APROXIMATE ALGORITHMS FOR NONLINEAR FILTERING OF CHAOS*
ALGORITMOS DE APROXIMACIÓN PARA FILTRADO NO LINEAL DE CAOS

Valeri Kontorovich[1]
Zinaida Lovtchikova[2]

Recibido: Mayo 2009
Aprobado: Julio 2009

Abstract

The paper is dedicated to the description of two approximate methods of non-linear filtering algorithms for signals of Lorenz, Chua and Rössler attractors to provide real time filtering solutions for scenarios with low Signal Noise Ratios (SNR). For those cases the method of the Global (Integral) Approximation of the a-posteriori Probability Density Function (PDF) is considered. Some asymptotical solutions are presented as well.

Keywords
Nonlinear filtering, Global Approximation, Probability Density Function, attractors, stationary conditions, asymptotical solutions.

Resumen
Este documento está dedicado a la descripción de dos métodos de aproximación de algoritmos de filtrado no lineales para señales de Lorentz y atractores Chua y Rössler, ofreciendo soluciones de filtrado en tiempo real para escenarios con bajas tasas de señal de ruido (RSR). Para esos casos, se considera el método de la Aproximación Global (Integral) de la a posteriori función de densidad de probabilidad (FDP). También se presentan algunas soluciones asintóticas.

Palabras clave
Filtrado no lineal, Aproximación Global, Función de Densidad de Probabilidad, atractores, condiciones estacionarias, soluciones asintóticas.

1. INTRODUCTION

During the last decade, chaotic signals (chaotic models) were widely applied for different purposes in the electrical engineering field. Meanwhile the statistical descriptions of chaos haven't been widely developed so far. For this reason in [4], authors proposed the so-called "degenerated cumulant equations" approach in order to offer a rather simple and adequate method of statistical analysis for chaos, suitable for applications.

* This work was supported by the Intel Corporation through the "IntelVK" grant.
1 Doctor of Science, Electrical Engineering Department, Communications Section, Cinvestav-IPN (México D.F.). Correo: valeri@cinvestav.mx
2 Doctor of Science, Engineering and Advanced Technology Interdisciplinary Professional Unit, Upiita-IPN. Correo: alovtchikova@ipn.mx

Actually, chaos can be mathematically represented in a way:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t)) \tag{1}$$

where $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, $\boldsymbol{x} \in R^n$, $f(\cdot)$ is a deterministic vector function of $\boldsymbol{x}$ with the components $[f_1(\boldsymbol{x}),\ldots f_n(\boldsymbol{x})]^T$

Equation (1) can be transformed in the stochastic form: Stochastic Differential Equation or SDE [4]:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t)) + \varepsilon\xi(t) \tag{2}$$

where $\xi(t)$ is a vector of "weak" white noise with the matrix of intensities $\varepsilon = [\varepsilon_{ij}]^{nxn}$ with its elements tending to zero. Concrete type of the vector function *f(x)* in (1) and (2) describes different nonlinear dissipative systems (strange attractors), which generates almost randow chaotic signals [1]. In the following are considered only widely used attractors so as: Lorenz, Chua and Rössler [4].

It is worth to mention, that the SDE representation of chaos is rather opportunistic: in our case it allows to apply the well developed and well known theory of the nonlinear filtering for Markov processes [9], [6].

One can find in [9], [6], that if the received signal is:

$$\boldsymbol{y}(t) = \boldsymbol{s}(t, \boldsymbol{x}(t)) + \boldsymbol{n}_0(t) \tag{3}$$

where $\boldsymbol{y}(t)$ is a vector of the received signal, $s(t,\boldsymbol{x})$ is vector function of the desired signal, which depends to the Markov process $\boldsymbol{x}(t)$ (message), and $\boldsymbol{n}_0(t)$ is a vector of the additive white noises with the intensity matrix $\boldsymbol{N}_0$ (all with dimension "m").

