ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 271-276

Instabilities in Self–Organized Feature Maps with Short Neighbourhood Range

R. Der and M. Herrmann
Universität Leipzig, Institut für Informatik
PSF 920, D–04009 Leipzig, F. R. Germany

Abstract

We report a new kind of phase transition in Kohonen's self–organized feature maps with short neighbourhood range by means of analytical as well as numerical studies. Below a critical value of the neighbourhood width $\sigma^* \approx 0.49$ the symmetry of the globally ordered state is broken, i.e. a phase tranisition to a different ordering occurs. This is predicted theoretically by a stability analysis in terms of a Langevin approach to the dynamics of the system. Simulations justify the prediction that hexagon–shaped patterns emerge. Implications for the ordering properties of SOFM's are discussed.

1 Introduction

Self–organized feature maps generated by Kohonen's algorithm play an important role in many applications of neural nets. Their capability of generating a topology preserving vector quantization of real world input spaces is of interest in data reduction, feature extraction, and in developing low–dimensional internal representations of high–dimensional input spaces. Quite generally, Kohonen's feature maps may be considered as a non–linear principal component analyzer. However, in certain situations these properties may break down. If there is a topological mismatch between the input space and the net topology, the mapping tries to compromise between the two topologies. This effect is well known and has been predicted by Ritter et al. [6] by means of a linear stability analysis. The present authors [2] developed a quantitative description of this effect in terms of the Ginzburg–Landau approach borrowed from statistical physics.
In the present paper we report a new kind of phase transition that surprisingly occurs even in the case of a *topological agreement* between the input space and net topology. The origin of this effect is — as in the case of the topological mismatch — connected with the fixation of the network structure. If mapping a square–shaped input space the usual arrangement of the neurons in the neuron space is on a square lattice with a square elementary cell. The effect to be studied in the following arises from the conflict between square domains favoured by the

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 271-276

arrangement of neurons in the input space and the fact that the reconstruction error is reduced if the domain approaches a circular shape.
The equilibrium configuration of the map corresponds to the payoff between the elastic forces arising from the neighbourhood cooperation between the neurons and the tendency to decrease the reconstruction error. For a large neighbourhood range $\sigma \gg 1$ the elasticity effects are dominant and favour a square shaped domain, whereas for smaller σ these are counterbalanced by the reconstruction error effect. For $\sigma \leq \sigma^* \approx 0.49$ the latter effect dominates which is reflected by a series of phase transitions that eventually results in an equilibrium configuration corresponding to hexagon–shaped domains of the neurons. This *reordering* transition is not to be confounded with the "*melting*" transition occurring in the case of both high values of the learning parameter $\epsilon \gtrsim 0.5$ and small values of σ, where *any* order is disrupted.

2 Kohonen's algorithm

For the purpose of demonstrating the above mentioned effect we consider a Kohonen map [5] of a two–dimensional input space to a two–dimensional square lattice of the neurons.
More specifically there are N^2 neurons situated at sites $\vec{r} = (r_x, r_y);\ r_x, r_y \in \{1, \ldots, N\}$ that receive randomly chosen inputs $\vec{v}$ are from the square $[0,1]\times[0,1]$ via connections $\vec{w}_{\vec{r}} \in \mathcal{R}^2$. The latter evolve according to Kohonen's learning rule that is taken as usual

$$\Delta\vec{w}_{\vec{r}}(t) = -\epsilon h_{\vec{r},\vec{r}^*}(\vec{w}_{\vec{r}} - \vec{v}) \tag{1}$$

where $\vec{r}^*$ denotes the best matching (*winner*) neuron, i.e. the one with $||\vec{w}_{\vec{r}} - \vec{v}|| \geq ||\vec{w}_{\vec{r}^*} - \vec{v}||\ \forall \vec{r} \neq \vec{r}^*$. The neighbourhood function of width σ measuring the degree by which a neuron in the vicinity of the winner participates in the learning step is a Gaussian depending on the Euclidean distance $||\vec{r} - \vec{r}^*||$ in the network space.

$$h_{\vec{r},\vec{r}^*} = \exp\left(-\frac{||\vec{r} - \vec{r}^*||^2}{2\sigma^2}\right) \tag{2}$$

