3rd International Symposium on Physics in Signal & Image Processing, pp. 89–92, Grenoble, France, 29–31 Jan. 2003.

NONLINEAR ESTIMATION FROM QUANTIZED SIGNALS: QUANTIZER OPTIMIZATION AND STOCHASTIC RESONANCE.

David ROUSSEAU [1], *G. V. ANAND* [2], *François CHAPEAU-BLONDEAU* [1]

[1] Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), Université d'Angers,
62 avenue Notre Dame du Lac, 49000 ANGERS, France.

[2] Department of Electrical Communication Engineering, Indian Institute of Science,
Bangalore 560012, India.

ABSTRACT

We consider a parameter estimation task performed on a signal buried in noise by means of a quantized representation by a two-level quantizer of the signal-plus-noise mixture. By considering the Fisher information, we show that the performance for estimation can be maximized by an optimal choice of the quantization threshold. At the optimal threshold, we quantify the loss in performance when the analog input is replaced by its one-bit representation for estimation, and demonstrate the existence of conditions where this loss is very small. If constraints require us to work with a fixed threshold, we establish that noise addition at the input prior to quantization (a form of stochastic resonance) offers another means for maximizing the performance. For illustration, we derive the maximum likelihood estimator for estimation of a constant signal from the quantized data. We involve this estimator in adaptive schemes able to bring the quantizer to operate in optimal conditions, either by varying the quantization threshold at fixed noise level, or by increasing the noise level at fixed threshold. Examples are given with noises belonging to the family of generalized Gaussians, which occur in particular in ocean acoustics and sonar applications.

1. INTRODUCTION

For digital processing, analog signals have to be transformed into quantized data. Quantizers are intrinsically nonlinear devices with thresholds in their response. For (multi)sensor systems having to cope with limited time and resources for data processing, storage and communication, quantization may be desirable over a small number of levels, even down to two levels [1]. In such conditions, it is very important to optimize the quantization process in such a way that the quantized data are maximally efficient for the subsequent processing that is contemplated. In this paper, we address the optimization of two-level quantizers for purposes of parameter estimation based on the quantized data. We examine both threshold adjustment and stochastic resonance as two independent means for optimizing the performance of the quantizers.

Stochastic resonance has recently emerged as a possibility of exploiting the noise to control, and improve, the response of nonlinear systems in definite conditions [2, 3, 4, 5]. In a typical instance of stochastic resonance, a signal–noise mixture interacts with a nonlinear system to elicit a response whose quality is assessed by a specific measure of performance. Stochastic resonance takes place when the value of this measure of performance can be improved by means of an increase in the level of the noise. Various forms of stochastic resonance have already been observed, depending on the type of the signal, of the noise, of the nonlinear system and the measure of performance involved. Here, we demonstrate the possibility of stochastic resonance with threshold quantizers, to optimize the performance for parameter estimation.

2. THE ESTIMATION PROBLEM

An unknown parameter a is attached to a signal $s(t)$ corrupted by a noise $\eta(t)$. The signal-plus-noise mixture $s(t) + \eta(t) = x(t)$ is observed by means of a two-level quantizer with threshold θ, realizing

$$y(t) = \mathrm{sign}[s(t) + \eta(t) - \theta] = \pm 1 \; . \qquad (1)$$

Sampling of $y(t)$ at N distinct times t_j provides N data points $y_j = y(t_j)$. From the data set $\boldsymbol{y} = (y_1, \dots y_N)$ the parameter a is to be estimated. The noise samples $\eta(t_j)$ are assumed independent and identically distributed with cumulative distribution $F_\eta(u)$ and probability density $f_\eta(u) = dF_\eta/du$.

