UNIVERSITE PARIS 1 PANTHEON SORBONNE

CENTRE DE RECHERCHE

S.A.M.O.S

STATISTIQUE APPLIQUEE ET MODELISATION STOCHASTIQUE

Convergence of the one-demensional Kohonen algorithm

M. Benaïm, J.C. Fort, G. Pagès

Prépublication du SAMOS n° 74
Mai 1997

90, rue de Tolbiac - 75634 PARIS CEDEX 13

Convergence of the One-dimensional Kohonen Algorithm

Michel BENAÏM
Univ. Toulouse (France)

Jean-Claude FORT
Univ. Nancy I & SAMOS-Paris I (France)

Gilles PAGÈS
Univ. Paris 12 & Paris 6, URA 224 (France)

Abstract

We show in a very general framework the *a.s.* convergence of the 1-dimensional Kohonen algorithm – after self-organization – to its unique equilibrium when the learning rate decreases to 0 in a suitable way. The main assumption is a log-concavity of the stimuli distribution which includes all the usual (truncated) probability distributions (uniform, exponential, gamma distribution with parameter ≥ 1, ...). For the constant step algorithm, the weak convergence of the invariant distributions toward the equilibrium as the step goes to 0 is established too. The main ingredients of the proof are the Poincaré-Hopf Theorem and a result by M. Hirsch about the convergence of cooperative dynamical systems.

1 Introduction.

In 1982 T. Kohonen (see [14]) introduced what he called the Self Organizing Maps (S.O.M) which are very simple models of some neural behavior building sensory maps in various parts of the sensorymotor cortex (for instance the so-called retinotopic maps) The immediate interest that it raised was probably due to its very simplicity. The mathematical investigations began with a paper ([14]) by T. Kohonen himself who gave a sketch of proof of the self-organizing property in the one dimensional case with uniformly distributed stimuli. Actually it turned out that, even in this case, the mathematical treatment was not easy. M. Cottrell and J.C. Fort provided in this particular case a rigorous proof of self-organization and of conditional *a.s.*

convergence when the step is decreasing (see [5]). These results were then extended to more general stimuli distributions by C. Bouton and G. Pagès in [3]. Recently J.C. Fort and G. Pagès carried out a detailed study of the behavior of the so-called O.D.E of the algorithm. Namely, they proved that all possible equilibrium points were asymptotically stable. The result was obtained under some *log*-concave assumption on the density of the stimuli distribution. This assumption turned out to be quite classical in Information Theory to get the uniqueness of optimal quantizers (see *e.g.* [4]). Very recently A. Sadeghi first pointed out the link between cooperative dynamical systems and the Kohonen S.O.M. in [20]. This is a decisive remark allowing to use very powerful results in the fields of differential equations. Meanwhile, some multidimensional results – in the weak sense – were obtained either by physicist methods (see H. Ritter & K. Schulten [19]) or in a more rigorous setting (see [7], [6]). Nevertheless, the only strong approximation results – including this paper – concern the 1-dimensional case. This illustrates the gap between the simplicity of the definition and implementation of the S.O.M. which made their success among the users and the difficulty of its mathematical treatment.

In this paper, gathering all these previous results and combining the celebrated Hopf-Poincaré theorem on the vector fields with a beautiful result of M.W. Hirsch about the cooperative dynamical systems we derive a quite general result of *a.s.* convergence after self-organization for the Kohonen maps. Namely, there exists a unique equilibrium point for the Kohonen one dimensional S.O.M. and that it converges weakly or *a.s.* toward this unique equilibrium point.

As a conclusion, we point out the general feature of our proof by stating an abstract result for stochastic cooperative algorithms.

The rest of the paper is divided in four parts. In section 2 is recalled the definition of the Kohonen algorithm, the existing results and finally our main result. Section 3 is a technical one where we prove the asymptotical stability of the equilibrium points under a weaker assumption than in [7]. Uniqueness of the equilibrium point is proved in the section 4. In section 5 we show that the ODE is strongly cooperative and the proof of weak and *a.s.* convergence is carried out in section 6.

2 Definition and results

2.1 Definition of the algorithm

First we shortly recall the basic definitions and known results about the Kohonen one-dimensional algorithm.

A set I of *units* is given, identified to $I := \{1, 2, \cdots, n\}$. The set I set is endowed with a neighborhood structure defined through a neighborhood function $\sigma : \mathbb{Z} \to [0, 1]$. It satisfies : $\sigma(0) := 1$, $\sigma(k) = \sigma(-k)$ and σ restricted to $\mathbb{N}$ is nonincreasing. For every i, $j \in I$, $\sigma(i - j) = \sigma(|i - j|)$ measures the strength of the "connection" between i and j. To each unit i is associated an initial weight X_i^0. Let $X^0 := (X_i^0)_{1\leq i\leq n}$ be the initial weight vector. The Kohonen algorithm adapts this weight vector using a sequence of *stimuli* with two goals : first to organize the weights according to σ and then to quantize the stimuli space according to their distribution. In the special case of one dimension, a weight vector is *organized* once $i \mapsto X_i$ is monotone. For a more detailed introduction to mathematical aspects of the Kohonen algorithm, see [7].

The stimuli consist in an i.i.d. sequence $(\omega^t_{t\geq 1}$ of $[0,1]$-valued random variables with a *continuous* distribution μ. Let $X^t := (X_i^t)_{1\leq i\leq n}$ denotes the weight vector at time t. At time $t+1$ the algorithm is recursively defined in two phases :

(i) COMPETITIVE PHASE : computation of the winning unit

$$i^{t+1} := i(\omega^{t+1}, X^t) = \underset{k\in I}{\operatorname{argmin}} |\omega^{t+1} - X_k^t|. \tag{1}$$

(ii) COOPERATIVE PHASE :

$$\forall j \in \{1, 2, \cdots, n\}, \quad X_j^{t+1} = X_j^t - \varepsilon_{t+1}\sigma(i^{t+1} - j)(X_j^t - \omega^{t+1}) \tag{2}$$

where $(\varepsilon_t)_{t\geq 1}$ is a sequence of $]0,1[$-valued real numbers. ε_t is the *learning rate* – or step – at time t.
$(X^t)_{t\in\mathbb{N}}$ is a Markov chain, homogeneous iff $\varepsilon_t = \varepsilon > 0$. Furthermore, if $x \in D := \{y \in [0,1]^n / y_i \neq y_j \text{ if } i \neq j\}$ then, $\mathbb{P}_x$-*a.s.*, $X^t \in D$ for every $t \in \mathbb{N}$. Thus, as soon as $X^0 \in D$ *a.s.*, the algorithm is *a.s.* well defined.

As, in the one dimensional setting, self organization means that the weight vector has monotone components, we define the two sets of possible organized states, say :

noindent $F_n^+ := \{x \in [0,1]^n, 0 < x_1 < x_2 < \cdots < x_n < 1\}$,

$F_n^- = \{x \in [0,1]^n, 0 < x_n < x_{n+1} < \cdots < x_1 < 1\}$

and $F_n := F_n^+ \cup F_n^-$.

2.2 Previous results and notations

The first studies were mainly devoted to proving the self organizing property. The results are the following :

SELF-ORGANIZATION (*a*) If $\sigma := \mathbf{1}_{\{|k|\leq 1\}}$ ("2 neighbors", see [5] or [3]) or if, more generally, $k \mapsto \sigma(k)$ is non increasing on NN (see [6] or [7]), then F_n^+ and F_n^- are absorbing sets.
(*b*) If the step is constant ($\varepsilon_t = \varepsilon$), then the hitting time τ_{F_n} of X^t in F_n is $\mathbb{P}_x$-*a.s.* finite and admits an exponential moment, uniformly w.r.t. $x \in [0,1]^n$, under the various assumptions that follow :

– 2 neighbors and $\mu = U([0,1])$ (original result, established in [5]).

– 2 neighbors and the support of the continuous part of μ has a nonempty interior (see [3]). The supports of such distributions obviously can be not connected.

– The neighborhood function is decreasing on $\{0, 1, \ldots, n-1\}$ (see [6]).

These results cover almost all the cases of interest.

CONVERGENCE After self-organization the algorithm *a.s.* lives in one of the absorbing subsets $F_n^{\pm}$ of F_n. That is why *a.s.* convergence has always been investigated once the algorithm lives – for instance – in F_n^+ and – for classical reasons – when the step ε_t is decreasing.

As usual, most works rely on the well-known *ODE* method for stochastic algorithms or, namely in the uniform setting, on a Robbins-Sigmund martingale approach. So, let us introduce the continuous time *ODE* associated with the discrete time algorithm. It reads :

$$\begin{aligned} \dot{x} &= -h(x) \quad \text{where } h : D \longrightarrow \mathbb{R}^n \text{ reads} \\ \forall x \in D, \qquad h_i(x) &= \sum_{k=1}^{n} \sigma(i-k) \int_{]\tilde{x}_k, \tilde{x}_{k+1}]} (x_i - \omega)\mu(d\omega), \ 1 \le i \le n, \end{aligned} \tag{3}$$

where, by convention : $\quad \tilde{x}_1 = 0_-, \quad \tilde{x}_k = \dfrac{x_k + x_{k-1}}{2}, \ 2 \le k \le n, \quad \tilde{x}_{n+1} = 1_+$

in the following sense : every function $g : [0,1] \to \mathbb{R}$ is assumed to have 0 left limit at 0 and 0 right limit at 1 (that is $g(0_-) = g(1_+) = 0$).

Here are the main existing results.

