Extremum seeking based on a Hammerstein-Wiener representation

Christian G. Feudjio Letchindjio* Jean-Sebastien Deschenes** Laurent Dewasme* Alain Vande Wouwer*

* Automatic Control Laboratory, University of Mons, Belgium (e-mail: {ChristianGabin.FeudjioLetchindjio, Laurent.Dewasme, Alain.VandeWouwer}@umons.ac.be) ** Universite du Quebec a Rimouski, Canada (e-mail: Jean – sebastien_Deschenes@uqar.ca)

Abstract: This study is concerned with the development of an extremum seeking (ES) strategy based on recursive least square (RLS) for on-line estimation, and a regression model in the form of a Hammerstein-Wiener model. RLS usually provides a faster convergence than the classical bank of filter estimators, and the consideration of process dynamics allows to take account for the phase-shift and attenuation occurring when increasing the frequency of the dither signal. The resulting ES scheme achieves very significant improvement in convergence speed, as illustrated with a numerical example, and a more realistic application to micro-algae cultures in a photobioreactor in simulation.

Keywords: Real-time optimization, recursive least squares, process control, biotechnology, micro-algae.

1. INTRODUCTION

Extremum-Seeking (ES) control is a Real-Time Optimization (RTO) technique that aims at driving an objective function to its optimum (maximum or minimum) by adjusting on-line the process input. In response to an excitation provided by a dither signal, an estimator is used to extract gradient information from the output signal. This estimation is then driven to zero (in average) thanks to an optimizer. ES has been the subject of intense research in the last decades (see for instance (Dochain et al., 2011) for a survey), and of a few real-life experimental studies (Wang et al., 2000; Leyva et al., 2006; Deschênes et al., 2012).

One of the earlier forms of ES (Ariyur and Krstic, 2003) is based on a Bank Of Filters (BOF), which is designed in order to extract information about the objective function gradient. Fig 1 shows the ES scheme described by the following equations :

$$y = f(\hat{u} + Asin(\omega t)) \tag{1a}$$

$$\hat{u} = k\xi \tag{1b}$$

$$\xi = (y - \eta) Asin(\omega t)$$
(1c)

$$\dot{\eta} = -\omega_h \eta + \omega_h y \tag{1d}$$

where

- $u = \hat{u} + A \sin(\omega t)$: the input signal applied to the process, including a sinusoidal dither signal,
- \overline{y} : the measurable objective function,
- $y \eta$: the filtered signal at the output of the high-pass filter,
- δ : the demodulated signal,

- $\hat{\xi} \approx \frac{1}{k} \frac{\partial \hat{u}}{\partial t}$: a proportional gradient estimation, obtained as the output of a low-pass filter,
- \hat{u} : the estimated input.



Fig. 1. ES with BOF.

The cut-off pulsation of the low pass filter ω_l and high-pass filter ω_h should be lower than the dither signal pulsation ω . Furthermore the dither signal pulsation ω should be small with respect to the process dynamics and its amplitude sufficient to ensure catching gradient information. This illustrates the three-scale time separation (Ariyur and Krstic, 2003) where :

- the process has the fastest dynamics, and is approximated as a static map,
- the dither signal has intermediate dynamics,
- the gradient estimator has the slowest dynamics.

A suitable choice of the algorithm parameters ensures an exponential convergence in a $O(\omega + A)$ neighborhood of the optimum u^* , thus, of y to the maximum $y^* = f(u^*)$ (Ariyur and Krstic, 2003).

However, this time-scale separation implies relatively slow convergence, which makes the algorithm of little use in some practical situations. On the other hand, increasing the dither signal frequency introduces phase-shift and amplitude attenuation due to process dynamics. Some solutions have been proposed to compensate for this behavior in (Deschenes and St-Onge, 2013; Sharafi et al., 2013; Moase and Manzie, 2011; Krstic, 2000) but their implementation is complex. In (Dewasme et al., 2011, 2012; Chioua et al., 2016) a linear static map between the input and output is still assumed, but the bank of filter is replaced by a continuous Recursive Least Square (RLS) for the gradient estimation. These studies show that the RLS scheme improves the speed of convergence over the classical BOF configuration.

