1), \dots, \gamma_N(Y_N)), \gamma_1(Y_1), \dots, \gamma_N(Y_N), \theta)|\theta] d\theta$$

The optimal quantizers are those which minimize this cost function  $J(\gamma)$ . For a given  $\gamma_0(\cdot)$ , the problem can be stated as,

$$\gamma^* = \arg \min_{\gamma \in \Gamma} J(\gamma) \quad (1)$$

## 3. OPTIMALITY CONDITIONS FOR CONDITIONALLY INDEPENDENT OBSERVATIONS

In this section, we provide optimality conditions for quantizers for an arbitrary cost function under the assumption of conditionally independent observations.

This work was supported in part by ARO under Award W911NF-12-1-0383, AFOSR under Award FA9550-10-1-0458 and NSF under Award 1218289

dependent observations. We first provide a proposition in Sec. 3.1 which will be used for deriving the optimality conditions. The results in this section are derived using an approach similar to [11].

### 3.1. Preliminaries

Let  $\theta$  be a random parameter to be estimated with prior pdf  $p(\theta)$  and let  $X$  be a random variable, taking values in a set  $\mathcal{X}$ , with known conditional distribution given  $\theta$ . Let  $D$  be some positive integer, and let  $\Delta$  the set of all functions  $\delta : \mathcal{X} \rightarrow \{1, \dots, D\}$ . Consistent with our earlier terminology, we shall call such functions quantizers.

**Proposition 3.1** *Let  $Z$  be a random variable taking values in a set  $\mathcal{Z}$  and assume that, conditioned on  $\theta$ ,  $Z$  is independent of  $X$ . Let  $F : \{1, \dots, D\} \times \mathcal{Z} \times \Theta \rightarrow \mathcal{R}$  be a given cost function. Let  $\delta^*$  be an element of  $\Delta$ . Then  $\delta^*$  minimizes  $E[F(\delta(X), Z, \theta)]$  over all  $\delta \in \Delta$  if and only if*

$$\delta^*(X) = \arg \min_{d=1, \dots, D} \int_{\Theta} a(\theta, d) p(\theta|X) d\theta \quad \text{with probability 1} \quad (2)$$

where

$$a(\theta, d) = E[F(d, Z, \theta)|\theta] \quad \forall \theta, d \quad (3)$$

**Proof** The minimization of  $E[F(\delta(X), Z, \theta)]$  over all  $\delta \in \Delta$  is equivalent to requiring that  $\delta(X)$  minimize  $E[F(d, Z, \theta)|X]$ , over all  $d \in \{1, \dots, D\}$ , with probability 1. The expression being minimized can be re-written as  $E[E[F(d, Z, \theta)|\theta, X]|X]$  which by conditional independence of  $X$  and  $Z$ , is equal to

$$E[E[F(d, Z, \theta)|\theta, X]|X] = \int_{\Theta} E[F(d, Z, \theta)|\theta] p(\theta|X) d\theta \quad (4)$$

Therefore, conditional independence decouples the design of  $\delta^*(X)$  from  $Z$ , i.e.,  $\delta^*(X)$  depends on  $Z$  only through  $a(\theta, d)$ . ■

### 3.2. Optimality Conditions

The optimal strategy which minimizes  $J(\gamma)$  satisfies the following necessary conditions presented in Proposition 3.2:

**Proposition 3.2** *For  $i \neq 0$  and suppose that  $\gamma_j \in \Gamma_j$  has been fixed for all  $j \neq i$ . Then  $\gamma_i$  minimizes  $J(\gamma)$  over the set  $\Gamma_i$  only if*

$$\gamma_i(Y_i) = \arg \min_{d=1, \dots, D} \int_{\Theta} a(\theta, d) p(\theta|Y_i) d\theta \quad \text{with probability 1} \quad (5)$$

where for any  $\theta$  and  $d$ ,

$$a(\theta, d) = E[C(U_0, U_1, \dots, U_{i-1}, d, U_{i+1}, \dots, U_N, \theta)|\theta] \quad (6)$$

and where each  $U_i$ ,  $i \neq 0$  is a random variable defined by  $U_i = \gamma_i(Y_i)$  and  $U_0 = \gamma_0(Y_0, U_1, \dots, U_{i-1}, d, U_{i+1}, \dots, U_N)$ .

