Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Correction for Non-identical Units in Multi-unit Optimization ⋆

Lyne Woodward, Michel Perrier, Bala Srinivasan

Department of Chemical Engineering,
École Polytechnique Montréal,
Montréal, Canada, H3C 3A7

Abstract: Extremum-seeking control is a class of methods where real-time static optimization of a dynamic system is achieved by controlling the gradient. Estimation of the gradient is a key issue in extremum-seeking methods. In multi-unit optimization, the gradient is obtained using finite difference of the outputs of multiple identical units driven with inputs that are offset by a design parameter. However, if the units are not identical, the stability of such a scheme is in question and the point where the scheme converges could be far from the desired optimum. This paper proposes correctors to compensate for the differences between the units. It is shown that the scheme with correctors is locally asymptotically stable and converges to the desired optimum for all the units. The multi-unit optimization method with correctors is applied to bioreactors producing green fluorescent protein.

Keywords: Real-time optimization, Extremum-seeking control, Multi-unit optimization, Bioreactors, Convergence, Asymptotic stability

1. INTRODUCTION

Real-time optimization is a valuable tool to bring and maintain a process at its optimal point of operation. The classical approach consists of two steps (Marlin and Hrymak, 1997): First, a model of the process is used to numerically calculate the optimum and secondly, the model is updated using the available measurements and the updated model is used for re-optimization.

Extremum-seeking is an alternate approach (Sternby, 1980; Blackman, 1962), where optimization is achieved by following the necessary conditions of optimality, i.e., in an unconstrained case, controlling the gradient to zero. The different extremum-seeking control methods differ from one another in the way in which the gradient is estimated. For gradient estimation, perturbation methods (Krstic and Wang, 2000) can be used, when measurements of the performance criterion are available. If only auxiliary measurements are available, a model-based gradient estimation approach is needed (Guay and Zhang, 2003).

The multi-unit optimization method (Srinivasan, 2007) is a recent technique for gradient estimation. Here, multiple identical units are driven with inputs that are offset by a pre-fixed value. The gradient is obtained as a finite difference of the outputs of the units. This method has shown faster convergence than the perturbation method mainly because perturbations are in the "units" dimension instead of time dimension. Such a scheme requires several identical units and examples of such processes can be found in the domains of micro-array reactors, multiple production lines, and arrays of fuel cells.

⋆ This work was supported in part by the National Sciences and Engineering Research Council of Canada

However, the fact that the units have to be identical is a very strong assumption which does not depict the reality. The convergence of the multi-unit method with non-identical units is mainly determined by the sign of the offset used. Even if the convergence is assured, the point of convergence can be far from the real optimum (Woodward *et al.*, 2007).

The main goal of this paper is to show that the optima of the multiple non-identical units can be reached by the introduction of correctors within the multi-unit scheme. Local stability and the fact that the units converge to their respective optima are guaranteed.

In Section 2, the multi-unit optimization with identical units is reviewed. Section 3 gives the convergence conditions of the multi-unit scheme with non-identical units. Section 4 presents the improved multi-unit optimization method using correctors. The scheme is then applied to two non-identical bioreactors in Section 5. Finally, conclusions appear in Section 6.

2. MULTI-UNIT OPTIMIZATION

2.1 Problem formulation

Consider a dynamic system with state $x \in \Re^n$, input $u \in \Re^m$ that has to be operated so as to minimize a convex function $J(x,u)$ at steady state. The problem is shown below:

$$\min_u \; J(x,u) \tag{1}$$

$$s.t. \quad \dot{x} = F(x,u) \equiv 0 \tag{2}$$

where $F(x,u)$ is the function describing the dynamics of the system, which is assumed to be stable. The necessary

conditions of optimality are given by :

$$\frac{dJ}{du} = \frac{\partial J}{\partial u} - \frac{\partial J}{\partial x}\left(\frac{\partial F}{\partial x}\right)^{-1}\frac{\partial F}{\partial u} = 0 \tag{3}$$

As in the steepest descent method for numerical optimization (Nocedal and Wright, 1999), extremum-seeking makes the process evolve in the opposite direction of the gradient. But instead of using the iteration index as in numerical methods of optimization, the iterations evolve in real time. The extremum-seeking control law is given by :

$$\dot{u} = -k\left(\frac{dJ}{du}\right)^{T} \tag{4}$$

The key problem is the estimation of the gradient, which could be addressed using several methods (Krstic and Wang, 2000; Guay and Zhang, 2003). The multi-unit method provides an estimate of the gradient by finite differences as will be shown next.

