Noise-aided nonlinear Bayesian estimation

François Chapeau-Blondeau
Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France
(Received 15 May 2002; published 11 September 2002)

Estimation on a noisy signal observed by a nonlinear sensor taking the form of a threshold quantizer is considered. The optimal Bayesian estimator with minimal error is derived in this nonlinear setting. The existence of conditions where the performance of this estimator can be improved by raising the level of noise is established, both theoretically and numerically. These results constitute a different instance of the nonlinear phenomenon of stochastic resonance for signal enhancement by noise.

DOI: 10.1103/PhysRevE.66.032101 PACS number(s): 05.40.−a, 02.50.−r

I. INTRODUCTION

Stochastic resonance is a nonlinear phenomenon of noise-aided signal transmission displaying very attractive potentialities for nonlinear signal processing [1]. It has been reported to occur, under various forms, in a variety of nonlinear systems, including electronic circuits [2–4], optical devices [5–8], and neurons [9–11]. Nonlinear transmission of periodic signals has been quantified by signal-to-noise ratios improvable by addition of noise [12,13]. For nonlinear transmission of aperiodic or random signals, correlation measures [14] or information-theoretic quantities [15–17] have been shown to be improvable by noise. Signal detection performances enhanced by noise have been reported in Refs. [18–20]. Estimation performances, essentially for estimating the value of a deterministic constant signal, and quantified through the Fisher information or the Cramér-Rao bound, have been shown improvable by noise [21–23].

Here, we extend the conditions under which a form of stochastic resonance can be obtained. We consider an estimation task in a Bayesian framework. We seek to estimate a random parameter (amplitude, frequency, phase,...) belonging to an information signal corrupted by noise and observed through a nonlinear sensor. We derive the optimal Bayesian estimator with minimal error in this nonlinear setting, and demonstrate the possibility of improving its performance by raising the level of noise.

II. BAYESIAN ESTIMATION

An unknown parameter ν is attached to a signal $s_\nu(t)$ corrupted by a noise $\eta(t)$. An observable signal $y(t)$, related to the signal-noise mixture, is available for measurement, with the aim of estimating ν from $y(t)$. In a Bayesian framework [24], the possible values for ν are distributed according to the prior probability density function (PDF) $p_\nu(\nu)$. Observation of $y(t)$ at N distinct times t_j provides N data points $y_j = y(t_j)$, for $j=1$ to N. Once $\mathbf{y}=(y_1,\ldots,y_N)$ is observed, a posterior PDF $p(\nu|\mathbf{y})$ for the parameter ν can be defined. A mean square error in the estimation follows as the expectation (conditioned by observation $\mathbf{y}$)

$$\mathcal{E}=E((\nu-\hat{\nu})^2|\mathbf{y})=\int(\nu-\hat{\nu})^2 p(\nu|\mathbf{y})d\nu. \tag{1}$$

Error $\mathcal{E}$ of Eq. (1) can equivalently be expressed as

$$\mathcal{E}=[\hat{\nu}-E(\nu|\mathbf{y})]^2+\mathrm{var}(\nu|\mathbf{y}), \tag{2}$$

with $E(\nu|\mathbf{y})=\int \nu p(\nu|\mathbf{y})d\nu$ and $\mathrm{var}(\nu|\mathbf{y})=\int[\nu-E(\nu|\mathbf{y})]^2 p(\nu|\mathbf{y})d\nu$.