3. FISHER INFORMATION

To assess the performance in the estimation task, we turn to the Fisher information, which sets a bound to the efficacy of any conceivable unbiased estimator, this applying especially to the asymptotic behavior of the maximum likelihood estimator [6]. For the two-level quantizer of Eq. (1), the Fisher information J_{out} contained in one sample $y(t_j)$ about the parameter a is defined as

$$J_{\mathrm{out}} = \sum_{z=-1,1} \frac{1}{\Pr\{y=z\}} \left(\frac{\partial}{\partial a} \Pr\{y=z\} \right)^2 \; . \qquad (2)$$

At a given time t_j, this quantity can be explicitly evaluated under the form

$$J_{\rm out} = \left[\frac{1}{q_1(t_j)} + \frac{1}{1-q_1(t_j)} \right] \left[\frac{\partial q_1(t_j)}{\partial a} \right]^2 \quad (3)$$

where $q_1(t_j) = \Pr\{y(t_j) = 1\}$. We also have $q_1(t_j) = \Pr\{s(t_j) + \eta(t_j) > \theta\}$, which here with $s(t_j)$ deterministic gives $q_1(t_j) = \Pr\{\eta(t_j) > \theta - s(t_j)\}$, amounting to $q_1(t_j) = 1 - F_\eta[\theta - s(t_j)]$. Then, for the derivative we have

$$\frac{\partial q_1(t_j)}{\partial a} = f_\eta[\theta - s(t_j)] \, \frac{\partial s(t_j)}{\partial a} \ , \quad (4)$$

and Fisher information $J_{\rm out}$ of Eq. (3) results as

$$J_{\rm out} = \left[\frac{\partial s(t_j)}{\partial a} \right]^2 \frac{f_\eta^2[\theta - s(t_j)]}{F_\eta[\theta - s(t_j)]\{1 - F_\eta[\theta - s(t_j)]\}} \ . \quad (5)$$

A natural approach suggested by Eq. (5) is then to seek to adjust the threshold θ of the quantizer so as to maximize $J_{\rm out}$ in definite conditions for the observed signal $s(t)$ and for the noise $\eta(t)$. For example, we consider estimation of an unknow constant $s(t) \equiv a$ corrupted by a generalized Gaussian noise of exponent $\alpha > 0$, with the probability density $f_\eta(u) = f_0(u/\sigma_\eta)/\sigma_\eta$, where the standardized density [7]

$$f_0(u) = A \exp(-|bu|^\alpha) \quad (6)$$

with $A = (\alpha/2)[\Gamma(3/\alpha)]^{1/2}/[\Gamma(1/\alpha)]^{3/2}$ and $b = [\Gamma(3/\alpha)/\Gamma(1/\alpha)]^{1/2}$. Such conditions with generalized Gaussian noise are specially relevant for ocean acoustics and sonar applications for instance [8]. Figure 1 represents the evolutions of $J_{\rm out}$ as a function $v = \theta - a$ for different values of the exponent α of the generalized Gaussian noise.

Figure 1 reveals that there is an optimal value $v_{\rm opt}$ of the difference $v = \theta - a$ that maximizes the Fisher information $J_{\rm out}$, and $v_{\rm opt}$ ceases to be zero as soon as $\alpha > 2$. For the familly of generalized Gaussian noises, Fig. 2 shows $v_{\rm opt}$ as a function of exponent α.

A further characterization of the quantizer performance, is provided by the ratio $J_{\rm out,max}/J_{\rm in}$, where $J_{\rm out,max}$ is the maximum value of $J_{\rm out}$ obtained at the optimum $v_{\rm opt}$, while $J_{\rm in}$ is the Fisher information contained in one input sample $a + \eta(t_j) = x(t_j)$ about parameter a. With $\eta(t)$ the generalized Gaussian noise, one obtains

$$J_{\rm in} = \frac{2A\alpha b}{\sigma_\eta^2} \ \Gamma\left(\frac{2\alpha - 1}{\alpha} \right) \ . \quad (7)$$

Figure 3 represents an example of the evolution of $J_{\rm out,max}/J_{\rm in}$ as a function of the exponent α of the generalized Gaussian noise $\eta(t)$, for a given value of its rms amplitude σ_η. The evolution in Fig. 3 shows that the ratio $J_{\rm out,max}/J_{\rm in}$ can come close to unity, especially in the region of the Gaussian case $\alpha = 2$. For the Laplacian case $\alpha = 1$, this ratio is even found equal to one, for any value of σ_η. This behavior reveals possibilities of very limited reduction of performance when the analog input is replaced by a one-bit representation for estimation.

