HAL open science

Adjustable Quantizers for Joint Estimation of Location and Scale Parameters

Rodrigo Cabral Farias, Jean-Marc Brossier

To cite this version:

Rodrigo Cabral Farias, Jean-Marc Brossier. Adjustable Quantizers for Joint Estimation of Location and Scale Parameters. ICASSP 2013 - 38th IEEE International Conference on Acoustics, Speech and Signal Processing, May 2013, Vancouver, Canada. pp.SPTM-P13.7. hal-00875935

HAL Id: hal-00875935

https://hal.science/hal-00875935v1

Submitted on 23 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ADJUSTABLE QUANTIZERS FOR JOINT ESTIMATION OF LOCATION AND SCALE PARAMETERS

Rodrigo Cabral Farias, Jean-Marc Brossier

GIPSA-lab, Images and Signal Department
University of Grenoble and Grenoble Institute of Technology
11, rue des Mathématiques, BP 46, 38402 Saint Martin d'Hères, France

ABSTRACT

An adaptive algorithm to estimate jointly unknown location and scale parameters of a sequence of symmetrically distributed independent and identically distributed random variables using quantized measurements from a quantizer with adjustable input gain and input offset is presented. The asymptotic variance of estimation is obtained, simulations under Cauchy and Gaussian distributions are presented to validate the asymptotic results and they are compared to the continuous optimal estimator performance.

***Index Terms*—** Parameter estimation, quantization, adaptive algorithm.

1. INTRODUCTION

Sensor networks became an important domain of research recently due to advances in sensor and communication technology. Its potential applications include military sensing, traffic surveillance, industrial automation, environmental monitoring and many others [1].

With the increasing number of sensors in sensing systems, constraints as bandwidth and complexity that were commonly neglected might now be taken into account in the design of the system. One simple way to treat these constraints is to consider quantization of the sensor output measurements.

When considering the problem of optimal quantization and reconstruction of the measurements, well known results can be found in the literature [2]. However, in sensing systems the problem to be solved is normally the optimization of the performance of estimation of a parameter from noisy quantized measurements.

Main results on estimation from quantized measurements can be found in [3], where the behavior of the Cramér–Rao lower bound (CRB) on the variance of estimation of a constant parameter based on uniformly quantized noisy measurements was studied for different types of quantizer input offset, it was shown that a good type of offset should be based on feedback from the quantizer output. An interesting result was that in the binary measurement case with Gaussian distributed noise, the optimal quantizer threshold should be exactly placed at the parameter. This result motivated the development of adaptive schemes in [4] and [5] for placing the thresholds of a sensor network with binary quantizers around the parameter to be estimated, thus enhancing the performance of estimation based on the measurements from the network. This was done by recursively placing the thresholds of the sensor network at the last estimate of the parameter.

In this paper, an adaptive algorithm for estimating a constant location parameter from multiple bit quantized noisy measurements is proposed. As in practice the scale factor of the noise is unknown, the proposed algorithm will jointly estimate the constant location parameter and the noise scale factor by using measurements from a quantizer with adjustable input gain and input offset. For simplification purposes, the noise will be considered to be symmetrically distributed. The problem treated here can also be viewed as the estimation of the parameters of a symmetrically distributed location-scale model based on quantized measurements from an adjustable quantizer.

Differently from [4] and [5], where complex maximum likelihood estimators are used and only binary quantization is considered, the algorithm proposed here has low complexity and is designed for multibit quantization. Also, as a main difference from [3], where location parameter estimation is studied, the problem treated here includes the joint estimation of the scale parameter.

After stating the problem and the adaptive algorithm form, the estimation performance in terms of asymptotic estimation variance will be analyzed using adaptive algorithms theory. The optimal quantizer and algorithm parameters will be obtained and simulation results will be presented under Gaussian and Cauchy noise with a comparison with the best continuous measurement estimation performance.

The authors would like to thank the EBW II program for funding this work.

2. PROBLEM STATEMENT AND ADAPTIVE ALGORITHM

A sequence of independent and identically distributed (i.i.d.) random variables Y_k with marginal cumulative distribution function (CDF) $F\left(\frac{y-x}{\delta}\right)$ are quantized with an adjustable quantizer resulting in a sequence of discrete observations i_k, where k is the sample index. The pair of parameters (x, δ) is unknown and the objective is to estimate it based on the quantized observations.