**Proof** Observe that the minimization is of

$E[C(U_0, U_1, \dots, U_{i-1}, \gamma_i(Y_i), U_{i+1}, \dots, U_N, \theta)|\theta]$  over  $\gamma_i \in \Gamma_i$  where  $U_0 = \gamma_0(Y_0, U_1, \dots, U_{i-1}, \gamma_i(Y_i), U_{i+1}, \dots, U_N)$ . This is of the form considered in Proposition 3.1 where  $X = Y_i$ ,  $d = \gamma_i(X) = \gamma_i(Y_i)$ ,  $Z$  is the random vector given by  $Z = (Y_0, U_1, \dots, U_{i-1}, U_{i+1}, \dots, U_N)$  and  $F(d, Z, \theta) = C(U_0, U_1, \dots, U_{i-1}, \gamma_i(Y_i), U_{i+1}, \dots, U_N, \theta)$ . The result follows from Proposition 3.1 and yields a person-by-person optimal solution. ■

Proposition 3.2 provides the necessary conditions for optimal quantizers for an arbitrary cost function  $C(\cdot)$ . However, the optimization problem is difficult to solve in general due to the complexity of solving  $N$  simultaneous optimization problems. For the remainder of the paper, we consider the design of optimal quantizers for a specific cost function namely the Mean-Square Error (MSE), i.e.,  $C(\hat{\theta}, U_1, \dots, U_N, \theta) = E[(\hat{\theta} - \theta)^2]$  where  $\hat{\theta} = \gamma_0(Y_0, \gamma_1(Y_1), \dots, \gamma_N(Y_N))$  and  $\theta$  is the true parameter.

## 4. QUANTIZERS FOR CONDITIONALLY UNBIASED AND EFFICIENT ESTIMATORS

In this section, we find the optimal quantizers in distributed estimation for estimators which are efficient and conditionally unbiased. By conditionally unbiased, we mean  $E_{x|\theta}[\hat{\theta}] = \theta$  for all  $\theta$ . The motivation behind such an analysis is that most of the widely used estimators, among them maximum likelihood estimator and maximum a posteriori estimator, are asymptotically unbiased and efficient. In such a scenario, the cost function (MSE) becomes the variance of the estimator which attains the Posterior Cramér-Rao Lower Bound (PCRLB). Therefore, the optimization problem can now be formulated as the minimization of PCRLB, or equivalently, the maximization of posterior Fisher Information. Since  $\gamma_0(\cdot)$  is assumed to be a fixed efficient, conditionally unbiased estimator, the optimization is now performed over  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N)$ . While our results hold for any estimator that achieves the PCRLB, the design methodology also applies to cases where no efficient estimator exist; the optimization is therein on performance bounds that are not necessarily attainable.

**Proposition 4.1** *Let  $\Gamma$  denote the set of all possible strategies for the distributed estimation problem with identical and conditionally independently distributed sensor observations and  $\Gamma^I$  denote the set of all strategies in which all peripheral sensors use identical quantizers. If an efficient and unbiased estimator exists at the Fusion Center, there is no loss in estimation performance (characterized by the Mean Square Error of the estimator) by restricting the search space of optimal strategy to  $\Gamma^I$ . In other words, if an efficient and conditionally unbiased estimator exists at the FC, there exists an optimal strategy for which all the peripheral sensors use identical quantization rules.*

**Proof** The posterior Fisher Information under the conditional independence assumption is given by

$$F(\gamma) = -E_{\theta, \mathbf{U}, Y_0} [\nabla_{\theta} \nabla_{\theta}^T \ln p(\mathbf{U}, Y_0, \theta)] \quad (7)$$

$$= -E_{\theta, \mathbf{U}} [\nabla_{\theta} \nabla_{\theta}^T \ln p(\mathbf{U}|\theta)] - E_{\theta, Y_0} [\nabla_{\theta} \nabla_{\theta}^T \ln p(Y_0|\theta)] - E_{\theta} [\nabla_{\theta} \nabla_{\theta}^T \ln p(\theta)] \quad (8)$$

$$= F_D + F_0 + F_P \quad (9)$$

where  $F_D$ ,  $F_0$  and  $F_P$  represent the local sensor data's contribution, FC data's contribution and prior's contribution to  $F$  respectively.

