

# Extension of the Global Optimization Using Multi-unit Extremum Seeking Control for Noisy Scalar Systems

F. Esmailzadeh Azar, M. Perrier

**Abstract**—An efficient global optimization method based on multi-unit extremum seeking has been proposed recently for noise-free nonlinear continuous static, scalar systems. Herein, a monotonically decreasing offset is introduced between the inputs of two identical units and the estimated gradient by finite difference is controlled to zero by an integrator. This paper is the extension of the mentioned algorithm to globally optimize scalar objective functions with the presence of noise. The algorithm is applied on several noisy global optimization test problems to show the capability of this methodology.

## I. INTRODUCTION

The global optimization of a nonlinear function in the presence of noise is a challenging though an interesting problem. The effect of stochastic noise can be demonstrated in an optimization problem in different ways. For example, the imperfect convergence of a numerical simulation can severely affect the calculation of the objective function [11]. Noise can produce bias and topographical inaccuracy in the objective function evaluation, affecting the accuracy and convergence of the optimization problem. The continuous increase of computational power during recent years has renewed the interest in global optimization of noisy functions.

Stochastic approximation methods are usually applied for local optimization of noisy functions. For global optimization of such problems a combination of a stochastic approximation method with a random multi-start algorithm may be employed [15]. A recursive algorithm which uses a switching method between the stochastic approximation and the random search is proposed in [7]; a good switching rule and an efficient method to reduce the noise effect is the key point when using such a combination. A stochastic globally optimizing neural network called a “noise annealing neural network” is another attempt to solve the nonlinear constrained global optimization problems in the presence of noise [2]. This method is an extension of the canonical

Farhad Esmailzadeh Azar was a Ph.D. Graduate and a Research Assistant at École Polytechnique Montréal, Département of Chemical Engineering, Montréal, Canada, H3C 3A7. He is now a R&D Senior Control Systems Engineer working for Covidien (Puritan Bennett), 2101 Faraday Avenue, Carlsbad, California 92008, USA. E-mail: farhad.esmailzadeh-azar@polymtl.ca

Michel Perrier is Professor and Director of the Department of Chemical Engineering, École Polytechnique Montréal, Canada, H3C 3A7. Email: Michel.perrier@polymtl.ca

nonlinear programming neural network. Another alternative is based on a statistical model of an objective function. In this framework, the objective function of a mono-variable optimization problem is modeled as a standard probability model (Wiener process) which is corrupted with independent Gaussian noise [3]. Finally, a novel Monte-Carlo technique based on a black-box approach has been proposed by Ramani et al. [12]. This method is based on optimizing the parameters of an arbitrary denoising algorithm by assessing the true mean-squared-error (MSE) from the measured noisy data.

Extremum-seeking methods use different gradient estimation strategies for their convergence [1,8,10]. These schemes typically find the local optimum by controlling the gradient to zero. The multi-unit optimization method is a real-time extremum seeking control strategy where the gradient is computed based on the finite difference between a set of parallel units which operate with input values differing by a constant, pre-fixed offset [13]. The estimated gradient converges to zero by an integrator. On the other hand, a global extremum-seeking method for a restricted class of noise-free scalar maps has been proposed by Tan et al. [14] where a sinusoidal temporal perturbation with a pre-fixed amplitude is added to the input and its amplitude is reduced to zero. Here, the gradient is obtained as a correlation between the inputs and the outputs. As an alternative Azar et al. [6] proposed a deterministic global optimization method for a general class of noise-free scalar systems based on multi-unit extremum seeking control where the pre-fixed offset between the inputs of two identical units is reduced to zero. Simultaneously, the gradient evaluated using the finite difference between two units operating with decreasing offset is controlled to zero by an integrator. The units are assumed to be identical meaning that there is no difference between the operating curves of their static maps. The comparison of the global optimization of scalar systems by multi-unit extremum seeking using genetic algorithms, simulated annealing and DIRECT algorithms showed better results for the multi-unit method in terms of fewer numbers of function evaluations.

DIRECT (Dividing RECTangles) is a deterministic global optimization algorithm based on a space-partitioning scheme for bound-constrained problems. A simple adaptation of this algorithm for noisy functions has been developed by Deng et al. [4]. In this method, multiple function evaluations are replicated at each point and an average is taken to reduce functional uncertainty. Bayesian sample information is employed to determine appropriate numbers of replications.