(*a*) If μ is continuous, the mean function of the algorithm, denoted by $-h$ (see (3)), admits a continuous extension on $\overline{F}_n^+$ and h has at least an equilibrium point (*i.e.* a zero) x^* in $\overline{F}_n^+$

(*b*) If μ is continuous, if $\text{supp}\mu = [0,1]$ and if the neighborhood function σ satisfies

$$(H_\sigma) \equiv \text{there exists } k_0 \le \tfrac{n-1}{2}, \text{ such that } \sigma(k_0+1) < \sigma(k_0)$$

then the equilibrium set $\{h = 0\}$ is included in F_n^+.

(*c*) If μ admits a density function $f > 0$ on $]0,1[$ satisfying :

$$(\mathcal{L}_\mu) \equiv \begin{cases} \bullet\ f \text{ is strictly } log\text{-concave on }]0,1[\\ \bullet \text{ or } f \text{ is } log\text{-concave on }]0,1[\text{ and } f(0_+) + f(1_-) > 0 \end{cases},$$

then h is Lipschitz on $\overline{F}_n^+$ and all equilibrium points are stable (*i.e.* have a stable attracting basin in the Kushner & Clark sense, (see [16] or subsection 6.2 for some background)). Consequently there are finitely many of them.

(d) If $\sum_t \varepsilon_t = +\infty$ and $\sum_t \varepsilon_t^2 < +\infty$ (the so-called "decreasing step" assumption), and

– if the assumptions of item (c) are fulfilled then the algorithm *conditionally* converges *in the K&C sense* toward an equilibrium (see [16] or subsection 6.2 for some background).

– if $\mu := U([0,1])$ then h admits a unique equilibrium x^* in F_n^+ and $X^t \to x^*$ $\mathbb{P}_x$-*a.s.*.

Claim (a) is established in [3] for the "2-neighbor" setting and in [7] in full generality. Claim (b) in its general form is due to Sadhegi in [20]. Claims (c) and (d) are proved in [7] under a slightly more restrictive assumption. The next section of the paper is devoted to extend it whenever (a) is established under (H_σ).

A convention on log-concave densities on $[0,1]$: A non negative log-concave density function defined on $[0,1]$ is continuous on $(0,1)$ has a right limit at 0 and a left limit at 1. So, one may assume that such a function is *continuous* on the closed unit interval by setting, if necessary, $f(0) := f(0_+)$ and $f(1) := f(1_-)$.

2.3 The main result

The aim of this whole work is to prove the following

Theorem 1 *If μ satisfies condition $(\mathcal{L}_\mu)$, if σ satisfies (H_σ) for some $k_0 < \frac{n-1}{2}$, if $X^0 \in \overline{F}_n^+$ then X^t converges $\mathbb{P}_x$ a.s. to the unique equilibrium point x^*, unique zero of h in $\overline{F}_n^+$.*

To prove Theorem 1, we proceed in three steps :

1- first we prove that there is a unique equilibrium point x^* for the *ODE* $\dot{x} = -h(\sigma, x)$,

2- then we verify the assumptions of Hirsch's Theorem about the strongly monotone dynamical systems,

3- and finally we apply the celebrated Kushner & Clark Theorem to conclude.

Notice that we will need some extension of the definition of the algorithm in order to include the case where X^t belongs to ∂F_n^+. This will be done when needed (see formula (9) below).

The first step essentially relies on the asymptotical stability of all equilibrium points under the assumption (H_σ) (see section 3).

3 Stability of the equilibrium points

Here we extend the result in [7] from the stronger assumption $\sigma(3) < \sigma(0)$ to the weaker assumption (H_σ).

Proposition 1 *If assumption $(\mathcal{L}_\mu)$ and (H_σ) hold then any equilibrium point x^* (lies in F_n^+ and) is stable in the following sense*

$$\forall\, \lambda \in \mathrm{Sp}(\nabla h(x^*)), \qquad \Re(\lambda) > 0$$

where $\mathrm{Sp}(A) := \{\text{eigenvalues of } A\}$. *$x^*$ is then stable in the K&C sense (see section 6.2).*

So, the basic ingredient of this section is the expression of the gradient ∇h when the distribution μ has a continuous density function f on $(0,1)$. This computation was carried out in [7]. W.l.g. on the density function f, we will adopt throughout the text the following convention :

$$f(0_-)=0 \text{ (hence } f(\tilde{x}_1)=0) \text{ and } f(1_+)=0 \text{ (hence } f(\tilde{x}_{n+1})=0).$$

Lemma 1 *If the density f is continuous on $(0,1)$ then h is continuously differentiable on F_n^+ and $\nabla h(x) := Diag[\zeta_1, \cdots, \zeta_n] + [\alpha_{ij}]_{1\le i,j\le n}$ where*

$$\zeta_i := \sum_{k=1}^{n} \sigma(i-k) \int_{\tilde{x}_k}^{\tilde{x}_{k+1}} f(\omega)d\omega,\ 1 \le i \le n, \tag{4}$$

$$\alpha_{ij} := \frac{\sigma(i+1-j)-\sigma(i-j)}{2}(x_i-\tilde{x}_j)f(\tilde{x}_j)+\frac{\sigma(i-j)-\sigma(i-j-1)}{2}(x_i-\tilde{x}_{j+1})f(\tilde{x}_{j+1}).$$

The rest of the section is devoted to the proof of Proposition 1. Actually it is essentially an improvement of that in [7]. Before going into technicalities, let us say that the approach basically consists in showing that the real parts of the eigenvalues of the $\nabla h(x^*)$ are positive. It relies on the the celebrated Gershgorin Lemma on matrices with dominating diagonal and one of its variant (that can be found *e.g.* in [7]).

Lemma 2 *Let $A := [a_{ij}]_{1\le i,j\le n}$ be a real valued matrix.*
(a) Classical Gershgorin Lemma *(see e.g. [11] for a proof) : if A satisfies*
$\forall\, i \neq j,\ a_{ij} \le 0$ *and* $\forall\, i, \sum_j a_{ij} > 0,$

then : $$\forall\, \lambda \in \mathrm{Sp}(\nabla h(x^*)), \qquad \Re(\lambda) > 0.$$

(b) Extended Gershgorin Lemma : *(see e.g. [7] for a proof) : Assume there is some $p \in \{1, \cdots, n-1\}$ satisfying*
(i) $\forall\, i \neq j,\ a_{ij} \le 0$ *and* $\forall\, i, \sum_j a_{ij} \ge 0,$

(ii) $a_{i,i\pm p} < 0$ *whenever* $i \pm p \in \{1, \ldots, n\}$,
(iii) $\exists\, i_1, \cdots, i_p \in \{1, \cdots, n\}$ *s.t.* $i_k \equiv k$ *mod. p* *and* $\sum_j a_{i_k j} > 0$.
Then $\forall\, \lambda \in \mathrm{Sp}(\nabla h(x^*)), \quad \Re(\lambda) > 0.$

Following Lemma 2, it turns out that our main task is to investigate, for every $x \in F_n^+$, the sign of components of $\nabla h(x)$ and of their sum $L_i(x) := \sum_{j=1}^{n} \frac{\partial h_i}{\partial x_j}(x)$.

Lemma 3 *(a)* $\forall\, x \in F_n^+$, $\forall\, i \neq j$, $\frac{\partial h_i}{\partial x_j}(x) \leq 0$,
(b) If $f > 0$ *on the open unit interval* $(0,1)$ *and assumption* (H_σ) *holds, then*

$$\forall\, i \in \{1, \ldots, n\},\ \forall\, x \in F_n^+,\ \frac{\partial h_i}{\partial x_{i\pm(k_0+1)}}(x) < 0 \quad \text{where} \quad \sigma(k_0+1) < \sigma(k_0)$$

(At least one of these partial derivatives does exist).

Proof : One first notices that matrix a in equation (4) is the sum $a := [a_{ij}] + [a_{i,j+1}]$ where $a_{ij} = \frac{\sigma(i+1-j) - \sigma(i-j)}{2}(x_i - \tilde{x}_j) f(\tilde{x}_j)$. Now $a_{ij} \leq 0$ (and $a_{i1} = a_{i,n+1} = 0$) which implies item (a).
Claim (b) follows from the fact that

$$j = i + k_0 + 1 \implies \sigma(i+1-j) - \sigma(i-j) = \sigma(k_0) - \sigma(k_0+1) \neq 0.$$
$$j = i - (k_0 + 1) \implies \sigma(i-j) - \sigma(j-i+1) = \sigma(k_0+1) - \sigma(k_0) \neq 0.$$

This completes the proof. $\diamond$

Proof of Proposition 1 (step 1) : As we are only concerned here by the gradient at x^*, we may use the equilibrium equation $h(x^*) = 0$:

$$x_i^* = \frac{\sum_{j=1}^{n} \sigma(i-j) \int_{\tilde{x}_j^*}^{\tilde{x}_{j+1}^*} \omega f(\omega) d\omega}{\sum_{j=1}^{n} \sigma(i-j) \int_{\tilde{x}_j^*}^{\tilde{x}_{j+1}^*} f(\omega) d\omega},\ 1 \leq i \leq n. \tag{5}$$

Thus, plugging (5) into the expression (4) of $\nabla h(x^*)$ finally yields

$$\forall\, 1 \leq i \leq n, \quad L_i(x^*) := \frac{D_i(x^*)}{\sum_{j=1}^{n} \sigma(i-j) \int_{\tilde{x}_j^*}^{\tilde{x}_{j+1}^*} f(\omega) d\omega}, \tag{6}$$

where $D_i(x^*)$ is given (keeping in mind the conventions $\tilde{x}_1 = 0_-$, $\tilde{x}_{n+1} = 1_+$) by

$$D_i(x^*) := \left(\sum_{j=1}^{n}\sigma(i-j)\int_{\tilde{x}_j^*}^{\tilde{x}_{j+1}^*} f(\omega)d\omega\right)^2 - \sum_{j=1}^{n-1}(\sigma(i-(j+1))-\sigma(i-j))f(\tilde{x}_{j+1}^*)\sum_{k=1}^{n}\sigma(i-k)\int_{\tilde{x}_k^*}^{\tilde{x}_{k+1}^*}(\omega-\tilde{x}_{j+1}^*)f(\omega)d\omega. \tag{7}$$

Equation (7) shows that the problem is equivalent to specifying if $D_i(x^*)$ is 0 or not and, if not, its sign.