The contribution of this work is to extend these latter results by considering a dynamic representation of the plant in the form of a Hammerstein-Wiener model. Using a recursive RLS estimator, the speed of convergence of the ES scheme is significantly improved, as illustrated with an application to the optimization of the production of microalgae in a continuous photo-bioreactor (PBR).

This paper is organized as follows. The next section presents the ES scheme modifications, including the use of RLS instead of BOF, and the representation of the process dynamics by a Hammerstein-Wiener model. Section 3 presents the application of the resulting ES scheme to production optimization in a micro-algae PBR. Finally, the last section draws some conclusions.

2. ES WITH RLS ESTIMATION AND HAMMERSTEIN-WIENER MODELING

The BOF in Fig 1 is replaced by a RLS algorithm with forgetting factor (see Fig 2). The static map is approximated by the regression form :

$$y = \phi^T \theta + \nu(t) \tag{2}$$

where y is the output, ϕ a known vector of explanatory variables, θ is the parameter vector, $\nu(t)\sim \mathcal{N}(0,\sigma^2)$.

The following objective function is minimized:

$$J(t) = \sum_{i=1}^{t} \lambda^{t-i} [y(i) - \phi(i-1)^T \theta(t)]^2 r^{-1} + [\theta(t) + \theta(0)]^T P(0)^{-1} [\theta(t) + \theta(0)]$$
(3)

The recursive solution is (Landau and Dugard, 1986) :

$$e(t) = y(t) - \phi(t-1)^T \theta(t-1)$$
(4)

$$P(t) = \frac{1}{\lambda} \left[P(t-1) - \frac{P(t-1)\phi(t-1)\phi(t-1)^T P(t-1)}{\lambda r + \phi(t-1)^T P(t-1)\phi(t-1)} \right]$$
(5)

$$\theta(t) = \theta(t-1) + P(t)\phi(t-1)e(t)$$
(6)

The forgetting factor $0 < \lambda \leq 1$ determines how fast past data are disregarded, and typically $0.8 \leq \lambda \leq 0.99$. The "memory length" of the estimator is given by (Åström and Wittenmark, 1995):

$$N_0 = \frac{2}{1 - \lambda} \tag{7}$$

r > 0 is the variance estimate; with no model mismatch, we set $r = E[\nu^2(t)]$.



Fig. 2. ES with RLS.

For gradient estimation in a ES scheme, the model is chosen as:

$$y(t) = b + m u(t) \tag{8}$$

with $\phi^T = [1, u(t)]^T$ and $\theta = [b, m]^T$.

RLS (see Fig 2) provides estimates (\hat{m}, \hat{b}) of (m, b) where $\hat{m} \approx \frac{\partial y}{\partial u}$, the gradient estimation, is pushed to zero in average by the integrator.

Note that the dither signal is used to ensure persistency of excitation. (Åström and Wittenmark, 1995) has shown that a sinusoid is persistently exciting of order 2 and thus allows the identification of at least 2 parameters. As a general rule, a minimum of $\frac{n}{2}$ distinct sinusoids is necessary for the identification of n parameters (Landau and Dugard, 1986) (e.g for MISO Extremum Seeking) .

2.1 BOF versus RLS: a simple numerical example

Consider a process described by :

$$y(u) = b + mu \tag{9}$$

with m = -1, b = 6, and the regression model defined as in (8), thus with a structure perfectly matching the process. The input u as well as output y(u) are assumed measurable and the objective function to be maximized is of the form (this form corresponds to the productivity of a bioreactor):

$$h(u) = u y(u) = bu + mu^2$$
 (10)

At the maximum, $u^* = 3$ and $h^* = 9$ as illustrated in Figure 3.

Since y(u) and u are available, the gradient of the objective function can be computed as:

$$\frac{\partial h(u)}{\partial u} = \frac{\partial u y(u)}{\partial u} = y(u) + u \frac{\partial y(u)}{\partial u} = y(u) + um \quad (11)$$

The selected parameters for BOF and RLS (with sampling time $T_s = 0.01 \ h$) are provided in Table 1. Simulation results in Fig 4 shows that ES with RLS is faster than with BOF. Whereas RLS estimates the objective function gradient, BOF only provides a proportional estimate of this gradient (Ariyur and Krstic, 2003). Furthermore, increasing the integrator gain in the BOF scheme in order to speed up the convergence may result in system instability. Indeed, from averaging theorem, there exists an upper bound on k depending on the perturbation pulsation (Ariyur and Krstic, 2003).