This paper discusses the extension of the global optimization of scalar systems by multi-unit extremum seeking control to handle noisy objective functions. In the noise-free global optimization by multi-units, the offset between the inputs and the estimated gradient can be decreased to zero linearly, exponentially or in any other manner [6]. In the proposed algorithm, it is shown that the global optimum of a noisy function can be obtained if the gains of the adaptation laws for the offset and estimated gradient are changed adaptively. The noise considered in this study is Gaussian noise. The key contribution lies in formulating the adaptation law of the offset between two units to increase the replication of function evaluations adaptively where the values of their objective functions are sufficiently close. Accordingly, an average is taken to reduce functional uncertainty produced by noise. Furthermore, the algorithm is automatically relaxed when the values of the objective function of two units are sufficiently far. Therefore, there is no need to replicate multiple function evaluations even if the impact of noise is significant. The proposed method is suitable for mono-variable black-box (i.e. simulation-based) optimization problems when the function is observed with noise. Black-box models describe process behavior when closed-form equations are not available. This method only requires the sampled data from the objective function assuming that it is corrupted by Gaussian noise. It does not need any information about the functional form of the noise-free objective function.

The outline of the paper is as follows. Section II provides an overview of global optimization using multi-unit extremum seeking controllers for the unconstrained noise-free scalar systems. Section III presents the extension of the method to optimization of scalar systems in the presence of noise. In section IV, the established algorithm is applied on a number of noisy global optimization problems and the effects of different parameters on the convergence of the method are outlined. The conclusions of the paper are presented in section V.

## II. GLOBAL OPTIMIZATION USING MULTI-UNIT EXTREMUM SEEKING CONTROL FOR NOISE-FREE SCALAR SYSTEMS

The global optimization using multi-units is an extremum seeking technique that estimates the gradient by the finite difference of the outputs of two identical units where the inputs differ by a decreasing pre-fixed offset  $\Delta$ . An integral controller then forces the gradient to zero converging to a local optimum of the objective function [13]. The extension of the local extremum seeking control using multi-units to global optimization of the noise-free static nonlinear and continuous scalar systems was developed by Azar et al. [6]. The main idea is to decrease the offset to zero monotonically. The schematic is presented in figure 1.

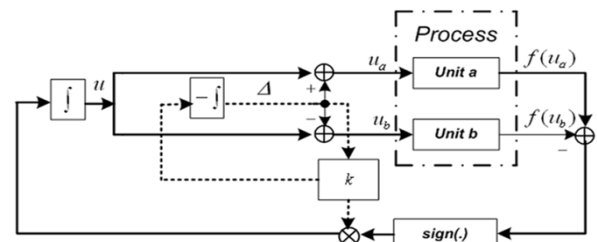


Figure 1. Global extremum-seeking control with multi-units (noise-free systems)

The update equations and adaptation laws for global optimization (maximization) of a scalar and continuous function  $f(u)$  are given by,

$$u_a = u + \Delta \quad , \quad u_b = u - \Delta \quad (1)$$

$$\dot{u} = k\Delta \text{sign}(f(u + \Delta) - f(u - \Delta)) \quad , \quad u(0) = u_0 \quad (2)$$

$$\dot{\Delta} = -k\Delta \quad , \quad \Delta(0) = \Delta_0 > 0 \quad (3)$$

Depending on the rate of convergence  $k$ , the offset  $\Delta$  decreases to zero monotonically.  $\Delta$  can be decreased to zero in any predefined fashion; the exponential adaptation has been considered here. However, the absolute value of its dynamics must be equal to the absolute value of the dynamics of  $u$ . This way, the dynamics of each unit is formulated in a way that at each instant, the movement of each unit is towards a better local operating point on the nonlinear map  $f(u)$ . The basic condition for this algorithm to converge to the global optimum is  $|u_0 - u^{**}| \leq \Delta_0$ .  $u_0$  is the initial input value and  $u^{**}$  is the global optimum of the nonlinear map. Since the location of  $u^{**}$  is not known a priori, the above condition will be satisfied by choosing an initial value large enough for  $\Delta_0$ . The proof of convergence for this algorithm has been provided using the mathematical contradiction formalism [6]. This algorithm was also extended to constrained problems where a switching control strategy is employed [5].

