STOCHASTIC EXTREMUM SEEKING IN THE PRESENCE OF CONSTRAINTS

F. Coito[1], J. M. Lemos[2], S. S. Alves[3]

[1] *Faculdade de Ciências e Tecnologia - Universidade Nova de Lisboa. Email: fjvc@fct.unl.pt*
[2] *INESC ID-Lisboa. Email: jlml@inesc-id.pt*
[3] *Instituto Superior Técnico -Universidade Técnica de Lisboa. Email: salves@alfa.ist.utl.pt*

Abstract: The problem of adaptive minimization of globally unknown functions under constraints on the independent variable is considered in a stochastic framework. The main contribution of this paper consists in the extension of the CAM algorithm to vector problems. By resorting to the ODE analysis for analyzing stochastic algorithms and singular perturbation methods, it is shown that the only possible convergence points in the vector case are the constrained local minima. Simulations for dimension 2 problems illustrate this result.

Keywords: Adaptive control; constraint satisfaction; optimal search techniques; convergence analysis; singular perturbation method

1. INTRODUCTION

There are engineering optimization problems in which the global form of both the cost function and the constraints are unknown. In these extremum seeking problems, when the independent variable is settled to a specific value, the corresponding value of the function can be read and the decision whether the constraints are or are not being violated can be made. Although these extremum seeking methods have already been the subject of early literature in Adaptive Systems – see (Ariyur, 2003) for a review – they are receiving increasing interest in recent literature. (Zhang and Guay, 2003; Guay, *et al.*, 2003; Peterson and Stefanopoulou, 2004).

This kind of problems are solved in (Wellstead and Scotson, 1990; Bozin and Zarrop, 1991) by using a self-tuning extremum seeker in which the cost function is locally approximated by a quadratic function and no constraints are assumed in the independent variable. The contribution of this work consists in the extension of the above algorithm by incorporating constraints and the use of vector independent variables. As will be explained, this is achieved by solving the equation expressing the Kuhn-Tucker complementary condition using a stochastic approximation scheme.

The paper is organized as follows: First the problem to solve is formulated. Then, an algorithm, hereafter referred to as the CAM algorithm (Constrained Adaptive Minimization) is given for solving the problem. By using the ODE method for analyzing stochastic algorithms (Ljung, 1977), together with singular perturbation techniques for ordinary differential equations (Kokotovic, *et al.*, 1986), the CAM algorithm is analyzed, characterizing its possible points of convergence as the constrained minima. Finally, several simulation examples are presented.

2. PROBLEM FORMULATION

Let $y(\cdot)$ be a differentiable function of $\Re^2$ in $\Re$. Consider the following problem

Problem 1 Find $\mathbf{x}^* = \begin{bmatrix} x_1^* & x_2^* \end{bmatrix}^T$ such that $y(\mathbf{x}^*)$ is minimum, subject to the set of constraints

$$\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0} \tag{1}$$

where $\mathbf{g} \in \Re^n$ and $\mathbf{0}$ is the null vector. □

According to the Kuhn-Tucker theorem, Problem 2 is equivalent to the following

Problem 2 Define the Lagrangean function

$$£(\mathbf{x},\boldsymbol{\rho}) \overset{\Delta}{=} y(\mathbf{x}) + \boldsymbol{\rho}^T \mathbf{g}(\mathbf{x}) \tag{2}$$

Find the $\mathbf{x}^*$ minimizing $£(\mathbf{x},\boldsymbol{\rho}^*)$, in which $\boldsymbol{\rho}^*$ is a vector of Lagrange multipliers, satisfying the Kuhn-Tucker complementary condition;

$$\boldsymbol{\rho}^* \overset{\bullet}{\times} \mathbf{g}(\mathbf{x}^*) = \mathbf{0} \tag{3}$$

where $\overset{\bullet}{\times}$ is the *term-by-term* multiplication. □

Hereafter, the following assumption is supposed to hold:

H0. The global form of functions $y(\cdot)$ and $\mathbf{g}(\cdot)$ is unknown and may be possibly time varying. However, for each $\mathbf{x}$, $y(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ may be observed, possibly corrupted by observation noise. □

3. THE CAM ALGORITHM

The algorithm that solves Problem 2 must accomplish two tasks: the adjustment of the Lagrange multipliers $\boldsymbol{\rho}$ in order to fulfill the Kuhn-Tucker complementary condition (3) and, once $\boldsymbol{\rho}$ is settled, to adjust $\mathbf{x}(t)$.

