# Analyzing the Formation of Structure in High-Dimensional Self-Organizing Maps Reveals Differences to Feature Map Models

Maximilian Riesenhuber[1], Hans-Ulrich Bauer[2] and Theo Geisel[2]

[1] Center for Biological & Computational Learning and Department of Brain & Cognitive Sciences, MIT, Cambridge, MA 02142, USA
[2] Max-Planck-Institut für Strömungsforschung, Göttingen and SFB 185 "Nichtlineare Dynamik", Univ. Frankfurt, 60054 Frankfurt, Germany

**Abstract.** We present a method for calculating phase diagrams for the high-dimensional variant of the Self-Organizing Map (SOM). The method requires only an ansatz for the tesselation of the data space induced by the map, not for the explicit state of the map. Using this method we analyze two recently proposed models for the development of orientation and ocular dominance column maps. The phase transition condition for the orientation map turns out to be of different form than of the corresponding low-dimensional map.

## 1 Introduction

The high-dimensional self-organizing map (SOM) and the low-dimensional self-organizing feature (SOFM) map have been used to model a variety of self-organizational paradigms in the brain and in technical applications [1]. In both variants stimuli from an input space are mapped to a lattice of output elements (neurons), each characterized by a position $\mathbf{r}$ in the lattice plus a receptive field $\mathbf{w_r}$. A stimulus $\mathbf{v}$ is mapped onto that neuron $\mathbf{s}$ whose receptive field $\mathbf{w_s}$ matches $\mathbf{v}$ best. This amounts to a winner-take-all rule, i.e. a strong lateral nonlinearity.

In the SOM (as well as in previous [2] or more recent [3] formulations of map formation processes) stimuli are normalized patterns of activity in a high-dimensional space (eg. images on a discretized retina). The map results as a stationary state of a self-organization process, which successively changes all receptive fields $\mathbf{w_r}$,

$$\Delta\mathbf{w_r} = \epsilon h_{\mathbf{rs}}(\mathbf{v} - \mathbf{w_r}), \qquad h_{\mathbf{rs}} = e^{-\frac{||\mathbf{r}-\mathbf{s}||^2}{2\sigma^2}}, \tag{1}$$

following the presentation of stimuli $\mathbf{v}$. $\epsilon$ controls the size of learning steps. The neighborhood function $h_{\mathbf{rs}}$ enforces neighboring neurons to align their receptive fields, thereby imposing the property of topography on the SOM. The best-matching neuron $\mathbf{s}$ for a particular stimulus is determined by

$$\mathbf{s} = \arg\max_{\mathbf{r}}(\mathbf{w_r} \cdot \mathbf{v}). \tag{2}$$

In the low-dimensional SOFM the full stimulus distribution $\mathbf{v}$ and the receptive field distributions $\mathbf{w_r}$ are replaced by (low-dimensional) features $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{w}}_\mathbf{r}$ (eg. center of gravity-coordinates) which can be extracted from $\mathbf{v}$ and $\mathbf{w_r}$ by application of a linear operator $R$,

$$\tilde{\mathbf{v}} = R(\mathbf{v}), \qquad \tilde{\mathbf{w}}_\mathbf{r} = R(\mathbf{w_r}). \tag{3}$$

Application of the operator $R$ to the learning rule (1) yields (exploiting the linearity of $R$ and using (3)) the SOFM learning rule

$$R(\Delta \mathbf{w_r}) = R(\epsilon h_{\mathbf{rs}}(\mathbf{v} - \mathbf{w_r})) = \epsilon h_{\mathbf{rs}}(\tilde{\mathbf{v}} - \tilde{\mathbf{w}}_\mathbf{r}). \tag{4}$$

So the features $\tilde{\mathbf{w}}_\mathbf{r}$ obey dynamics of the same form as the dynamics for the full receptive field parameters $\mathbf{w_r}$.

However, as the dot product is not a useful measure in the feature space, the mapping rule (2) has to be changed for the SOFM. Here, the best-matching neuron $\mathbf{s}$ is determined as the neuron whose receptive field has the smallest Euclidean distance to the stimulus.

These different mapping rules would yield identical results when operating on the same vectors $\mathbf{w}$ and $\mathbf{v}$ only if the vectors were square normalized (whereas in the SOM, they are sum-normalized). Even more important, the distance measures operate on vectors in different spaces which can yield different best-matching neurons even if the normalization issue is ignored (illustrated in Fig. 1). Therefore, the two versions of the algorithm can possibly deliver different structure formation results when applied to analogous systems.