## III. CONSTRUCTION OF THE ALGORITHM FOR GLOBAL OPTIMIZATION OF NOISY SCALAR SYSTEMS

Consider the global optimization of a noisy scalar, static, non-convex continuous, nonlinear function as follows,

$$\begin{aligned} \max_{u \in R} \quad & y = E[f(u) + \xi(\omega)] = F^N \\ \text{st.} \quad & L \leq u \leq U \end{aligned} \quad (4)$$

$f(u)$  is an unknown noise-corrupted function and is considered as a black-box. The function  $y = f(u)$  may have multiple local maxima,  $u_m^*$ ,  $m = 1, 2, \dots, n$ , but a unique global maximum,  $u^{**}$ .  $\xi(\omega)$  is a random variable which affects the sampled output response of the function  $f(\cdot)$  [9].  $F^N$  is the notation for noisy functional value of  $f(u)$ .  $L$  and  $U$  are the lower and upper bounds for the input  $u$ , respectively.

### A. Schematic diagram

In this paper, the noise-free global optimization algorithm using multi-units is modified to achieve global optimization of noisy scalar functions. The schematic of the noise-corrupted extremum seeking control is presented in Figure 2. The proposed strategy is explained in the following sections.

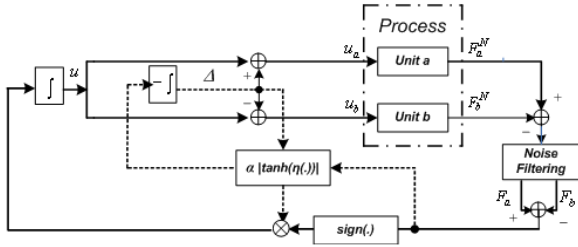


Figure 2. Global extremum-seeking control with multi-units (noisy systems)

### B. Adaptive offset value $\Delta$

The main contribution lies in the fact that the offset parameter between the two units is reduced to zero based on the tangent hyperbolic function. In the new algorithm, the adaptation laws are based on the same equations (2), (3), and the adaptation gain is given by,

$$k = \alpha |\tanh(\eta(F_a - F_b))| + \varepsilon \quad (5)$$

$F_a$  and  $F_b$  are the noise-filtered functional values of  $F_a^N$  and  $F_b^N$  respectively and will be introduced in the next section.  $k$  determines the rate of convergence of  $\Delta$  and  $\alpha$  is a coefficient which effects this rate of convergence. The larger the adaptation gain  $k$ , the smaller will be the time constant of the system.  $\alpha$  and  $\eta$  are two prefixed tuning parameters which affect the precision of the algorithm. The larger the value of  $\alpha$ , the faster will be the integration. However, regarding the simplicity of the algorithm  $\alpha = 1$  is chosen in the rest of this paper.  $\varepsilon$  is a certain small positive value. The smaller the difference between  $F_a$  and  $F_b$ , the smaller is their Tangent hyperbolic, and  $k$  more closely approximates the Sign function. The Sign function belongs to the tangent hyperbolic family: from Figure 3, the Sign function is just a special case of the tangent hyperbolic function ( $\eta = \infty$ ). Depending on the ‘‘Lipschitzian’’ characteristic of the static maps, the adaptation laws may introduce stiffness in the integration process. The stiffer the integration, the more time would be required for convergence and the more samples would be required. An alternative solution to overcome this stiffness is to use another tuning parameter  $\eta$ .

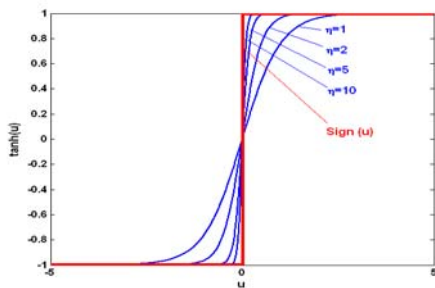


Figure 3. Tangent hyperbolic function

The lower the value of  $\eta$ , the faster will be the integration. However, a low value of  $\eta$  might lead to a situation where

the global optimum is missed and the algorithm converges to a local optimum. It will be demonstrated that the number of noisy function evaluations and the convergence speed of the system depend on the relaxation of the adaptation law for the integration of offset value  $\Delta$ . The numbers of replications of function evaluation depends on the stiffness of the objective function, the tuning parameters of the algorithm and the impact of the noise.