The components of $\vec{w}_{\vec{r}}$ may be viewed as the coordinates of the image of the neuron at lattice site $\vec{r}$. This defines the map of the lattice of neurons into the input space. The algorithm subdivides the input space into a disjoint set of domains of the neurons (*Voronoi tessellation*) the topological order of the input space being conserved as far as possible.
The short–range neighbourhood function restricted to nearest neighbours is defined by

$$h_{\vec{r}\vec{r}^*} = \begin{cases} \exp\left(-\frac{||\vec{r}-\vec{r}^*||^2}{2\sigma^2}\right) & \text{if } ||\vec{r} - \vec{r}^*|| \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Provided that $\sigma \ll 1$ the neighbourhood function (3) is a very good approximation of the long–range type defined in (2). The effects of the inclusion of non–nearest neighbours with Gaussian weights are marginal at the values of $\sigma \lesssim 0.5$ which are relevant here.

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 271-276

On the other hand, short–range interactions become in the limit $\sigma \to \infty$

$$h_{\vec{r}\vec{r}^*} = \begin{cases} 1 & \text{if } \|\vec{r} - \vec{r}^*\| \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Short–range interactions of this type have been studied in [6] (see also [1]).

3 Analysis of short–range SOFM's

We apply a Langevin approach for the analysis of the dynamics of the map from a two–dimensional square–shaped input set onto a square array of neurons in the vicinity of the presupposed stationary state $\vec{w}_{\vec{r}}^{(0)} = \vec{r}$. Using deviations $u_{\vec{r}} = \vec{w}_{\vec{r}} - \vec{w}_{\vec{r}}^{(0)}$ the dynamics of the map is analyzed in terms of the Fourier amplitudes

$$\vec{u}_{\vec{k}} = \frac{1}{N} \sum_{\vec{r}} e^{i\vec{k}\vec{r}} u_{\vec{r}} \tag{5}$$

Kohonen's algorithm can be rewritten as a generalized Langevin equation [2] which in the linear region takes the simple form

$$\frac{\partial}{\partial t} \vec{u}_{\vec{k}} = -\mathbf{B}(\vec{k}) \vec{u}_{\vec{k}} + \vec{f}_{\vec{k}}(t) \tag{6}$$

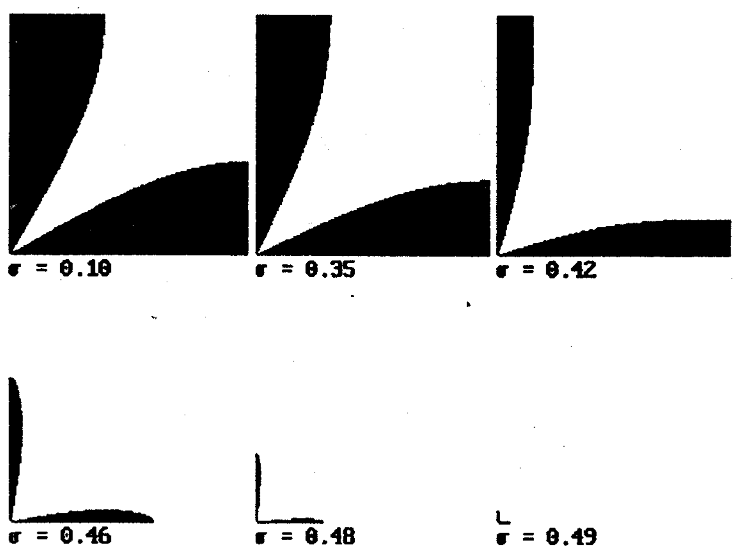


Figure 1: *Ranges of stability for various values of σ. The eigenvalues of the matrix $\mathbf{B}(\vec{k})$ (6) determine the stability of the corresponding modes $\vec{u}_{\vec{k}}$. Black regions correspond to instable modes, where at least one of the eigenvalues is negative.*

where $\vec{f}_{\vec{k}}(t)$ is the noise and $\mathbf{B}$ is a matrix of generalized friction coefficients. The stability of the modes is governed by the signs of the eigenvalues of the matrix $\mathbf{B}$. If at least one of these eigenvalues, which have been calculated analytically, is

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 271-276

less than zero the corresponding mode is instable. In Figure 1 a stability plot is given for various values of σ. The regions of instability increase with decreasing values of σ, cf. Figure 2. If $\sigma > \sigma^* \approx 0.49$ the stationary state is stable with respect to any wavelength. If $\sigma^* > \sigma > \sigma^\times \approx 0.45$ patterns of relatively long wavelength arise. Evidently, unidirectional modes, i.e. with either k_x or k_y being small, are preferred.