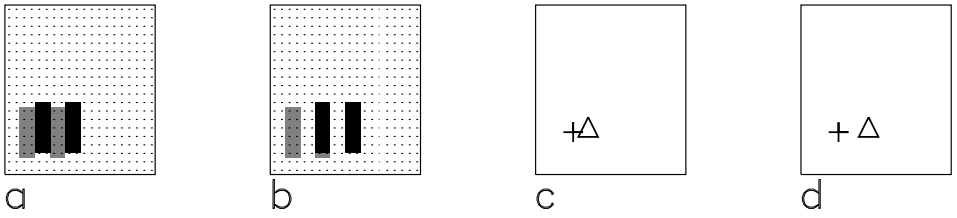


**Figure 1**: SOMs (**a,b**) and SOFMs (**c,d**) can yield different winner neurons when stimulated with analogous stimuli. Consider a square input space, discretized by a square lattice for the SOM case. The stimulus is a double bar, as indicated as the (combined) gray regions. Two exemplary neurons have receptive fields indicated by the (combined) black regions in (a) and (b), resp. The stimulus-receptive field overlap is larger in **b**, so this neuron will be the winner. In (c,d), stimuli and receptive fields are represented by their centers of gravity (crosses and triangles, resp.). Now, the neuron depicted in **c** (and **a**) is the winner.

## 2 Evaluating tesselations of data sets

So far, a mathematical analysis of SOM map development process has only been achieved for the feature map variant [9] but not for the high-dimensional

SOMs. Here, we propose to make an ansatz for plausible states of the map, to evaluate a distortion function for these states and to determine the state of lowest distortion. Since such an ansatz in terms of the explicit weight vectors $\mathbf{w_r}$ is difficult, if not impossible to make, we consider a new distortion measure

$$E_{\mathbf{v}} = \sum_{\mathbf{r}} \sum_{\mathbf{r}'} \sum_{\mathbf{v}' \in \Omega_{\mathbf{r}'}} \sum_{\mathbf{v} \in \Omega_{\mathbf{r}}} (\mathbf{v}' - \mathbf{v})^2 e^{\left( - \frac{||\mathbf{r} - \mathbf{r}'||^2}{2\sigma^2} \right)}, \tag{5}$$

which requires only an ansatz for the stimulus space tesselations as given by the $\Omega_{\mathbf{r}}$ ($\Omega_{\mathbf{r}}$ denotes the set of stimuli $\mathbf{v}$ which are mapped onto node $\mathbf{r}$. By Eq. (2 and depends on the $\mathbf{w_r}$ in an implicit way). Under quite general assumptions, the minima of $E_{\mathbf{v}}$ coincide with those of the naive "energy function" for the SOM-model [7] in the limit of $\sigma \to 0$, and the deviations are small otherwise. A great advantage of this new distortion measure lies in the fact that the $\Omega_{\mathbf{r}}$-ansatz necessary for an evaluation of Eq. (5) is comparatively easy to make (see following examples). A more detailed account of this method can be found in forthcoming publications [5, 6].

## 3 Results

Using Eq. (5), we now analyze two SOM-models for the development of visual maps, one for the development of orientation maps, in which the emergence of oriented receptive fields is contingent upon the presence of oriented stimuli, and one for the development of ocular dominance maps. Finally we make a brief reference to a third model in which oriented receptive fields arise from non-oriented stimuli, a symmetry breaking phenomenon which cannot be described in feature map approximation but only in high-dimensional SOMs.

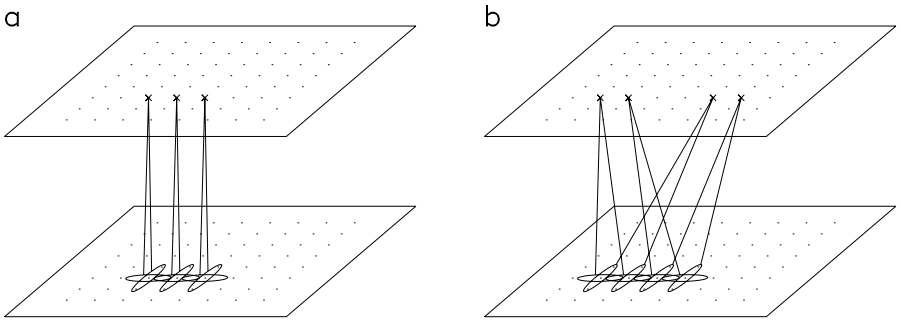


**Figure 2**: SOM-model for the formation of orientation maps. Oriented stimuli in a retinal layer are mapped to a sheet of cortical neurons. **a:** Stimuli with different orientations, but located at the same retinal position, are mapped to the same cortical neuron: non-oriented receptive fields. **b:** Same orientation, different retinal positions: oriented receptive fields.