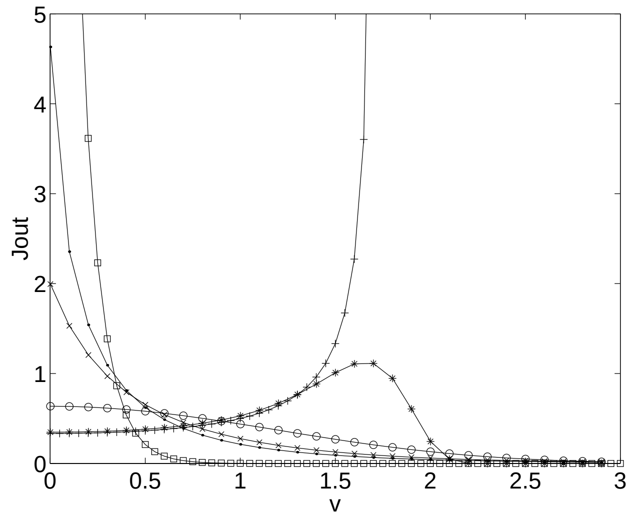

Fig. 1. Output Fisher information $J_{\rm out}$ from Eq. (5) as a function of $v = \theta - a$, when $\eta(t)$ is a zero-mean generalized Gaussian noise with rms amplitude $\sigma_\eta = 1$ and exponent $\alpha = 1/2$ (□), $\alpha = 3/4$ (·), $\alpha = 1$ (×), $\alpha = 2$ (∘), $\alpha = 10$ (∗) and $\alpha = \infty$ (+).

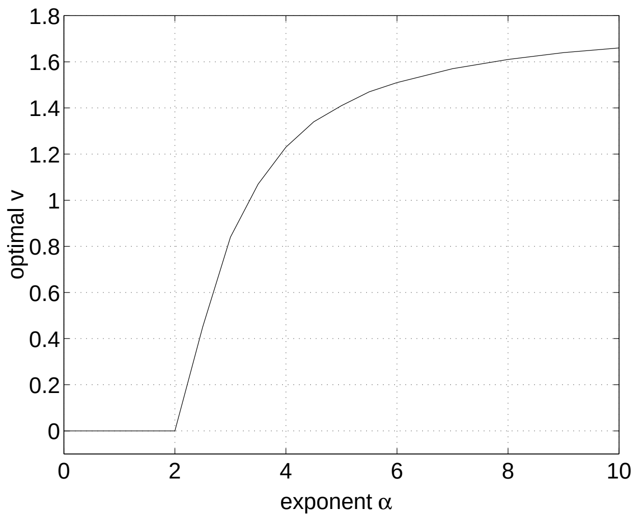

Fig. 2. Optimal value $v_{\rm opt}$ maximizing $J_{\rm out}$ of Eq. (5) as a function of the exponent α of the zero-mean generalized Gaussian noise $\eta(t)$ with rms amplitude $\sigma_\eta = 1$.

