Self-Organisation in the One-Dimensional SOM with a Reduced Width Neighbourhood

John A. Flanagan*

Circuits and Systems Group
Electrical Engineering Department
Swiss Federal Institute of Technology Lausanne
CH-1015 Lausanne, Switzerland
e-mail : flanagan@ims.dmt.epfl.ch

Abstract

Previous analyses of the the one dimensional SOM have revealed certain conditions which must be satisfied before probability one self-organisation of the neuron weights can be proved. One general proof relies on the fact that the neighbourhood function is such that every neuron weight is updated at each iteration. The result is two intervals of non zero probability of the input signal are required to prove self-organisation. By reducing the width of the neighbourhood function every neuron is not necessarily updated at each iteration. By introducing an extra condition on the support of the probability distribution of the input signal, probability one self-organisation of the neuron weights can be proved for a neighbourhood function whose width is greater than half the number of neurons. This extra condition is specified in terms of the need for a third interval of non zero probability. The proof is extendable to higher dimensions.

1 Introduction

The self-organising property of Kohonen's Self-Organising Map (SOM) [4] has found a large area of application and hence an analysis and understanding of its self-organising abilites are important. This property has been widely studied in [2], [1] [3]. In [3] a proof of self-organisation was given for one and higher dimensional maps, in a general form which imposed minimal conditions on the probability distribution of the input signal. The proof in itself was quite straightforward due to the fact that the width W of the neighbourhood function was greater than the number of neurons in any row of the network. The result; all the neuron weights were updated at each iteration. This allows for a global form of self-organisation. However when the conditions on the neighbourhood function are changed, that is its width is reduced, such a proof becomes more difficult, and the specifications on the probability distribution and support of the input signal become more stringent to ensure probability one convergence to an organised configuration. In this work we will look at the problem of proving self-organisation for a one dimensional SOM with a neighbourhood function where $N' \leq W < N$, N the total number of neurons and $N' = N/2+1$ for N even and $N' = (N+1)/2$ for N odd.

[0]*Currently with the IMS-DMT, EPFL, CH-1015 Lausanne, Switzerland

We consider the SOM with neuron weights x_i, $i = 1, \ldots, N$. After an arbitrary initialisation a random input $\omega(t)$ with probability distribution function $P(\omega)$ at time t is presented to the network. A *winner* neuron $v(t)$ is chosen such that $\mid x_v(t) - \omega(t) \mid \leq \mid x_j(t) - \omega(t) \mid, \forall 1 \leq j \leq N$. The neuron weights are then updated as

$$x_i(t+1) \;=\; x_i(t) + \alpha(t)h(i,v)(\omega(t) - x_i(t)) \quad \forall 1, \leq j \leq N \tag{1}$$

$0 < \alpha_m \leq \alpha(t) < 1$ is a gain function, and $h(i,v) = h(\mid i - v \mid)$ is the neighbourhood function. In this case h is such that

- $h(i,i) = 1$
- $h(i, j \pm 1) \leq h(i,j) + \phi$, $\phi > 0$
- $h(i, i \pm W) = h_m > 0$
- $h(i,j) = 0$, $j > \mid i \pm W \mid$

From this define the relationship between two variables γ and κ by

$$\gamma < \frac{\kappa \alpha_m \phi}{1 - \alpha_m h_m} \tag{2}$$

The stochastic process to be examined $\mathbf{X} = (x_1, \ldots, x_N)^T$ is the vector of neuron weights on a common probability space $(\Psi, \mathcal{F}, \pi)$.

In the next sections a self-organising theorem will be specified for this case and a detailed outline of the proof of this theorem will follow in the remainder. Some of the propositions and proofs are given in the Appendix.

2 A Theorem for Self-Organisation

The absorbing nature of the organised configuration in the one dimensional SOM is well known [2], thus when the neuron weights are in the set A where

$$A = \{\mathbf{X} : x_1 < x_2 < \ldots < x_N \textit{ or } x_N < x_{N-1} < \ldots < x_1\} \tag{3}$$

they remain there. The first entry time of $\mathbf{X}$ into the configuration A is τ_A. Now define three sets A_j, A_k, A_m on the interval $[0,1]$ where μ_r, ν_r are respectively the minimum and maximum value of a set A_r and where it is true that $\mu_k - \nu_j \geq \kappa$, $\mu_m - \nu_k \geq \kappa$. Define C to be the smallest interval containing A_j, A_m. Figure 1 shows A_j, A_k etc. Now the theorem for self-organisation can be stated as