The adjustable quantizer Q is characterized by an adjustable sequence of input offsets b_k, input gains $\frac{1}{\Delta_k}$ and a static vector of thresholds $\boldsymbol{\tau}$ defining its N_I quantization intervals

$$\boldsymbol{\tau} = \left[\tau_{-\frac{N_I}{2}} \ \cdots \ \tau_{-1} \ \tau_0 \ \tau_1 \ \cdots \ \tau_{\frac{N_I}{2}}\right],$$

its input-output relation is given by

$$\begin{aligned} i_k &= Q\left(\frac{Y_k - b_k}{\Delta_k}\right) = Q\left(Z_k\right) = \\ &= \begin{cases} i, \text{ for } Z_k \geq \tau_0 \text{ and } Z_k \in \left[\tau_{i-1}, \tau_i\right), \\ i, \text{ for } Z_k < \tau_0 \text{ and } Z_k \in \left[\tau_i, \tau_{i+1}\right). \end{cases} \end{aligned} \tag{1}$$

The amplitude resolution of the quantizer can be enhanced adaptively at each sample by using the last parameter estimates

$$b_k = \hat{X}_{k-1}, \quad \Delta_k = c_\delta \hat{\delta}_{k-1}. \tag{2}$$

The constant c_δ is a free parameter that can be used, *e.g.* in optimal uniform quantization, when the dynamical input range of the quantizer is fixed. A choice for the estimation algorithm that allows for the online update of the estimates is

$$\begin{bmatrix} \hat{X}_k \\ \hat{\delta}_k \end{bmatrix} = \begin{bmatrix} \hat{X}_{k-1} \\ \hat{\delta}_{k-1} \end{bmatrix} + \frac{\boldsymbol{\Gamma}}{k} \hat{\delta}_{k-1} \begin{bmatrix} \eta_x\left(i_k\right) \\ \eta_\delta\left(i_k\right) \end{bmatrix} \tag{3}$$

where $\boldsymbol{\Gamma}$ is a 2×2 matrix of gains, $\eta_x[i]$ and $\eta_\delta[i]$ are sequences of N_I update coefficients $\left\{\eta_x\left[-\frac{N_I}{2}\right] \ldots \eta_x\left[\frac{N_I}{2}\right]\right\}$ and $\left\{\eta_\delta\left[-\frac{N_I}{2}\right] \ldots \eta_\delta\left[\frac{N_I}{2}\right]\right\}$. The advantages of this algorithm are its low complexity and the fact that it belongs to a general class of adaptive algorithms for which the performance was studied in [6].

Some assumptions on the noise distribution, quantizer thresholds and update coefficients will now be stated for simplification purposes:

- Y_k has a locally Lipschitz continuous CDF and its probability density function (PDF) $\frac{1}{\delta} f\left(\frac{y-x}{\delta}\right)$ is an even function strictly decreasing w.r.t. $|y - x|$.
- The quantizer has symmetric thresholds $\tau_i = -\tau_{-i}$ with $\tau_0 = 0$ and $\tau_{\frac{N_I}{2}} = +\infty$.
- The update coefficients have odd symmetry w.r.t. i in the case of η_x, $\eta_x[i] = -\eta_x[-i]$, and they have even symmetry in the case of η_δ, $\eta_\delta[i] = \eta_\delta[-i]$. The coefficients for positive i will be denoted in vector form by $\boldsymbol{\eta}_x = \left[\eta_x[1] \ldots \eta_x\left[\frac{N_I}{2}\right]\right]^T$ and $\boldsymbol{\eta}_\delta = \left[\eta_\delta[1] \ldots \eta_\delta\left[\frac{N_I}{2}\right]\right]^T$.

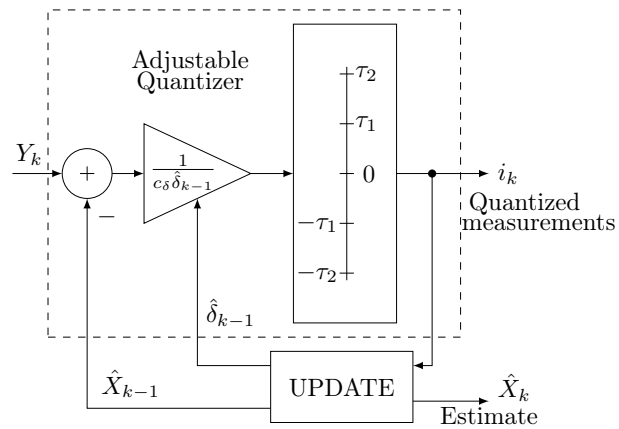

Fig. 1. Scheme representing the adjustable quantizer. The offset and gain are adjusted dynamically using the estimates while the quantizer thresholds are fixed.