Then the a-posteriori PDF $W_{PS}(\boldsymbol{x},t)$ of the process *x(t)* follows the so called i*n*tegral-differential Stratonovich-Kushner Equation (SKE), [6], [9]:

$$\frac{\partial W_{PS}(\boldsymbol{x},t)}{\partial t} = -div\,\hat{\boldsymbol{\pi}}(\boldsymbol{x},t) + \\ + \frac{1}{2}\left[\boldsymbol{F}(\boldsymbol{x},t) - \langle F(\boldsymbol{x},t)\rangle\right] W_{PS}(\boldsymbol{x},t) \tag{4}$$

where $\langle F(\boldsymbol{x},t)\rangle = \int_{-\infty}^{\infty} \boldsymbol{F}(\boldsymbol{x},t) W_{PS}(\boldsymbol{x},t) d\boldsymbol{x}$

$$\boldsymbol{F}(\boldsymbol{x},t) = [\boldsymbol{y}(t) - s(\boldsymbol{x},t)]^T \boldsymbol{N}_0^{-1} [\boldsymbol{y}(t) - s(\boldsymbol{x},t)]$$

$\hat{\pi}(\boldsymbol{x},t)$ is a "probabilistic *flow*" with the components:

$$\boldsymbol{\pi}_i(\boldsymbol{x},t) = f_i(\boldsymbol{x},t) W_{PS}(\boldsymbol{x},t) - \frac{1}{2}\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[\varepsilon_{ij} W_{PS}(\boldsymbol{x},t)\right]$$

T is the symbol of transposition.

It is well known, that with the exception of some few special cases [2], SKE doesn't provide with exact solutions. Though, the majority of the nonlinear filtering algorithms are approximate ones. Mainly those approximate algorithms are considered as extended Kalman filtering (EKF) algorithms or their modifications [2], [8]; as a matter of fact EKF and its modifications are really applicable only for high Signal Noise Ratios (SNR) scenarios.

In the following will be considered the low SNR cases for nonlinear real-time filtering of chaos, which requires some special attempts. First of all, hereafter is proposed in case of low SNR to apply for $W_{PS}(\boldsymbol{x}, t)$ approximation the method of Global or Integral Approximation. Second, it was proposed to reduce the dimension of SDE (2) by its statistically equivalent one-dimensional SDE (SDE-1). The last ones can be synthesized with the help of methodology of [7]. Third, for SDE-1 can be utilized the asymptotic algorithms in the low SNR sense.

In this paper we'll assume that the subject of the filtering is one (for example, the first) component of each strange attractor. Keeping in mind that attractors of interest have $n = 3$ in (1), (2), this first component is called as observable one. Though, one has to find the statistically equivalent SDE-1 for the three-dimensional SDE (2), where the observable component has the same statistical properties as SDE-1.

In this way in the following is done a reduction of dimensionality of the SDE (2). So, the received signal is:

$$\boldsymbol{y}(t) = \boldsymbol{x_1}(t) + \boldsymbol{n}_0(t) \tag{5}$$

where $\boldsymbol{x}_1(t)$ is the observable component of attractor (desired signal); $\boldsymbol{n}_0(t)$ is an additive white noise. To obtain statistical characterization of the Lorenz, Chua and Rössler attractors, one have to refer to [4] and to the Table 1.

The paper is organized in the following way. At section 2 one can find a brief description of the Integral or Global Approximation principles. Section 3 is completely dedicated to the application of Integral (Global) Approximation to $W_{PS}(\boldsymbol{x},t)$. Here also are presented asymptotical algorithms for case of low SNR. Conclusions are presented in Section 4.

2. INTEGRAL OR GLOBAL APPROXIMATION FOR $W_{PS}(x,t)$.

The main challenge of the Integral (Global) Approximation is to avoid the "local" estimations of maximums of $W_{PS}(\boldsymbol{x},t)$ of the filtered process $\boldsymbol{x}(t)$. The latter is typical for EKF algorithms which are applying mainly Gaussian approximations of $W_{PS}(\boldsymbol{x},t)$, etc. In contrary to it, Integral or Global approach is for successful "complete" approximation of $W_{PS}(\boldsymbol{x},t)$ including its "tails".