Since $\mathrm{var}(\nu|\mathbf{y})$ in Eq. (2) is non-negative and independent of $\hat{\nu}$, the optimal Bayesian estimator that minimizes error $\mathcal{E}$ comes out as

$$\hat{\nu}_B=E(\nu|\mathbf{y})=\int \nu p(\nu|\mathbf{y})d\nu, \tag{3}$$

and its performance is measured by the minimal error

$$\mathcal{E}_B=\mathrm{var}(\nu|\mathbf{y})=\int[\nu-E(\nu|\mathbf{y})]^2 p(\nu|\mathbf{y})d\nu. \tag{4}$$

A model of how the observation $\mathbf{y}$ is produced in relation to the parameter ν and to the noise spoiling the observation, allows one to define the PDF $p(\mathbf{y}|\nu)$ of observing $\mathbf{y}$ given ν. With the prior information summarized by $p_\nu(\nu)$, the Bayes rule then provides access to the posterior PDF under the form

$$p(\nu|\mathbf{y})=\frac{p(\mathbf{y}|\nu)p_\nu(\nu)}{p(\mathbf{y})}, \tag{5}$$

with the PDF $p(\mathbf{y})=\int p(\mathbf{y}|\nu)p_\nu(\nu)d\nu$.

For any given observation $\mathbf{y}$, the optimal Bayesian estimator $\hat{\nu}_B$ of Eq. (3) achieves the minimum $\mathcal{E}_B$ of Eq. (4) of the error $\mathcal{E}$ of Eq. (1). Consequently, $\hat{\nu}_B$ also achieves the minimum $\overline{\mathcal{E}_B}$ of error $\mathcal{E}$ averaged over every possible observation $\mathbf{y}$, i.e.,

$$\overline{\mathcal{E}_B}=\int \mathrm{var}(\nu|\mathbf{y})p(\mathbf{y})d\mathbf{y}, \tag{6}$$

where $\int\cdot d\mathbf{y}$ stands for the N-dimensional integral $\int\cdots\int\cdot dy_1\cdots dy_N$.

We shall now address a specific estimation problem amenable to this general Bayesian estimation procedure. We shall consider nonlinear conditions of observation of the signal-noise mixture. In such case, we shall show that the optimal estimator of Eq. (3) displays a performance, measured by Eq. (4) or Eq. (6), that can be improved by raising the level of the noise.

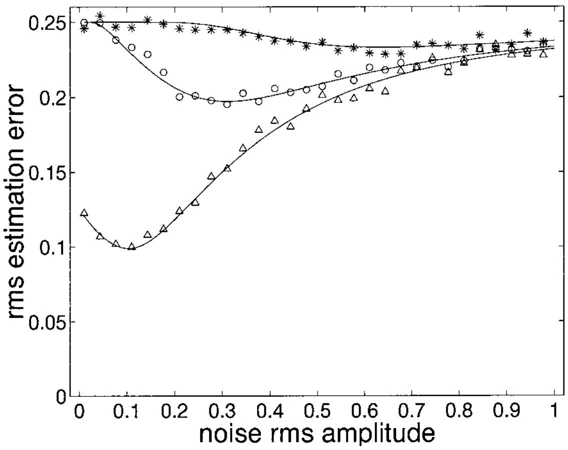

FIG. 1. rms estimation error $\overline{\mathcal{E}}_B^{1/2}$ from $N=4$ data points y_j as a function of the rms amplitude σ_η of the noise $\eta(t)$ chosen zero-mean Gaussian. The signal is $s_\nu(t)\equiv\nu$, with $p_\nu(\nu)$ uniform of mean 1 and standard deviation 0.25. The solid lines are the theory from Eq. (6). The discrete points are from Monte Carlo estimation by Eq. (3) with (*) $\theta=0$, (○) $\theta=0.5$, (△) $\theta=1$.