2.2 The multi-unit scheme

The various units are operated with input values that are offset by a prefixed value. In order to simplify the presentation of the method, the single input case would be considered. So, two identical units are required and the scheme is presented in Fig. 1. The first unit is operated at the input value $u_1 = u - \frac{\Delta}{2}$ while the second unit is operated with input $u_2 = u + \frac{\Delta}{2}$. Then, the gradient is estimated by :

$$\hat{g}(u) = \frac{J_2(x_2, u_2) - J_1(x_1, u_1)}{\Delta} \tag{5}$$

The extremum-seeking control law (4) is then applied:

$$\dot{u} = -k\hat{g}^{T}(u) \tag{6}$$

Let u^* be the equilibrium point where the multi-unit optimization algorithm converges. This means that the inputs of the two units would converge to $u_1^* = u^* - \frac{\Delta}{2}$ and $u_2^* = u^* + \frac{\Delta}{2}$.

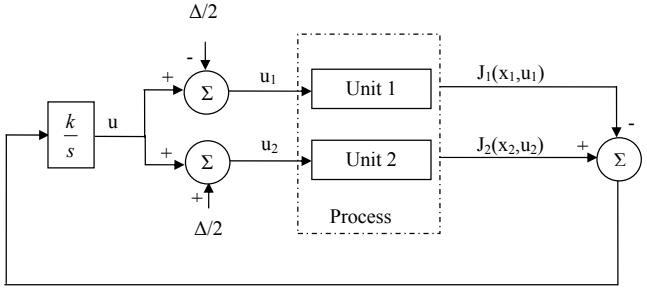

Fig. 1. Schematic for multi-unit optimization

All units follow the same control law and always keep an input offset of Δ from each other. The convergence of this scheme to a ball around the optimum has been proven despite the errors caused by the dynamics (which is assumed to be stable) and the error due to finite differences.

3. MULTI-UNIT OPTIMIZATION WITH NON-IDENTICAL UNITS

The main limitation of the multi-unit scheme as presented in the previous section is the requirement of identical units. The extension to non-identical units is essential to make this method more realistic.

3.1 Characterization of the difference between the units

The differences between the various units can manifest in the following ways: differences in dynamics, differences in static responses, or differences in disturbance effects. The case considered for analysis in this paper assumes the following: (i) the system has only one input, (ii) the dynamics are the same and are very fast compared to the optimization time-scale, i.e. the process can be considered quasi-static, (iii) no noise effects are considered, and (iv) the functions are convex.

Under these conditions, without loss of generality, the static characteristics of the two units are represented using $J_1(u_1)$ and $J_2(u_2)$. Let the relationship between the two static maps is given by:

$$J_2(u) = J_1(u + \beta) + \gamma + \bar{J}(u) \tag{7}$$

where $\beta = u_1^{opt} - u_2^{opt}$ and $\gamma = J_2(u_2^{opt}) - J_1(u_1^{opt})$, u_1^{opt} and u_2^{opt} are the optima of the first and second units, respectively.

From this definition, by evaluating the function and its derivative at u_2^{opt} it can be seen that

$$\bar{J}(u_2^{opt}) = 0, \qquad \left.\frac{\partial \bar{J}}{\partial u}\right|_{u_2^{opt}} = 0 \tag{8}$$

So, in the neighborhood of the optimum, it can be assumed that $\bar{J} = 0$. The transformation is to bring the second unit to the same coordinates as that of the first by a translation on the input and one of the output.

3.2 Convergence with Non-identical Units

As shown in (Woodward *et al.*, 2007), when multi-unit optimization is applied to processes with non-identical units the following can be inferred.