Let us assume, that $W_{PS}(\boldsymbol{x},t)$ can be represented in the way:

$$W_{PS}(\boldsymbol{x},t) = W_{PS}(\boldsymbol{x},\alpha(t)) \tag{6}$$

Where $\alpha(t)$ is an unknown vector of approximation parameters.

Then, applying the well know Kullback measure as an approximation criteria, one can obtain the following equation for the unknown vector α:

$$\dot{\alpha} = \left\langle L_{PR}^{+}\{\boldsymbol{h}(\boldsymbol{x},t)\}\right\rangle + \boldsymbol{V}^{-1}(t)\left\langle \boldsymbol{h}(\boldsymbol{x},t)\boldsymbol{F}(\boldsymbol{x},t)\right\rangle \tag{7}$$

where ; $\boldsymbol{h}(\boldsymbol{x},t) = \dfrac{\partial \ln W_{PS}(\boldsymbol{x},\alpha(t))}{\partial\alpha}$

$$\boldsymbol{V}(t) = -\int_{-\infty}^{\infty}\left[\frac{\partial \ln W_{PS}(\boldsymbol{x},\alpha(t))}{\partial\alpha}\right]^{T} W_{PS}(\boldsymbol{x},\alpha(t))dx = -\left\langle\frac{\partial^2 W_{PS}(\boldsymbol{x},\alpha(t))}{\partial\alpha\partial\alpha^{\mathrm{T}}}\right\rangle$$

$$L^{+}\{\ \} = f_i(\boldsymbol{x})\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{j=1}^{n}\varepsilon_{ij}\frac{\partial^2}{\partial x_i \partial x_j}$$

The details of this development for (7) can be found at [3], [10]. Let us represent (6) in the following way:

$$W_{PS}(\boldsymbol{x},\alpha(t)) = Ce^{\left\{\sum_{p=1}^{K}\alpha_p(t)\varphi_p(\boldsymbol{x})+\varphi_0(\boldsymbol{x})\right\}} \tag{8}$$

Where **α** *(t)* is a vector of sufficient statistics for $W_{PS}(\cdot)$; $\{\varphi_P(\boldsymbol{x})\}$ are the set of orthogonal multidimensional functions: Hermite, Laguerre, etc.; C is a normalization constant.

One can presume from (7) and (8), that the filtering algorithm for $n = 3$ might be rather complex and not adequate for the real time solutions. In the next section will be presented an approach how to reduce those complexities.

3. INTEGRAL APPROXIMATION FOR $W_{PS}(x,t)$.

Let us consider Lorenz, Chua and Rössler attractors. As it follows from [4], and can be seen from the Table 1, marginal PDF's of the components for Lorenz attractor are practically Gaussian, or it's orthogonal representation, [7] has a Gaussian kernel PDF; for Rössler attractor orthogonal representation with the Gaussian kernel PDF is also valid for "*x*" and "*y*" components of the attractor. The opposite situation takes place for Chua attractor (Table 1): it can be seen that this attractor represents a clearly non-Gaussian case.