Fig. 3. Static Map: Evolution of the cost function h with respect to the input u.

Parameters	BOF	RLS
A	0.4	0.4
k	2	2
ω	1	1
ω_h	0.95	-
ω_l	0.9	_
λ	-	0.99
r	-	1
P_0	-	1e3

Table 1. BOF and RLS parameters



Fig. 4. Comparison of RLS and BOF schemes

2.2 Hammerstein-Wiener modeling

Consider now the case where the output static map y(u) is followed by a linear filter (strictly proper and stable), e.g, a 1st or a 2nd order transfer function. This leads to the classical Hammerstein-Wiener representation of Fig 5. From a practical point of view, the transfer function can represent system and/or sensor dynamics.

Fig. 5. Hammerstein-Wiener Model

Measurements are usually collected at discrete times, and discrete-time transfer functions derived with the matched pole-zero method, are considered for first- and secondorder systems

$$G_1(p) = \frac{1}{1 + \tau p} \longrightarrow G_1(z) = \frac{K_1}{z - \alpha}$$
(12)

with $\alpha = e^{-\frac{T_s}{\tau}}$ and $K_1 = 1 - \alpha$

$$G_2(p) = \frac{1}{(1+\tau_1 p)(1+\tau_2 p)} = \frac{\gamma_2}{p^2 + \gamma_1 p + \gamma_2}$$
$$\longrightarrow G_2(z) = \frac{(1-\alpha_1)(1-\alpha_2)z}{(z-\alpha_1)(z-\alpha_2)} = \frac{K_2 z}{z^2 + \beta_1 z + \beta_2} \quad (13)$$

with $\alpha_1 = e^{-\frac{T_e}{\tau_1}}$, $\alpha_2 = e^{-\frac{T_e}{\tau_2}}$, $\beta_1 = -(\alpha_1 + \alpha_2)$, $\beta_2 = \alpha_1 \alpha_2$, $K_2 = (1 - \alpha_1)(1 - \alpha_2) = 1 + \beta_1 + \beta_2$.

Figure 6 displays the modified ES scheme assuming 1^{st} order dynamics. The model outputs read:

$$x(t) = b + m u(t)$$

$$y(t) = K_1 x(t) + \alpha y(t-1)$$
(14)
(15)

or,

$$y(t) = K_1 m u(t) + \alpha y(t-1) + K_1 b$$
(16)
$$y(t) = \phi^T \theta$$

where

$$\phi^T = \{1, y(t-1), u(t)\}^T \text{ and } \theta = \{K_1 b, \alpha, K_1 m\}^T.$$

The RLS provides estimates $\widehat{\theta}_{i=1,2,3}$ from which we deduce successively:

•
$$\widehat{\alpha} = \widehat{\theta}_2$$
,
• $\widehat{K}_1 = 1 - \widehat{\alpha}$,
• $\widehat{b} = \frac{\widehat{\theta}_1}{\widehat{K}_1}$
• $\widehat{m} = \frac{\widehat{\partial x}}{\partial u} = \frac{\widehat{\theta}_3}{\widehat{K}_1}$

For a 2^{nd} order transfer function,

$$y(t) = K_2 [x(t-1)] - \beta_1 y(t-1) - \beta_2 y(t-2)$$
(17)

Then, (17) and (14):

$$y(t) = K_2 m [u(t-1)] - \beta_1 y(t-1) - \beta_2 y(t-2) + K_2 b$$
(18)
$$y(t) = \phi^T \theta$$

where $\phi^T = \{1, -y(t-1), -y(t-2), u(t-1)\}^T$ and $\theta = \{K_2b, \beta_2, \beta_1, K_2m\}^T$.