### C. Noise filtering

The noise-filtered functional values of  $F_a$  and  $F_b$  are described as follows,

$$F_{aj} = \begin{cases} \frac{F_{aj}^N + F_{a(j-1)}}{2} & \text{if } |F_{a(j-1)} - F_{b(j-1)}| \leq \varepsilon \\ F_{aj}^N & \text{otherwise} \end{cases} \quad (6)$$

$$F_{bj} = \begin{cases} \frac{F_{bj}^N + F_{b(j-1)}}{2} & \text{if } |F_{a(j-1)} - F_{b(j-1)}| \leq \varepsilon \\ F_{bj}^N & \text{otherwise} \end{cases} \quad (7)$$

$j$  is an indicator that shows the number of function evaluations or sampling in the integration of the adaptation laws. In other words,  $F_{aj}^N$  and  $F_{bj}^N$  are the sequential noisy function values ( $j=1,2,\dots$ ).  $F_{a0} = F_{a1}^N$ ,  $F_{b0} = F_{b1}^N$  and  $\varepsilon$  is a certain small positive value. After each sampling of certain parts of the nonlinear curve, its average with the previous samples is computed. Therefore, a better approximation of the real functional value is considered by averaging the outputs. In this way the impact of the noise on the objective function is reduced. The key point here is that the averaging of the noisy samples is not a continuous part of the algorithm at every given sample of the nonlinear curve. However, the averaging of the samples is automatically done only from certain functional values where the outputs of the two units are very close. This method of averaging prevents an excessive number of function evaluations in the algorithm. Note that the location of the mentioned segments can vary depending on the initial value of the offset parameter  $\Delta_0$ . The offset value  $\Delta$  evolves depending on the difference value between two noise-filtered functional values. Similar to the noise-free scenario, the length of step taken by the inputs is determined not by the gradient but by the variation of the offset parameter ( $\Delta$ ). The key difference in optimization of noisy functions is that the offset parameter does not decrease just monotonically exponentially. In the proposed algorithm,  $\Delta$  varies adaptively based on the tangent hyperbolic of the difference value between two noise-filtered outputs. If this difference is greater than  $\varepsilon$ , it would not be worth taking further function evaluations. In this case, the same noisy functional value is used in the adaptation gain (5). Therefore, the dynamics of the unit with a better noisy function value remains zero until the other one arrives at a better operating point. The offset value between the inputs of two units keeps shrinking continuously. As soon as the difference between the noise-filtered objective functions becomes smaller than  $\varepsilon$ , the convergence rate of the offset value  $\Delta$  slows down.

This rate of slowdown is controlled by varying the gain parameter  $k$ . When the difference between  $F_a$  and  $F_b$  becomes smaller than  $\varepsilon$ ,  $\Delta$  decreases slower, and thereby more function values are sampled from the two units and the more averaging occurs using these samples. This reduces the uncertainty of the noisy function.

#### D. Convergence to the global optimum

The proposed algorithm follows the spirit of noise-free global optimization. With the modifications to the noise-free algorithm, it can be shown that no preconditions on the nonlinear noisy function are required to achieve the global maximum except for continuity. Similarly, the initial condition and the initial value of the offset parameter should be chosen such that the global maximum lies in the initial interval  $[u_b(0), u_a(0)]$ . Note that the proof of convergence to the unconstrained noisy global optimum is based on the fact that the global maximum always lies within the interval  $[u_b(t), u_a(t)]$  for all  $t$ . The proof of the algorithm follows the argument of the scalar case in [6]. Although the existence of noise in the system makes the context of the algorithm different, however, the proof is quite similar. In noise-free algorithm, it can be shown that the global optimum cannot be removed from the interval  $[u_b(t), u_a(t)]$ . As the offset interval shrinks to zero, since the feasible global optimum is trapped in the interval, both units converge to the global optimum. For noisy systems, because of presence of noise, it generally cannot be guaranteed that the above scheme is indeed global. However, in the limiting case, it can be shown that this algorithm is capable of avoiding the local noisy optima and converging to a small neighborhood of the global optimum.

#### IV. ILLUSTRATIVE EXAMPLES

**Example 1:** Consider the following nonlinear static map (shown in Figure 4) with a unique global maximum at  $u^{**} = 1.68$  and several local optima at other points [6].