Theorem 1 *Given the three sets A_j, A_k, A_m of width γ on $[0,1]$ and three open balls $B_j \subset A_j$, $B_k \subset A_k$, $B_m \subset A_m$ such that for $\epsilon > 0$, $\int_{B_j} P(\omega)d\omega \geq \epsilon$, $\int_{B_k} P(\omega)d\omega \geq \epsilon$, $\int_{B_m} P(\omega)d\omega \geq \epsilon$ and arbitrary $\mathbf{X}(0)$ then for some $T < \infty$ and $\delta > 0$*

$$\pi_{\mathbf{x}^0}(\{\psi \in \Psi : \tau_A \leq T\}) \;\geq\; \delta \tag{4}$$

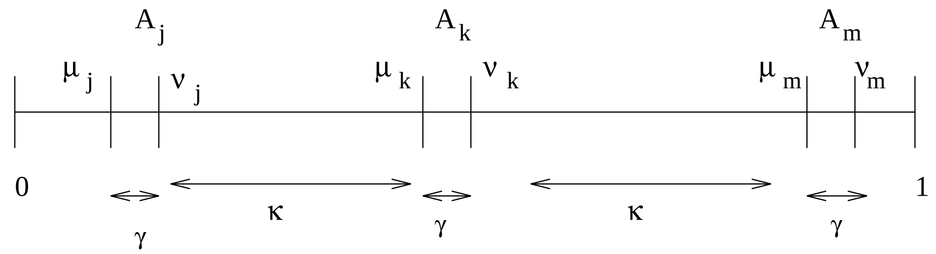

Figure 1: The different sets described for the proof of self-organisation.

Proof : The proof of the theorem comes from the results of Lemmas 1, 2, 3. Lemma 1 shows how to take $\mathbf{X}(0) \in C$ to the set D in a finite time with non zero probability. Lemma 2 shows how to take $\mathbf{X}$ from the set D to the set A also in a finite time with a non zero probability. The final Lemma 3 shows that for $\mathbf{X}(0) \notin C$ how the organising process results in a finite time with non zero probability of all the neuron weights being in C. Hence the proof of the theorem. □

An intuitive outline of the proof is as follows. First it is assumed that all the neurons are in C, and that in a finite time more than half the neuron weights are in A_j by choosing $\omega(t)$ in A_j. When in this configuration $\omega(t)$ is chosen to be in A_m until such a time that once again more than half the neuron weights are in A_m. This time however there is a certain level of organisation which has been reached. These conditions are described in Lemma 2. From this it can be shown that the neuron weights can be brought to a partially organised configuration D, defined later, where for example the first q neuron weights are in A_j the last p weights are in A_m and the weights in between are organised with respect to these. From this configuration it is shown how in two iterations, by choosing the correct winner neuron, that the neuron weights can be completely organised. To show that all neuron weights can be brought to C use is made of the organised state of a subset of neuron weights and then it must be true that from this condition it is possible to update all neuron weights a finite number of times with non zero probability.

3 Proving Self-Organisation

Given an interval A_j of width γ on $[0, 1]$, and an initial random distribution of the neuron weights $\mathbf{X}(0)$. Define D_i^j for each $i \in [1, N]$ as

$$D_i^j = \{\mathbf{X} : x_p \in A_j,\ p = i - W, i - W + 1, \ldots, i + W - 1, i + W\} \tag{5}$$

A direct result of this definition is that

Proposition 1 *For every $D_i^u, i \in [1, N]$ then there are three possible configurations of* $\mathbf{X}$.

$$\begin{aligned}
C_1^u &= \{\mathbf{X} :\ x_p \in A_u\ p = 1, \ldots, N\} \\
C_{2q}^u &= \{\mathbf{X} :\ x_p \in A_u\ p = 1, \ldots, q > W \textit{ and } \exists r > W \textit{ s.t. } x_r \notin A_u\} \\
C_{3q}^u &= \{\mathbf{X} : x_p \in A_u,\ p = N, \ldots, q < W \textit{ and } \exists r < W \textit{ s.t. } x_r \notin A_u\}
\end{aligned}$$

We can now state the following, for $\epsilon > 0$.