To this end, we introduce some auxiliary functions φ_i^n, $1 \le i \le n$ on $\overline{F}_{n+1}^+$:

$$\forall u \in \overline{F}_{n+1}^+,\ \varphi_i^n(u) := \psi_i^n(u) - \sum_{k=1}^{n+1}\tau(k,i)f(u_k)\sum_{j=1}^{n}\sigma(i-j)\int_{u_j}^{u_{j+1}}(\omega-u_k)f(\omega)d\omega$$

$$\text{where}\quad \begin{cases} \tau(k,i) & := \sigma(i-k)\mathbf{1}_{\{k\le n\}} - \sigma(i+1-k)\mathbf{1}_{\{k\ge 2\}} \\ \psi_i^n(u) & := \left(\displaystyle\sum_{k=1}^{n}\sigma(i-k)\int_{u_k}^{u_{k+1}} f(\omega)d\omega\right)^2. \end{cases}$$

Equation (7) now reads, for every $i \in \{1,\dots,n\}$,

$$\begin{aligned} D_i(x^*) &= \varphi_i^n(\tilde{x}^*) + f(0)\sigma(i-1)\sum_{k=1}^{n}\sigma(i-k)\int_{\tilde{x}_k^*}^{\tilde{x}_{k+1}^*}\omega f(\omega)d\omega \\ &\quad + f(1)\sigma(i-n)\sum_{k=1}^{n}\sigma(i-k)\int_{\tilde{x}_k^*}^{\tilde{x}_{k+1}^*}(1-\omega)f(\omega)d\omega. \end{aligned}$$

(still using the conventions $\tilde{x}_1 = 0_-$ and $\tilde{x}_{n+1} = 1_+$, $f(0) := f(0_+)$, $f(1) := f(1_-)$). ◇

To get the sign of the φ_i^n functions we will specify the (partial) sign structure of some of their partial derivatives. The next lemma relies on the well-known fact that a (positive) log-concave function has right derivatives (a little algebra is left to the reader).

Lemma 4 *If the density f is* log-*concave, φ_i^n has right first partial derivatives on F_{n+1}^+ and some right cross second partial derivatives, namely*
(a) For every $\ell \in \{1,\dots,n+1\}$,

$$\frac{\partial_+\varphi_i^n}{\partial u_\ell}(u) = -\psi_i^n(u)f(u_\ell)\tau(\ell,i) + \left(\sum_{k=1}^{n+1}\tau(k,i)f(u_k)(u_\ell-u_k)\right)f(u_\ell)\tau(\ell,i) - \cdots$$

$$\tau(\ell,i)f'_+(u_\ell)\left(\sum_{j=1}^{n}\sigma(i-j)\int_{u_j}^{u_{j+1}}(\omega-u_\ell)f(\omega)d\omega\right)$$

where the subscript "+" denotes right derivative.

(b) For every ℓ, $m \in \{1, \ldots, n+1\}$, $\ell \neq m$,

$$\frac{\partial^2_+ \varphi_i^n}{\partial u_\ell \partial u_m}(u) = \tau(\ell, i)\tau(m, i)(u_\ell - u_m)\left(f'_+(u_m)f(u_\ell) - f'_+(u_\ell)f(u_m)\right). \tag{8}$$

The following two lemmas show that the above partial derivatives (equation (8)) are sufficient to specify the sign of φ_i^n. They are proved in [7].

Lemma 5 *(a)* $\forall u \in [0, 1], \quad \varphi_i^n(u, \cdots, u) = 0,$

$\forall u \in (0, 1), \forall \ell \in \{1, \cdots, n+1\}, \quad \dfrac{\partial_+ \varphi_i^n}{\partial u_\ell}(u, \cdots, u) = 0.$

(b) Assume that $f > 0$ *on* $(0, 1)$*. If* $\log f$ *is concave and* $u \in \overline{F}^+_{n+1} \cap (0, 1)^{n+1}$ *then the symmetrical matrix* $\left[\dfrac{\partial^2_+ \varphi_i^n}{\partial u_\ell \partial u_m}(u)\right]_{\ell, m \in \{1, \cdots, n+1\}}$ *has a sign structure given by :*

$$\begin{array}{c} \\ \\ \\ \\ i+1 \\ \\ \\ \\ \end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & & & i+1 & & & n+1 \end{array} \\
\left[\begin{array}{ccccccc}
\times & & \geq 0 & \vdots & & & \\
 & \times & & \vdots & & \leq 0 & \\
\geq 0 & & \times & \vdots & & & \\
\cdots & \cdots & \cdots & \times & \cdots & \cdots & \cdots \\
 & & & \vdots & \times & & \geq 0 \\
 & \leq 0 & & \vdots & & \times & \\
 & & & \vdots & \geq 0 & & \times
\end{array}\right]
\end{array},$$

the sign of the diagonal terms being unknown a priori.

Lemma 6 *Assume that* $f > 0$ *on* $(0, 1)$ *and* $\log f$ *is concave. Then*
(a) $\forall u \in \overline{F}^+_{n+1}, \ \varphi_i^n(u) \geq 0.$
(b) If $\log f$ *is strictly concave and* $\sigma(1) < \sigma(0)$ *then*

$$\forall u \in \overline{F}^+_{n+1}, \ u_i < u_{i+1} \Longrightarrow \varphi_i^n(u) > 0.$$

Proof of Proposition 1 (step 2) :
LOG-CONCAVE ASSUMPTION : For a proof that x^* belongs to F_n^+, see [20] or Lemma 8(c) below. The proof relies on equation (7) :

$$\begin{aligned}
D_i(x^*) = \underbrace{\varphi_i^n(\tilde{x}^*)}_{\geq 0} \ &+ \ f(0)\sigma(i-1)\sum_{k=1}^{n}\sigma(|k-i|)\int_{\tilde{x}_k^*}^{\tilde{x}_{k+1}^*} \omega f(\omega)d\omega \\
&+ \ f(1)\sigma(n-i)\sum_{k=1}^{n}\sigma(|k-i|)\int_{\tilde{x}_k^*}^{\tilde{x}_{k+1}^*}(1-\omega)f(\omega)d\omega.
\end{aligned}$$

It follows that all the $D_i(x^*)$'s are non-negative and, furthermore,

$$f(0) > 0 \implies D_i(x^*) > 0,\ 1 \leq i \leq k_0 + 1,$$
$$f(1) > 0 \implies D_i(x^*) > 0,\ n - k_0 \leq i \leq n,$$

since $\sigma(k_0 + 1) < \sigma(k_0)$. Lemma 3 implies that $\frac{\partial h_i}{\partial x_{i\pm(k_0+1)}} < 0$. Hence, as $L_j(x^*) > 0$ iff $D_j(x^*) > 0$, one derives from Lemma 2(a) (with $p = k_0 + 1$) that all the eigenvalues of $\nabla h(x^*)$ have positive real parts.

STRICT LOG-CONCAVE ASSUMPTION : The aim is now to fulfill the classical Gershgorin Lemma. Let k_0 be defined by assumption (H_σ). Then $\sigma(0) = \ldots = \sigma(k_0) > \sigma(k_0 + 1)$.

Step 1 ($k_0 = 0$) : then $\sigma(1) < \sigma(0)$ since $\tilde{x}_i^* < \tilde{x}_{i+1}^*$, Lemma 6(b) implies $D_i(\tilde{x}^*) > 0$ for every i.

Step 2 (*general case*) : One sets $\sigma_0(i) = \sigma(i + k_0)$ for every $i \in \{0, \ldots, n - k_0\}$ (and $\sigma_0(n - k_0 + 1) = 0$). One checks after some tedious but elementary computations that, for every $u := (u_1, \ldots, u_{n+1}) \in \overline{F}^+_{n+1}$,

$$\begin{aligned}
\forall i \in \{1, \ldots, k_0\},\ \varphi_i^n(u) &= \varphi_{0,1}^{n-(i+k_0)}(u_1, u_{i+k_0+1}, \ldots, u_{n+1}) \\
\forall i \in \{k_0 + 1, \ldots, n - (k_0 + 1)\},\ \varphi_i^n(u) &= \varphi_{0,i-k_0}^{n-2k_0}(u_1, \ldots, u_{i-k_0}, u_{i+k_0+1}, \ldots, u_{n+1}), \\
\forall i \in \{n - k_0, \ldots, n\},\ \varphi_i^n(u) &= \varphi_{0,i-k_0}^{i-k_0+1}(u_1, \ldots, u_{i-k_0}, u_{n+1}),
\end{aligned}$$

where $\varphi_{0,i}^n$ denote the φ_i^n function related to the σ_0 neighborhood function (and its restrictions to smaller unit sets). Finally, case 1 yields the expected result, that is,

$$\forall u \in F^+_{n+1}, \qquad \varphi_i^n(u) > 0,$$

which in turn implies, still using equation (7), that, $D_i(x^*) > 0, 1 \leq i \leq n$.