$$f(u) = -3u^4 + 64\sin^2(u^3) + 12u^2 + 4u - 80 \quad (8)$$

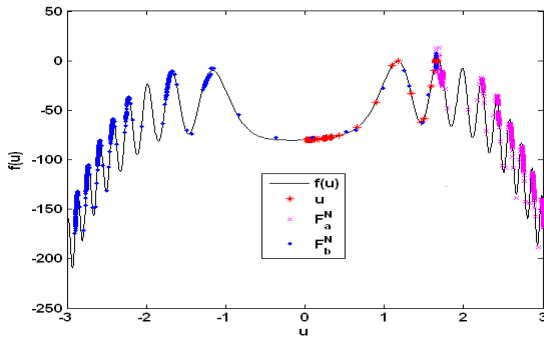


Figure 4. Static noisy nonlinear map for example 1

Gaussian noise with zero mean value and variance of 40 and 10 affect the outputs of units  $a$  and  $b$  respectively. The proposed global optimization algorithm is applied to optimize this nonlinear noisy system. The initial inputs  $u_0 = 0.1$  and  $\Delta_0 = 3$  were considered such that the global

maximum along with several other local maxima lie in the interval. The parameters used were  $\alpha=1$ ,  $\eta=1$ ,  $\varepsilon=0.01$ . The time evolution of the input  $u$  and  $\Delta$  are shown in Figure 5. The offset parameter  $\Delta$  evolves depending on the difference between functional values  $F_a$  and  $F_b$ . The slower  $\Delta$  decreases, the more function evaluations are required by the algorithm. It is interesting to see that most of the function evaluations of the nonlinear noisy curve occur when  $F_a - F_b$  is smaller than  $\varepsilon$  (Figure 4). These are the areas where the noise filtering is more concentrated by the algorithm. This is an advantage as long as the algorithm relaxes at unimportant segments of the nonlinear curve.

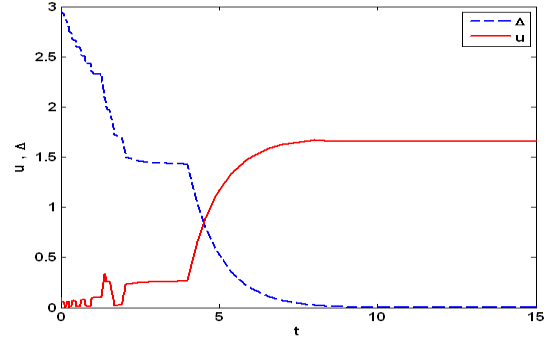


Figure 5. Evolution of the inputs for example 1

The evolution of the inputs to the two units,  $u_a$  and  $u_b$  are shown in Figure 6. Similar to the noise-free algorithm [6], most of the time only one input evolves and the other, which has a higher objective function value, is maintained constant. By comparing Figure 5 and 6, it can be seen that in time intervals where  $\Delta$  remains almost constant (such as  $2 < t < 4$ ),  $u_a$  and  $u_b$  remain constant and more function evaluations are required from the noisy system. These are the intervals where the noise filtering or averaging algorithm is more active. Note that the integration time ( $t$ ) is the time horizon over which the algorithm is integrated.

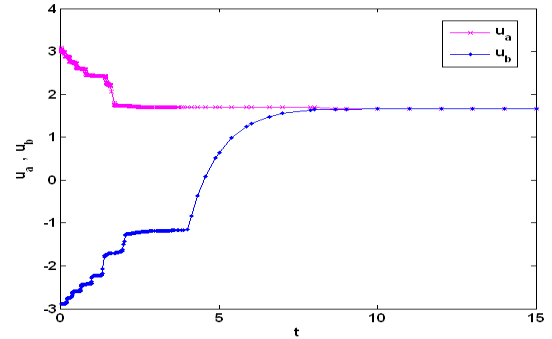


Figure 6. Evolution of  $u_a$  and  $u_b$

**Example 2:** Consider the following 6<sup>th</sup>-order polynomial [14] which has 3 maxima at  $u = -0.8985$ ,  $u = 0.5$ ,  $u = 0.8951$ , with  $u = -0.8985$  being the global maximum. Note that the maxima are very close in this example.

$$f(u) = -u^6 + \frac{1}{10}u^5 + \frac{623}{400}u^4 - \frac{659}{4000}u^3 - \frac{11287}{20000}u^2 + \frac{259}{4000}u + \frac{637}{20000} \quad (9)$$

Similar to example 1, Gaussian noise with zero mean value and variance of 0.01 and 0.02 affect the outputs of units  $a$  and  $b$ , respectively. The key inequality condition  $|u_0 - u^{**}| \leq \Delta_0$  is satisfied by choosing  $u_0 = 0.01$  and  $\Delta_0 = 1$ . The initial condition of  $u$  is chosen intentionally to be closer to the local maxima and further away from the global one. The other parameters used were  $\alpha=1$ ,  $\eta = 10$ ,  $\varepsilon = 0.01$  which results in the proposed algorithm converging to the global maximum at  $u^{**} = -0.8985$ . The static map is depicted in Figure 7.