The RLS provides estimates $\hat{\theta}_{i=1,2,3,4}$ from which we deduce successively:

$$\widehat{\beta}_2 = \widehat{\theta}_2, \widehat{\beta}_1 = \widehat{\theta}_3, \widehat{K}_2 = 1 + \widehat{\beta}_1 + \widehat{\beta}_2 \widehat{b} = \frac{\widehat{\theta}_1}{\widehat{K}_2} \widehat{m} = \frac{\widehat{\partial}x}{\partial u} = \frac{\widehat{\theta}_4}{\widehat{K}_2}$$

A gradient estimation can be computed following:

$$\begin{aligned}
\widehat{\frac{\partial h}{\partial u}} &= \frac{\widehat{\partial u.y}}{\partial u} = y + u \,\,\widehat{\frac{\partial y}{\partial u}} \\
&= y + u \,\,\widehat{\frac{\partial y}{\partial x}} \,\,\widehat{\frac{\partial x}{\partial u}} \\
&= y + u \,\,\widehat{\gamma}(t) \,\,\widehat{m}
\end{aligned} \tag{19}$$
(20)

where an estimation of $\gamma(t)$ is obtained as:

$$\widehat{\gamma}(t) = \frac{y(t)}{\widehat{m}u(t) + \widehat{b}}$$
(21)



Fig. 6. Discrete ES scheme with RLS estimator with forgetting factor.

 $\gamma(t)$ accounts for phase-shift and the attenuation introduced by the plant dynamics at higher frequencies.

Simulation results in Fig 7 show the results of ES applied to a first-order plant, either with a regression model including the process dynamics or omitting it. The parameters selected for the simulation are : $\omega = 1$, A = 0.1, $T_s = 0.01$, k = 0.1, r = 1e - 5, $P_0 = 1e3$. It is clear that the inclusion of the process dynamics ($\tau = 2.5$) in the regression model (16) allows the convergence to the optimum whereas performance of the loop deteriorates with a static regression model (8). Figure 8 shows that the plantmodel mismatch also results in a bias in the parameter estimates and especially in $m = \frac{\partial y}{\partial u}$. On the other hand, (16) provides perfect estimation of the parameters and thus of the gradient (Fig 8).

The combination of RLS with a regression model in the form of a Hammerstein-Wiener model, e.g., (16) and (18), provides an elegant and easy-to-implement ES strategy offering the possibility to increase the frequency of the dither signal and in turn increased speed of convergence. This later point is further illustrated with a real-life application, i.e., the optimization of the production of micro-algae in a continuous photo-bioreactor.

3. APPLICATION TO MICRO-ALGAE CULTURE PRODUCTION

The proposed ES scheme with RLS and Hammerstein-Wiener modeling is now applied to the maximization of the production in micro-algae cultures. Microalgae have a wide range of potential applications ranging from wastewater treatment (Mairet et al., 2011; Abdel-Raouf et al., 2012) to biofuel production (Christi, 2008), and it is of interest to drive the process towards an optimal region of operation, without developing precise prior knowledge about the process model, which would require intensive laboratory work to collect data and perform model identification. RTO such as ES provides an appealing alternative as it requires a priori no or few knowledge about the process dynamics.



Fig. 7. ES with RLS applied to a first-order plant. Red: optimum operation, Blue: regression model based on a Hammerstein-Wiener form (16), Black: omitting dynamics in the regression model as in (8) - $\tau = 2.5$ and $\omega = 1$.



Fig. 8. Parameter estimate convergence. Red: true parameters, Blue: estimates using the regression form (16), Black: estimates using (8) - $\tau = 2.5$ and $\omega = 1$.

In the present simulation case study, the underlying process is described by an extended Droop Model (Bernard et al.; Bernard and Rémond, 2012) identified for culture of *Isochrisys Galbana*. This model accounts for photoacclimation and photo-inhibition and has already been used by the authors in (Dewasme et al., 2017), where a classical BOF scheme is applied. Parameters are presented in Table 2.