3.1 Adjustment of the Lagrange multiplier

Following the development in (Lemos, 1992), $\boldsymbol{\rho}$ is adjusted according to a gradient minimization scheme:

$$\boldsymbol{\rho}(t) = \boldsymbol{\rho}(t-1) + \varepsilon\gamma(t-1)\boldsymbol{\rho}(t-1) \overset{\bullet}{\times} \mathbf{g}(\mathbf{x}(t)) \tag{4}$$

where ε is a vanishing small parameter and $\{\gamma(t)\}$ is a sequence of positive gains satisfying:

i. $\sum_{t=1}^{\infty} \gamma(t) = \infty$

ii. $\exists p : \sum_{t=1}^{\infty} \gamma^p(t) < \infty$

iii. $\{\gamma(t)\}$ is a decreasing sequence

iv. $\lim_{t\to\infty} \sup \frac{1}{\gamma(t)} - \frac{1}{\gamma(t-1)} < \infty$

Remark 1: These are technical assumptions needed to perform a convergence analysis using the ODE method for analyzing stochastic algorithms (Ljung, 1977). In particular i) is needed to ensure that any point of the space can be attained. While ii) implies that γ should tend to zero, in practice this may not be a good option. Indeed, if the constraints or the cost function are slowly time-varying, the objective is to track a moving minimum, and convergence to a constant point is undesirable. □

3.2 Adaptive optimization

H1. It is assumed that, close to $\mathbf{x}^*$, the Lagrangean function $£(\mathbf{x},\boldsymbol{\rho}^*)$ may be approximated by a quadratic function:

$$L(t) \overset{\Delta}{=} £(\mathbf{x}(t),\boldsymbol{\rho}) = £^* + \left[\mathbf{x}(t) - \mathbf{x}^*\right]^T \mathbf{A}\left[\mathbf{x}(t) - \mathbf{x}^*\right] + \bar{e}(t) \tag{5}$$

in the sequel it will be assumed $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$ to be symmetric, which does not affect the problem generality. $\mathbf{A}$, $£^*$ and $\mathbf{x}^*$ are unknown parameters, which depend on the value of $\boldsymbol{\rho}$; $\bar{e}$ is a residue.

Define the increments:

$$\Delta L(t) \overset{\Delta}{=} L(t) - L(t-1) \tag{6}$$

$$\Delta x_i(t) \overset{\Delta}{=} x_i(t) - x_i(t-1) \quad ; \quad i = 1,2 \tag{7}$$

$$\Delta x_i^2(t) \overset{\Delta}{=} x_i^2(t) - x_i^2(t-1) \quad ; \quad i = 1,2 \tag{8}$$

$$\Delta[x_1 x_2](t) \overset{\Delta}{=} x_1(t) \cdot x_2(t) - x_1(t-1) \cdot x_2(t-1) \tag{9}$$

Then equation (5) may be written as

$$\Delta L(t) = \begin{bmatrix} \theta_1 & \cdots & \theta_5 \end{bmatrix} \begin{bmatrix} \Delta x_1(t) \\ \Delta x_2(t) \\ \Delta x_1^2(t) \\ \Delta x_2^2(t) \\ \Delta[x_1 x_2](t) \end{bmatrix} + e(t) \tag{10}$$

where

$$\theta_1 = -2a_{11}x_1^* - 2a_{12}x_2^* \tag{11}$$

$$\theta_2 = -2a_{22}x_2^* - 2a_{12}x_1^* \tag{12}$$

$$\theta_3 = a_{11} \qquad \theta_4 = a_{22} \qquad \theta_5 = 2a_{12} \tag{13}$$

and $e(t) \overset{\Delta}{=} \bar{e}(t) - \bar{e}(t-1)$ is assumed to be an uncorrelated zero mean stochastic sequence such that all moments exist.