Proposition 2 *If the interval A_j contains an open ball B_j such that $\int_{B_j} P(\omega)d\omega \geq \epsilon$ and $\tau_{D_i^j}$ is the first entry time of* $\mathbf{X}$ *into D_i^j for some i, then*

$$\pi_{\mathbf{x}^0}(\{\psi \in \Psi : \tau_{D_i^j} \leq T\}) \ \geq \ \delta \tag{6}$$

Proof : The proof of this proposition is the same as that of Theorem 1 in [3]. The only observation to make is that for any winner neuron at t then one of the neurons 1 or N must be updated at t. This ensures the result. □

This proposition forms the basis of going to an organised configuration and will be used repeatedly in what follows. Consider now the two intervals A_j and A_m. Define what could be called a partially organised set of neurons, $D = D_{1,N} \bigcup D_{N,1}$ as follows

Definition 1

$$\begin{aligned} D_{1,N} \ &= \ \{\mathbf{X} : 1 \leq p < q \leq N;\ x_i \in A_j,\ i = 1, \ldots, p;\ \nu_k < x_{p+1} < x_{p+2} \\ &< \ldots < x_q < \nu_m,\ i = p+1, \ldots, q;\ x_i \in A_m,\ i = q, \ldots, N, \} \end{aligned}$$

$$\begin{aligned} D_{N,1} \ &= \ \{\mathbf{X} : 1 \leq p < q \leq N;\ x_i \in A_j\ i = N, \ldots, q;\ \nu_k < x_{q-1} < x_{p-2} < \\ &\ldots < x_p < \mu_m,\ i = q-1, \ldots, p;\ x_i \in A_m,\ i = q, \ldots, 1\} \end{aligned}$$

If τ_D is the first entry time of $\mathbf{X}$ into D then the following can be stated.

Lemma 1 *Given the two sets A_j, A_M and two open balls $B_j \subset A_j$ and $B_m \subset A_m$ such that for $\epsilon > 0$, $\int_{B_j} P(\omega)d\omega \geq \epsilon$ and $\int_{B_m} P(\omega)d\omega \geq \epsilon$ and $\mathbf{X}(0) \in C$ then,*

$$\pi_{\mathbf{x}^0}(\{\psi \in \Psi : \tau_D \leq T\}) \ \geq \ \delta \tag{7}$$

Proof : From Proposition 2 it is shown that $\exists t_1 < \infty$ for which $\mathbf{X}(t_1) \in D_r^j$. From Proposition 1 there are three possible different configurations. If $\omega(t) \in A_m, t_1 \leq t \leq t_2$ then from proposition 2 $\exists t_2 < \infty$ s.t. $\mathbf{X}(t_2 - 1) \notin D_s^k$ and $\mathbf{X}(t_2) \in D_s^k$ in one of the 3 confgurations for $D_s^m$of Proposition 1, but having some extra characteristics now to be described. First define the following set as $\Gamma_q = \Gamma_q^{1,N} \bigcup \Gamma_q^{N,1}$

$$\Gamma_q^{1,N} \ = \ \{\mathbf{X} : x_p \in A_m\ p > q,\ \mu_j < x_1 < x_2, \ldots < x_q < \nu_m\}$$

$$\Gamma_q^{N,1} \ = \ \{\mathbf{X} : x_p \in A_m\ p < q,\ \mu_j < x_N < x_{N-1}, \ldots < x_q < \nu_m\}$$

If $\mathbf{X}(t_1) \in C_1^j$ then $\mathbf{X}(t_2) \in C_1^m$, Γ_q *or* D. The fact that $\mathbf{X}(t_2) \in \Gamma_q^{1,N}, D$ is possible can be seen as follows. From proposition 3 it is seen that $x_w(t_1+1) > x_{w-1}(t_1+1) > \ldots > x_{v(t_1)-W}(t_1+1)$. If $v(t_1) - W \leq 1$ then using Proposition 4 it is seen that $\mathbf{X}(t_2)$ may be in $\Gamma_q^{1,N}$. If $v(t_1) - W > 1$ then using Proposition 4 it is seen that $\mathbf{X}(t_2)$ may be in $\Gamma_q^{1,N}$ or D. If $\mathbf{X}(t_2) \in C_1^m$ then either $x_N(t_2) < x_i(t_2)\ \forall i \neq N$ or $x_1(t_2) < x_i(t_2)\ \forall i \neq 1$. This can be seen also from the application of propositions 3 and 4. For $\mathbf{X}(t_2) \in C_1^m, \Gamma_q$ then for $\omega(t_2) \in A_j$ then $v(t) = 1$ *or* N. For $\omega(t) \in A_j, t > t_2$ then $\mathbf{X}(t_3) \in D$ for some $t_3 < \infty$. The results of this also follow from Propositions 3 and 4.