This time the classical Gershgorin Lemma and Lemma 3(a) complete the proof : the eigenvalues of $\nabla h(x^*)$ have positive real part, so x^* is stable. $\diamond$

4 Uniqueness of the equilibrium point

We recall the celebrated Poincaré-Hopf's Theorem (see [12]) :

Theorem 2 *(Poincaré-Hopf) Let V be a C^0 manifold with boundary ∂V and let f be a C^0 vector field on V pointing (strictly) outside V on ∂V, having finitely many zeros (necessarily) inside V. Then the sum of all the Morse indices of the zeros of f is equal to the Euler characteristics of V.*

Theorem 3 *If μ satisfies condition $(\mathcal{L}_\mu)$, if σ satisfies (H_σ) then the ODE has a unique equilibrium point.*

Lemma 7 *If μ is continuous and supp(μ) = [0, 1] and if the neighborhood function σ is decreasing on $\{0, \ldots, n-1\}$, the vector field h is pointing (strictly) outside $\overline{F}_n^+$ on ∂F_n^+.*

Proof : Let $x \in \partial F_n^+$ and the integral interval $[i_0, i_1]$ be a cluster of x that is $x_{i_0} = x_{i_0+1} = \ldots = x_{i_1}$. Then $\tilde{x}_i = x_{i_0}$, $i_0+1 \leq i \leq i_1$. A little algebra yields, that, for every $\ell \in \{i_0+1, \ldots, i_1\}$

$$\begin{aligned} h_\ell(x) - h_{i_0}(x) &= \underbrace{(\sigma(0) - \sigma(\ell - i_0))}_{>0} \underbrace{\int_{\tilde{x}_{\ell+1}}^{\tilde{x}_{\ell+1}} (x_{i_0} - \omega)\mu(d\omega)}_{<0 \text{ if } \ell = i_1,\ 0 \text{ otherwise}} \\ &+ \sum_{i=i_1+1}^{n} \underbrace{(\sigma(i-\ell) - \sigma(i-i_0))}_{>0} \underbrace{\int_{\tilde{x}_i}^{\tilde{x}_{i+1}} (x_{i_0} - \omega)\mu(d\omega)}_{<0} \\ &+ \sum_{i=1}^{i_0} \underbrace{(\sigma(\ell - i) - \sigma(i_0 - i))}_{<0} \underbrace{\int_{\tilde{x}_i}^{\tilde{x}_{i+1}} (x_{i_0} - \omega)\mu(d\omega)}_{>0} < 0. \end{aligned}$$

This shows that the vector field $h(x)$ is pointing outside the vector subspace vec$\{e_\ell - e_{i_0},\ i_0 < \ell \leq i_1\}$ where $(e_1, \ldots, e_n)$ denotes the canonical basis of $\mathbb{R}^n$. The same property holds for every cluster of $x \in \partial F_n^+$, subsequently $h(x)$ is pointing outside (all) the tangent hyperplane(s) of the convex set $\overline{F}_n^+$ at x as long as $0 < x_1 le \ldots le x_n < 1$. The cases $x_1 = 0$ or $x_n = 1$ follow the same way round. $\diamond$

Proof of theorem 3 : The canonical closed simplex $\overline{F}_n^+$ is homeomorphic to the unit closed disk $\overline{D}_n$: namely, let $x_0 \in F_n^+$ be a fixed point. As $\overline{F}_n^+$ is a closed convex set, we can define

$$\lambda(x) := \sup\{s \geq 0 \ / \ x_0 + s(x - x_0) \in \overline{F}_n^+\}.$$

The map $\psi(x) := \frac{x - x_0}{\lambda(x)}$ is then an homeomorphism from $\overline{F}_n^+$ onto the disk $\overline{D}_n$. Thus, the Euler characteristics of $\overline{F}_n^+$ is 1. We know from Proposition 1 (Section 3) that every zero x^* of h have a stable attracting basin such that all the eigenvalues of $\nabla h(x^*)$ have positive real part. Hence, $\det(\nabla h(x^*))$ is positive and the Morse index of x^* is one.

Case 1 (σ decreasing) : It straightforwardly follows from Lemma 7 that the sum of the Morse indices of the zeros of h is 1. This means that there is exactly one equilibrium point for the O.D.E..

Case 2 (extension to general σ functions) : At this stage we need to mention the σ-dependency of the mean function h and we write $h(\sigma, x)$ instead of $h(x)$. The function $h : \mathbb{R}^n \times F_n^+ \longrightarrow F_n^+$ is clearly C^1 and its partial derivative $\dfrac{\partial h(\sigma, x^*)}{\partial x} = \nabla_x h(\sigma, x^*)$

is invertible at each equilibrium "couple" (σ, x^*) such that σ satisfies (H_σ) . Thus the implicit functions Theorem applies.

Let σ_0 be a given non-increasing neighborhood function that fulfills (H_σ) and assume there are two different solutions x_1^* and x_2^* of the equilibrium equation $h(\sigma_0, x) = 0$ (necessarily in F_n^+, see claim (b) of uniqueness and convergence, subsection 2.2). There exist an open neighborhood $\mathcal{V}(\sigma_0)$ of σ_0 in $\mathbb{R}^n$, two open neighborhoods $\mathcal{W}(x_i^*)$ of x_i^* and two C^1 functions ϕ_i defined on $\mathcal{W}(x_i^*)$, $i=1,2$ such that, for every $i=1,2$:

$$h(\sigma, x) = 0, \ (\sigma, x) \in \mathcal{V}(\sigma_0) \times \mathcal{W}(x_i^*) \quad \text{iff} \quad x = \phi_i(\sigma).$$

First note that the open set $\mathcal{V}(\sigma_0) \cap \{\sigma \,/\, k \mapsto \sigma(k) \text{ is decreasing}\}$ is nonempty since $\sigma_0 \in \overline{\{\sigma \,/\, k \mapsto \sigma(k) \text{ is decreasing}\}}$. Now, let $(\sigma_n)_{n\geq 1}$ be a sequence in $\mathcal{V}(\sigma_0) \cap \{\sigma \,/\, k \mapsto \sigma(k) \text{ is decreasing}\}$ that converges to σ_0. Then $\phi_i(\sigma_n)$ converges toward x_i^*. Now, the uniqueness result in case 1 implies that $\phi_1(\sigma_n) = \phi_2(\sigma_n)$ for every $n \geq 1$, which, in turn, implies that $x_1^* = x_2^*$. This yields a contradiction. So Theorem 3 is proved. $\diamond$

5 Cooperative dynamical systems and applications to S.O.M.

We follow here the definitions and results of M. Hirsch about cooperative dynamical systems. Although "cooperation" and "competition" are two keywords in the very definition of the Kohonen S.O.M. – and actually in many Neural models – the link beyond the coincidence of the words with the cooperative systems has been made very recently. A. Sadeghi noticed that the non-negativity of off-diagonal terms of $-\nabla h$ is exactly the basic definition of a cooperative dynamical system.

Unfortunately, *all* the assumptions required to apply Hirsch's convergence flow Theorem are not satisfied in full generality by the Kohonen S.O.M. maps. Namely the kind of "strong" irreducibility assumption is not, except when σ does decrease. So we had to slightly weaken this assumption in the original Hirsch's Theorem to apply it under the standard "probabilistic" irreducibility assumption for matrices with non-negative coefficients.

5.1 General background

Definition 1 *An autonomous dynamical system $\dot{x} = f(x)$ defined on a convex open set W of $\mathbb{R}^n$ is cooperative if the vector field f is C^1 and its gradient matrix $\nabla f(x)$ has non-negative off-diagonal entries, that is* $\frac{\partial f_i}{\partial x_j}(x) \geq 0$, $i \neq j$.

Definition 2 *(a) A $n \times n$ matrix A is irreducible if one of the following equivalent properties holds*

(a) for every nonempty proper subset I of $\{1,\dots,n\}$ there is an $i \in I$ and $j \in^c I$ such that $a_{ij} \neq 0$.

(b) for every couple $(i,j) \in \{1,\dots,n\}^2$, there exists $r(i,j) \in \mathbb{N}^*$ such that $(|A|^{r(i,j)})_{ij} \neq 0$ where $|A| := [|a_{ij}|]$.

Put some way round, this definition means that there is a path from i to j on the graph defined by the nonzero coefficients of A (the edge $k \to l$ exists iff $A_{kl} \neq 0$). Item (a) is the original definition used by Hirsch in [13] and Smith [21] while item (b) is the usual probabilistic definition of irreducibility of stochastic matrices.

The *vector order* on $\mathbb{R}^n$ is denoted $\leq$ where $x \leq y$ means that $x_i \leq y_i$ for all i. If $x \leq y$ and $x \neq y$ then we write $x < y$. If $x_i < y_i$ for all i then we write $x << y$.

The following theorem, originally due to Hirsch, was first proved for a more restrictive irreducibility assumption ([12]). It can be found in the recent monograph by H. L. Smith (see [21]) with the standard probabilistic definition of irreducibility.

Theorem 4 *(Hirsch) Let $\dot{x} = f(x)$ be a cooperative dynamical system defined on the open set W. If $\nabla f(x)$ is irreducible then the flow $\{\Phi_t\}$ ([1]) induced by f is well defined and has positive partial derivatives : $\frac{\partial \Phi_{t,i}}{\partial x_j}(x)$ for all $t > 0$ and every $1 \leq i, j \leq n$.*

Then, he derives

Theorem 5 *(Hirsch) Let W be a convex open set. If the flow $\{\Phi_t\}$ of $\dot{x} = f(x)$ is well defined and has positive derivatives then it is strongly monotone on W in the following sense*

$$x < y \Longrightarrow \Phi_t(x) << \Phi_t(y)$$

for all $t > 0$ i.e. $\Phi_{t,i}(x) < \Phi_{t,i}(y)$ for every i.

We may now state a last theorem by Hirsch in the specific case where the dynamical system does have a *unique* equilibrium point x^* in W.