No.	Name of the Strange attractor	$\boldsymbol{g}(\boldsymbol{x})$	ε	$W_{PR}(x_i)$	Comments
1	Lorenz, $n = 3$ $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$	$\begin{cases} \sigma(x_2 - x_1) \\ Rx_1 - x_2 - x_3x_1 \\ x_1x_2 - Bx_3 \end{cases}$ $\sigma, R, B \geq 0$	$\varepsilon_{11} = \varepsilon \to 0$ $\varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23} =$ $\varepsilon_{21} = \varepsilon_{32} = \varepsilon_{33} = 0$	$x_1 \sim W_G(\cdot)$ $x_2 \sim W_G(\cdot)$ $x_3 \sim W_G(\cdot)$	Normalized dates, $W_G(\cdot)$ – Gaussian PDF
2	Chua, $n = 3$ $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$	$\begin{cases} \beta_1(x_2 - x_1) - \alpha h(x_1) \\ \beta_2(x_1 - x_2) + \beta_4 x_3 \\ -\beta_3 x_2 \end{cases}$ $\beta_1 - \beta_4 \geq 0, \alpha < 0$	$\varepsilon_{11} = \varepsilon \to 0$ $\varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23} =$ $\varepsilon_{21} = \varepsilon_{32} = \varepsilon_{33} = 0$	$x_1 \sim Const \exp(p_1 x_1^2 - q_1 x_1^4)$ $x_2 \sim W_G(\cdot)$ $x_3 \sim Const \exp(p_3 x_3^2 - q_3 x_3^4)$ $p_1, p_2, q_1, q_2 > 0$	Normalized dates, $p_1 \sim 3{,}5$ $p_3 \sim 3{,}5$ $q_1 \sim 1{,}5$ $q_3 \sim 2{,}5$
3	Rössler, $n = 3$ $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$	$\begin{cases} -x_2 - x_3 \\ x_1 + ax_2 \\ b + x_3x_1 - x_3C \end{cases}$ $a, b, c, \geq 0$	$\varepsilon_{11} = \varepsilon \to 0$ $\varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23} =$ $\varepsilon_{21} = \varepsilon_{32} = \varepsilon_{33} = 0$	$x_1 \sim W_G(x_1)\left[1 + \frac{\gamma_3^{(1)}}{3!} H_3(x_1) + \gamma_4^{(1)} H_4(x_1)\right]$ $x_2 \sim W_G(x_2)\left[1 + \frac{\gamma_3^{(2)}}{3!} H_3(x_2) + \gamma_4^{(2)} H_4(x_2)\right]$ $x_3 \sim W_G(\cdot)$ with $R_{33} < 1$	Normalized dates, $\gamma_3^{(1)} \sim 0{,}2$ $\gamma_4^{(1)} \sim 0{,}6$ $\gamma_3^{(2)} \sim 0{,}2$ $\gamma_4^{(2)} \sim 0{,}6$

Table 1. Strange Attractors and its Properties.

Next, when SNR is low, then the influence of the second summand in SKE (4) at $W_{PS}(\boldsymbol{x},t)$ is low as well, and for the first approximation it is possible to assume that the marginal a-posteriori PDF's are close to their a-priori shapes. Therefore, it is feasible that EKF algorithms will be rather adequate for both high and low SNR scenarios for Lorenz and Rössler attractors, but not for Chua attractor.

Now, let us consider Chua attractor with the Integral (Global) Approximation for the a-posteriori PDF, assuming (Table 1) that first component has a symmetric $W_{PS}(\boldsymbol{x}_1,t)$. Supposing that $\{\varphi_i(\boldsymbol{x}_i)\}^K$ are polynomials of Hermite and $K = 4$, from (8) it follows:

$$W_{PS}(x_1,t) = Ce^{\{\alpha_1(t)H_1(x_1)+\alpha_2(t)H_2(x_1)+\alpha_3(t)H_3(x_1)+\alpha_4(t)H_4(x_1)\}} \quad (9)$$

With the help of definition of the Hermite polynomials, [7] one can get for (9):

$$W_{PS}(x_1,t) = Ce^{[-\alpha_2(t)-3\alpha_4(t)]}e^{\{Ax+Bx^2+Cx^3+Dx^4\}} \quad (10)$$

where $A=\alpha_1(t)-3\alpha_3(t); B=\alpha_2(t)-6\alpha_4(t); C=\alpha_3(t); D=\alpha_4(t)$

As $\{\boldsymbol{\alpha}_i(t)\}^4$ are sufficient statistics for $W_{PS}(\boldsymbol{x}_1,t)$ and invoking the symmetry and normalization conditions for a-posteriori PDF, one can get:

$$A = C = 0, C_1(\alpha_1,\alpha_4) = C\cdot e^{\{-\alpha_2(t-3\alpha_4(t)\}}$$
$$W_{PS}(x_1,t) = C_1(\alpha_2,\alpha_4)\cdot e^{\{Bx^2-Dx^4\}} \quad (11)$$