The estimation scheme is depicted in Fig. 1, where the UPDATE block is the estimation algorithm.

3. ESTIMATION PERFORMANCE

The analysis of the algorithm will be done using the results of general adaptive algorithms theory presented in [6, Chap. 3]. The estimation performance will be analyzed in terms of the mean error and the asymptotic covariance matrix of estimation error.

The estimator mean can be approximated using the following ordinary differential equation (ODE)

$$\frac{d}{dt}\begin{bmatrix} \hat{x} \\ \hat{\delta} \end{bmatrix} = \boldsymbol{\Gamma}\mathbf{h}\left(\hat{x}, \hat{\delta}\right), \tag{4}$$

the relation between continuous and discrete time is $t_k = \sum_{j=1}^{k} \frac{1}{j}$ and $\mathbf{h}$ is the following mean vector field:

$$\mathbf{h}\left(\hat{x}, \hat{\delta}\right) = \mathbb{E}\begin{bmatrix} \hat{\delta}\eta_x\left(Q\left(\frac{y-\hat{x}}{c_\delta \hat{\delta}}\right)\right) \\ \hat{\delta}\eta_\delta\left(Q\left(\frac{y-\hat{x}}{c_\delta \hat{\delta}}\right)\right) \end{bmatrix} = \tag{5}$$

$$= \hat{\delta}\begin{bmatrix} \sum_{i=1}^{\frac{N_I}{2}} \eta_x[i]\left\{F_d\left(i, \hat{x}, x, \hat{\delta}, \delta\right) - F_d\left(-i, \hat{x}, x, \hat{\delta}, \delta\right)\right\} \\ \sum_{i=1}^{\frac{N_I}{2}} \eta_\delta[i]\left\{F_d\left(i, \hat{x}, x, \hat{\delta}, \delta\right) + F_d\left(-i, \hat{x}, x, \hat{\delta}, \delta\right)\right\} \end{bmatrix}$$

where the expectation is w.r.t. F, the second equality comes from the symmetry assumptions and F_d is

$$F_d = \begin{cases} F\left(\frac{\tau_i c_\delta \hat{\delta}}{\delta} + \frac{\hat{x}-x}{\delta}\right) - F\left(\frac{\tau_{i-1} c_\delta \hat{\delta}}{\delta} + \frac{\hat{x}-x}{\delta}\right) \\ \qquad\qquad \text{if } i \in \left\{1, \cdots, \frac{N_I}{2}\right\}, \\ F\left(\frac{\tau_{i+1} c_\delta \hat{\delta}}{\delta} + \frac{\hat{x}-x}{\delta}\right) - F\left(\frac{\tau_i c_\delta \hat{\delta}}{\delta} + \frac{\hat{x}-x}{\delta}\right) \\ \qquad\qquad \text{if } i \in \left\{-1, \cdots, -\frac{N_I}{2}\right\}, \end{cases} \tag{6}$$

This result is valid under the condition that $\mathbf{h}$ is locally Lipschitz continuous, this condition is satisfied by the continuity assumption on F.

The conditions on the mean convergence of the algorithm are then conditions on the global asymptotic stability of the point $\left(\hat{x} = x, \hat{\delta} = \delta\right)$. One necessary condition for asymptotic stability is that the true parameters must be an equilibrium point of the ODE, which means that $\mathbf{h}\left(\hat{x} = x, \hat{\delta} = \delta\right)$ must be the zero vector. From the symmetry assumptions, the following is obtained

$$\mathbf{h}\left(\hat{x} = x, \hat{\delta} = \delta\right) = \begin{bmatrix} 0 \\ 2\boldsymbol{\eta}_\delta^T \mathbf{F}_d^{vec} \end{bmatrix}, \tag{7}$$

where $\mathbf{F}_d^{vec}$ is the vector $\left[F_d[1], \cdots, F_d\left[\frac{N_I}{2}\right]\right]^T$ whose elements $F_d[i] = F_d(i, x, x, \delta, \delta)$ do not depend on the parameters. Then, the condition for the parameters to be the equilibrium point is

$$\boldsymbol{\eta}_\delta^T \mathbf{F}_d^{vec} = 0. \tag{8}$$

Other conditions are necessary for the mean convergence of the algorithm, these conditions can be found by the analysis of the ODE using Lyapunov theory. The analysis of these other conditions will not be detailed here and under the assumptions already stated and the constraint on $\boldsymbol{\eta}_\delta$ given in (8), it will be assumed that the algorithm converges in the mean to the true parameters.