Since the prior's contribution to  $F$  given by  $F_P$  and FC's contribution given by  $F_0$  are independent of  $\gamma$ , the optimization problem can be re-stated as

$$\gamma^{opt} = \arg \max_{\gamma \in \Gamma} F_D = \arg \min_{\gamma \in \Gamma} E_{\theta, \mathbf{U}} \left[ \frac{\partial^2 \ln p(\mathbf{U}|\theta)}{\partial \theta^2} \right] \quad (10)$$

As the sensor observations  $(Y_1, \dots, Y_N)$  are conditionally independent and the quantizers  $\gamma_i$  are independent of each other, the

quantizer outputs are also conditionally independent, i.e.  $\ln p(\mathbf{U}|\theta) = \sum_{i=1}^N \ln p(U_i|\theta)$ . The objective function now becomes

$$E_{\theta, \mathbf{U}} \left[ \frac{\partial^2 \ln p(\mathbf{U}|\theta)}{\partial \theta^2} \right] = \sum_{i=1}^N E_{\theta, U_i} \left[ \frac{\partial^2 \ln p(U_i|\theta)}{\partial \theta^2} \right] \quad (11)$$

The solution to this problem is

$$\gamma^{opt} = \arg \min_{\gamma \in \Gamma} \sum_{i=1}^N E_{\theta, U_i} \left[ \frac{\partial^2 \ln p(U_i|\theta)}{\partial \theta^2} \right] \quad (12)$$

which can be decoupled into  $N$  optimization problems given by

$$\gamma_i^{opt} = \arg \min_{\gamma_i \in \Gamma_i} E_{\theta, U_i} \left[ \frac{\partial^2 \ln p(U_i|\theta)}{\partial \theta^2} \right] \quad \text{for } i = 1, \dots, N \quad (13)$$

Since the above  $N$  optimization problems are identical when all the peripheral sensors have identical statistics, they give identical solutions. Therefore, for an efficient and conditionally unbiased estimator, no loss is incurred when identical quantizers are used at the local sensors. ■

Proposition 4.1 states that we can constrain all peripheral sensors to use the same quantization rule, without increasing the MSE of the efficient, conditionally unbiased estimator. Furthermore, this optimal quantizer can be found by solving the optimization problem in (13). In the next section, we find this optimal quantizer for location parameter estimation under an additive noise model.

## 5. THE LOCATION PARAMETER ESTIMATION PROBLEM

In this section, we consider the location parameter estimation problem where the observations are corrupted by independent and identically distributed (i.i.d) additive noise with pdf  $p_W(w)$ .

$$Y_i = \theta + W_i \quad \text{for } i = 1, \dots, N \quad (14)$$

where  $\theta \sim p_\Theta(\theta)$  and  $W_i$  is the i.i.d noise. The local sensors process their own observations locally before sending the processed one-bit data ( $U_i$  for  $i = 1, \dots, N$ ) to the FC. The FC then estimates  $\theta$  from  $\mathbf{U} = [U_1 \dots U_N]$  and  $Y_0$ . As shown in the previous sections, for an efficient and unbiased estimator at the FC, the optimal quantizers are identical. Let the quantizer be represented by  $\gamma(Y)$  which maps the data  $Y_i$  to one of the two bit values  $\{0, 1\}$ . We represent the quantizers probabilistically as

$$\gamma(Y_i) = P(U_i = 1|Y_i) \quad (15)$$

Thus,  $\gamma(Y_i)$  denotes the probability with which the  $i^{th}$  local sensor sends a '1' to the FC given its observation,  $Y_i$ . Stochastic quantizers are employed here as they cover a wide range of possible quantizers including both the threshold quantizers and the dithering quantizers.

### 5.1. Posterior Cramér Rao Lower Bound

For the location parameter estimation problem,  $F$  from (9) is the posterior Fisher Information [12] which is a function of the prior distribution  $p_\Theta(\theta)$ , the quantizer  $\gamma(Y)$  and the noise pdf  $p_W(w)$ . It is given as

$$F(p_\Theta, \gamma, p_W) = F_D + F_0 + F_P \quad (16)$$

where  $F_D$ ,  $F_0$  and  $F_P$  are as defined before.