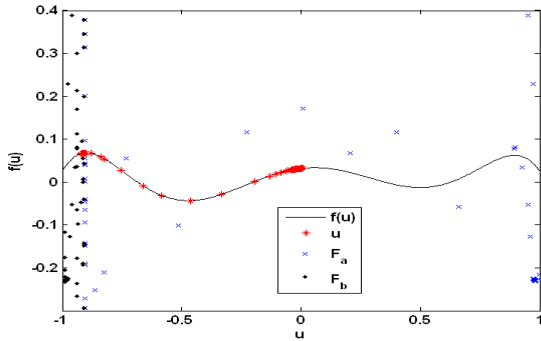


Figure 7. Static nonlinear map of example 2

Figure 8 shows the evolution of the offset value  $\Delta$  and the gain value  $k$ . It can be seen that in the time interval ( $8 < t < 18$ ), the rate of change of  $\Delta$  increases suddenly because of a rapid change in its adaptive gain  $k$ . This is attributed to the relaxation of the algorithm in this interval. On the other hand, in the time intervals  $t < 8$  and  $t > 18$  the rate of change of the offset parameter slows down. As a result, more function evaluations are required from the noisy system when the algorithm converges to the global optimum (Fig 7).

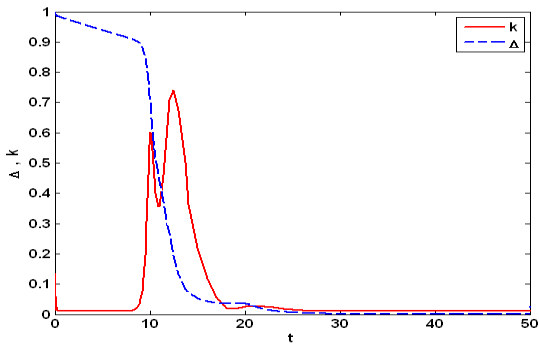


Figure 8. Evolution of  $\Delta$  and its adaptive gain  $k$  for example 2

**Example 3:** Consider the sum of exponentials in (10) [14]. This example has two local maxima at  $u = 0$  and  $u = 15$ . The global maximum is at  $u = 15$ .

$$f(u) = e^{\frac{1}{1+0.2u^2}} + e^{\frac{1}{1+5(u-15)^2}} \quad (10)$$

The static map is given in Figure 9. The Gaussian noise which affects the outputs of units  $a$  and  $b$  has zero mean value and the variance of 0.02 and 0.03, respectively. The proposed noisy extremum seeking algorithm successfully converges to the global maximum with the following parameters:  $\alpha=1$ ,  $\eta = 1$ ,  $\varepsilon = 0.1$ ,  $u_0 = 3$  and  $\Delta_0 = 25$ .

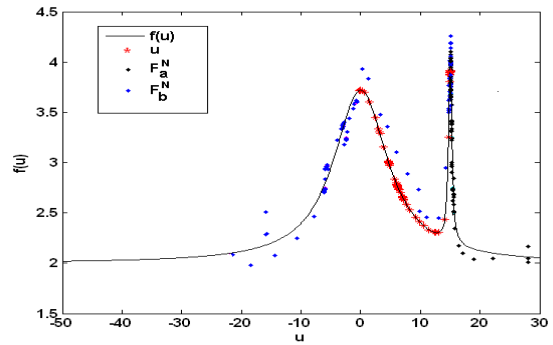


Figure 9. Static nonlinear map of example 3

In this example, both the noise-free algorithm (presented in section II) and the proposed algorithm were applied. The evolution of the offset parameter  $\Delta$  using the noise-free algorithm is always monotonically exponential whereas using the noisy algorithm its rate of change varies depending on the adaptive gain  $k$  (Figure 10).