Asymptotic results for adaptive algorithms with decreasing gains presented in [6, pp. 110–113] can be applied to (3) to get the asymptotic performance of the estimator. It can be shown under all the assumptions above that the estimation error $\boldsymbol{\epsilon}_k$ tends in distribution to a zero mean Gaussian random variable as follows

$$k^{\frac{1}{2}} \boldsymbol{\epsilon}_k \underset{k\to\infty}{\rightsquigarrow} \mathcal{N}(0, \mathbf{P}), \tag{9}$$

where $\mathbf{P}$ is the covariance matrix given by the optimal gain $\boldsymbol{\Gamma}^\star$. The matrices $\mathbf{P}$ and $\boldsymbol{\Gamma}^\star$ are the following:

$$\mathbf{P} = \frac{\delta^2}{2} \begin{bmatrix} \frac{\boldsymbol{\eta}_x^T \mathbf{F}_d \boldsymbol{\eta}_x}{\left(\boldsymbol{\eta}_x^T \mathbf{f}_d^{(x)}\right)^2} & 0 \\ 0 & \frac{\boldsymbol{\eta}_\delta^T \mathbf{F}_d \boldsymbol{\eta}_\delta}{\left(\boldsymbol{\eta}_\delta^T \mathbf{f}_d^{(\delta)}\right)^2} \end{bmatrix}, \tag{10}$$

$$\boldsymbol{\Gamma}^\star = \frac{1}{2} \begin{bmatrix} -\frac{1}{\boldsymbol{\eta}_x^T \mathbf{f}_d^{(x)}} & 0 \\ 0 & -\frac{1}{\boldsymbol{\eta}_\delta^T \mathbf{f}_d^{(\delta)}} \end{bmatrix}, \tag{11}$$

where $\mathbf{F}_d$ is a diagonal matrix $\mathbf{F}_d = \operatorname{diag}\left[\mathbf{F}_d^{vec}\right]$ and $\mathbf{f}_d^{(x)} = \left[f_d^{(x)}[1], \cdots, f_d^{(x)}\left[\frac{N_I}{2}\right]\right]^T$, $\mathbf{f}_d^{(\delta)} = \left[f_d^{(\delta)}[1], \cdots, f_d^{(\delta)}\left[\frac{N_I}{2}\right]\right]^T$ are the derivatives in vector form of the quantizer output probabilities $F_d\left(i, \hat{x}, x, \hat{\delta}, \delta\right)$ when $\hat{x} = x$ and $\hat{\delta} = \delta$:

$$f_d^{(x)} = f(\tau_i) - f(\tau_{i-1}), \tag{12}$$

$$f_d^{(\delta)} = c_\delta\left[\tau_i f(\tau_i) - \tau_{i-1} f(\tau_{i-1})\right]. \tag{13}$$

Minimization of the estimation variance can be done through the minimization of the terms of the diagonal of $\mathbf{P}$ w.r.t. $\boldsymbol{\eta}_x$ and $\boldsymbol{\eta}_\delta$. The minimization problems can be solved separately. In the case of the optimization w.r.t. $\boldsymbol{\eta}_\delta$, the equilibrium constraint (8) has to be taken into account. The optimal $\boldsymbol{\eta}_x$ can be found by using the Cauchy-Schwarz inequality, while the optimal $\boldsymbol{\eta}_\delta$ are obtained by casting the constrained minimization problem as a modified eigenvalue problem solved in [7].

The optimal coefficients are

$$\boldsymbol{\eta}_x = -\mathbf{F}_d^{-1} \mathbf{f}_d^{(x)} \tag{14}$$

$$\boldsymbol{\eta}_\delta = -\mathbf{F}_d^{-1} \mathbf{f}_d^{(\delta)} + \mathbb{1} \mathbf{f}_d^{(\delta)} = -\mathbf{F}_d^{-1} \mathbf{f}_d^{(\delta)}, \tag{15}$$

where $\mathbb{1}$ is a squared matrix with ones. The second equality comes from the fact that the sum of $\mathbf{f}_d^{(\delta)}$ is zero.