If $\mathbf{X}(t_1) \in C^j_{w2}$ (similar arguements apply for $\mathbf{X}(t_1) \in C^j_{w3}$) then $\mathbf{X}(t_2) \in C^m_1, C^m_{q2}, \Gamma^{1,N}_q, D$. The fact that $\mathbf{X}(t_2) \in \Gamma^{1,N}_q, D$ is possible can be seen as follows. From proposition 3 it is seen that $x_w(t_1+1) > x_{w-1}(t_1+1) > \ldots > x_{v(t_1)-W}(t_1+1)$. If $v(t_1) - W \leq 1$ then using Proposition 4 it is seen that $\mathbf{X}(t_2)$ may be in $\Gamma^{1,N}_q$. If $v(t_1) - W > 1$ then using Proposition 4 it is seen that $\mathbf{X}(t_2)$ may be in $\Gamma^{1,N}_q$ or D.

If $\mathbf{X}(t_2) \in C^m_1$ then the same proof as for $\mathbf{X}(t_1) \in C^j_1$ can be applied choosing $\omega(t) \in A_j, t > t_2$. If $\mathbf{X}(t_2) \in C^m_{q2}$ then it must be true that for $\omega(t_2) \in A_j$ that $v(t_2) > v(t_1)$. This is true from the fact that $x_{v(t_1)}(t_2) \in A_m$. Now for $t_2 \leq t \leq t_3 < \infty$ and $\omega(t) \in A_j$ once again we have the same arguement. If at time t_3, $\mathbf{X}(t_3) \in C^j_{s2}$ then once again it is true that $v(t_3) > v(t_2)$ for $\omega(t_3) \in A_m$. Continuing this arguement for some $t_N < \infty$ then $\mathbf{X}(t_N) \in C^{j\ or\ m}_1, C^{j\ or\ m}_{q2}, \Gamma^{1,N}_q, D$. For $\mathbf{X}(t_N) \in C^{j\ or\ m}_{q2}$ then $v(t_N) = N$ and the result follows. □

Lemma 2 *Given the three sets A_j, A_k, A_m of width γ on $[0,1]$ and three open balls $B_j \subset A_j$, $B_k \subset A_k$, $B_m \subset A_m$ such that for $\epsilon > 0$, $\int_{B_j} P(\omega)d\omega \geq \epsilon$, $\int_{B_k} P(\omega)d\omega \geq \epsilon$, $\int_{B_m} P(\omega)d\omega \geq \epsilon$ and $\mathbf{X}(0) \in D$ then,*

$$\pi_{\mathbf{x}^0}(\{\psi \in \Psi : \tau_A \leq T\}) \ \geq \ \delta \tag{8}$$

Proof : If $\mathbf{X}(0) \in D$ then assume that it is in $D_{1,N}$ (Similar arguements apply for $\mathbf{X}(0) \in D_{N,1}$). Now we consider the neuron N'. If $\mathbf{X}(0) \in D_{1,N}$ and $x_{N'} < \mu_k$ then for $\omega(0) \in A_k$, $v(0) \geq N'$. Thus at $t = 1$ $x_j < x_{j-1} < \ldots < x_N$ for $j = v(0) - W$. Now for $t \geq 1$ $\omega(t) \in A_m$ until such a time $t_1 < \infty$ where $x_{N'}(t_1 - 1) < \nu_k$ and $x_{N'}(t_1) \geq \nu_k$. Thus for $\omega(t_1) \in A_k$ then $v(t_1) \leq N'$ and $x_1(t_1+1) < x_2(t_1+1) < \ldots < x_N(t_1+1)$. This must be true as for the time $t \leq t_1$ we use the absorbing nature of the one dimensional SOM to show that the neurons $N', \ldots, N$ remain organised. Hence we can define $T = t_1 + 1$ and $\delta = \epsilon^{t_1+1}$. □

Lemma 3 *Given the three sets A_j, A_k, A_m of width γ on $[0,1]$ and three open balls $B_j \subset A_j$, $B_k \subset A_k$, $B_m \subset A_m$ such that for $\epsilon > 0$, $\int_{B_j} P(\omega)d\omega \geq \epsilon$, $\int_{B_k} P(\omega)d\omega \geq \epsilon$, $\int_{B_m} P(\omega)d\omega \geq \epsilon$. If τ_C is the first entry time of $\mathbf{X}$ into C from any initial $\mathbf{X}(0)$ then*

$$\pi_{\mathbf{x}^0}(\{\psi \in \Psi : \tau_C \leq T\}) \ \geq \ \delta \tag{9}$$