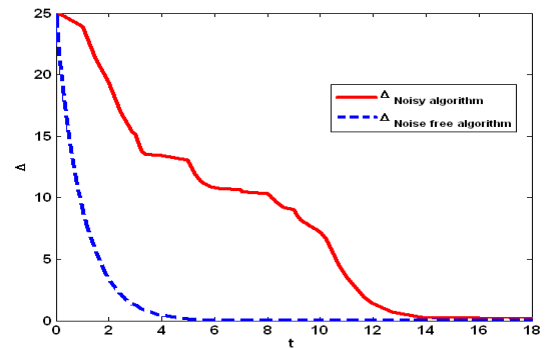


Figure 10. Evolution of  $\Delta$  for noisy and noise-free algorithms in example 3

The evolution of the unit inputs  $u_a$  and  $u_b$  for both algorithms is depicted in Figure 11. It can be seen that the noisy algorithm (red line) converges to the global optimum of the system at  $u^{**} = 15$  whereas the noise-free algorithm (blue dotted line) converges to another point. Both algorithms are independent from the input initial conditions.

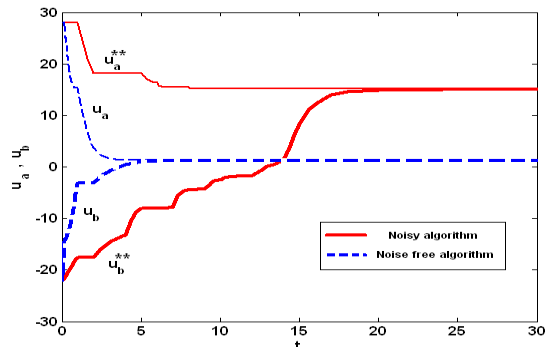


Figure 11. Evolution of  $u_a^{**}$ ,  $u_b^{**}$  for noisy algorithm and  $u_a$ ,  $u_b$  for noise-free algorithm in example 3

**Example 4:** The last example is a nonlinear system which has several equal valued and symmetric local optima [6]. This example has a global maximum that occurs at  $u^{**} = 0$  where its derivative is discontinuous.

$$f(u) = -|u + 3\sin(u)| \quad (11)$$

The mean value of the Gaussian noise which affects the outputs of the two units was zero with variances of 10 and 20. The optimization and control parameters were:  $\alpha=1$ ,  $\eta=1$ ,  $\varepsilon=0.1$ ,  $u_0=-10$  and  $\Delta_0=30$ . The extended noisy algorithm converges to the global optimum despite of discontinuous derivatives. It can be seen from Figure 12 that the most number of function evaluations and sample averaging are required at the global optimum by the algorithm.

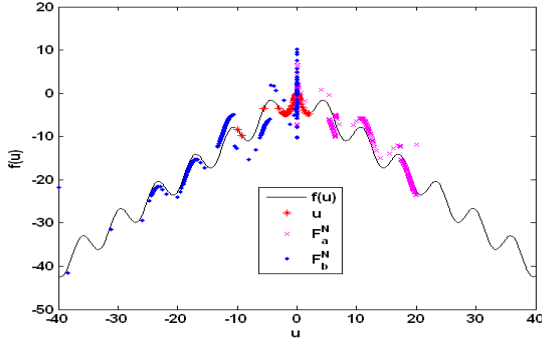


Figure 12. Static nonlinear map of example 4

An appropriate choice of  $\eta$  is crucial in converging to the global optimum. If a low value of  $\eta$  is chosen, i.e.  $\eta=0.1$ , the algorithm is misled and converges to an optimum which is not global ( $u=0.7947$ ). Using a large value  $\eta=100$  leads to the larger number of function evaluations ( $N=168$ ) in contrast to ( $N=102$ ) in the former case with  $\eta=1$ . Increase in the number of function evaluations can be attributed to the increase in stiffness of the integration of the offset parameter  $\Delta$ . If too small a value of  $\eta$  is chosen, then the global optimum will be missed. As a result, a low value of  $\eta$  indicates that the adaptation gain  $k$  is less sensitive to the difference of functional values between the two units. Thus, the value of  $\eta$  should be chosen as a compromise between the sensitivity of the algorithm to its stiffness and integration time. Figure 13 shows the evolution of  $f(u)$ ,  $F_a^N$  and  $F_a$  for example 4 in a given time interval. It can be seen that most of the time, the noise-filtered function  $F_a$  (blue line) is closer to the real function  $f(u)$  compared to the noisy function evaluation  $F_a^N$  (red dotted line). Here, the variance of the Gaussian noise has been decreased 100 times (i.e. 0.1 and 0.2) in order to show this effect more clearly.