- Stability is guaranteed iff Δ is chosen such that
$$\Delta(\Delta + \beta) > 0$$
- The equilibrium point can be approximated by:
$$u^* \simeq \frac{u_1^{opt} + u_2^{opt}}{2} - \frac{\gamma}{(\Delta + \beta)\frac{\partial J_1^2}{\partial u^2}}$$

4. USE OF CORRECTORS IN MULTI-UNIT OPTIMIZATION WITH NON-IDENTICAL UNITS

In the previous section, it was seen that the equilibrium point can be quite far away from the real optimal values, especially when $\gamma \neq 0$. To avoid such an occurrence, it will be shown in this section that correctors could be employed to push each of the units to their respective optima.

The proposed corrections are motivated by (7). A first correction, $\hat{\beta}$, is applied to the input of the second unit

u_2 so as to translate the curves along the input axis. A second correction, $\hat{\gamma}$, applied onto the objective function of the second unit, J_2, translates the curves along the output axis so as to eliminate the bias in the estimation of the gradient.

To derive the update laws for the estimates, the core idea is to alternate between the multi-unit method and two different perturbation signals as shown in Fig. 2. Both perturbation signals are periodic with period $T = T_1 + T_2$.

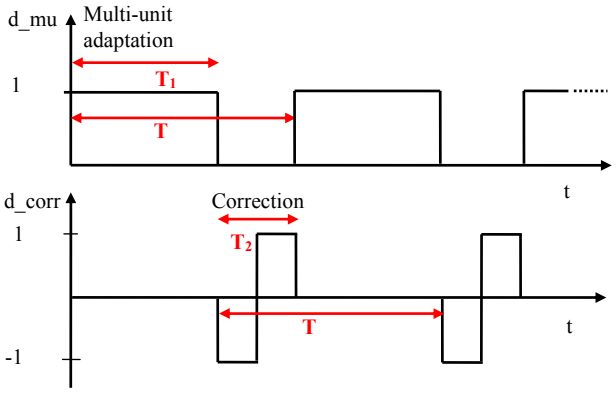

Fig. 2. Perturbation signals for multi-unit with correctors

The adaptation laws are formulated based on the following ideas. In the correction phase the difference between the two inputs, Δ, is removed. So, the two units act at the same operating point (corrected by $\hat{\beta}$ if any). Then, the corrected output values should be equal, if the vertical shift ($\hat{\gamma}$) is computed correctly. So, the difference between the corrected outputs provides the adaptation law for $\hat{\gamma}$. Also herein, the gradients should be the same if the horizontal shift ($\hat{\beta}$) is compensated. The difference in gradients is computed using the standard perturbation method and $\hat{\beta}$ is adapted to push this difference to zero.

4.1 Scheme with Correctors

Consider two units such that $J_2(u) = J_1(u+\beta)+\gamma$ and J_1 is convex. Let the inputs of the two units be synchronized as follows:

$$u_1 = u - \frac{\Delta}{2} d_{mu} + a\ d_{corr} \tag{9}$$

$$u_2 = u + \frac{\Delta}{2} d_{mu} + a\ d_{corr} - \hat{\beta} \tag{10}$$

with the multi-unit adaptation law being:

$$\dot{u} = -\frac{k_{mu}}{\Delta}(J_2 - J_1 - \hat{\gamma})\ d_{mu} \tag{11}$$

Also, consider the following adaptation laws for the correctors $\hat{\beta}$ and $\hat{\gamma}$:

$$\dot{\hat{\beta}} = -\frac{k_\beta}{a}(J_2 - J_1 - \hat{\gamma})\ d_{corr} \tag{12}$$

$$\dot{\hat{\gamma}} = -k_\gamma (J_2 - J_1 - \hat{\gamma})\ (1 - d_{mu}) \tag{13}$$

where, k_{mu}, k_β and k_γ are positive constants.

The averaged dynamics are derived next. As the perturbation signals of Fig. 2 are periodic signals with period T, averaging as explained in (Khalil, 1996) can be applied to the system (11)-(13). The averaged system is given by,

$$\dot{\hat{\beta}} = -\frac{k_\beta}{Ta} I_\beta, \quad I_\beta = \int_0^T D\ d_{corr}\ dt \tag{14}$$

$$\dot{\hat{\gamma}} = \frac{-k_\gamma}{T} I_\gamma, \quad I_\gamma = \int_0^T D\ (1 - d_{mu})\ dt \tag{15}$$