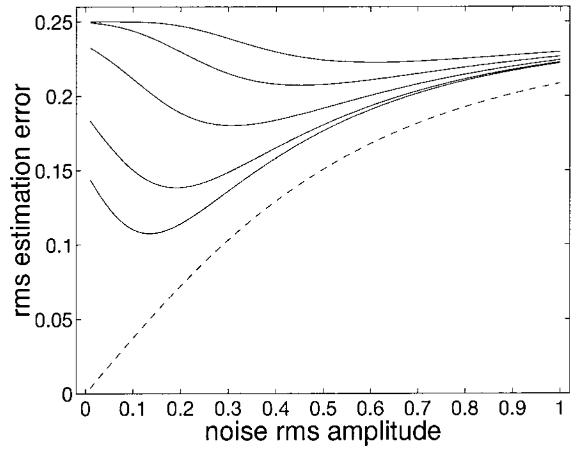

FIG. 2. rms estimation error $\overline{\mathcal{E}}_B^{1/2}$ of Eq. (6) from $N=7$ data points y_j as a function of the rms amplitude σ_η of the noise $\eta(t)$ chosen zero-mean Gaussian. The signal is $s_\nu(t)\equiv\nu$, with $p_\nu(\nu)$ Gaussian of mean 1 and standard deviation 0.25. For the solid lines, from top to bottom: $\theta=0$, 0.25, 0.5, 0.75, 1. The dashed line is for the linear optimal estimator operating directly on the signal-plus-noise mixture instead of its quantized version.

III. NOISE-AIDED ESTIMATION

The observation of the signal-plus-noise mixture $s_\nu(t)+\eta(t)$ is realized through a memoryless nonlinearity as

$$y(t)=g(s_\nu(t)+\eta(t)). \tag{7}$$

Various forms of the nonlinearity $g(\cdot)$ could lead to the possibility of a noise-enhanced performance in the estimation, for instance, multilevel quantizers. For a simple illustration of this possibility, we take the nonlinearity $g(\cdot)$ as a two-level quantizer with threshold θ, giving

$$y(t)=\mathrm{sgn}[s_\nu(t)+\eta(t)-\theta]=\pm 1. \tag{8}$$

The noise $\eta(t)$ is assumed stationary, white, with cumulative distribution function $F_n(u)$. In this case, the conditional PDF factorizes as $p(\mathbf{y}|\nu)=\Pi_{j=1}^N p(y_j|\nu)$, with the PDF

$$p(y_j|\nu)=\Pr\{y_j=-1|\nu\}\delta(y_j+1)+\Pr\{y_j=1|\nu\}\delta(y_j-1). \tag{9}$$

One has the probability $\Pr\{y_j=-1|\nu\}=\Pr\{s_\nu(t_j)+\eta(t_j)<\theta\}$, which amounts to $\Pr\{y_j=-1|\nu\}=F_\eta(\theta-s_\nu(t_j))$. In the same way, $\Pr\{y_j=1|\nu\}=1-\Pr\{y_j=-1|\nu\}=1-F_\eta(\theta-s_\nu(t_j))$. The above expressions enable, through Eq. (5), explicit calculation of the optimal estimator $\hat{\nu}_B$ of Eq. (3) and its performance of Eqs. (4) or (6).

We first consider the case of a constant signal $s_\nu(t)=\nu$, $\forall t$, to be estimated. In this case, Fig. 1 represents the rms estimation error $\overline{\mathcal{E}}_B^{1/2}$ computed from Eq. (6), with Gaussian noise $\eta(t)$, when the value of ν has a uniform prior PDF. The results of Fig. 1 clearly reveal a possibility of reducing the estimation error by increasing the noise level, down to a minimal error occurring for a non-zero optimal noise level. A Monte Carlo simulation of the estimation scheme has also been realized numerically. A large number of trials (10^3 for each noise level σ_η) of the optimal estimation through $\hat{\nu}_B$ of Eq. (3) have been generated and the average estimation error $\overline{\mathcal{E}_B}$ has been evaluated as an empirical average. In Fig. 1, the noise-aided estimation is confirmed by both the theory and the simulation.

The conditions of Fig. 1 are merely illustrative. Noise-aided estimation is preserved in other conditions, as further exemplified by Fig. 2, which shows the nonmonotonic evolution of the estimation error with Gaussian noise $\eta(t)$ and a Gaussian prior PDF for ν.