It is worth to mention that for the case of low SNR (11) coincides with the a-priori PDF $W_{PR}(\boldsymbol{x}_1,t)$ for Chua attractor (Table 1). Now, from (7) it follows:

$$\langle f(x_1)\,\varphi_i'(x_1)\rangle + \frac{\varepsilon}{2}\langle\varphi_i''(x_1)\rangle + \langle h_i(x_1)F(t,x_1)\rangle = 0 \quad (12)$$

where $i = 2, 4$.

Statistically equivalent SDE-1 with PDF (11) can be found from [7] (chapter 7) with

$$f(x_1) = \varepsilon\left(Bx_1 - 2Dx_1^3\right)$$

Then, for $i = 2$, one get ($\varepsilon \to 0$) a following equation:

$$2\varepsilon\left[B\langle x_1^2\rangle - 2D\langle x_1^4\rangle\right] = \frac{y^2(t)}{N_0}\left[\langle x_1^2\rangle - 1\right] \quad (13)$$

$$\text{where } \langle x_1^{2m}\rangle = \frac{\Gamma\left(m+\frac{1}{2}\right)D_{-m-\frac{1}{2}}(-\delta)}{\sqrt{\pi}D_{-\frac{1}{2}}(-\delta)(2D)^{\frac{m}{2}}} \quad (14)$$

$m= 1,2,\ldots$; $D(\cdot)$ is function of parabolic cylinder $\delta = \frac{B}{\sqrt{2D}}$.

Analogically for i = 4 ($\varepsilon \to 0$), it yields:

$$\varepsilon\left\{4\langle x^4\rangle\left(B-6D-12B\langle x^2\rangle-8D\langle x^6\rangle\right)\right\}+$$
$$+6\varepsilon\langle x^2\rangle = \frac{y^2(t)}{N_0}\left[\langle x^4\rangle-6\langle x^2\rangle+3\right]+$$
$$+\frac{1}{N_0}\left[\langle x^6\rangle-6\langle x^4\rangle+3\langle x^2\rangle\right] \quad (15)$$

Assuming that in (13) ÷ (15) $y^2(t)$ tends to its stationary value $\bar{y}^2(t)$ while $t \to \propto$, and substituting into (13) ÷ (15), one can get nonlinear algebraic equations for stationary parameters $\bar{\boldsymbol{\alpha}}_2, \bar{\boldsymbol{\alpha}}_4$, which are obviously related to the variance and fourth moment (cumulant) of $W_{PS}(\boldsymbol{x}_1)$.

Therefore, $\bar{\alpha}_2$ can be used as a measure of the filtering accuracy, being calculated with influence of the fourth a-posteriori moment (cumulant). The similar approach was already used at [5].

The asymptotical filtering algorithm for $\boldsymbol{x}_{I^\Delta}(t) = \boldsymbol{x}(t)$ of Chua attractor in discrete time can be represented using [9], [8]:

$$\hat{x}_{i+j} = \hat{x}_j + T_0 f\left(\hat{x}_j\right) + \\ + \sigma^2_{\varepsilon_j} \frac{d}{dx_{j+1}} \ln W_{PS}\left[\left(y_{j+1} - x_{j+1}\right)\right]_{x_{j+1}=\hat{x}} \qquad (16)$$

Where T_0 is a sampling interval, σ^2_ε is a-posteriori variance.