$$\dot{u} = \frac{-k_{mu}}{T\Delta} I_u, \quad I_u = \int_0^T D\ d_{mu}\ dt \tag{16}$$

where $D = (J_2(u_2) - J_1(u_1) - \hat{\gamma})$. As d_{corr} is equal to zero during $[0, T_1]$ and d_{mu} is equal to zero during $[T_1, T]$, $I_\beta = \int_{T_1}^T D\ d_{corr} dt$, $I_\gamma = \int_{T_1}^T D\ dt$ and $I_u = \int_0^{T_1} D\ dt$. The average equilibrium point is represented as $(\hat{\beta}^e, \hat{\gamma}^e, u^e)$.

Noting that $J_2(u) = J_1(u+\beta)+\gamma$, the second order Taylor series expansion of D around u^e,

$$D = \gamma - \hat{\gamma} + \left.\frac{\partial J_1}{\partial u}\right|_{u^e} (u_2 + \beta - u_1) + \left.\frac{\partial^2 J_1}{\partial u^2}\right|_{\bar{u}^e} (u_2 + \beta - u_1)(u_1 + u_2 + \beta - 2u^e) \tag{17}$$

The expansion is in fact exact if the second derivative is computed at an intermediate point $\bar{u}^e$. In the interval $[T_1, T]$, $(u_2 + \beta - u_1) = \beta - \hat{\beta}$ and $(u_1 + u_2 + \beta - 2u^e) = \beta - \hat{\beta} + 2a d_{corr} + 2(u - u^e)$. So, I_β is given by

$$I_\beta = \int_{T_1}^T \left(\gamma - \hat{\gamma} + \left.\frac{\partial J_1}{\partial u}\right|_{u^e} (\beta - \hat{\beta})\right) d_{corr} + \left.\frac{\partial^2 J_1}{\partial u^2}\right|_{\bar{u}^e} (\beta - \hat{\beta})(\beta - \hat{\beta} + 2(u - u^e)) d_{corr} + \left.\frac{\partial^2 J_1}{\partial u^2}\right|_{\bar{u}^e} 2a(\beta - \hat{\beta}) d_{corr}^2 dt = 0 \tag{18}$$

Looking at the d_{corr} signal, it can be seen that the its integral in the interval $[T_1, T]$ is zero and that the integral of d_{corr}^2 is T_2. So,

$$I_\beta = T_2 \left.\frac{\partial^2 J_1}{\partial u^2}\right|_{\bar{u}^e} 2a(\beta - \hat{\beta}) \tag{19}$$

Through a similar reasoning, it can be shown that

$$I_\gamma = T_2 \left(\gamma - \hat{\gamma} + \left.\frac{\partial J_1}{\partial u}\right|_{u^e} (\beta - \hat{\beta}) + \left.\frac{\partial^2 J_1}{\partial u^2}\right|_{\bar{u}^e} (\beta - \hat{\beta})(\beta - \hat{\beta} + 2(u - u^e))\right) \tag{20}$$

In the interval $[0, T_1]$, $(u_2 + \beta - u_1) = \beta - \hat{\beta} + \Delta$ and $(u_1 + u_2 + \beta - 2u^e) = \beta - \hat{\beta} + 2(u - u^e)$. So, I_u is given by

$$I_u = T_1 \left(\gamma - \hat{\gamma} + \left.\frac{\partial J_1}{\partial u}\right|_{u^e} (\beta - \hat{\beta} + \Delta) + \left.\frac{\partial^2 J_1}{\partial u^2}\right|_{\bar{u}^e} (\beta - \hat{\beta} + 2(u - u^e))(\beta - \hat{\beta} + \Delta)\right) \tag{21}$$

4.2 Equilibrium with Correctors

It is shown next that the desired optimum points, i.e $u_1^e = u_1^{opt} \pm \Delta$ and $u_2^e = u_2^{opt} \pm \Delta$ are indeed achieved.

Theorem 4.1. The equilibrium of scheme (9)-(13) on the average, i.e. the equilibrium of (14)-(16), $(\hat{\beta}^e, \hat{\gamma}^e, u^e)$ is given by $\hat{\beta}^e = \beta$, $\hat{\gamma}^e = \gamma$ and $u^e = u_1^{opt}$, where u_1^{opt} is the real optimal point of operation of unit 1.