$$X = \mu (Q, I, \theta) X - DX - RX$$

$$\dot{S} = -\rho (S, Q) X + D (S_{in} - S)$$

$$\dot{Q} = \rho (S, Q) - \mu (Q, I, \theta) Q$$

$$\dot{I}^* = \delta \mu (Q, I, \theta) (\bar{I} - I^*)$$
(22)

where D, X, S, Q, I^* respectively represent the dilution rate, the biomass concentration, the substrate concentration, the internal quota concentration and the irradiance at which the micro-algae are photo-acclimated. The growth rate is assumed to be a Haldane function of the incident irradiance:

$$\mu(Q, I, \theta) = \mu_{max} \frac{I}{K_{sI} + I + \left(\frac{I^2}{K_{iI}}\right)} \left(1 - \frac{Q_{min}}{Q}\right)$$
(23)

and the uptake rate:

$$\rho(S,Q) = \rho_{max} \frac{S}{S + K_S} \left(1 - \frac{Q}{Q_{max}} \right)$$
(24)

Further details can be found in (Bernard et al.; Bernard and Rémond, 2012).

Table 2. Model parameter values for IsochrisysGalbana

Parameter	Value
ρ_{max}	$0.0730 \ gN.gC^{-1}d^{-1}$
K_S	$0.0012 \ gN.m^{-3}$
μ_{max}	$1.7000 \ d^{-1}$
Q_{min}	$0.0500 \ g N.g C^{-1}$
Q_{max}	$0.25 \ gN.gC^{-1}$
δ	1
R	$0.0081 \ d^{-1}$
K _{sI}	$1.4 \ \mu mol \ m^{-2}s^{-1}$
K _{iI}	295 $\mu mol \ m^{-2}s^{-1}$

The dilution rate and biomass are assumed measurable on-line. The production is defined as h = DX and the gradient is estimated using (19) following the scheme presented in Fig. 6 with (u, y) = (D, X). Based on open-loop simulation, the regression form used for this simulation is (16) assuming 1^{st} order dynamics. The sampling time is taken as $T_s = 0.01 \ day = 14.4 \ min$, the dither signal is defined by A = 0.05, $\omega = 2$, the parameters of RLS are $\lambda = 0.95$, r = 1e - 5, $P_0 = 1e3 \ I_{3\times3}$, and the integrator gain is k = 0.025. The incident light intensity is switched from 100 $\mu mol/(m^2.s)$ to 200 $\mu mol/(m^2.s)$ after 60 days and switched back to 100 $\mu mol/(m^2.s)$ at $t = 120 \ days$.

Starting with $D = 0.1 \ day^{-1}$, simulation results (Fig. 9 and 10) show a convergence to the maximum in 20 dayswhereas, in a previous study (Dewasme et al., 2017), more than 200 days were needed with a BOF scheme (the best set of parameters found was $\omega = 0.17$, $\omega_l = 0.95\omega$, $\omega_h =$ $0.9, \omega \ k = 40, \ A = 0.01$). Furthermore Fig 11 illustrates the ability of the proposed scheme to seek the maximum of production despite changes in operating conditions.

4. CONCLUSION

A simple extremum seeking strategy is proposed in this study, which includes a recursive least square estimator and a regression model in the form of a Hammerstein-Wiener model. The algorithm is developed in the case of first- or second-order dynamics, which is sufficient in most practical cases. The inclusion of process dynamics in the regression model dramatically improves the speed of convergence with respect to a classical bank-of-filter (BOF) approach. Especially, the application to the production maximization in micro-algae cultures shows a speed-up by a factor of 10, which makes the strategy of practical use (whereas previous results with BOF were difficult to achieve in real-life experiments due to the very long convergence time). On-going work entails the experimental validation with a lab-scale photo-bioreactor.



Fig. 9. ES with RLS scheme 6 applied to *Isochrysis* Galbana - State evolution



Fig. 10. ES with RLS scheme (as in Fig. 6) applied to cultures of *Isochrysis Galbana* - Production and dilution rate evolution - Red: optimum for I = $100\mu \ mol/(m^2.s)$ and Blue: $I = 200\mu \ mol/(m^2.s)$

ACKNOWLEDGEMENTS

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office. The authors also gratefully acknowledge the support of the "commission mixte Quebec - Wallonie-Bruxelles" for making possible several research stays at UMONS and UQAR.