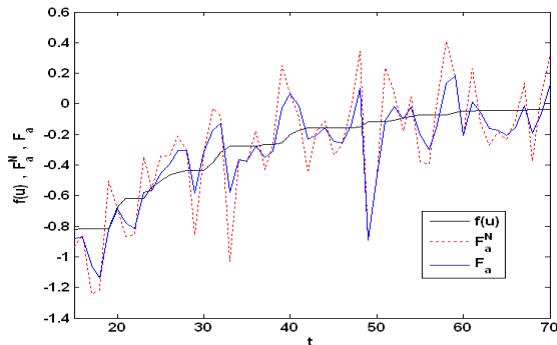


Figure 13. Evolution of  $f(u)$  and  $F_a^N$  and  $F_a$  in example 4

## V. CONCLUSION

The model-free, unconstrained global optimization of nonlinear static and scalar systems using the multi-unit method was extended to noisy systems where the offset parameter between the inputs of the two units was systematically reduced to zero. The main contribution of the proposed strategy was to change the rate of change of the offset parameter adaptively instead of monotonically in the noise-free algorithm. Furthermore, a noise filtering algorithm was used in order to reduce the impact of the noise and functional uncertainty. The next steps of this research framework are to develop the proposed algorithm to multi-input and constrained noisy optimization problems.

## REFERENCES

- [1] K.B. Ariyur and M. Krstic, "Real-time Optimization by Extremum-Seeking Control," John Wiley, New York, 2003.
- [2] J.Z. Biró, T. Koronkai, and Trón, "A Noise Annealing Neural Network for Global Optimization," *IEEE International Conference on Neural Networks ICNN 96*, Washington, USA, 1996, pp. 513-518.
- [3] J.M. Calvin and A. Žilinskas, "One Dimensional Global Optimization for Observations with Noise," *Computers and Mathematics with Applications*, vol.50, pp. 157-169, 2005.
- [4] G. Deng and M.C. Ferris, "Extension of the Direct Optimization Algorithm for Noisy Functions," *Proceedings of the Winter Simulation Conference, USA, 2007*, pp. 497-504.
- [5] F. Esmailzadeh Azar, M. Perrier and B. Srinivasan, "Real-time Global Optimization of Constrained Non-linear Systems using Multiple Unit Optimization Scheme," *8th World Congress of Chemical Engineering*, Montreal, Canada, August 23-27, 2009.
- [6] F. Esmailzadeh Azar, M. Perrier and B. Srinivasan, "A Global Optimization Method Based on Multi-unit Extremum-seeking for Scalar Nonlinear Systems," *Journal of Computers and Chemical Engineering*, vol. 35, pp. 456-463, 2011.
- [7] H.T. Fang and H.F. Chen, "Almost Surely Convergent Global Optimization Algorithm Using Noise-Corrupted Observations," *Journal of Optimization Theory and Applications*, vol. 104, No. 2, pp. 343-376, 2000.
- [8] M. Guay, D. Dochain and M. Perrier, "Adaptive extremum seeking control of continuous stirred tank bioreactors with unknown growth kinetics," *Automatica*, vol. 40, pp. 881-888, 2004.
- [9] R.V. Hogg and A.T. Craig, "Introduction to Mathematical Statistics", Section 3.3., 4th edition. Macmillan, New York, 1978.
- [10] M. Leblanc, "Sur l'Électrification des Chemins de Fer au Moyen de Courants Alternatifs de Fréquence Élevée," *Revue générale de l'électricité*, vol. 12, no. 8, pp. 275-277, 1922.
- [11] V.J. Romero, "Efficient Global Optimization Under Conditions of Noise and Uncertainty - A Multi-Model Multi-Grid Windowing Approach," *3rd World Conference of Structural and Multidisciplinary Optimization*; Buffalo, NY, 1999.
- [12] S. Ramani, T. Blu, and M. Unser, "Blind Optimization of Algorithms for Signal Denoising by Monte-Carlo SURE," *IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*, Las Vegas, USA, 2008, pp. 905-908.
- [13] B. Srinivasan, "Real-time Optimization of Dynamic Systems using Multiple Units," *International Journal of Robust and Nonlinear Control*, vol. 17, pp. 1183-1193, 2007.
- [14] Y. Tan, D. Nešić, I.M.Y. Mareels and A. Astolfi, "On Global Extremum Seeking in the Presence of Local Extrema," *45th IEEE Conference on Decision and Control*, San Diego, CA, USA, 2006.
- [15] A. Žilinskas and D. Kučinskas, "Implementation and Testing of an Algorithm for Global Optimizations with Noise," *International Conference on Computer Systems and Technologies - CompSysTech*, 2004, pp. 1-6.