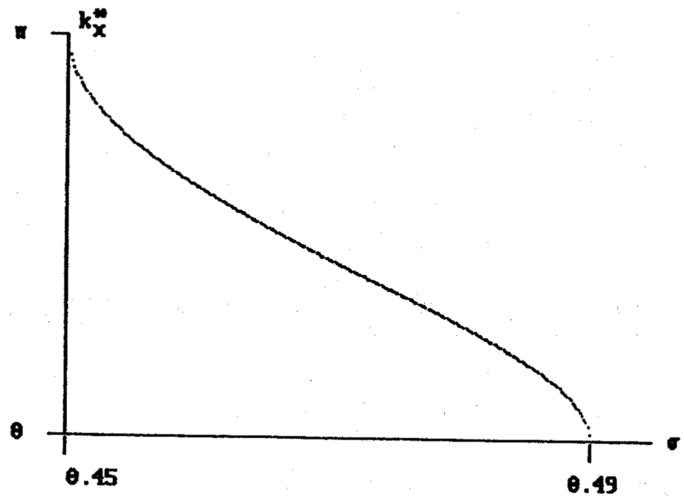


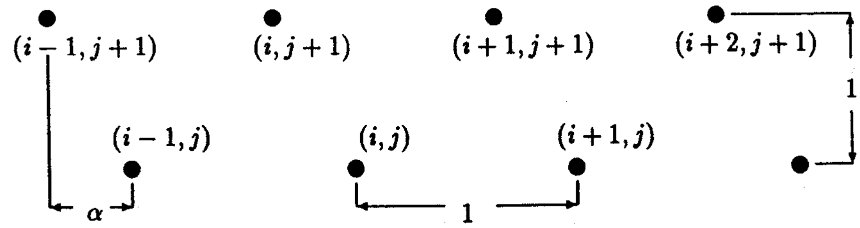
Figure 2: *Depicted is the instable wavenumber $\vec{k}_c$ with maximal absolute value versus σ. $1/\|\vec{k}_c\|$ is the minimum length scale of induced instabilities. It is expected that the observed structures are preferentially produced on this length scale.*

4 Improvement of the reconstruction error

At low σ the equilibrium properties of the map are largely dominated by the algorithms tendency to minimize the reconstruction error.
The averaged mean square distance from the centroid in a regular square network with lattice constant unity and homogeneous input distribution is $E_G = \frac{1}{6}$.
In the case of idealized deformations (cf. below) of the square lattice the error obeys