REFERENCES

- Abdel-Raouf, N., Al-Homaidan, A., and Ibraheem, I. (2012). Microalgae and wastewater treatment. Saudi Journal of Biological Sciences, 19, 257–275.
- Ariyur, K.B. and Krstic, M. (2003). Real-time Optimization by Extremum-seeking Control. John Wiley & Sons, INC, wiley-interscience edition.
- Åström, K. and Wittenmark, B. (1995). Adaptive control.



- Fig. 11. ES with RLS scheme (as in Fig. 6) applied to cultures of *Isochrysis Galbana* Production evolution wrt dilution Rate Red: steady-state map for $I = 100\mu \ mol/(m^2.s)$ and Blue: $I = 200\mu \ mol/(m^2.s)$
- Bernard, O., Masci, P., and Sciandra, A. (2009). A photobioreactor model in nitrogen limited conditions. 6th Vienna International Conference on Mathematical Modelling.
- Bernard, O. and Rémond, B. (2012). Validation of a simple model accounting for light and temperature effect on microalgal growth. *Bioresource Technology*, 123, 520– 527.
- Chioua, M., Srinivasan, B., Guay, M., and Perrier, M. (2016). Performance improvement of extremum seeking control using recursive least square estimation with forgetting factor. *IFAC-PapersOnLine*, 49(7), 424–429.
- Christi, Y. (2008). Biodiesel from microalgae beats bioethanol. *Trends Biotechnol*, 26(3), 126–131.
- Deschenes, J.S. and St-Onge, P.N. (2013). Achievable performances for basic perturbation-based extremum seeking control in wiener-hammerstein plants. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, 2991–2998. IEEE.*
- Deschênes, J.S., St-Onge, P.N., Collin, J.C., and Tremblay, R. (2012). Extremum seeking control of batch cultures of microalgae nannochloropsis oculata in pre-industrial scale photobioreactors. *IFAC Proceedings Volumes*, 45(15), 585–590.
- Dewasme, L., Vargas, A., Moreno, J.A., and Vande Wouwer, A. (2012). Real-time optimization of a fed-batch bioreactor with substrate inhibition using extremum-seeking. In *Control Applications (CCA)*, 2012 IEEE International Conference on, 615–620. IEEE.
- Dewasme, L., Letchindjio, C.G.F., Zuniga, I.T., and Vande Wouwer, A. (2017). Micro-algae productivity optimization using extremum-seeking control. In Control and Automation (MED), 2017 25th Mediterranean Conference on, 672–677. IEEE.
- Dewasme, L., Srinivasan, B., Perrier, M., and Vande Wouwer, A. (2011). Extremum-seeking algorithm design for fed-batch cultures of microorganisms with overflow metabolism. *Journal of Process Control*, 21(7), 1092– 1104.
- Dochain, D., Perrier, M., and Guay, M. (2011). Extremum seeking control and its application to process and reac-

tion systems: A survey. Mathematics and Computers in Simulation, 82(3), 369–380.

- Krstic, M. (2000). Performance improvement and limitations in extremum seeking control. Systems and Control Letters, 39, 313–326.
- Landau, I.D. and Dugard, L. (1986). Commande adaptative. Aspects practiques et theoriques. Masson, Paris.
- Leyva, R., Alonso, C., Queinnec, I., Cid-Pastor, A., Lagrange, D., and Martínez-Salamero, L. (2006). Mppt of photovoltaic systems using extremum-seeking control. *IEEE transactions on aerospace and electronic systems*, 42(1), 249–258.
- Mairet, F., Bernard, O., Ras, M., Lardon, L., and Steyer, J.P. (2011). Modeling anaerobic digestion of microalgae using adm1. *Bioresource Technology*, 102(13), 6823– 6829.
- Moase, W.H. and Manzie, C. (2011). Fast extremumseeking on hammerstein plants. *IFAC Proceedings Vol*umes, 44(1), 108–113.
- Sharafi, J., Moase, W.H., Shekhar, R.C., and Manzie, C. (2013). Fast model-based extremum seeking on hammerstein plants. In *Decision and Control (CDC)*, 2013 IEEE 52nd Annual Conference on, 6226–6231. IEEE.
- Wang, H.H., Yeung, S., and Krstic, M. (2000). Experimental application of extremum seeking on an axial-flow compressor. *IEEE Transactions on Control Systems Technology*, 8(2), 300–309.