As the  $N$  observations are conditionally independent, we have

$$F_D = N E_\theta [I(\theta)] \quad (17)$$

where  $I(\theta)$  is the Fisher Information (FI) for a single observation.

Let  $g(\theta)$  denote the probability that the quantizer output is '1' given the true value of  $\theta$

$$g(\theta) = P(U_i = 1|\theta) = E_{W_i}[\gamma(\theta + W_i)] \quad (18)$$

$$= \int_Y \gamma(y) p_W(y - \theta) dy \quad (19)$$

For a binary quantizer, the FI is given by [3]

$$I(\theta) = \frac{(g'(\theta))^2}{g(\theta)(1 - g(\theta))} \quad (20)$$

where  $g'(\theta)$  represents the first derivative of  $g(\theta)$  with respect to  $\theta$ . From (16), (17) and (20), the posterior FI is given by

$$F(p_\Theta, \gamma, p_W) = N \int_\theta \frac{(g'(\theta))^2}{g(\theta)(1 - g(\theta))} p_\Theta(\theta) d\theta + F_0 + F_P \quad (21)$$

### 5.2. Optimal Quantizer Design

The optimal quantizer  $\gamma^*(y)$  minimizes the PCRLB or, equivalently, maximizes  $F(p_\Theta, \gamma, p_W)$ . Since  $F_0$  and  $F_P$  are independent of the quantizer, the optimization problem can be stated as

$$\gamma^*(\cdot) = \arg \max_{\gamma(\cdot)} F_D = \arg \max_{\gamma(\cdot)} \int_\theta \frac{(g'(\theta))^2}{g(\theta)(1 - g(\theta))} p_\Theta(\theta) d\theta \quad (22)$$

This problem can be solved by observing that the objective function  $F_D$  depends on  $\gamma(\cdot)$  only through  $g(\theta)$  given in (19) which can be re-written as  $g(\theta) = (\gamma(y) * p_W(-y))(\theta)$  where '\*' represents the convolution operation. Transforming this into frequency domain using the Fourier Transform, we get  $G(f) = H(f)P_W(-f)$  where  $G(f)$ ,  $H(f)$  and  $P_W(f)$  are the Fourier transforms of  $g(\cdot)$ ,  $\gamma(\cdot)$  and  $p_W(\cdot)$  respectively. Therefore, given the noise pdf  $p_W(\cdot)$ , the quantizer  $\gamma(\cdot)$  can be found (if it exists) as

$$\gamma(y) = F^{-1} \left[ \frac{G(f)}{P_W(-f)} \right] \quad (23)$$

where  $F^{-1}$  is the Inverse Fourier transform.

The problem now reduces to that of finding the optimal  $g(\theta)$  to maximize the integrand in (22). Note that this optimal  $g^*(\theta)$  is independent of the noise pdf  $p_W(w)$ . Upon obtaining  $g^*(\theta)$ , the optimal quantizer for a given noise pdf can then be designed using (23). Therefore, the optimization in (22) can be re-stated as

$$g^*(\cdot) = \arg \max_{g(\cdot)} \int_\theta I(\theta) p_\Theta(\theta) d\theta \quad (24)$$

where  $I(\theta)$  is given in (20).

**Proposition 5.1** *Given the prior distribution  $p_\Theta(\theta)$ , the optimal  $g^*(\theta)$  can be found by solving the following differential equation*

$$p_\Theta(\theta)(g'(\theta))^2(1 - 2g(\theta)) = 2g(\theta)(1 - g(\theta))(g''(\theta)p_\Theta(\theta) + g'(\theta)p'_\Theta(\theta)) \quad (25)$$

where ' and '' denote respectively the first and the second derivatives with respect to  $\theta$ .