Theorem 6 *(Hirsch) Assume that the flow $\{\Phi_t\}$ is strongly monotone on a convex open set W and lives in W. If every orbit has compact closure in W and if there is only one equilibrium point $x^* \in W$, then*

$$\forall x \in W, \qquad \lim_{t \to +\infty} \Phi_t(x) = x^*.$$

[1]The flow $\Phi_t(x)$ is defined as a mapping from $W \times \mathbb{R}_+ \longrightarrow W$ satisfying the relation : $\forall t \in \mathbb{R}_+, \forall x \in W, \dfrac{\partial \Phi_t(x)}{\partial t} = f(\Phi_t(x))$.

5.2 Application to the Kohonen S.O.M.

To apply Theorem 6 to the Kohonen dynamical system, we need to prove that the $ODE \equiv \dot{x} = (-h)(x)$ is cooperative, irreducible (probabilistic sense) and that all trajectories issued from $\overline{F}_n^+$ live in F_n^+ whenever $t > 0$ along with all their limit points. Hence the theorem will work for trajectories starting from ∂F_n^+ which will be crucial for the transfer of the converging property to the algorithm.

Proposition 2 (a) *If the distribution μ has a continuous density on $(0,1)$ and the neighborhood function σ is non-increasing, then the Kohonen S.O.M. ODE is cooperative on F_n^+.*
(b) *Assume that $(\mathcal{L}_\mu)$ and (H_σ) hold. Then, $\nabla h(x)$ is irreducible at every $x \in F_n^+$ iff assumption (H_σ) holds for some $k_0 < \frac{n-1}{2}$.*

Proof : (a) straightforwardly follows from the expression of $\nabla h(x)$ provided in Lemma 1 and its sign structure from Lemma 3.
(b) $(\Rightarrow)$ If $n = 2k_0+1$ (and $\sigma(k_0) = \sigma(0) = 1$) the non diagonal terms of the (k_0+1)-th line of $\nabla h(x)$ are all zero. Hence $\nabla h(x)$ cannot be irreducible.
$(\Leftarrow)$ assume now that $k_0 < \frac{n-1}{2}$. Set

$$A := [a_{ij}]_{1\leq i,j\leq n},\ a_{ij} := |\sigma(|i+1-j|) - \sigma(|i-j|)|\,\mathbf{1}_{\{2\leq j\leq n\}}.$$

On F_n^+, the irreducibility of $\nabla h(x)$ is clearly equivalent to that of $B := [a_{ij} + a_{i,j+1}]_{1\leq i,j\leq n}$ where $a_{i,n+1} := 0$.

Now, $a_{ij} = |\sigma(k_0+1) - \sigma(k_0)| > 0$ whenever $j \geq 2$ and $i - j = k_0$ or $-(k_0+1)$ and is always non-negative. Let $i \in \{1,\ldots,n\}$.

$$\text{if}\quad i \leq n-(k_0+1) \text{ then}\qquad A^2_{i,i+1} \geq a_{i,\underbrace{i+k_0+1}_{\leq n}}\, a_{i+k_0+1,\underbrace{i+1}} > 0$$

$$\text{if}\quad i \geq n-(k_0+1) \text{ then}\qquad A^2_{i,i+1} \geq a_{i,\underbrace{i-k_0}_{\geq 2}}\, a_{i-k_0,i+1} > 0.$$

Subsequently, if $i_0,\ j_0 \in \{1,\ldots,n\}$, $i_0 < j_0$, the above sign structure of A implies that

$$i_0 \xrightarrow{A^2} i_0+1 \xrightarrow{A^2} \ldots \xrightarrow{A^2} j_0.$$

Symmetrically, one shows that $A' := [a_{i,j+1}]$ satisfies $A'^2_{i,i-1} > 0$ which in turn implies that, if $i_0,\ j_0 \in \{1,\ldots,n\}$, $i_0 > j_0$, then A'^2 will lead from i_0 to j_0 in $i_0 - j_0$ shots.

The irreducibility of B follows from the obvious inequality $B_{ij} \geq \max(A_{ij}, A'_{ij})$. $\diamond$

Remark : The above proof shows that a $\{1,\ldots,n\}$-valued Markov chain jumping with positive probability from i to $i+k_0$, $i+k_0+1$, $i-k_0$ or $i-(k_0+1)$ (when it is possible) is irreducible.

Our last task in that subsection is to check that the flow of the ODE has no limiting point on ∂F_n^+ (even when it starts from it).

Proposition 3 *Assume that the distribution μ has a bounded density, continuous and positive on $(0,1)$([2]) and that (H_σ) holds. For every $x \in F_n^+$, the set of the limiting values of the flow $\{\Phi_t(x)\}$ of the S.O.M. ODE is a compact connected set of F_n^+.*

This proposition relies on a "parting" lemma for the mean function h and a modified version of the algorithm : we need to define an extension of the *algorithm* itself when starting from the boundary of F_n^+ that complies with the continuous extension of the mean function h on ∂F_n^+ (see equation (3)) : so far the algorithm is only defined when $x \in D$.

EXTENSION OF THE ALGORITHM ON ∂F_n^+ : At time $t+1$, one has

(i) *Competitive phase :* computation of the winning unit :

$$\text{If } \begin{cases} \omega^{t+1} \geq \operatorname{argmin}_{1\leq k\leq n}|\omega^{t+1} - X_k^t| & \text{then } i(\omega^{t+1}, X^t) := \max \operatorname{argmin}_{1\leq k\leq n}|\omega^{t+1} - X_k^t|, \\ \omega^{t+1} \leq \operatorname{argmin}_{1\leq k\leq n}|\omega^{t+1} - X_k^t| & \text{then } i(\omega^{t+1}, X^t) := \min \operatorname{argmin}_{1\leq k\leq n}|\omega^{t+1} - X_k^t| \end{cases} \tag{9}$$

(ii) *Cooperative phase :* unchanged.

This new algorithm $\mathbb{P}_x$-*a.s.* coincides with the original one (see Lemma 8) whenever $x \in F_n^+$ ($F_n^+ \subset D$ and $\partial F_n^+ \subset {}^cD$). Furthermore, one readily checks the following important facts

Proposition 4 (a) *The closed simplex $\overline{F}_n^+$ is left stable by the extended algorithm (9).*

(b) *The Markov transitions defined by equation (9) that is*

$$P_{t,t+1}(f)(x) := \int_{[0,1]} f(x - \varepsilon_{t+1}H(x,\omega))\mu(d\omega)$$

are Feller on $\overline{F}_n^+$, whenever the stimuli distribution μ is continuous (i.e. weights no single point).

(c) *The mean function of the extended algorithm, still denoted $h(x) = \mathbb{E}_x(H(x,\omega^1))$, coincides with the continuous extension of the original algorithm on $\overline{F}_n^+$.*

Remark : One must take care that such an extension is only valid on $\overline{F}_n^+$. The algorithm has a Feller extension $\overline{F}_n^-$ which is not compatible with (9). No Feller version of the Kohonen S.O.M. can be defined on the original state space $[0,1]^n$ as it is contradictory with self-organization.

[2]This is especially satisfied when μ has a log-concave density

Lemma 8 *(Parting lemma) Assume that* (H_σ) *holds, that* μ *is continuous with* $supp(\mu)=[0,1]$.

(a) Set $\mathcal{E}(x) := \mathrm{card}\{x_i,\ 1\le i\le n\}\in\{1,\dots,n\}$ *(number of distinct values of the components of the vector* x *so that* $\mathcal{E}(x)<n$ *iff* $x\in\partial F_n^+$*). Then, for every* $x\in\partial F_n^+$, $\mathcal{E}(X^1)\ge\mathcal{E}(x)$ $\mathbb{P}_x$*-a.s. and* $\mathbb{P}_x(\mathcal{E}(X^1)>\mathcal{E}(x))>0$. *Furthermore,* $\mathcal{N}_a:=\mathrm{card}\{i\,/\,x_i=a\}$, $a=0,1$, *satisfy the same property.*

(b) For every $x\in\partial F_n^+$, *for every* $i\in\{1,\dots,n\}$ *such that* $x_i=x_{i+1}$, $h_i(x)-h_{i+1}(x)\ge 0$ *and there exists at least one* i_0 *such that* $x_{i_0}=x_{i_0+1}$ *and* $h_{i_0}(x)-h_{i_0+1}(x)>0$.

(c) The set of zeros of the continuous extension of h *on* $\overline{F}_n^+$ *is contained in* F_n^+.

Proof : (a) For the sake of understanding we will only deal with $\mathcal{E}$. The (preliminary) treatment of $\mathcal{N}_0$ and $\mathcal{N}_1$ is in fact simpler. So, we assume that $x\in\partial F_n^+\cap(0,1)^n$ is fixed. $\mathbb{P}_x$-*a.s.*, $X^1\in\overline{F}_n^+\cap(0,1)^n$, hence $\mathcal{E}(X^1)\ge\mathcal{E}(x)$, $\mathbb{P}_x$-*a.s.*.

Let $[i_1,i_2]$ be an integral interval whose components made up a full cluster *i.e.* $x_{i_1-1}<x_{i_1}=x_{i_2}<x_{i_2+1}$.

Case 1 $(i_2-i_1\ge k_0+2)$ *:* One may assume w.l.g. that $x_{i_2}<1$. Then, picking up the stimulus ω^1 in $(x_{i_2},\tilde{x}_{i_2+1})$ will split the cluster $[i_1,i_2]$ into two smaller clusters : $\mathcal{E}(X^1)>\mathcal{E}(x)$ on that event whose probability is not zero since $\mathrm{supp}(\mu)=[0,1]$.