Adjustment of the threshold θ of the quantizer, according to the behavior of the Fisher information $J_{\rm out}$ of Eq. (5), thus offers a possibility for optimizing the performance for estimation. Now if constraints exist on the quantizer that require operating with a fixed nonadjustable threshold θ, we shall demonstrate that another possibility is available for optimizing the perfomance measured by the Fisher information. This possibility is to increase the noise rms amplitude σ_η through noise injection at the input prior to quantization. Figure 4

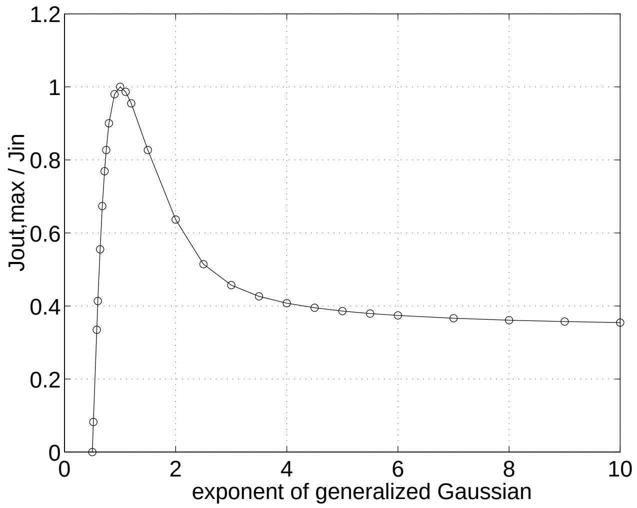

Fig. 3. Ratio of Fisher information $J_{\mathrm{out,max}}/J_{\mathrm{in}}$ as a function of the exponent α of the zero-mean generalized Gaussian noise $\eta(t)$ with rms amplitude $\sigma_\eta = 1$.

shows examples of evolution of the Fisher information J_{out} as the noise level σ_η is increased, at different values of the quantization threshold θ. When the threshold is optimal, J_{out} undergoes a monotonic decay as σ_η is raised, meaning that noise always degrades performance for the optimal quantizer. Yet, when the quantizer is not optimal, Fig. 4 shows that enhancement of the noise level σ_η can sometimes increase J_{out}, which translates in improved performance in the estimation task. This is a form of stochastic resonance, where injection of noise improves the performance.

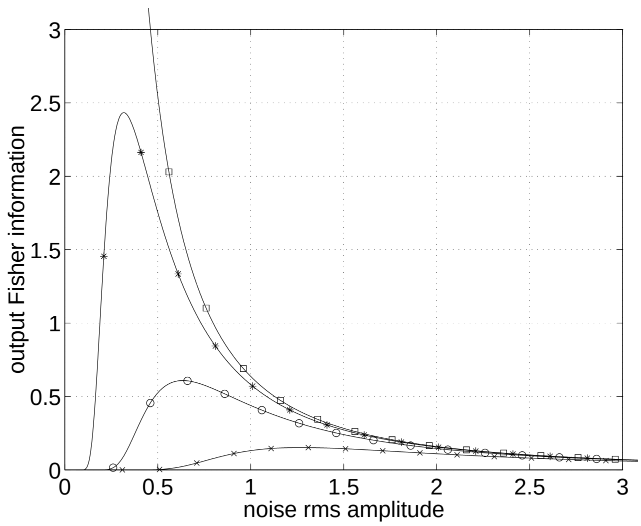

Fig. 4. Output Fisher information J_{out} as a function of the rms amplitude σ_η of the zero-mean Gaussian noise $\eta(t)$, for $v = v_{\mathrm{opt}} = 0$ ($\square$), $v = 0.5$ ($*$), $v = 1$ ($\circ$) and $v = 2$ ($\times$).

In addition to its quantitative assessment of Fig. 4, this favorable action of the noise can be qualitatively understood. When the signal by itself is not well positioned in relation to the fixed (nonoptimal) quantization threshold, noise injection has the ability to move the operating zone of the quantizer towards a region more favorable to the efficient quantization of the noisy input. This translates into the possibility of increasing J_{out} and also $J_{\mathrm{out}}/J_{\mathrm{in}}$, as the noise level σ_η is increased, as illustrated in Fig. 4.