For estimation of a constant signal $s_\nu(t)\equiv\nu$ with a prior PDF $p_\nu(\nu)$ symmetric about its prior mean $E(\nu)$, it can be verified that the estimation error $\overline{\mathcal{E}_B}$ remains unchanged when the threshold θ is changed from $E(\nu)+h$ to $E(\nu)-h$, for any h. As a consequence, $\overline{\mathcal{E}_B}$ viewed as a function of θ at any fixed noise level σ_η, has an extremum at $\theta=E(\nu)$. This extremum is the minimum occurring at $\theta=E(\nu)=1$ in Figs. 1 and 2, which corresponds to the lowest curve $\overline{\mathcal{E}}_B^{1/2}$ as a function of σ_η and associated with $\theta=E(\nu)=1$. In such conditions, for estimation of a constant signal $s_\nu(t)\equiv\nu$ from a quantized signal-plus-noise mixture, the optimal location of the quantization threshold is thus $\theta=E(\nu)$, i.e., it is when the quantization threshold is located at the prior mean that the estimation error $\overline{\mathcal{E}_B}$ is minimal. But, moreover, when this optimal threshold is implemented, the results of Figs. 1 and 2 clearly show that further benefit can be obtained by raising the level of the noise σ_η, over some ranges of σ_η. In other words, the optimal Bayesian estimator operating on the quantized signal-plus-noise mixture, can have a performance improvable by addition of noise.

However, if the complete signal-plus-noise mixture $s_\nu(t)+\eta(t)$ is available for estimation (instead of its quantized version), then, in general, the performance of the optimal "linear" Bayesian estimator will be better than that of the optimal "nonlinear" estimator after quantization, and this performance will usually undergo a monotonic degradation

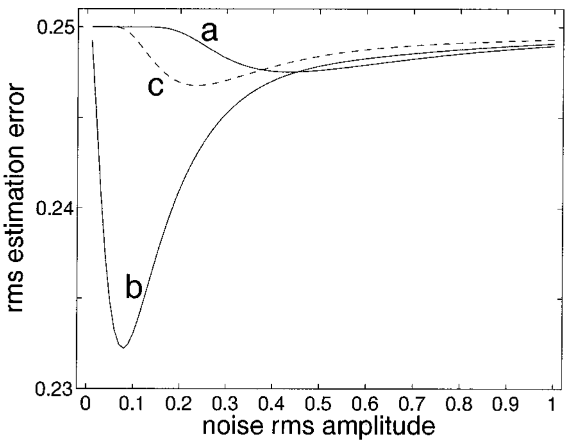

FIG. 3. rms estimation error $\overline{\mathcal{E}}_B^{1/2}$ of Eq. (6) as a function of the rms amplitude σ_η of the noise $\eta(t)$ chosen zero-mean Gaussian. The signal is $s_\nu(t)=\exp(-\nu t)\cos(2\pi t/0.2)$, with $p_\nu(\nu)$ uniform of mean 1 and standard deviation 0.25. The $N=6$ data samples are equispaced with time step 0.1 from $t_1=0$ to $t_6=0.5$. The quantization threshold is $\theta=0$ (a), $\theta=0.55$ (b), $\theta=1.1$ (c).

as the noise level σ_η is raised. This is exemplified in Fig. 2 with Gaussian noise $\eta(t)$ and a Gaussian prior $p_\nu(\nu)$. Nevertheless, the linear estimator operates on a continuous (analog) representation of the data, or with practical hardware, on a 16- or 12- or 8-bit representation. This is to be contrasted with the much parsimonious one-bit representation per data point used by the nonlinear estimator. If some measure of hardware requirement is included in the evaluation of the performance, the interest of the nonlinear estimator becomes manifest.

Noise-enhanced performance in Bayesian estimation from quantized data is also possible for other types of signal $s_\nu(t)$. For illustration, Fig. 3 considers the case of a damped sinusoid $s_\nu(t)=\exp(-\nu t)\cos(2\pi t/0.2)$. Figure 3 represents the rms error $\overline{\mathcal{E}}_B^{1/2}$ for the optimal Bayesian estimation of the damping factor ν from data quantized through Eq. (8). Conditions demonstrating the possibility of a noise-improved performance in the estimation are shown in Fig. 3.