Defining

$$\boldsymbol{\theta}^* = [\theta_1 \quad \cdots \quad \theta_5] \tag{14}$$

$$\boldsymbol{\varphi}(t) = [\Delta x_1(t) \quad \Delta x_2(t) \quad \Delta x_1^2(t) \quad \Delta x_2^2(t) \quad \Delta[x_1 x_2](t)]^T \tag{15}$$

(10) yields

$$\Delta L(t) = \boldsymbol{\theta}^* \boldsymbol{\varphi}(t) + e(t) \tag{16}$$

which constitutes a linear regression model in which $\boldsymbol{\theta}^*$ is the vector of coefficients to estimate and $\boldsymbol{\varphi}$ is the data vector.

The vector $\boldsymbol{\theta}^*$ may be estimated using a recursive least-squares algorithm, and the value of $\mathbf{x}$ that minimizes $L(\mathbf{x})$ is given by:

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} \dfrac{2\theta_1\theta_4 - \theta_2\theta_5}{\theta_5^2 - 4\theta_3\theta_4} \\ \dfrac{2\theta_2\theta_3 - \theta_1\theta_5}{\theta_5^2 - 4\theta_3\theta_4} \end{bmatrix} \tag{17}$$

3.3 The CAM algorithm

Combining both the above procedures results in the following *Constrained Adaptive Minimization* (CAM) algorithm:

1. Apply $\mathbf{x}(t)$ to the system and measure $y(t)$ and $\mathbf{g}(\mathbf{x}(t))$
2. Adjust the Lagrange multiplier vector according to equation (4).
3. Using equation (2) build the Lagrangean function associated with the current Lagrange multiplier vector and the current value $y(t)$.
4. Compute the increments (6-9).
5. Using a RLS algorithm update the estimates of $\boldsymbol{\theta}$ in the model (16).
6. Update the estimates according to

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \dfrac{2\theta_1\theta_4 - \theta_2\theta_5}{\theta_5^2 - 4\theta_3\theta_4} \\ \dfrac{2\theta_2\theta_3 - \theta_1\theta_5}{\theta_5^2 - 4\theta_3\theta_4} \end{bmatrix} + \boldsymbol{\eta}(t) \tag{18}$$

7. Increment the time and go back to step 1.

4. ODE ANALYSIS

The CAM algorithm is now analyzed using the ODE method for analyzing stochastic algorithms (Ljung, 1977) and singular perturbation theory for ordinary differential equations (Kokotovic, *et al.*, 1986).

The algorithm is associated with the following set of differential equations:

$$\frac{d\,\boldsymbol{\rho}(t)}{d\,t} = \varepsilon \dot{\boldsymbol{\rho}}(t) \times \mathbf{g}(\mathbf{x}(t)) \qquad \frac{d\,\boldsymbol{\theta}(t)}{d\,t} = R^{-1}\mathbf{f}(\boldsymbol{\theta},\boldsymbol{\rho}) \tag{19}$$

where

$$R \overset{\Delta}{=} E[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^T(t)] \qquad \mathbf{f}(\boldsymbol{\theta},\boldsymbol{\rho}) \overset{\Delta}{=} E\{\boldsymbol{\varphi}(t)[\Delta L(t) - \boldsymbol{\varphi}^T(t)\boldsymbol{\theta}]\} \tag{20}$$

Define the functions $\mathbf{G}(\boldsymbol{\theta},\boldsymbol{\rho})$ and $\mathbf{H}(\boldsymbol{\theta},\boldsymbol{\rho})$

$$\mathbf{G}(\boldsymbol{\theta},\boldsymbol{\rho}) \overset{\Delta}{=} R^{-1}\mathbf{f}(\boldsymbol{\theta},\boldsymbol{\rho}) \qquad \mathbf{H}(\boldsymbol{\theta},\boldsymbol{\rho}) \overset{\Delta}{=} \dot{\boldsymbol{\rho}} \times \mathbf{g}(\mathbf{x}(t)) \tag{21}$$