A difficulty for drawing full benefit of these two available optimization strategies (threshold adjustment or noise injection) is that in general the optimal threshold and optimal noise level are dependent upon the signal $s(t)$ which contains the unknown parameter a (e.g. $\theta_{\mathrm{opt}} = v_{\mathrm{opt}} + a$ when $s(t) \equiv a$). In practice, some prior knowledge may exist on the feasible values or range for the parameter to be estimated, and this will be useful to configure the quantization threshold or noise level close to optimum. If this prior knowledge is expressable by a prior probability density for a, both optimal threshold and optimal noise level can be determined to maximize the expected or average Fisher information over this prior density. Another approach, that we shall present, is to devise an adaptive scheme that would bring the quantizer to operate in optimal conditions through interaction with a definite estimator.

4. ADAPTIVE ML ESTIMATION

The behavior of the Fisher information characterizes, through the Cramér-Rao inequality, the overall best performance that can be expected for any unbiased estimator of a. Both possibilites for maximizing this performance, through threshold adjustment or through noise injection, are available in principle for any type of parametric dependence of $s(t)$ upon a. Next, to illustrate the theory, we again consider the example of estimation of a constant signal $s(t) \equiv a$ by means of the maximum likelihood (ML) estimator $\widehat{a}_{\mathrm{ML}}$.

When estimation is performed from N samples $(y_1, \dots y_N) = \boldsymbol{y}$ at the output of the two-level quantizer of Eq. (1), the likelihood $L(\boldsymbol{y}; a)$ is expressable as

$$L(\boldsymbol{y}; a) = \prod_{j=1}^{N} \Pr\{y_j | a\} \ , \tag{8}$$

where $\Pr\{y_j = -1|a\} = F_\eta(\theta - a)$ and $\Pr\{y_j = 1|a\} = 1 - F_\eta(\theta - a)$. Maximization of $L(\boldsymbol{y}; a)$ with respect to a yields the ML estimator under the form

$$\widehat{a}_{\mathrm{ML}} = \theta - F_\eta^{-1}\left(\frac{1-\overline{y}}{2}\right) \ , \tag{9}$$

with the sample mean $\overline{y} = N^{-1}\sum_{j=1}^{N} y_j$.

At large N, the mean squared error of $\widehat{a}_{\mathrm{ML}}$ is $1/(NJ_{\mathrm{out}})$ (respectively $1/(NJ_{\mathrm{in}})$) when estimation is made from the output $y(t)$ (respectively the input $s(t) + \eta(t)$). Thus the performance of the ML estimator from the output of the quantizer benefits from the possibilites of improvement reported in the previous Section with the Fisher information, through threshold adjustment or through noise injection.

To bring the ML estimator to operate in optimal conditions of quantization as derived from Eq. (5) and expressed by curves like in Fig. 2, we devise an adaptive algorithm. In definite noise conditions, this algorithm iteratively adjusts the quantization threshold so as to reach the optimum of Fig. 2.

This algorithm performs block processing of the data. From the current block and current threshold, an estimate is obtained from the ML estimator of Eq. (9). With this current estimate and the appropriate curve like in Fig. 2, a value of the threshold for optimal quantization is deduced. This threshold value is then used for quantization of the next block of data, and the process is iterated. It can be proven that this procedure converges to the performance of the best ML estimator operating in optimal conditions of quantization. Futhermore, we show that in the regime of operation of this procedure, in the vicinity of the optimal quantization conditions, the nonlinear function $F_\eta^{-1}(.)$ of Eq. (9) can be linearized so as to give

$$\widehat{a}_{\rm ML} \approx \theta - v_{\rm opt} + \frac{1}{2f_\eta(v_{\rm opt})}\left[\overline{y} + 2F_\eta(v_{\rm opt}) - 1\right] , \quad (10)$$

which gets more accurate as $|a - \widehat{a}_{\rm ML}| \ll \sigma_\eta$. An expression like Eq. (10), linear in the sample mean $\overline{y}$, is specially useful for fast real-time processing with limited computational resources.