**Proof** Define  $K(\theta) = I(\theta)p_{\Theta}(\theta)$  as the function of  $\theta$  which is the integrand in (24). The optimization problem presented in (24) is a typical variational calculus problem and it can be solved using the Euler-Lagrange equation [13] stated below

$$\frac{\partial K}{\partial g} = \frac{d}{d\theta} \frac{\partial K}{\partial g'} \quad (26)$$

From the expression of  $I(\theta)$  given in (20), we have

$$\frac{\partial K}{\partial g} = \frac{(g')^2 p_{\Theta}(1-2g)}{(g-g^2)^2} \quad (27)$$

and

$$\frac{\partial K}{\partial g'} = \frac{2g'p_{\Theta}}{(g-g^2)} \quad (28)$$

Differentiating (28) with respect to  $\theta$  and using (26), we get the desired result. ■

As can be seen from (25), the differential equation can be solved for a given prior  $p_{\Theta}(\theta)$ . After finding this optimal  $g^*(\theta)$ , the optimal quantizer  $\gamma^*(x)$  can be found for a given noise pdf  $p_W(w)$  using (23).

### 5.3. Example: Uniform prior

In this section, we consider a special case of  $\theta$  being uniformly distributed and find the optimal  $g^*(\theta)$ .

**Proposition 5.2** Given that  $\theta$  is uniformly distributed in  $[\theta_{min}, \theta_{max}]$ , the solution to the optimization problem in (24),  $g^*(\theta)$  is given by

$$g^*(\theta) = \frac{1}{2} \left[ 1 + \sin \pi \left( \frac{\theta - \theta_{min}}{\theta_{max} - \theta_{min}} - \frac{1}{2} \right) \right], \quad \theta \in [\theta_{min}, \theta_{max}] \quad (29)$$

**Proof** The proposition can be proved either by using (25) or using the fact that the prior is uniformly distributed and therefore  $p_{\Theta}(\theta)$  is independent of  $\theta$ . This simplifies the Euler-Lagrange equation of (26) to  $I - g' \frac{\partial I}{\partial g'} = k$ , where  $k$  is a constant. Using (20) and without loss of generality, assuming the boundary conditions as  $g(\theta_{min}) = 0$  and  $g(\theta_{max}) = 1$ , we obtain  $g_s(\theta)$  given by (29) as a stationary point. It can be verified that this stationary point is a maximum as it satisfies the second order necessary conditions for a maximum given by Osgood [14] and also observe that  $g^*(\theta)$  cannot be a minima as  $L(g^*) = \frac{\pi^2}{4} > L(g_1) = 0$  where  $L(g(\cdot)) = \int_{\theta} I(\theta)p_{\Theta}(\theta)$  and  $g_1(\theta) = \text{const}$ . Therefore, the optimal  $g^*(\theta)$  is given by (29). ■

It is interesting to observe that the same result was obtained by Chen and Varshney [3] using the minimax CRLB as the performance metric for a distributed estimation problem with deterministic unknown parameter  $\theta$ . This result is expected as when the FC has a non-informative prior (uniform prior), there is no Fisher information provided by the prior ( $F_P = 0$ ). This implies that using the minimax CRLB or PCRLB (which is the average CRLB in the case of uniform prior) would give the same optimal result.

Without loss of generality, let  $\theta_{min} = -1$  and  $\theta_{max} = 1$ . The optimal  $g^*(\theta)$  given in (29) becomes

$$g^*(\theta) = \frac{1}{2} \left[ 1 + \sin \frac{\pi\theta}{2} \right], \quad \text{for } \theta \in [-1, 1] \quad (30)$$

#### 5.3.1. Noiseless observations

The performance limit of this distributed estimation problem under the Bayesian criterion can be characterized by observing the performance when the observations are noiseless. When these observations at the local sensors prior to quantization are noiseless, i.e., the observation model is perfect,  $p_W(w) = \delta(w)$ . The optimal quantizer, for this case, is given by the sine quantizer

$$\gamma^*(y) = \frac{1}{2} \left[ 1 + \sin \frac{\pi y}{2} \right], \quad \text{for } y \in [-1, 1] \quad (31)$$

In this case, the posterior Fisher information is  $F = \frac{\pi^2}{4}$  and the PCRLB is  $\frac{4}{\pi^2}$ . This represents the performance limit under the Bayesian criteria for the distributed location parameter estimation problem with uniform prior.