Making use of (25-26) and changing the time scale by $\tau = \varepsilon t$, equations (21-22) may then be written in the standard form for singular perturbation analysis:

$$\frac{d\,\boldsymbol{\rho}(\tau)}{d\,\tau} = \mathbf{H}(\boldsymbol{\theta},\boldsymbol{\rho}) \qquad \varepsilon\frac{d\,\boldsymbol{\theta}(\tau)}{d\,\tau} = \mathbf{G}(\boldsymbol{\theta},\boldsymbol{\rho}) \tag{22}$$

According to the ODE theory exposed in (Ljung, 1977), the only possible convergence points of the CAM algorithm are the equilibrium points of (27-28), such that the Jacobian matrix

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial \mathbf{H}}{\partial \boldsymbol{\rho}} & \dfrac{\partial \mathbf{H}}{\partial \boldsymbol{\theta}} \\ \dfrac{\partial \mathbf{G}}{\partial \boldsymbol{\rho}} & \dfrac{\partial \mathbf{G}}{\partial \boldsymbol{\theta}} \end{bmatrix} \tag{23}$$

has all its eigenvalues in the left complex half-plane.

H2. The disturbance signal $\boldsymbol{\eta}$ in (18) ensures the persistent excitation requirement, i.e. $E[\boldsymbol{\varphi}(\tau)\boldsymbol{\varphi}^T(\tau)]$ is full rank. □

H3. The function $\mathbf{G}(\overline{\boldsymbol{\theta}},\boldsymbol{\rho}^*)$ has isolated real roots □

Proof of H3: If assumption H1 holds (which implies that equation (16) is valid together with assumption H2, then:

$$\begin{aligned} &\mathbf{G}(\boldsymbol{\theta},\boldsymbol{\rho}) = R^{-1}E\{\boldsymbol{\varphi}(\tau)[\Delta L(\tau) - \boldsymbol{\varphi}^T(\tau)\tau\boldsymbol{\theta}]\} = \\ &E^{-1}[\boldsymbol{\varphi}(\tau)\boldsymbol{\varphi}^T(\tau)]\{E[\boldsymbol{\varphi}(\tau)\boldsymbol{\varphi}^T(\tau)](\overline{\boldsymbol{\theta}} - \boldsymbol{\theta}) + E[\boldsymbol{\varphi}(\tau)e(t)]\} \end{aligned} \tag{24}$$

which yields:

$$\mathbf{G}(\boldsymbol{\theta},\boldsymbol{\rho}) = \mathbf{0} \Rightarrow E^{-1}[\boldsymbol{\varphi}(\tau)\boldsymbol{\varphi}^T(\tau)]E[\boldsymbol{\varphi}(\tau)\boldsymbol{\varphi}^T(\tau)](\overline{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \mathbf{0} \Rightarrow \boldsymbol{\theta} = \overline{\boldsymbol{\theta}} \tag{25}$$
□

The equilibrium points of (27-28) are characterized by one of the following conditions:

A-equilibria

$$\boldsymbol{\rho} = \mathbf{0} \qquad f(\boldsymbol{\theta},\mathbf{0}) = \mathbf{0} \tag{26}$$

B-equilibria

$$\mathbf{g}(\mathbf{x}) = \mathbf{0} \quad \text{and thus } \boldsymbol{\rho} = \boldsymbol{\rho}^* \qquad f(\boldsymbol{\theta},\boldsymbol{\rho}^*) = \mathbf{0} \tag{27}$$

4.1 Analysis of the A-equilibria

If (32-33) holds the constrained minimum equals the unconstrained minimum. The constrained minimum is therefore interior to the region defined by the set of constraints (1)

Provided the persistent excitation requirement holds, as $\frac{\partial \mathbf{H}}{\partial \boldsymbol{\theta}} = \mathbf{0}$, the Jacobian matrix (23) becomes lower triangular and its eigenvalues are the ones of $\frac{\partial \mathbf{H}}{\partial \boldsymbol{\rho}} = [\mathbf{g}(\mathbf{x})]_D$ and $\frac{\partial \mathbf{G}}{\partial \boldsymbol{\theta}} = -\mathbf{I}$, where $\mathbf{I}$ is the diagonal unit matrix and $[\mathbf{g}(\mathbf{x})]_D$ is a diagonal matrix whose elements are the $g_i(\mathbf{x})$. As $\rho_i=0$ which implies $g_i(\mathbf{x})<0$, all the Jacobian eigenvalues have negative real parts.