Thus the optimal $\mathbf{P}$ and $\boldsymbol{\Gamma}^\star$ are

$$\mathbf{P} = \delta^2 \boldsymbol{\Gamma}^\star = \frac{\delta^2}{2} \begin{bmatrix} \frac{1}{\mathbf{f}_d^{(x)^T} \mathbf{F}_d^{-1} \mathbf{f}_d^{(x)}} & 0 \\ 0 & \frac{1}{\mathbf{f}_d^{(\delta)^T} \mathbf{F}_d^{-1} \mathbf{f}_d^{(\delta)}} \end{bmatrix}. \tag{16}$$

Note that the asymptotic variances are equal to the CRB for estimating the parameters based on quantized measurements, when the quantizer offset and gains are placed exactly at x and $c_\delta \delta$. It is interesting that the optimal coefficients do not depend on x or δ, they depend only on the normalized CDF and PDF and on the quantizer parameters $\boldsymbol{\tau}$ and c_δ, thus they can be implemented easily with a lookup table. Also, notice that $\boldsymbol{\Gamma}^\star$ and $\mathbf{P}$ are diagonal matrices, indicating that both algorithm and performance for estimating the location and the scale parameters are decoupled. This is a behavior that is also present in the standard continuous version of this problem, where the CRB also decouples.

The asymptotic performance can still be optimized w.r.t. $\boldsymbol{\tau}$ and c_δ. The optimization w.r.t. $\boldsymbol{\tau}$ is a difficult multidimensional minimization problem and it will not be treated here. In what follows, it will be considered that the quantizer is uniform with positive thresholds given by

$$\boldsymbol{\tau}^+ = \begin{bmatrix} \tau_1 = 1 & \cdots & \tau_{\frac{N_I}{2}-1} = \frac{N_I}{2} - 1 & \tau_{\frac{N_I}{2}} = \infty \end{bmatrix}^T. \tag{17}$$

Thus the only free parameter for optimization will be c_δ.

4. SIMULATION

The algorithm will now be simulated to validate the theoretical results, the simulation will be focused on the performance

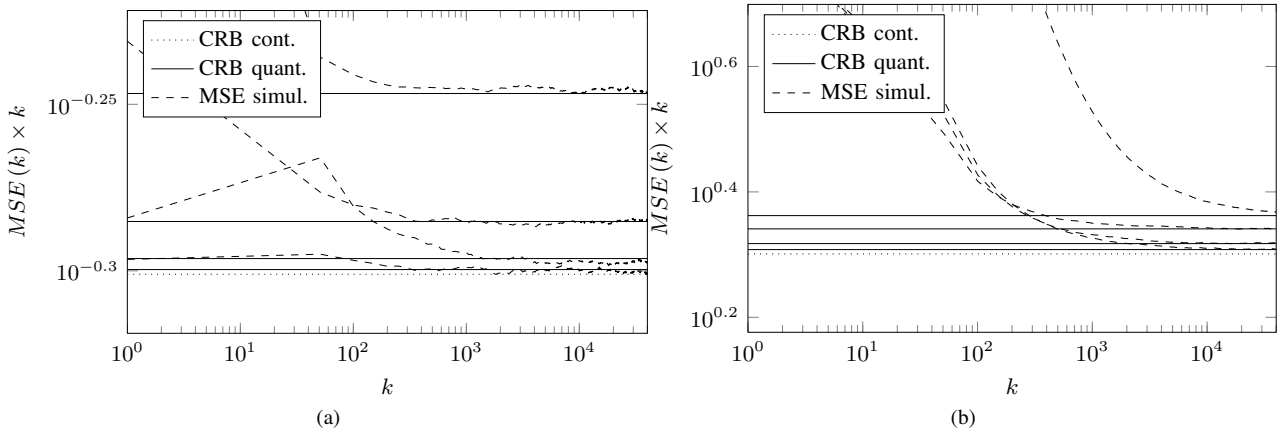

Fig. 2. Normalized (multiplied by k) Cramér–Rao bound for continuous and quantized measurements and normalized mean squared error ($MSE(k) \times k$) for the adaptive algorithm, the numbers of quantization intervals are $N_I = \{4, 8, 16, 32\}$ and the noise is Gaussian distributed in (a) and Cauchy distributed in (b). In both cases $\delta = 1$, $\hat{\delta}_0 = 2$, $x = 0$ and $\hat{X}_0 = 1$. The curves with lower asymptotic values correspond to larger numbers of quantization intervals.