$$E_G(\alpha) = \frac{2 - \alpha^2 + 2\alpha^3 - \alpha^4}{12} \tag{7}$$

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 271-276

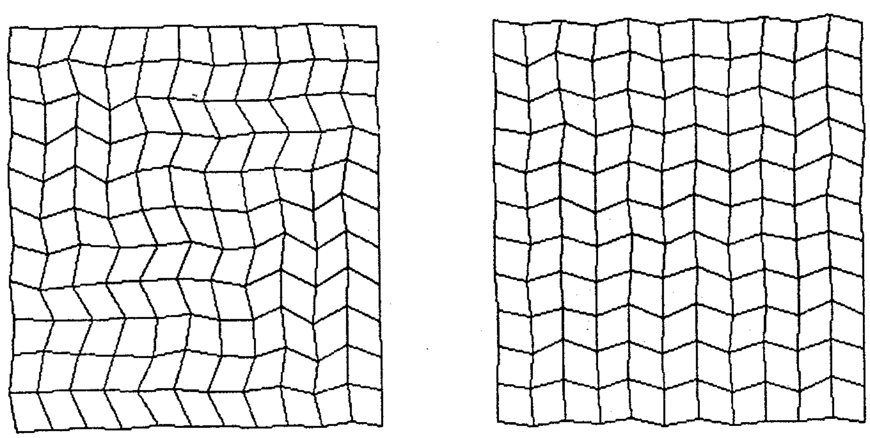


In the relevant interval $\alpha \in [0,1]$ the minimum error $E_G(\alpha) = \frac{31}{192}$ is achieved at $\alpha = \frac{1}{2}$, i.e. hexagonal structures (which are non-regular due to an elongation of the hexagons in one direction), and $E_G(\alpha)$ decreases monotonously when going from the square lattice to the hexagonal one. Hence, a phase transition between phases of different symmetry is to be expected if σ decreases below $\sigma^\times \approx 0.45$.

5 Numerical results

Simulations of (1) have been performed at various network sizes ($N^2 = 4^2 \ldots 64^2$) and at $\sigma = 0.2$. Since a whole band of modes is instable for this value of σ. symmetry breaking is to be expected due to the selection of a small number of different $\vec{k}$.

If, in particular, $\sigma < 0.45$ either of the modes with $(k_x, 0)$, $(0, k_y)$ succeeds to become the principal mode, i.e. neighbouring rows (or columns) interlace, cf. Figure 3b. In our simulations, however, deviations from the square lattice occur domainwise either in vertical or horizontal direction, cf. Figure 3a. The arousal of a global domain is expected to take place only at very large times. On the other hand, if the network is initialized close to a state with unidirectional symmetry breaking it remains in this state.



a) *b)*

Figure 3: *a) Domainwise symmetry breaking. b) Global symmetry breaking.*

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 271-276

6 Conclusion

The analytical as well as the numerical results presented so far are of importance for both the theoretical understanding and practical use of Kohonen maps. A first point to be addressed are implications to the property of topological ordering which has been proved for one–dimensional Kohonen nets in Ref. [4]. We have shown that in the two–dimensional case the existence of instabilities for small values of σ excludes the possibility of globally regular ordering in the parameter region $\sigma < \sigma^*$ in the case that the neurons are arranged on square lattice. Therefore, $\sigma > \sigma^*$ is suggested as a necessary condition for convergence to a globally ordered state in two and (for an accordingly changed value of σ^* also for) higher dimensions.
Of more practical relevance is the predicted behaviour at small values of σ. Although the reconstruction error further decreases with σ decreasing below σ^*, the topography of the map is deteriorated, where, practically, faults between domains of unidirectional symmetry breaking are more obstaculous than slight deviations in the domain of a neuron. Our results imply that hexagonal neural connectivities which where mentioned already in [5] are more stable at small values of σ than usually applied square connectivities.

Acknowledgement

THE REPORTED RESULTS ARE BASED ON WORK DONE IN THE PROJECT LADY SPONSORED BY THE BFMT (F. R. GERMANY) UNDER GRANT 01 IN 106B/3.

References

[1] I. Csabai, T. Geszti, G. Vattay (1992): Criticality in the one–dimensional Kohonen neural map. *Physical Review A* **46**, R6181–R6184.

[2] R. Der, M. Herrmann (1993) Phase Transitions in Self–Organized Feature Maps. In: S. Gielen, B. Kappen: *Proc. ICANN'93, Amsterdam.* Springer, London, Berlin, Heidelberg. p. 597–600. To appear in *Phys. Rev. E.*

[3] R. Der, Th. Villmann (1993) Dynamics of self–organized feature mapping. In: J. Mira, J. Cabestany, A. Prieto: *New trends in neural computation.* LNCS 686. Springer, Berlin, Heidelberg, New York. p. 312–315.

[4] E. Erwin, K. Obermayer, K. Schulten (1992) Self–organizing maps: Stationary states, metastability and convergence rate. *Biol. Cybern.* **67** 35–45. & Self–organizing maps: Ordering, convergence properties and energy functions. *Biol. Cybern.* **67** 47–55.

[5] T. Kohonen (1984) *Self–organization and associative memory.* Springer Series in Information Science **8**, Springer, Berlin, Heidelberg, New York.

[6] H. Ritter, T. Martinetz, K. Schulten (1992) *Neural computation and self–organizing maps.* Addison Wesley, Reading, Massachusetts.