Thus the only possible convergence points are solutions of Problem 1.

4.2 Analysis of the B-equilibria

If (34-35) holds the constrained minimum is different from the unconstrained minimum, being located on the boundary of the region defined by (1). In this case $\frac{\partial \mathbf{H}}{\partial \boldsymbol{\theta}}$ is no longer null. Thus, the Jacobian matrix is not lower triangular, and the analysis from the previous section does not hold.

Making use of the singular perturbation theory (Kokotovic, *et al.*, 1986), assuming that the parameter ε in (4) is vanishing small (22) may be seen as the slow and fast subsystems, respectively.

Assume that H3 holds and consider the boundary layer correction $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta} - \bar{\boldsymbol{\theta}}$ whose dynamics is

$$\frac{d\tilde{\boldsymbol{\theta}}}{d\tau} = \frac{1}{\varepsilon}\mathbf{G}\left(\tilde{\boldsymbol{\theta}}, \boldsymbol{\rho}^*\right) \tag{28}$$

<u>H4.</u> Assume that $\tilde{\boldsymbol{\theta}}(\tau) = \mathbf{0}$ is an equilibrium point of (28), asymptotically stable, uniformly in $\boldsymbol{\rho}^*$, and that $\boldsymbol{\theta}(0) - \bar{\boldsymbol{\theta}}(0)$ belongs to its domain of attraction. □

Proof of H4: It follows from

$$\frac{d\tilde{\boldsymbol{\theta}}}{d\tau} = \frac{1}{\varepsilon}R^{-1}E\left\{\boldsymbol{\varphi}(t)\left[\boldsymbol{\varphi}^T(t)\left(\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) + e(t)\right]\right\} = -\frac{1}{\varepsilon}R^{-1}R\tilde{\boldsymbol{\theta}} = -\frac{1}{\varepsilon}\tilde{\boldsymbol{\theta}} \tag{29}$$
□

<u>H5.</u> The eigenvalues of $\frac{\partial \mathbf{G}}{\partial \boldsymbol{\theta}}$, calculated for $\varepsilon=0$, have strictly negative real part. □

Proof of H5: Observe that

$$\frac{\partial \mathbf{G}}{\partial \boldsymbol{\theta}} = \frac{\partial}{\partial \boldsymbol{\theta}}R^{-1}\mathbf{f}\left(\bar{\boldsymbol{\theta}}, \boldsymbol{\rho}^*\right) = \frac{\partial}{\partial \boldsymbol{\theta}}R^{-1}E\left\{\boldsymbol{\varphi}(\tau)\left[\Delta L(\tau) - \boldsymbol{\varphi}^T(\tau)\boldsymbol{\theta}\right]\right\} \tag{30}$$

From (5) $\varepsilon=0 \Rightarrow \Delta L(\tau) = \boldsymbol{\varphi}^T(\tau)\bar{\boldsymbol{\theta}} + e(t)$, yielding

$$\frac{\partial \mathbf{G}}{\partial \boldsymbol{\theta}} = -\frac{\partial}{\partial \boldsymbol{\theta}}R^{-1}E\left\{\boldsymbol{\varphi}(\tau)\boldsymbol{\varphi}^T(\tau)\right\}\tilde{\boldsymbol{\theta}} = -\frac{\partial}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\theta}} = -\mathbf{I} \tag{31}$$

□

Since these assumptions hold, Tikhonov's theorem (Kokotovic, *et al.*, 1986) allows to conclude the following proposition:

<u>Proposition 1:</u>

As seen in (Kokotovic, *et al.*, 1986), if assumptions H4 and H5 hold, then

$$\boldsymbol{\rho} = \boldsymbol{\rho}^* + O(\varepsilon) \qquad \boldsymbol{\theta} = \bar{\boldsymbol{\theta}}(t) + \tilde{\boldsymbol{\theta}}(t) + O(\varepsilon) \tag{32}$$

hold for all $t\geq 0$ and, further, there is $t_1\geq 0$ such that the approximation

$$\boldsymbol{\theta} = \bar{\boldsymbol{\theta}}(t) + O(\varepsilon) \tag{33}$$

holds for $t\geq t_1>0$. □

Proposition 1 states that the only possible convergence points of the CAM algorithm are the constrained minima of the optimization problem 1.