for the estimation of x. As it was mentioned, the quantizer is uniform and c_δ will be chosen so as to minimize the variance of estimation of x. As this is a scalar problem, it can be solved by an exhaustive search using a fine grid. After finding the optimal c_δ, the other parameters of the algorithm $\mathbf{\Gamma}$, $\boldsymbol{\eta}_x$ and $\boldsymbol{\eta}_\delta$ can be evaluated using the information from the measurement distribution.

The Gaussian and Cauchy distributions will be used for modeling the measurements. The Gaussian distribution can be used to model a sequence of measurements of an unknown constant (x), where the measurements are corrupted by thermal noise with unknown scale factor, while the Cauchy distribution can be used to model measurements corrupted by noise with outliers. Their PDFs are given by

$$\frac{1}{\delta} f_G\left(\frac{y-x}{\delta}\right) = \frac{1}{\delta\sqrt{\pi}} e^{-\left(\frac{y-x}{\delta}\right)^2}, \tag{18}$$

$$\frac{1}{\delta} f_C\left(\frac{y-x}{\delta}\right) = \frac{1}{\delta\pi\left(1+\left(\frac{y-x}{\delta}\right)^2\right)}. \tag{19}$$

The algorithm was simulated for 5×10^5 blocks with 4×10^4 samples each. The simulated mean squared error (MSE) for the estimation of the location parameter was evaluated by calculating the mean of the squared error for each sample. Other simulation parameters are $\delta = 1$, $\hat{\delta}_0 = 2$, $x = 0$, $\hat{X}_0 = 1$ and $N_I = \{4, 8, 16, 32\}$. For comparison purposes, the CRB for the estimation of x based on continuous measurements

$$CRB = \frac{\delta^2}{k}\left[\int_{\mathbb{R}} \frac{\left(\frac{df(y)}{dy}\right)^2}{f(y)} \mathrm{d}y\right]^{-1} \tag{20}$$

was also evaluated for Gaussian and Cauchy distributions and they are respectively given by $CRB_G = \frac{1}{2}\frac{\delta^2}{k}$ and $CRB_C = 2\frac{\delta^2}{k}$. The results of the simulation are shown in Fig. 2, where the MSE was normalized by k and the logarithm scale is used in both axis for better visualization. It can be observed that after a time transient, the simulated performance becomes very close to the asymptotic theoretical results, also it can be seen that the gain in performance when increasing N_I is very small even for a small number of quantization intervals ($N_I = 8$ or 16), and that the gap between the performance given by $N_I = 32$ and the continuous measurement bound is negligible.

5. CONCLUSIONS

In this article, an algorithm for estimating the location (x) and scale (δ) parameters of a symmetrically distributed sequence of i.i.d. variables based on quantized measurements from a quantizer with adjustable input offset and gains was proposed. The algorithm was chosen to be a low complexity adaptive algorithm for which theoretical results of its performance could be obtained in terms of its mean error and asymptotic error covariance. It was shown that the asymptotic variance of estimation was equal to the CRB for measurements from a static quantizer with input offset and gains given by x and $c_\delta\delta$ and this was verified through simulation using Gaussian and Cauchy distributions for the measurements. It was observed that with only a few quantization intervals ($N_I = 16$ and 32) the asymptotic estimation MSE is very close to the continuous measurement CRB. This indicates that it is not necessary to use high resolution quantizers when a large block of samples is used.

When the distribution of the variables is characterized by location, scale and shape parameters (*e.g.* the generalized Gaussian distribution), a possible extension of this work would be to estimate jointly the additional shape parameter.

6. REFERENCES

[1] C.Y. Chong and S.P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[2] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Springer, 1992.

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

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

[5] J. Fang and H. Li, "Distributed adaptive quantization for wireless sensor networks: From delta modulation to maximum likelihood," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5246–5257, 2008.

[6] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive algorithms and stochastic approximations*, Springer-Verlag New York, Inc., 1990.

[7] G.H. Golub, "Some modified matrix eigenvalue problems," *Siam Review*, pp. 318–334, 1973.