This a-posteriori variance can be calculated through $\bar{\alpha}_2$ and $\bar{\alpha}_4$ (see above), but also might be found from the following equation:

$$\hat{\sigma}^2_{\varepsilon_{j+1}} = \hat{\sigma}^2_{\varepsilon_j} + 2\hat{\sigma}^2_{\varepsilon_j} f'\left(\hat{x}_j\right) T_0 + \\ + \hat{\sigma}^4_{\varepsilon_j} \frac{\partial^2}{\partial x^2_{j+1}} \ln W_{PS}\left[\left(y_{j+1} - x_{j+1}\right)\right]_{x_{j+1}=\hat{x}_j} \qquad (17)$$

If the SNR is low and $n_0(t)$ is a Gaussian additive white noise, then applying Taylor series expansion for the $\mathrm{In} W_{PS}(\cdot)$, with this asymptotic one can get:

$$\hat{x}_{j+1} = \hat{x}_j + T_0 f\left(\hat{x}_j\right) + \frac{\sigma^2_{\varepsilon_j}}{\sigma^2_n} n_{j+1}$$

$$\hat{\sigma}_{\varepsilon_{j+1}} = \hat{\sigma}^2_{\varepsilon_j} + 2\hat{\sigma}^2_{\varepsilon_j} \frac{\hat{f}(\hat{x}_j)T_0}{\varepsilon \hat{\sigma}^2_{\varepsilon_j}} (\hat{x}) T_0 \qquad (18)$$

In stationary conditions is:

$$\bar{\hat{\sigma}}^2_{\varepsilon_j} = \frac{\varepsilon T_0}{2 f'(\hat{x}_j)} = \frac{T_0}{2(B - 6D\hat{x}^2)} \qquad (19)$$

It can be seen from (19) that accuracy of the filtering depends on absolute value of $\hat{x}$.

This interesting issue follows from the dependence of $\bar{\hat{\sigma}}^2_\varepsilon$ on the derivative of the nonlinear drift $f'(\hat{x}_j)$. Moreover, as it was mentioned before, the value of $\bar{\hat{\sigma}}^2_{\varepsilon_j}$ can be additionally reduced by application of $\bar{\alpha}_4$. It was shown at [5] that if SNR is less than one, then $\bar{\hat{\sigma}}^2_{\varepsilon_j}$ can be reduced by two times if applying the fourth cumulant.

4. CONCLUSIONS

In this paper were presented two different approaches for the approximate nonlinear filtering for the low SNR scenarios: Integral (Global) Approximations for the a-posteriori PDF and asymptotic approach. It was shown that finally it is reasonable to aggregate both of them in order to improve the accuracy of filtering.

REFERENCES

[1] V. S. Anischenko et ál. «Statistical properties of dynamical chaos”. *Physics–Uspekhi*, 48(2) (2005): 151-166.

[2] I. Arasaratnan, S. Haykin, and R. Elliot. “Discrete-time non-linear filtering algorithms using Gauss-Hermite quadrature”. *Proceedings of the IEEE*, 96(5) (2007, may): 953-977.

[3] A. Jazwinski. *Stochastic Processing and filtering theory*. N.Y. Academic, 1970.

[4] V. Kontorovich, and Z. Lovtchikova. “Cumulant analysis of strange attractors. Theory and applications”, in *Recent advances in non-linear Dynamics and Synchronization (NDS-1)-Theory and applications*, chap. 5. 2009 (in publishing).

[5] V. Kontorovich. “Non-linear filtering for Markov stochastic processes using high-order statistics (HOS) approach”. *Non-linear Analysis. Theory, Methods and Applications*, 30(5) (1997): 3165-3170.

[6] H. Kushner. “Dynamic equations for optimal non-linear filtering”. *J. Differ. Eq.*, 3 (1971): 179-190.

[7] S. Primak, V. Kontorovich, and V. Lyandres. *Stochastic methods and their applications to communications: Stochastic Differential Equations Approach*. John Wiley & Sons, 2004.

[8] V. Pugachev, and I. Sinitsyn. *Stochastic Differential Systems. Analysis and Filtering*. John Wiley & Sons, 1987.

[9] R. Stratonovich. *Topics of the theory of random noise*, vols. 1 and 2. Gordon and Breach, 1963.

[10] V. Tikhonov, and V. Kharisov. *Statistical Analysis and Synthesis of Radio Devices and Radio Systems*. Radio i Sviaz, 1991.