5. SIMULATION RESULTS

The ODE analysis characterizes the possible convergence points of the CAM algorithm. Yet, it does not prove that the algorithm will actually converge.
In order to exhibit the algorithm convergence features, a number of simulations are presented.

5.1 Example 1

In this example Problem 1 is considered, in which

$$y(\mathbf{x}) = (\mathbf{x} - \mathbf{x_0})^T(\mathbf{x} - \mathbf{x_0}) \tag{34}$$

with $\mathbf{x_0} = [0.6 \quad 0.8]^T$, and constraints

$$\begin{aligned} g_1(\mathbf{x}) &= 3 - x_2 e^{x_1/3} \leq 0 \\ g_2(\mathbf{x}) &= -x_1 \leq 0 \\ g_3(\mathbf{x}) &= x_1 - 1 \leq 0 \\ g_4(\mathbf{x}) &= 2 - x_2 \leq 0 \\ g_5(\mathbf{x}) &= x_2 - 3 \leq 0 \end{aligned} \tag{35}$$

The identification is performed using RLS with exponential forgetting factor.
Figures 1 and 2 present the evolution of the optimum estimate towards the feasibility region. The constrained minimum is on the frontier of the region. Thus while the Lagrange multipliers related the in active constraints go to zero $(\rho_i \to 0)$, those related to active constraints converge to the optimum $\rho_j \to \rho^*$ (figure 3).

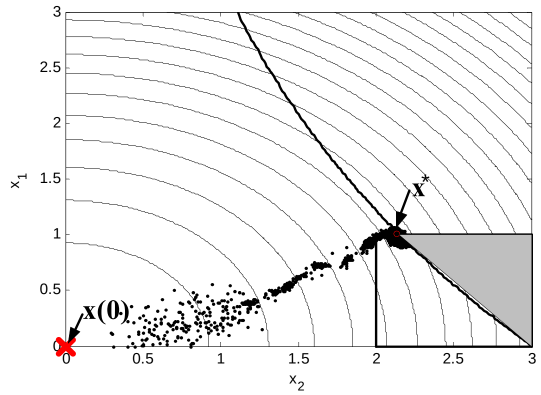

Fig. 1. Adaptive optimum search from example 1. The gray area is the feasibility region.

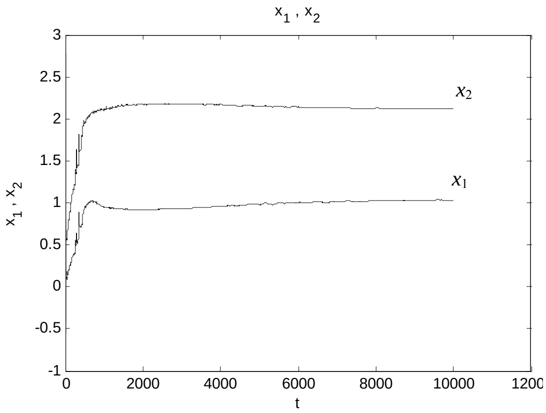

Fig. 2. Adaptive optimum search from example 1.

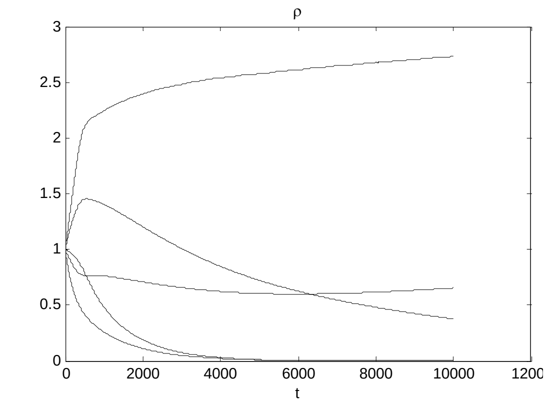

Fig. 3. Evolution of the Lagrange multipliers in example 1.