Case 2 $(i_2-i_1\le k_0+1)$ *:* If there is some cluster $[j_1,j_2]$ such that $i_1+k_0+1\le j_1\le i_2+k_0+1$, then, on the event $\{\omega^1\in(\tilde{x}_{j_1},x_{j_1})\}$, the cluster $[i_1,i_2]$ is split, hence $\mathcal{E}(X^1)>\mathcal{E}(x)$.

If no such cluster exists, that means that the indices $i_1+k_0+1,\dots,i_2+k_0+1$ (lower than n) are never the lowest point of their cluster. That means that they are all included in the same one, say $[k_1,k_2]$, $i_2<k_1\le i_1+k_0$. Now, picking up ω^1 in $(\tilde{x}_{i_1},x_{i_1})$ will split the cluster $[k_1,k_2]$... if it exists.

If $[k_1,k_2]$ does not exist, that means that $i_1+k_0>n$; but, then, $i_1>\frac{n+1}{2}>k_0+1$. The same method as above leads to splitting $[i_1,i_2]$ with positive probability using a "lower" cluster $[\ell_1,\ell_2]$ s.t. $i_1-(k_0+1)\le\ell_2\le i_2-(k_0+1)$ or to splitting a lower cluster $[m_1,m_2]$, $i_2-k_0\le m_2<i_1$ and $m_1\le\min(i_1-(k_0+1),1)$: this time $[m_1,m_2]$ necessarily exists and is a true cluster.

(b) The inequalities "$\ge$" on the boundary follow from the fact that the continuous extension of h on $\overline{F}_n^+$ satisfies *e.g.* $(Id-\frac{1}{2})h(\overline{F}_n^+)\subset\overline{F}_n^+$ which in turn follows from the stability of F_n^+ itself by the same application.

The existence of i_0 follows from claim (a). Indeed, as $\mathbb{P}_x(\mathcal{E}(X^1)>\mathcal{E}(x))>0$, there is at least one i_0 such that $x_{i_0}=x_{i_0+1}$, $X^1_{i_0}\le X^1_{i_0+1}$ and $\mathbb{P}_x(X^1_{i_0}<X^1_{i_0+1})>0$. These inequalities yield, after proper integration, that $h_{i_0+1}(x)-h_{i_0}(x)>0$.

(c) This point is obvious, given (b). $\diamond$

Proof of Proposition 3 : *Step 1 : (existence of the flow)* As, for every $\varepsilon\in(0,1)$ $Id-\varepsilon h$ leaves the convex compact set $\overline{F}_n^+$ stable, there is at least one $\overline{F}_n^+$-valued solution on $\mathbb{R}_+$ obtained *e.g.* by the Peano-Euler discretization method. The mean function h is Lipschitz on $\overline{F}_n^+$ since its gradient (see 1) is bounded on F_n^+. This implies global uniqueness for the ODE : the flow $\{\Phi_t\}$ of the ODE is well defined on the closure $\overline{F}_n^+$. As usual, the set of ω-limit points for the flow is a compact connected set of $\overline{F}_n^+$.

Our aim is to prove that, for every $x \in \overline{F}_n^+$, it actually lies in F_n^+

Step 2 (Φ_t instantly leaves ∂F_n^+) : Let $x \in \partial F_n^+$ being fixed and $\varepsilon > 0$. It follows from claim (b) in the above Lemma 8 that there is some i_0 such that $x_{i_0} = x_{i_0+1}$ and $\alpha_0 := (h_{i_0+1} - h_{i_0})(x) > 0$. Hence, there is some $\eta_0 > 0$ such that, for every $u \in (0, \eta_0]$, $\Phi_{u,i_0+1}(x) - \Phi_{u,i_0}(x) > \alpha_0 u$ and already parted components remain parted. One may assume w.l.g. that $\frac{\varepsilon}{n} \le \eta_0$, hence $\mathcal{E}(\Phi_{\frac{\varepsilon}{n}}(x)) > \mathcal{E}(x)$. Carrying on the process finally yields that there is some $k_0 \le n$ such that

$$\mathcal{E}(\Phi_\varepsilon(x)) = \mathcal{E}(\Phi_{\frac{k_0\varepsilon}{n}}(x)) = n$$

i.e. $\Phi_\varepsilon(x) \in F_n^+$. This is true for every $\varepsilon > 0$ which finally shows that

$$\forall x \in \partial F_n^+,\ \exists \eta_x > 0 \ \text{ such that } \ \forall u \in (0, \eta_x],\ \Phi_u(x) \in F_n^+.$$

Now, assume that there is some finite time, say u_0 such that $\Phi_{u_0}(x) \in \partial F_n^+$ and $\Phi_u(x) \notin \partial F_n^+$ for $u \in (0, u_0)$. One readily checks that for every i such that $\Phi_{u_0,i}(x) = \Phi_{u_0,i+1}(x)$, $h_i(\Phi_{u_0}(x)) = h_{i+1}(\Phi_{u_0}(x))$ which is impossible, still by claim (b) in Lemma 8. Finally, one gets

$$\forall x \in \partial F_n^+,\ \forall u \ge 0,\ \Phi_u(x) \in F_n^+.$$

Step 3 : Let $\varepsilon_0 > 0$. $x \mapsto \Phi_{\varepsilon_0}(x)$ is an homeomorphism from $\overline{F}_n^+$ onto its image, hence $\Phi_{\varepsilon_0}(\overline{F}_n^+)$ is homeomorphic to the closed disk $\overline{D}_n$, so is compact in F_n^+. Furthermore $\Phi_{\varepsilon_0}(\partial F_n^+)$ is parting F_n^+ into two connected components whose only one contains $\Phi_{\varepsilon_0}(F_n^+)$. Since the flow has an equilibrium point in F_n^+, it remains inside $\Phi_{\varepsilon_0}(F_n^+)$, thus $\Phi_{\varepsilon_0}(F_n^+)$ is a compact connected subset of F_n^+.
Consequently any positively invariant set of the flow is contain in this compact subset. The fact that the limiting values of any trajectory of the ODE is an invariant set of the flow completes the proof. $\diamond$

Collecting all the above results, and the Hirsch's theorem finally provides the following convergence result for the flow of the Kohonen S.O.M. ODE.

Theorem 7 *Assume that $(\mathcal{L}_\mu)$ and (H_σ) hold for some $k_0 < \frac{n-1}{2}$, then the flow $\{\Phi_t\}$ of the Kohonen S.O.M. ODE converges to its unique equilibrium point x^* in F_n^+.*

6 Convergence of the one dimensional Kohonen S.O.M.

6.1 Weak convergence of the invariant distributions

The extended algorithm defined by equation 9 being Feller, its constant step version $\varepsilon_t := \varepsilon \in (0,1)$ is an homogeneous Feller $\overline{F}_n^+$-valued Markov chain with transition $P^\varepsilon(x, dy)$. Subsequently, $\overline{F}_n^+$ being a compact set, for every $\varepsilon \in (0,1)$, this transition has an invariant distribution ν^ε. Furthermore, the compactness of $\overline{F}_n^+$ implies in turn that the set of probability measures $\{\nu^\varepsilon$ such that $\nu^\varepsilon P^\varepsilon = \nu^\varepsilon,\ \varepsilon \in (0,1)\}$ is tight for the weak convergence of probability measures.

One way to study the constant step setting for a stochastic algorithm is to investigate the asymptotic behavior of these invariant distributions ν^ε as $\varepsilon \to 0$.

Several theoretical results in that direction have been obtained (see [2] or [10]), namely that any limiting distribution of the ν^ε's is flow invariant for the *ODE*.

Applying these results provide the following weak convergence result.

Theorem 8 *Assume that $(\mathcal{L}_\mu)$ holds and so does (H_σ) for some $k_0 < \frac{n-1}{2}$. Let $(\nu^\varepsilon)_{\varepsilon\in(0,1)}$ be a family of invariant distributions for the extended algorithm.*
(a) For every $\varepsilon \in (0,1)$, $\nu^\varepsilon(F_n^+) = 1$; consequently, ν^ε is an invariant distribution for the original Kohonen S.O.M. algorithm as well.
(b) $\nu^\varepsilon \overset{(F_n^+)}{\Longrightarrow} \delta_{x^*}$ where $\overset{(F_n^+)}{\Longrightarrow}$ denotes the weak convergence of distribution on F_n^+.

Proof : (*a*) Let $\varepsilon > 0$. We will assume that $\nu^\varepsilon(\partial[0,1]^n) = 0$. The proof is the very same as that below using the $\mathcal{N}_0$ and $\mathcal{N}_1$ moduli instead of $\mathcal{E}$.

It follows from Lemma 8(*a*) that, for every $x \in \partial F_n^+$, $\mathcal{E}(x) \le \mathcal{E}(X^1)$ $\mathbb{P}_x$-*a.s.* and $\mathbb{P}_x(\mathcal{E}(X^1) > \mathcal{E}(x)) > 0$. Note that $\partial F_n^+ = \{\mathcal{E} \le n-1\}$.

$$\begin{aligned}
\nu^\varepsilon(\partial F_n^+) &= \int_{F_n^+} \nu^\varepsilon(dx)\mathbb{P}_x(\mathcal{E}(X^1) < \mathcal{E}(x)) + \sum_{k=1}^{n-2} \int_{\{\mathcal{E}=k\}} \nu^\varepsilon(dx)\, \underbrace{\mathbb{P}_x(\mathcal{E}(X^1) < k)}_{\le 1} \\
&\quad + \int_{\{\mathcal{E}=n-1\}} \nu^\varepsilon(dx)\, \underbrace{\mathbb{P}_x(\mathcal{E}(X^1) = \mathcal{E}(x))}_{<1} \\
&= \int_{\partial F_n^+} \nu^\varepsilon(dx)\rho(x)
\end{aligned}$$

where the function ρ is lower than 1. Finally $\nu^\varepsilon(\partial F_n^+) = 0$.