#### 5.3.2. Optimality of Threshold quantizers

Threshold quantizers are the most widely used quantizers due to their simplicity [15]. A threshold quantizer is given by

$$\gamma_T(y) = \begin{cases} 1, & \text{if } y \geq T \\ 0, & \text{otherwise} \end{cases} \quad (32)$$

An interesting question is to find the conditions on the noise pdf  $p_W(w)$  for which the threshold quantizers attain the performance limit as described in Sec. 5.3.1 which is the performance when the observations are noiseless (refer to the discussion after (31)). For the optimality condition to be satisfied, the threshold quantizer and the noise pdf should satisfy the following constraint

$$g^*(\theta) = \int_y \gamma_T(y) p_W(y - \theta) dy \quad (33)$$

$$= \int_{y=T}^{\infty} p_W(y - \theta) dy = 1 - P_W(T - \theta) \quad (34)$$

where  $P_W(w)$  is the cumulative distribution function of noise and  $g^*(\theta)$  is given by (31). Differentiating both sides and using the fact  $\frac{dP_W(w)}{dw} = p_W(w)$ , we get the sufficient condition for the threshold quantizer  $\gamma_T(y)$  to achieve performance limit when the noise pdf is

$$p_W(w) = \begin{cases} \frac{\pi}{4} \cos \frac{\pi}{2}(w - T), & \text{for } w \in [T - 1, T + 1] \\ 0, & \text{otherwise} \end{cases}$$

Threshold quantizers can still be optimal for a wide range of noise distributions but the performance limit can be reached only for the above noise pdf.

## 6. CONCLUSION

In this work, we have considered the problem of quantizer design for distributed estimation under the Bayesian criterion. We have found the optimality conditions for one-bit quantizers and showed that for conditionally unbiased efficient estimators, identical quantizers are optimal. For the location parameter estimation problem with a given prior distribution, we have found the optimal  $g^*(\cdot)$  as a solution of a differential equation. We have also characterized the performance limit when the prior distribution is uniform and found the sufficient condition on the noise distribution for which the threshold quantizers are optimal.

## 7. REFERENCES

- [1] W. Lam and A. R. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. Comm.*, vol. 41, no. 11, pp. 1602–1605, Nov. 1993.
- [2] S. Marano, V. Matta, and P. Willett, "Asymptotic design of quantizers for decentralized MMSE estimation," *IEEE Trans. Signal Process.*, vol. 55, no. 55, pp. 5485–5496, Nov. 2007.
- [3] H. Chen and P. K. Varshney, "Performance limit for distributed estimation systems with identical one-bit quantizers," *IEEE Trans. Signal Process.*, vol. 58, pp. 466–471, Jan. 2010.
- [4] —, "Nonparametric one-bit quantizers for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3777–3787, July 2010.
- [5] T. Wu and Q. Cheng, "One-bit quantizer design for distributed estimation under the minimax criterion," in *VTC Spring'10*, 2010, pp. 1–5.
- [6] S. Kar, H. Chen, and P. K. Varshney, "Optimal identical binary quantizer design for distributed estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3896–3901, July 2012.
- [7] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control Signals Systems*, vol. 1, no. 2, pp. 167–182, 1988.
- [8] P. Venkitasubramaniam, G. Mergen, L. Tong, and A. Swami, "Quantization for distributed estimation in large scale sensor networks," in *Int. Conf. on Intelligent Sensing and Information Process. (ICISIP 2005)*, Dec. 2005, pp. 121–127.
- [9] J. Xiao, A. Ribeiro, Z. Luo, and G. Giannakis, "Distributed Compression-Estimation Using Wireless Sensor Networks," *IEEE Signal Processing Magazine, Special issue on Distributed Signal Processing for Sensor Networks*, vol. 23, no. 4, pp. 27–41, July 2006.
- [10] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.
- [11] J. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. Poor and J. Thomas, Eds. Greenwich, CT: JAI Press, 1993.
- [12] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Piscataway, New Jersey: Wiley-IEEE Press, 2007.
- [13] B. Van Brunt, *The Calculus of Variations*. New York: Springer, 2004.
- [14] W. F. Osgood, "Sufficient conditions in the calculus of variations," *The Annals of Mathematics*, vol. 2, no. 1/4, 1900–1901.
- [15] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, Mar. 2006.