Proof:

The average equilibrium point $(\hat{\beta}^e, \hat{\gamma}^e, u^e)$ will satisfy $I_\beta = I_\gamma = I_u = 0$. Since the second derivative is positive by convexity, from (19), $I_\beta = 0$ leads to,

- $\hat{\beta}^e = \beta$

Using this in (20), $I_\gamma = 0$ leads to

- $\hat{\gamma}^e = \gamma$

Using the above two and noting that the term $u - u^e = 0$ at the equilibrium in (21), $I_u = 0$ gives $\frac{\partial J_1}{\partial u}\big|_{u^e} \Delta = 0$, which leads to

- $u^e = u_1^{opt}$, the optimum of J_1

□

4.3 Stability of the Scheme with Correctors

Next, the local stability of the averaged dynamics of the adaptation scheme is analyzed.

Theorem 4.2. The average of the scheme (9)-(13), i.e. (14)-(16), is locally asymptotically stable.

Proof:

The Jacobian of (14)-(16) evaluated at $(\hat{\beta}^e, \hat{\gamma}^e, u^e)$ is given by:

$$\mathcal{J} = \frac{1}{T} \begin{bmatrix} -2k_\beta T_2 K & 0 & 0 \\ 0 & -k_\gamma T_2 & 0 \\ -k_{mu} T_1 K & -\frac{k_{mu}}{\Delta} T_1 & -2T_1 k_{mu} K \end{bmatrix} \tag{22}$$

where $K \equiv \frac{\partial^2 J_1}{\partial u^2}\Big|_{\bar{u}^e}$. The eigenvalues of (22) are the three diagonal elements since it is lower triangular. As $k_{mu}, k_\beta, k_\gamma > 0$ by choice and $K > 0$ by convexity, the Jacobian matrix of (22) will be Hurtwitz. □

5. ILLUSTRATIVE EXAMPLE

5.1 Process Description

The system under study is the production of green fluorescent protein (GFP) by E. coli cells. The following kinetic model, based on glucose as growth-limiting substrate, presented in (Aucoin *et al.*, 2006), was used for the simulations:

$$\dot{X} = \mu X - \left(\frac{F}{V}\right) X \tag{23}$$

$$\dot{P} = (Y_{P/S}\mu + \beta)X - \frac{F}{V}P \tag{24}$$

$$\dot{S} = \left(\frac{F}{V}\right)(S_f - S) - \frac{\mu X}{Y_{X/S}} - \frac{(Y_{P/X}\mu + \beta)X}{Y_{P/S}} - m_s \left(\frac{S}{K_{s_m} + S}\right) X \tag{25}$$

where X is the biomass concentration, F the feed rate of the substrate into the bioreactor, μ the specific growth rate of the biomass, V the volume of the bioreactor, S the substrate concentration, S_f the concentration of the substrate inlet, P the concentration of GFP, $Y_{P/S}$ is the product yield on substrate coefficient, $Y_{P/X}$ the production yield on biomass coefficient, $Y_{X/S}$ the biomass yield on substrate coefficient, β the non-growth associated product formation constant, and m_S the maintenance coefficient. The Monod model is used for the expression of μ:

$$\mu = \frac{\mu_{max} S}{K_s + S} \tag{26}$$

where μ_{max} is the maximum growth rate constant and K_s a saturation constant.

The optimization problem is to maximize the quantity of GFP in the post-induction period by adjusting the substrate flow into the bioreactor:

$$\max_F \; FP \tag{27}$$

$$s.t. \quad (23), (24), (25) \equiv 0 \tag{28}$$

All the simulations results presented in this section have been done using the post-induction numerical values given in (Aucoin *et al.*, 2006) and presented in Table 1.

Table 1. Constant parameter values of the Bioreactor

m_s	0.0025	g S/(g X h)	K_S	0.4	g/L
$Y_{P/X}$	66.92	mg P/g X	S_f	60	g/L
$Y_{P/S}$	50	mg P/g S	K_{S_m}	0.04	g/L
β	0.1	mg P/(g X h)	V	20	L

The difference between the two units has been simulated by using different values for μ_{max} and $Y_{X/S}$. The initial values have been adjusted to simulate a steady-state for an averaged flow of 1 L/h for the overall process as shown in Table 2.