5.2 Example 2: fermentation process

This example considers the problem of optimizing agitation and aeration in a given fermentation process. The objective function is total electric power consumed for agitation, compression and refrigeration. The major constraint considered is to ensure that the dissolved oxygen concentration is above the critical value. This problem may be solved analytically when the process is well characterized (Alves and Vasconcelos, 1996). However, many process parameters have to be computed experimentally. This same problem may be tackled adaptively by the CAM algorithm whenever a *rough* feasibility area is known.

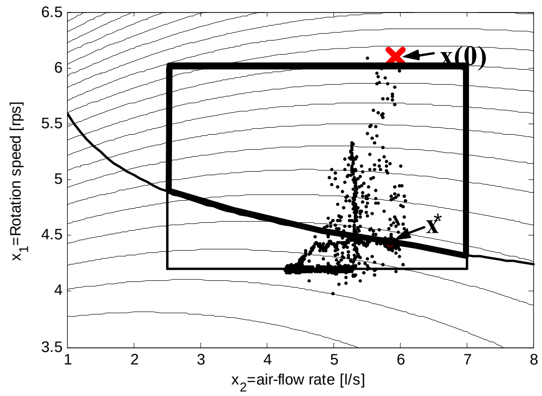

Fig. 4. Adaptive optimum search from example 2. The feasibility region lies inside the bold line.

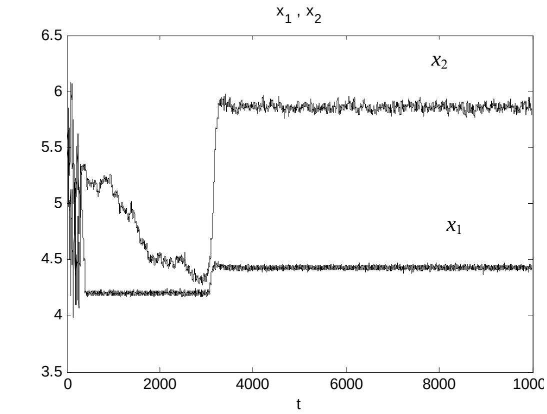

Fig. 5. Adaptive optimum search from example 2.

Simulation tests were performed using the model from (Alves and Vasconcelos, 1996) together with the CAM algorithm where:

y: electric power
g_1: dissolved oxygen concentration
g_2 and g_3: air-flow rate constraints
g_4 and g_5: agitator rotation speed

Results are presented in Figures 4 and 5.

Experiments using the updating scheme from equation (18) have shown that with this scheme assumption H1 and equation (16) would not hold. Thus in the experiment presented the updating scheme (18) was replaced by the following gradient scheme:

$$\mathbf{x}(t+1)=\mathbf{x}(t)-\delta\frac{\dfrac{\partial L}{\partial \mathbf{x}}}{\left\|\dfrac{\partial L}{\partial \mathbf{x}}\right\|} \qquad (36)$$

Comparing the results presented with those from (Alves and Vasconcelos, 1996) it is apparent that the algorithm converges towards the constrained minimum.

It should be noticed that in the example, appart from the constraints, no *a-priori* knowledge of the process was given to the algorithm.

5.3 Example 3: multiple local minima

The ODE analysis presented states that the convergence points are local minima from the constrained optimization problem. Thus, it is interesting to see what occurs when more than one minimum exists.

In this example the function to be minimized is given by:

$$y(\mathbf{x}) = 9 + \frac{9}{2}x_1 - 4x_2 + x_1^2 + 2x_2^2 - 2x_1x_2 + x_1^4 - 2x_1^2x_2 \tag{37}$$

and it is subject to the constraint

$$g(\mathbf{x}) = 24.25 - \left(x_1^2 + x_2^2\right) \leq 0 \tag{38}$$

Figure 6 presents the algorithm evolution when it starts from the initial point $\mathbf{x}(0)=[-1.9 \quad 7.95]^T$. It converges towards a local minimum, located at $\mathbf{x}^* = [-2 \quad 4.51]^T$, with a value of the objective function of 19.6.