Then, one readily checks that ν^ε is an invariant distribution for the original algorithm as both definitions coincide on F_n^+.

(*b*) It follows from Theorem 7 that the flow of the extended (Feller) algorithm *ODE* is converging to the unique equilibrium point x^*. Hence, applying the flow invariance

Theorem (see [2] or [10]) for the weak limiting values of the $(\nu^\varepsilon)_{\varepsilon\in(0,1)}$ straightforwardly yields that δ_{x^*} is the only possible weak limiting distribution for the family $(\nu^\varepsilon)_{\varepsilon\in(0,1)}$. This completes the proof. $\diamond$

6.2 *A.s.* convergence

In this section, we show how to transfer the result obtained on the *ODE* to the *a.s.* properties of the stochastic algorithm itself. This point is not trivial – actually this may not hold ! – as mentioned by several authors (see *e.g.* [1], [9]).

The algorithm reads in a more compact form

$$\begin{aligned} X^{t+1} &= X^t - \varepsilon_{t+1} H(X^t, \omega^{t+1}), && (10)\\ &= X^t - \varepsilon_{t+1} h(X^t) + \varepsilon_{t+1} \times \underbrace{(h(X^t) - H(X^t, \omega^{t+1}))}_{\text{martingale } L^\infty\text{-bounded increment}} && (11)\end{aligned}$$

where the reference filtration is $\mathcal{F}_t^\omega := \sigma(\omega^1, \ldots, \omega^t)$, ω^t i.i.d. sequence of μ-distributed r.v.

We apply in this special case the classical Kushner & Clark Theorem by checking that the whole state space $\overline{F_n}^+$ of the (extended) algorithm is a (compact) stable attracting basin for the algorithm in the Kushner & Clark sense.

6.2.1 More on the K& C Theorem

We first recall for the reader's convenience the well-known result by Kushner & Clark about conditional convergence of a $\mathbb{R}^n$-valued stochastic algorithm to an asymptotically stable equilibrium x^* of the *ODE*.

Definition 3 *An equilibrium point x^* of the mean function h is a stable equilibrium if it has a stable attracting basin Γ_{x^*} in the following sense : let $\{\Phi_t(x)\}$ denote the flow of the ODE starting from x. Γ_{x^*} is defined as a neighborhood of x^* satisfying*

$$\begin{cases} (i) & \forall x \in \Gamma_{x^*},\ \forall u \in \mathbb{R}_+,\quad \Phi_u(x) \in \Gamma_{x^*},\\ (ii) & \forall x \in \Gamma_{x^*}\quad \lim_{u\to+\infty} \Phi_u(x) = x^*,\\ (iii) & \begin{cases} \forall K \subset \Gamma_{x^*}, K \textit{ compact set, } \forall \varepsilon > 0,\ \exists\, \eta_{\varepsilon,K} > 0 \textit{ such that}\\ \forall x \in K,\ |x^0 - x^*| \le \eta_{\varepsilon,K} \quad \Longrightarrow \quad \sup_{u\in\mathbf{R}_+} |\Phi_u(x) - x^*| \le \varepsilon. \end{cases}\end{cases}$$

A sufficient condition for the existence of such a stable attracting basin Γ_{x^*} is the mean function h to be differentiable at x^* with a gradient $\nabla h(x^*)$ satisfying :

all the eigenvalues of $\nabla h(x^*)$ have positive real part.

It has already been shown in section 3 that this is satisfied in our problem by the unique equilibrium point of the *ODE*. We put in the lemma below a remark on the converging flows with unique stable equilibrium point whose proof is obvious in view of Definition 3.

Lemma 9 *Let $W \subset \mathbb{R}^d$ be a nonempty open set and let h be a continuous function defined on $\overline{W}$ such that*

(i) the flow $\{\Phi_t(x)\}$ of the ODE is defined on $\overline{W} \times \mathbb{R}_+$ and takes its values in $\overline{W}$ ([3]),

(ii) $\{h=0\} = \{x^*\}$ where x^* has a stable attracting basin Γ_{x^*},

(iii) For every $x \in \overline{W}$, the flow $\Phi_u(x) \to x^*$ as $u \to +\infty$,

then, the whole state space $\overline{W}$ is a stable attracting basin for x^*.

Proposition 5 *The whole state space $\overline{F_n}^+$ of the Kohonen S.O.M. algorithm (modified on ∂F_n^+ by formula (9) is a stable attracting basin.*

A useful probabilistic version of the celebrated Kushner & Clark Theorem (see [16] and [17] for the L^q extension) reads as follows

Theorem 9 *(Kushner & Clark) Assume that*

(i) STEP ASSUMPTION : $\sum_{t\geq 0} \varepsilon_t = +\infty$ *and* $\sum_{t\geq 0} \varepsilon_t^{1+\frac{q}{2}} < +\infty$ *for some* $q \geq 2$.

(ii) MARTINGALE PART ASSUMPTION : *the sequence of martingale increments $(\Delta M^t)_{t\geq 1}$ is L^q-bounded ($\Delta M^t := h(X^{t-1}) - H(X^{t-1}, \omega^t)$).*

(iii) ODE ASSUMPTION : *Let x^* be a zero of h and K be a compact subset of its attracting basin Γ_{x^*}.*

Then, the sequence $(X^t)_{t\geq 0}$ "conditionally" (a.s.) converges to x^* in the following sense

$$X^t \overset{t\to+\infty}{\longrightarrow} x^* \quad \text{on the event } A_K^{x^*} := \{(X^t)_{t\geq 0} \text{ is bounded and } X^t \in K \text{ infinitely often}\}.$$

¿From Lemma 9 and the Kushner & Clark Theorem, one derives the now obvious

Corollary 1 *If a stochastic algorithm $(X^t)_{t\geq 0}$ defined by the general formula (10) satisfies the above assumptions of both Lemma 9 and Theorem 9, then,*

$$\forall x \in F_n^+, \qquad \mathbb{P}_x\text{-a.s. (and in } L^q) \; X^t \longrightarrow x^* \text{ as } t \to +\infty.$$

Proposition 5 combined with the above Corollary 1 finally completes the proof of Theorem 1. Actually, we proved a slightly different result since our convergence result embodies the extended algorithm *i.e.* the extended algorithm converges to x^* $\mathbb{P}_x$-*a.s.*, for every $x \in \overline{F}_n^+$.

[3]This holds whenever h is Lipschitz and, for small enough $\varepsilon > 0$, $(Id - \varepsilon h)(\overline{W}) \subset \overline{W}$.

7 An abstract result as a (first) conclusion

If one looks carefully to the method of proof, one can easily derive a general theorem of *a.s.* convergence for stochastic *cooperative* algorithms. On a purely technical point of view, the result below is not powerful enough to embody our Theorem 1 in full generality. However, when the neighborhood function σ is (strictly) decreasing on $\{1,\ldots,n\}$, it works.

Consider on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a $\overline{C}$-valued stochastic algorithm by

$$X^{t+1} = X^t - \varepsilon_{t+1} h(X^t) + \varepsilon_{t+1}\Delta M^{t+1}. \qquad (12)$$

where $(\Delta M^t)_{t\geq 1}$ denotes a sequence of martingale increments. It may be seen as the normalized noise of the deterministic equation $x^{t+1} = x^t - \varepsilon_t h(x^t)$ whence the name of *average* or *mean function* for h.

The general result reads as follows

Theorem 10 *Let C be an open bounded convex set of $\mathbb{R}^n$ and $\overline{C}$ its compact closure. Consider a stochastic algorithm defined by equation (12) and ssume that the following assumptions hold :*

(i) EXISTENCE OF THE FLOW OF THE ODE : *h is Lipschitz on $\overline{C}$ and*

(ii) ASSUMPTIONS ON h : • $h \in C^1(C)$,

• *For every $x \in C$,* $\begin{cases} - & \nabla h(x) \text{ is cooperative (i.e. } \forall i \neq j,\ \dfrac{\partial h_i}{\partial x_j}(x) \leq 0), \\ - & \nabla h(x) \text{ is irreducible (in the usual probabilistic sense),} \end{cases}$

• *The zeros of h lie in C and at each equilibrium point $\nabla h(x)$ or its transpose has strictly dominating diagonal, that is*

$$\forall i \in \{1,\ldots,n\},\quad \sum_{j=1}^{n} \frac{\partial h_i}{\partial x_j}(x) > 0 \quad or \quad \forall j \in \{1,\ldots,n\},\quad \sum_{i=1}^{n} \frac{\partial h_i}{\partial x_j}(x) > 0$$

(this implies they are stable),

• $\forall x \in \partial C,\ \forall y \in C,\ (h(x)|y-x) < 0$ *(i.e. h is pointing outside C on ∂C),*

(iii) ASSUMPTION ON THE MARTINGALE PART : *the sequence of martingale increments $(\Delta M^t)_{t\geq 1}$ are L^q-bounded for some $q \geq 2$,*

(iv) STEP ASSUMPTION : $\sum_{t\geq 1} \varepsilon_t = +\infty$ *and* $\sum_{t\geq 1} \varepsilon_t^{1+\frac{q}{2}} < +\infty$ *(for the same $q \geq 2$),*

then

(a) $\{h=0\} = \{x^*\}$ *that is h has a single equilibrium point denoted x^*,*
and
(b) *for every $x \in \overline{C}$, $\mathbb{P}_x$-a.s. (and in L^q) $X^t \longrightarrow x^*$ as $t \to +\infty$.*

8 Conclusion about the Kohonen S.O.M.

This result almost ends the rigorous mathematical investigations about the *a.s.* convergence properties of the *one-dimensional* Kohonen S.O.M., even if some further work is to be carried out to solve the case of the $\frac{n-1}{2}$ nearest neighbor (n odd).