Table 2. Different parameter values and initial conditions of the two Bioreactors

	Unit 1	Unit 2	
μ_{max}	0.2	0.18	h^{-1}
$Y_{X/S}$	0.4	0.25	g X/g S
S_0	0.13	0.15	g/L
X_0	15.28	11.46	g/L
P_0	1053.4	761.27	mg/L
F_0	0.95	1.05	L/h

The static characteristic of the two units are shown in Fig. 3. The real optimal operation points of the units are

shown by stars on the graph. The optimal flow for unit 1 is 3.68 L/h while for unit 2 it is 3.32 L/h. Also, the optimal GFP production is 3561 mg/h for unit 1 and 2309 mg/h for unit 2. Note that $\beta = u^1_{opt} - u^2_{opt} = (3.68 - 3.32) = 0.36 L/h$ and $\gamma = J_2(u^2_{opt}) - J_1(u^1_{opt}) = -(3561 - 2309) = -1252 mg/h$.

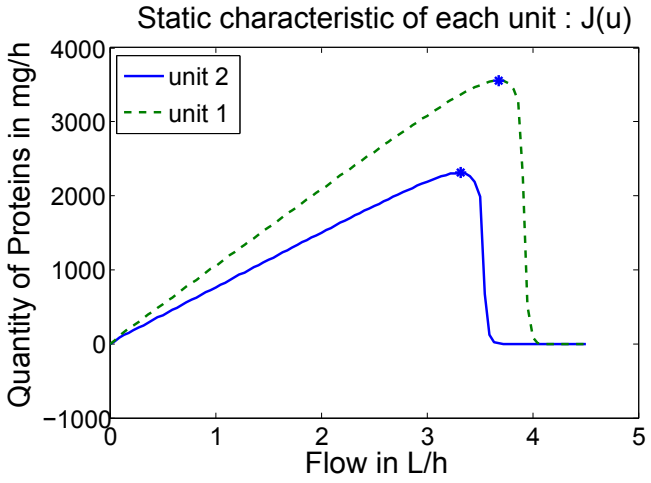

Fig. 3. Static characteristics of the two units

5.2 Convergence of the multi-unit optimization algorithm without correctors

The multi-unit optimization algorithm without any correctors for β and γ was first applied. It can be seen in Fig. 4 that the scheme converges quite rapidly but to a fairly suboptimal point. This simulation thus motivates the necessity of using correctors in the multi-unit scheme. A high value of $\Delta = 0.4\, L/h$ is used here, since lower values nearly push the inputs to zero.

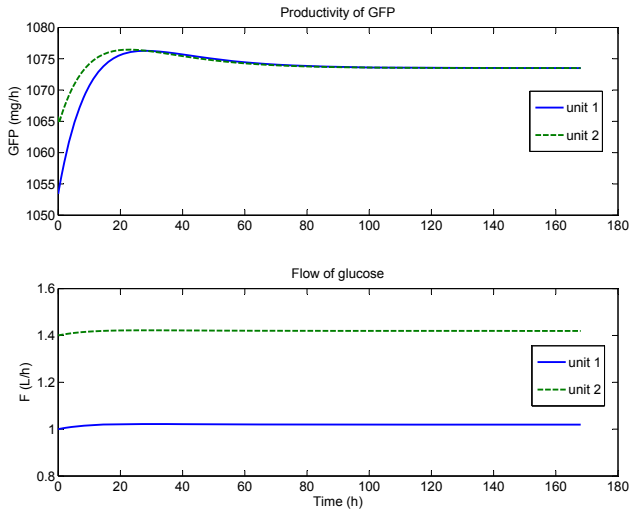

Fig. 4. Evolution of the two units under the multi-unit optimization algorithm with correctors – $\Delta = 0.4\, L/h$

5.3 Convergence of the multi-unit optimization algorithm with correctors

The multi-unit algorithm with correctors has been applied to the system using (9)-(13) and the tuning parameters presented in Table 3. Note that the signs of k_{mu} and k_γ have been inverted since we want to maximize FP.