In figure 7 the algorithm is started from a different initial point, $\mathbf{x}(0)=[0.198 \quad 6.95]^T$.

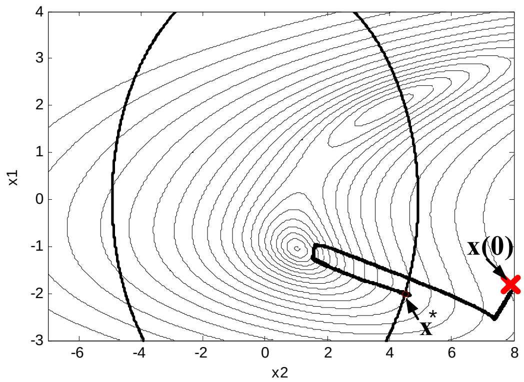

Fig. 6. Adaptive optimum search for Example 3. The algorithm converges to a local minimum.

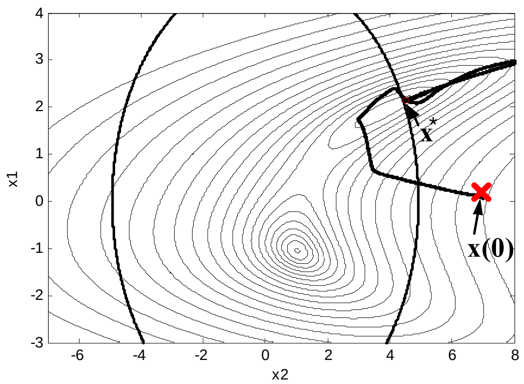

Fig. 7. Adaptive optimum search for Example 3. The algorithm converges to a local minimum.

In this case it converges to another local minimum located at $\mathbf{x}^*=[2.14 \quad 4.51]^T$, which corresponds to a value of the objective function of 1.21 (the absolute constrained minimum).

The minimum to which the algorithm converges depends on the initial point $\mathbf{x}(0)$, and in which domain of attraction it lies.

6. CONCLUSION

The problem of adaptive minimization of globally unknown functions under constraints on the independent variable was addressed in a stochastic framework. The CAM algorithm for vector problems was proposed. By resorting to the ODE analysis for analyzing stochastic algorithms and singular perturbation methods, it was shown that the only possible convergence points are the constrained local minima. A number of simulation results in 2 dimension were presented to illustrate this result.

REFERENCES

Alves, S. S. and J. M. T. Vasconcelos (1996). Optimisation of agitation and aeration in fermenters. *Bioprocess Engineering*, **14,** 119-123.

Ariyur, K. B. and M, Krstič (2003). *Real time optimization by estremum seeking control*. Wiley.

Bozin, A. S. and M. B. Zarrop (1991). Self-tuning extremum optimizer – Convergence and robustness properties". *Prep. ECC 91*, 672-677.

Guay, M., D. Dochain and M. Perrier (2003). Adaptive extremum seeking control of nonisothermal continuous stirred tank reactors with temperature constraints. *42nd IEEE-CDC Proceedings*, 4753-4758.

Kokotovic, P., H. K. Khalil and J. O'Reilly (1986). *Singular Perturbation Methods: Analysis and Design*. Academic Press.

Lemos, J. M (1992). Adaptive Optimization with Constraints. *Prep. 4th IFAC Symp. Adaptive Systems in Control and Signal Processing*, 53-58.

Ljung, L. (1977). Analysis of Recursive Stochastic Algorithms. *IEEE Transactions on Automatic Control*, **AC-22**, Nº4, 551-575.

Peterson, K. S. and A. G. Stefanopoulou (2004). Extremum Seeking Control for Soft Landing of an Electromechanical Valve Actuator. *Automatica*, **40**, Nº6, 1063-1069.

Wellstead, P. E. and P. G. Scotson (1990). Self-tuning extremum control. *IEE Proceedings*, **137**, Pt D. Nº 3, 165-175.

Zhang, T. and M. Guay (2003). Adaptive Extremum Seeking Control of Continuous Stirred-Tank Bioreactors. *AIChE Journal*, **49**, No. 1, 113-123.