Of course, even if one can get some satisfaction of having elucidated the one dimensional case, most of the people using the Kohonen S.O.M deal with multidimensional stimuli and generally with a 2-dimensional array of units. So the next challenge is to get at least some sufficient conditions to ensure the *a.s.* convergence of the Kohonen S.O.M in the higher dimensional cases. One way to proceed is to deeply investigate the behavior of the *ODE* since no hope of finding some nice absorbing classes is left (see [8]). Clearly the cooperativeness in the usual sense fail in this setting. Nevertheless the very deep and rich mathematical literature about differential equations should bring some light to break on through to the solution of this problem.

References

[1] M. Benaïm, A dynamical system approach to stochastic approximations, *SIAM Journal on Control and Optimization*, **34**, n^{0}2, 1996, pp. 437-472.

[2] M. Benaïm, Recursive algorithms, urn processes and chaining number of chain recurrent sets, to appear in *Ergodic Theory and Dynamical Systems*.

[3] C. Bouton, G. Pagès, Self-organization and convergence of the one-dimensional Kohonen algorithm with non-uniformly distributed stimuli, *Stochastic Processes and their Applications*, **47**, 1993, pp.249-274.

[4] P. Cohort, Sur l'unicité du quantifieur optimal (survey, in preparation), 1997.

[5] M. Cottrell and J.C. Fort, Étude d'un algorithme d'auto-organisation, *Annales de l' Institut Henri Poincaré, (série B)*, **23** , n^{0}1, 1987, pp.1-20.

[6] A. Flanagan, Self-organizing Neural Networks, Thèse 1306, Ecole Polytechnique de Lausanne, Lausanne, 1994.

[7] J.C. Fort and G. Pagès, On the *a.s.* Convergence of the Kohonen algorithm with a general neighborhood function, *The Annals of Applied Probability*, **5**, n^{0}4, 1995, pp.1177-1216.

[8] J.C. Fort, G. Pagès, About the Kohonen algorithm : strong or weak self-organization?, *Neural Networks*, **9**, n^0 5, 1996, pp.773-785.

[9] J.C. Fort and G. Pagès, Convergence of stochastic algorithms : from the Kushner & Clark theorem to the Lyapounov functional method, *Advances in Applied Probability*, **28**, n^o 4, 1996, pp.1072-1094.

[10] J.C. Fort and G. Pagès, Asymptotic behavior of the invariant distributions of a stochastic algorithms with constant step, submitted for publication, 1996.

[11] F.R. Gantmacher, *The theory of matrices*, vol. 1&2, Chelsea Publishing Company, New-York, 1959.

[12] M. Hirsch, Systems of differential equations which are competitive or cooperative II : convergence almost everywhere, *SIAM Journal Mathematical Analysis*, **16**, 1985, pp 423-439.

[13] M. Hirsch, Stability and convergence in strongly monotone dynamical systems, *J. Reine Angew. Math*, **383**, 1988, pp.1-53.

[14] T. Kohonen, Analysis of a simple self-organizing process, *Biol. Cybern.*, **44**, 1982, pp 135-140.

[15] T. Kohonen, *Self-organization and associative memory*, 3^{rd} edition Springer, Berlin, 1989.

[16] H. J. Kushner and D. S. Clark, *Stochastic approximation for constraint and unconstraint systems*, Appl. Math. Sci. Series, **26**, Springer, 1978.

[17] M. Métivier and P. Priouret, Théorèmes de convergence presque sûre pour une classe d'algorithmes stochastiques, *Probability Theory and Related Fieds*, **74**, 1987, pp 403-428.

[18] J.W. Milnor, *Topology from the differential viewpoint*, 2^{rd} edition the University Press of Virginia, 1969.

[19] H. Ritter, K. Schulten, Convergence properties of Kohonen's topology conserving maps : fluctuation, stability and dimension selection, *Biol. Cybern.*,**60**, 1988, pp 59-71.

[20] A. Sadeghi, Asymptotic behaviour of self-organizing maps with non-uniform stimuli distribution, to appear in *I.E.E.E. Transactions on Neural Networks*, 1997.

[21] H. L. Smith, *Monotone dynamical systems. An introduction to the theory of competitive and cooperative systems*, Mathematical Surveys and Monographs, Vol 41 AMS, 1995.

Prépublications du SAMOS (depuis 1995)

1995

40 - Claude BOUZITAT, Gilles PAGES
Pour quelques images de plus..., 12 p.

41 - Jean-Claude FORT, Gilles PAGES
About the Kohonen algorithm: Strong or Weak Self-organisation? 15 p.

42 - Marie COTTRELL, Patrick LETREMY
Classification et analyse des correspondances au moyen de l'algorithme de Kohonen: application à l'étude de données socio-économiques. 10 p.

43 - Joël CHADOEUF, Xavier GUYON, Jian-Feng YAO
Sur l'ergodicité de l'estimation par Restauration-Estimation de modèles incomplètement observés. 11 p.

44 - Jean-Gabriel ATTALI, Gilles PAGES
Approximation of functions by perceptrons, a new approach. 11 p.

45 - Marie COTTRELL, Bernard GIRARD, Yvonne GIRARD, Corinne MULLER, Patrick ROUSSET
Daily electrical power curves : classification and forecasting using a Kohonen map. 8 p.

46 - Fabienne COMTE, Cécile HARDOUIN
Regression on log-regularized periodogram for fractional models at low frequencies. 19 p.

47 - Fabienne COMTE, Cécile HARDOUIN
Regression on log-regularized periodogram under assumption of bounded spectral densities: the non fractional and the fractional cases. 14 p.

48 - Patrick ROUSSET
Prévision des courbes demi-horaires au moyen d'une classification de Kohonen. 25 p.

49 - Smail IBBOU et Marie COTTRELL
Multiple Correspondence analysis of a crosstabulations matrix using the Kohonen algorithm. 6 p.

50 - Philippe JOLIVALDT
Schémas de discrétisation pour la simulation et l'estimation d'un CAR(2): une étude expérimentale. 22 p.

51 - Philippe JOLIVALDT
Utilisation de méthodes implicites pour la simulation et l'estimation de modèles CAR(2) . 14 p.

1996

52 - Samuel BAYOMOG
Estimation of a Markov field dynamic. 14 p.

53 - Morgan MANGEAS et Jian-feng YAO
Sur l'estimateur des moindres carrés des modèles auto-régressifs fonctionnels. 19 p.

54 - Marie COTTRELL, Florence PIAT, Jean-Pierre ROSPARS
A Stochastic Model for Interconnected Neurons. 17 p.

55 - Marie COTTRELL, Jean-Claude FORT, Gilles PAGES
Two or three mathematical things about the Kohonen algorithm. 31 p.

56 - Marie COTTRELL, Bernard GIRARD, Patrick ROUSSET

Forecasting of curves using a Kohonen classification. 14 p.

57 - Jean-Claude FORT, Gilles PAGES
Quantization vs Organization in the Kohonen S.O.M. 5 p.

58 - Eric de BODT, Marie COTTRELL, Michel LEVASSEUR
Réseaux de neurones en finance.33 p.

59 - Marie COTTRELL, Eric de BODT, Emmanuel HENRION, Ismaïl IBBOU, Annick WOLFS, Charles Van WYMEERSCH
Comprendre la décision à l'aide d'une carte de Kohonen. Une étude empirique. 16 p.

60 - Marie COTTRELL, Eric de BODT
Understanding the leasing decision with the help of a Kohonen map. An empirical study of the Belgian market. 5p.

61 - Marie COTTRELL, Eric de BODT, Philippe GREGOIRE
The relation between interest rate shocks and the initial rate structure: an empirical study using a Kohonen map. 16p.

62 - Marie COTTRELL, Eric de BODT, Philippe GREGOIRE
A kohonen map representation to avoid misleading interpretation. 8p.

63 - Marie COTTRELL, Eric de BODT
Analyzing shocks on the interest rate structure with Kohonen map. 6p.

64 - Marie COTTRELL, Eric de BODT, Philippe GREGOIRE
Simulating interest rate structure evolution on a long term horizon. A kohonen map application. 5p.

65 - Fabienne COMTE, Cécile HARDOUIN
Log-regularized periodogram regression. 22p.

66 - Jian Feng YAO
Simulation et optimisation sous contrainte par une dynamique de Metropolis. 6p.

67 - Morgan MANGEAS, Jian-feng YAO
On least squares estimation for nonlinear autoregressive processes. 16p.

68 - Carlo GAETAN, Xavier GUYON
Simulation des modèles de Gibbs spatiaux par chaine de Markov. 28p.

69 - François GARDES, Patrice GAUBERT, Patrick ROUSSET
Cellulage de données d'enquêtes de consommation par une méthode neuronale. 41p.

70 - Serge IOVLEFF, José R. León
High-Frequency approximation for the Helmholtz equation :a probabilistic approach. 14p.

1997

71 - Xavier Guyon et Jian-feng Yao
On description of wrong parametrisation sets of a model . 23p.

72 - Sandie Souchet
Schéma de discrétisation anticipatif et estimation du paramètre d'une diffusion. 36p.

73 - M. Cottrell, J.C. Fort, G. Pagès
Theoretic aspects of the SOM algorithm 22p.

74 - M. Benaïm, J.C. Fort & G. Pagès

Convergence of the one-dimensional Kohonen algorithm. 23p.

75 - C. Bouton & G. Pagès

About the multidimensional competitive learning vector quantization algorithm with constant gain. 36p.

76 - M. Cottrell, P. Rousset

The Kohonen algorithm: a powerful tool for analyzing and reprenting multidimensional quantitative and qualitative data 12p.