Table 3. Tuning parameters of the multi-unit method with correctors

Δ	0.1 L/h	k_β	$2.15 \times 10^{-5}\ \frac{L^2}{h^2 mg}$
k_{mu}	$-1 \times 10^{-2}\ \frac{L^2}{h^2 mg}$	k_γ	-1 $\times 10^{-1}\ \frac{1}{h}$
T_1	20 h	T_2	40 h
a	0.05 L/h		

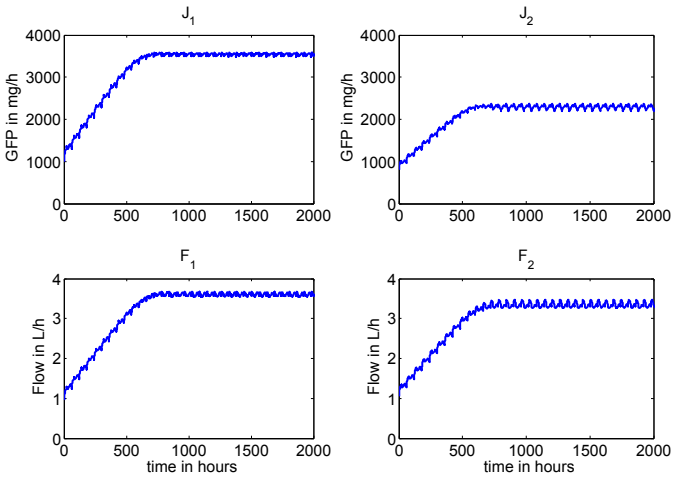

Fig. 5. Evolution of the two units under the multi-unit optimization algorithm with correctors

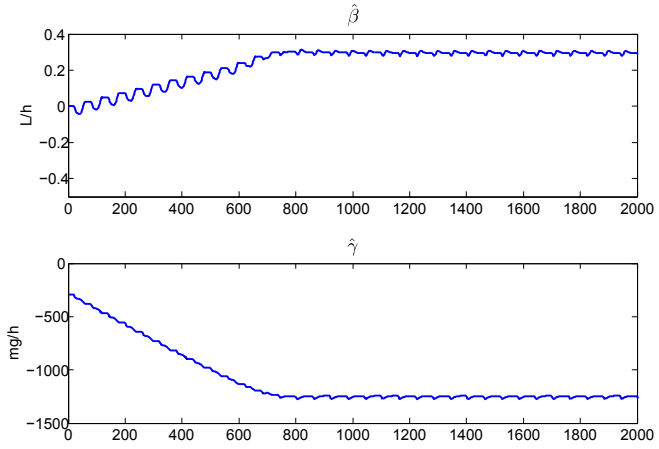

Fig. 6. Evolution of the corrections $\hat{\beta}$ and $\hat{\gamma}$

Fig. 5 shows the performance obtained by the bioreactors. The two units converge toward their respective optimum in about 800 hours. Comparing with the multi-unit scheme without correction it can be seen that the correctors slow down the adaptation by a factor of 10. However as will be seen next, the multi-unit scheme with correctors is still faster that the perturbation method.

5.4 Convergence of the perturbation method

As point of comparison, the perturbation method presented in (Krstic and Wang, 2000) has been applied to the same system with the parameters presented in Table 4. The schematic of the perturbation method is shown on Fig. 7.

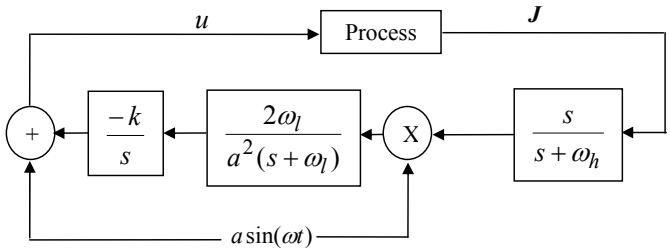

Fig. 7. Schematic for extremum-seeking using perturbations

Table 4. Tuning parameters of the perturbation method

a	0.1 L/h	ω_l	$0.006\,rad/h$
ω	$0.06\,rad/h$	ω_h	0.006 rad/h
k	1.9×10^{-4}		

As shown on Fig. 8, the perturbation method brings the process to it's optimal point in 4000 hours, 5 times slower than the multi-unit optimization method with correctors does.