A similar type of adaptive procedure can be envisaged to iteratively adjust the noise level prior to quantization, in case of a fixed nonadjustable threshold θ. As seen in the previous Section, for a given parameter value a and a fixed nonoptimal threshold θ, there usually exists a nonzero noise level $\sigma_{\rm opt}$ at which the estimation performance is maximized. If the level σ_η of the noise that pre-exists with the data is not too high, i.e. if $\sigma_\eta < \sigma_{\rm opt}$, addition of an extra noise $\xi(t)$ with variance $\sigma_\xi^2 = \sigma_{\rm opt}^2 - \sigma_\eta^2$ to the data prior to quantization, has the ability to improvre the estimation performance. Again, since $\sigma_{\rm opt}$ is usually dependent upon θ (known) and a (unknown), an adaptive procedure can be used to iteratively bring a given estimator like the ML estimator of Eq. (9), to operate at the optimal noise level $\sigma_{\rm opt}$. An adaptive algorithm performing block processing of the data is shown in Fig. 5.

5. CONCLUSION

In nonlinear processing, when conditions are sought to optimize the performance, it is interesting to realize that addition of noise may sometimes represent a favorable strategy. This is the phenomenon of stochastic resonance, which can occur under many different forms, in various types of nonlinear processing of signals. Here, we have shown the possibility of improving the performance for estimation by addition of noise, in specific conditions. Stochastic resonance is a phenomenon of relatively recent introduction, and probably many forms and properties of stochastic resonance remain to be uncovered and investigated.

Repeat
  $\sigma_{\rm opt} \longleftarrow \arg_\sigma \max J_{\rm out}(\sigma,\ v = \theta - \widehat{a}_{\rm ML})$
  $\sigma_\xi^2 = \sigma_{\rm opt}^2 - \sigma_\eta^2$
  For each j over the current block
    $x_j \longleftarrow x_j + \xi_j$
    $y_j = \mathrm{sign}(x_j - \theta)$
  End For
  $\overline{y}$ = sample mean of current block of y_j's
  $\widehat{a}_{\rm ML} = \theta - F_\eta^{-1}\left(\frac{1-\overline{y}}{2}\right)$
Until a stopping condition is true

Fig. 5. An adaptive algorithm performing block processing of the data to improve the performance of the ML estimator of Eq. (9) by addition of noise.

6. REFERENCES

[1] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, "Low-complexity digital encoding strategies for wireless sensors networks", in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, 1998, pp. 3273–3276.

[2] S. Mitaim and B. Kosko, "Adaptive stochastic resonance", *Proceedings of the IEEE*, vol. 86, pp. 2152–2183, 1998.

[3] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance", *Reviews of Modern Physics*, vol. 70, pp. 223–287, 1998.

[4] F. Chapeau-Blondeau, "Stochastic resonance and the benefit of noise in nonlinear systems", in *Noise, Oscillators and Algebraic Randomness: From Noise in Communication Systems to Number Theory*, M. Planat, Ed., pp. 137–155. Springer, Berlin, 2000.

[5] B. Andò and S. Graziani, *Stochastic Resonance: Theory and Applications*, Kluwer Academic Publishers, Boston, 2000.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Englewood Cliffs, 1993.

[7] F. Chapeau-Blondeau, "Nonlinear test statistic to improve signal detection in non-Gaussian noise", *IEEE Signal Processing Letters*, vol. 7, pp. 205–207, 2000.

[8] F. W. Machell, C. S. Penrod, and G. E. Ellis, "Statistical characteristics of ocean acoustic noise processes", in *Topics in Non-Gaussian Signal Processing*, E. J. Wegman, S. C. Schwartz, and J. B. Thomas, Eds., pp. 29–57. Springer, Berlin, 1989.