IV. CONCLUSION

The present study has addressed the situation of parametric Bayesian estimation based on data observed through nonlinear sensors, typically taking the form of threshold quantizers. Such conditions allowing parsimonious data representation are specially relevant for a number of existing and future multisensor networks or distributed intelligent systems. They make possible the optimization of speed and efficacy of processing with limited resources for data handling, storage, communication or energy supply [22]. In association with threshold adjustment at the quantizers, we have demonstrated that noise addition offers a complementary means that can be exploited to optimize the performance in estimation. Especially, we have shown the possibility of conditions where the quantizer with optimal threshold can be further improved by addition of noise. In other situations, adaptation to the optimal threshold may not be accessible, with a ''hard-wired'' threshold imposed by the physics of the sensor. This may be the case with neural systems that may use stochastic resonance to contribute to their high performances for information processing. In these conditions of limited flexibility of the sensors, the interest of considering noise addition to optimize the performance is even more manifest. The present results contribute to enlarging the potentialities of noise improvement by stochastic resonance and are useful to the progress in the understanding and control of information processing by nonlinear systems.

[1] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
[2] S. Fauve and F. Heslot, Phys. Lett. **97A**, 5 (1983).
[3] X. Godivier, J. Rojas-Varela, and F. Chapeau-Blondeau, Electron. Lett. **33**, 1666 (1997).
[4] D. G. Luchinsky, R. Mannella, P. V. E. McClintock, and N. G. Stocks, IEEE Trans. Circuits Syst., II: Analog Digital Signal Process. **46**, 1205 (1999).
[5] B. McNamara, K. Wiesenfeld, and R. Roy, Phys. Rev. Lett. **60**, 2626 (1988).
[6] M. I. Dykman *et al.*, Appl. Phys. Lett. **67**, 308 (1995).
[7] B. M. Jost and B. E. A. Saleh, Opt. Lett. **21**, 287 (1996).
[8] F. Vaudelle *et al.*, J. Opt. Soc. Am. B **15**, 2674 (1998).
[9] A. Bulsara *et al.*, J. Theor. Biol. **152**, 531 (1991).
[10] J. K. Douglass, L. Wilkens, E. Pantazelou, and F. Moss, Nature (London) **365**, 337 (1993).
[11] F. Chapeau-Blondeau, X. Godivier, and N. Chambet, Phys. Rev. E **53**, 1273 (1996).
[12] B. McNamara and K. Wiesenfeld, Phys. Rev. A **39**, 4854 (1989).
[13] F. Chapeau-Blondeau and X. Godivier, Phys. Rev. E **55**, 1478 (1997).
[14] J. J. Collins, C. C. Chow, and T. T. Imhoff, Phys. Rev. E **52**, R3321 (1995).
[15] A. R. Bulsara and A. Zador, Phys. Rev. E **54**, R2185 (1996).
[16] C. Heneghan *et al.*, Phys. Rev. E **54**, R2228 (1996).
[17] F. Chapeau-Blondeau, Phys. Rev. E **55**, 2016 (1997).
[18] M. E. Inchiosa and A. R. Bulsara, Phys. Rev. E **53**, R2021 (1996).
[19] V. Galdi, V. Pierro, and I. M. Pinto, Phys. Rev. E **57**, 6470 (1998).
[20] M. E. Inchiosa, J. W. C. Robinson, and A. R. Bulsara, Phys. Rev. Lett. **85**, 3369 (2000).
[21] M. Stemmler, Network Comput. Neural Syst. **7**, 687 (1996).
[22] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1998* (IEEE, Piscataway, 1998), pp. 3273–3276.
[23] P. E. Greenwood, L. M. Ward, and W. Wefelmeyer, Phys. Rev. E **60**, 4687 (1999).
[24] H. V. Poor, *An Introduction to Signal Detection and Estimation* (Springer, Berlin, 1994).