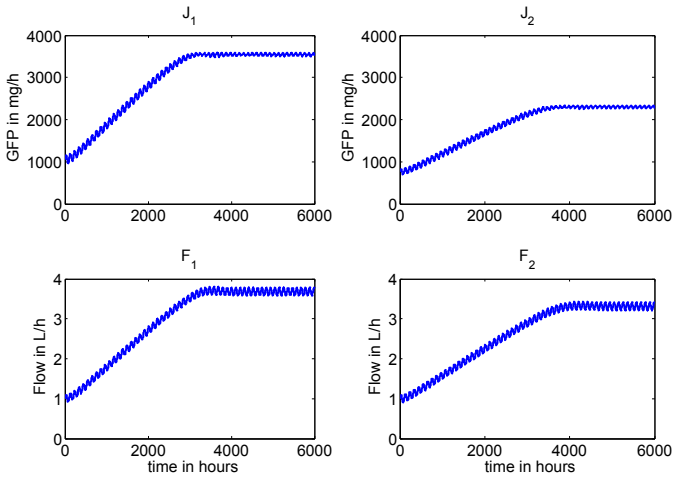

Fig. 8. Evolution of the two units under the perturbation optimization algorithm

Table 5. Performance of the perturbation and multi-unit methods

		Unit 1	Unit2	
Perturbation	J	3546	2297	mg/h
	F	3.68	3.31	L/h
Multi-unit	J	3545	2295	mg/h
	F	3.61	3.35	L/h
	$\hat{\beta}$	0.3		L/h
	$\hat{\gamma}$	-1252		mg/h

Table 5 gives the final operational values for both methods. This table along with Fig. 5 and Fig. 8 show that the values of the inputs are quite similar even if the variations in the performance index signals are slightly larger for the multi-unit method than the ones for the perturbation method. This is due to step changes imposed at $t = T_1$ and $t = T_2 + T_2$ as shown in Fig. 2). Since the system was non-minimum phase, these step changes introduced some oscillations at the output.

6. CONCLUSION

In this paper, an improved multi-unit scheme was presented introducing two correctors, in order to handle non-identical units. One corrector attenuates the effect of the difference in the optimal points of operation and the second one reduces the effect of the difference in the optimal values of the performance criteria. Local stability and that the scheme converges to the respective optima were demonstrated.

The ideas were illustrated with two bioreactors with differences both in static and dynamic characteristics. Both bioreactors converge to a neighborhood of their real optimum point. A comparison with the perturbation method demonstrates the rapidity of the multi-unit method. Future work will focus on issues such as multi-input systems, measurement noise, systems with inequality constraints, and differences in the dynamics.

REFERENCES

Aucoin, M.G., V. McMurray-Beaulieu, F. Poulin, E.B. Boivin, J. Chen, F.M. Ardelean, M. Cloutier, Y.J. Choi, C.B. Miguez and M. Jolicoeur (2006). Identifying conditions for inducible protein production in E.coli: Combining fed-batch and multiple induction. *Microbial Cell Factories* **5**(27), 1–13.

Blackman, P.F (1962). *An Exposition of Adaptive Control*. Chap. Extremum seeking regulators. J.H. Westcott ed.. The Macmillan Company. New York.

Guay, M. and T. Zhang (2003). Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. *Automatica* **39**(7), 1283–93.

Khalil, H.K. (1996). *Nonlinear Systems*. Prentice Hall.

Krstic, M. and H.-H. Wang (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica* **36**(4), 595–601.

Marlin, T.E. and A.N. Hrymak (1997). Real-time operations optimization of continuous processes. Vol. 316 of *AIChE Symposium Series*. p. 156.

Nocedal, J. and S. Wright (1999). *Numerical optimization*. Springer. New York.

Srinivasan, B. (2007). Real-time optimization of dynamic systems using multiple units. *International Journal of Robust and Nonlinear Control* **17**, 1183–1193.

Sternby, J. (1980). Adaptive control of extremum systems. *Canadian Journal of Chemical Engineering* pp. 151–160.

Woodward, L., M. Perrier and B. Srinivasan (2007). Convergence of multi-unit optimization with non-identical units: Application to the optimization of a bioreactor. Proceedings NOLCOS2007. pp. 832–837.