Proof : Only a sketch of this proof is given. It suffices to show that if $x_s(0) \notin C$ then it can be updated a finite number of times with non zero probability for $\omega(t) \in C$, the result then follows from Theorem 1 of [3]. This requires showing that a $\mid s - v(t) \mid \leq W$ finitely often with non zero probability. First consider $\mathbf{X}(t_1) \in D^j_r$ where from Proposition 2 $t_1 < \infty$. Thus for example neurons $x_1, x_2, \ldots, x_r \in A_j$ and $x_s(t_1) \notin C$. Consider the process described by Lemma 1 followed by the process described by Lemma 2 where all the neurons $1, \ldots, r$ can be shown to organise in a finite time at which point x_s may be in C. If not, from this partial organised configuration it is possible to show that neurons 1, r or s can be chosen as winner a finite number of times with non zero probability until $x_s \in C$. □

4 Conclusion

A proof of probability one convergence to an organised configuration from any random initial condition of the neuron weights is described in a one dimensional SOM. The main difference with the proof of [3] is the fact that the width of the neighbourhood function here is less than the number of neurons. The result is that an additional specification must be added for the support of the input signal. It is shown here that instead of two intervals of non-zero probability being separated by a certain distance as necessary for the case described in [3], that by reducing the width of the neighbourhood function, we need three intervals of non zero probability. The proof has been outlined here only for the one dimensional case but a similar approach would lead to a proof for higher dimensional cases in a manner similar to that in [3]. This approach also leaves the way open for a general analysis of self-organisation in the SOM.

References

[1] C. Bouton and G. Pagès. Self-organization and a.s. convergence of the one-dimensional Kohonen algorithm with non-uniformly distributed stimuli. *Stochastic Processes and Their Applications*, 47:249–274, 1993.

[2] M. Cottrell and J.-C. Fort. Étude d'un processus d'auto-organisation. *Annales de l'Institut Henri Poincaré*, 23(1):1–20, 1987. (in French).

[3] J. A. Flanagan. Self-organization in Kohonen's SOM. *Neural Networks*, 9(7):1185–1197, 1996.

[4] T. Kohonen. *Self-Organizing Maps.* Springer, Berlin, Heidelberg, 1995.

Appendix

Proposition 3 *If $x_i(t), x_j(t) \in (0,\gamma)$ and $\omega(t) > \gamma + \kappa$ such that $h(v(t),i) > h(v(t),j)$ then*

$$x_j(t+1) < x_i(t+1) \tag{10}$$

Proof : Consider $x_i(t+1) - x_j(t+1)$ given from equation 1 by

$$x_i(t+1) - x_j(t+1) \;=\; x_i(t) - x_j(t) + \alpha(t)(h(i,v)(\omega(t) - x_i(t)) - h(j,v)(\omega(t) - x_j(t)) \tag{11}$$

but $h(i,v) \geq h(j,v) + \phi$ thus

$$x_i(t+1) - x_j(t+1) \;\geq\; x_i(t) - x_j(t) + \alpha_m((h(j,v) + \phi)(\omega(t) - x_i(t)) - h(j,v)(\omega(t) - x_j(t)) \tag{12}$$

From this

$$x_i(t+1) - x_j(t+1) \;\geq\; (x_i(t) - x_j(t))(1 - \alpha_m h(j,v)) + \alpha_m \phi(\omega(t) - x_i(t)) \tag{13}$$

Given $\omega(t) - x_i(t) > \kappa$ and in the worst case $(x_i(t) - x_j(t)) = -\gamma$ then it is easy to derive the result using equation 2. □

Proposition 4 *Given $x_i(t) \geq x_j(t)$ and $\omega(t) > x_i(t)$ then for $h(i,v) > h(j,v), x_i(t+1) > x_j(t+1)$.*

Proof : if $x_j(t+1)$ and $x_i(t+1)$ are given by equation 1 then,

$$x_i(t+1) - x_j(t+1) \;=\; x_i(t) - x_j(t) + \alpha(t)[h(i,v)(\omega(t) - x_i(t)) - h(j,v)(\omega(t) - x_j(t))] \tag{14}$$

If $h(i,v) > h(j,v)$ then $h(i,v) = h(j,v) + \phi, \phi > 0$. The previous equation can be written as

$$x_i(t+1) - x_j(t+1) \;=\; (x_i(t) - x_j(t))(1 - h(j,v)) \;+\; (\omega(t) - x_i(t))\phi \tag{15}$$

This implies that $x_i(t+1) - x_j(t+1) > 0$. □