### 3.1 SOM Model for the Development of Orientation Maps

The first model is concerned with the development of orientation maps. Here the map input space is a two-dimensional sheet of retinal input channels, discretized as a $N \times N$-dimensional grid. Using ellipsoidal Gaussian activity distributions as stimuli (minor axis $\sigma_1$, major axis $\sigma_2 > \sigma_1$), simulations of this model led to maps with orientation preference for substantially elongated stimuli [4]. Using rather circular stimuli ($\sigma_2 \approx \sigma_1$) maps with neurons of no orientation preference were also observed [8]. To simplify calculations we restrict the number of stimulus orientations to two (horizontal and vertical), and the possible stimulus centers to the retinal channels (see Fig. 2).

For this restricted stimulus set, we have twice as many stimuli as neurons so the Voronoy sets $\Omega_{\mathbf{r}}$ contain two stimuli on average. What are sensible stimulus space tesselations which correspond to maps with or without orientation preference of the individual neurons? It is a sensible ansatz to assume that the non-oriented map is characterized by a tesselation of the stimulus set such that stimuli of both orientations, but centered at the retinal location, go to one neuron. In contrast we assume for the maps with orientation preference to have stimuli of the same orientation, but located at neighboring retinal positions in the sets $\Omega_{\mathbf{r}}$. For these different tesselations we can evaluate and compare the resp. values of our distortion measure. We obtain, in the limit of $\sigma_{1,2} \gg \sigma \gg 1$

$$\sigma_{2,crit} \approx \sigma_1 + \sqrt{3}\sigma. \tag{6}$$

The condition (6) for the break of symmetry from non-oriented to oriented receptive fields is very well corroborated by numerical simulations using the reduced stimulus set (Fig. 2), as well as by simulations with the full stimulus set (all orientations and positions). The additive relation between $\sigma_1$ and $\sigma_{2,crit}$ deviates from the multiplicative relation found for a corresponding SOFM [9].

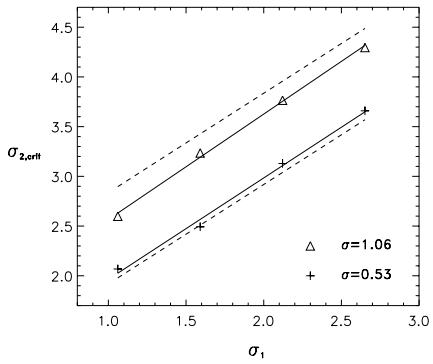


**Figure 3**: Critical value $\sigma_{2,crit}$ for the occurrence of an orientation map, as a function of $\sigma_1$, for two exemplary values of the map neighborhood function width $\sigma$. Symbols indicate the results of simulations of SOMs, the solid line is a fit to these four points, resp. The dashed lines show the corresponding analytic results (6).

This high-dimensional SOM-model for orientation map formation also formed the basis for the simulations in an accompanying paper which provides an explanation for the recently observed similarity of orientation maps which developed under uncorrelated stimulation conditions [10]

### 3.2 SOM-Model for the Development of Ocular Dominance Maps

Second, we analyze a SOM-model for the development of ocular dominance (OD) maps [11]. Two retinal layers for the two eyes are mapped to a cortical output layer. The stimuli are small patches of activity, located at the same (random) position in both input layers, with an amplitude ratio of $1 : c$ or $c : 1$. For $c \approx 1$, the weight vectors $\mathbf{w_r}$ develop symmetrically in both retinae: each $\Omega_\mathbf{r}$ contains two stimuli of opposite ocularity, but centered at the same retinal position (solution type **a**). With decreasing $c$, a symmetry breaking transition takes place, and neurons develop a preference for one or the other retina (ocular dominance). The $\Omega_\mathbf{r}$ now contain two stimuli of the same ocularity, but centered at different (neighboring) positions. Depending on the clustering pattern of same-ocularity neurons in the map, different types of solutions can be distinguished. Here we consider the following arrangements: a chequerboard-type alternation (type **b**), or bands of length $N$ in one direction, and widths 1 (**c**), 2 (**d**), or $N/2$ (**e**) in the orthogonal direction. Type **e** corresponds to a degenerate solution, which can be found in models [13], but not in the visual cortex.

Evaluating the distortions $E_\mathbf{v}$ for the different tesselations, we obtain the phase diagram depicted in Fig. 3a. For decreasing values of $c$, the preferred solution changes from no ocular dominance, via solutions with increasing band width, to the degenerate type **e** solution with only two OD bands. This transition scenario coincides with the findings of a recent neuroanatomical experiment [12]. Simulations corroborated the analytic results very well (Fig. 3b).

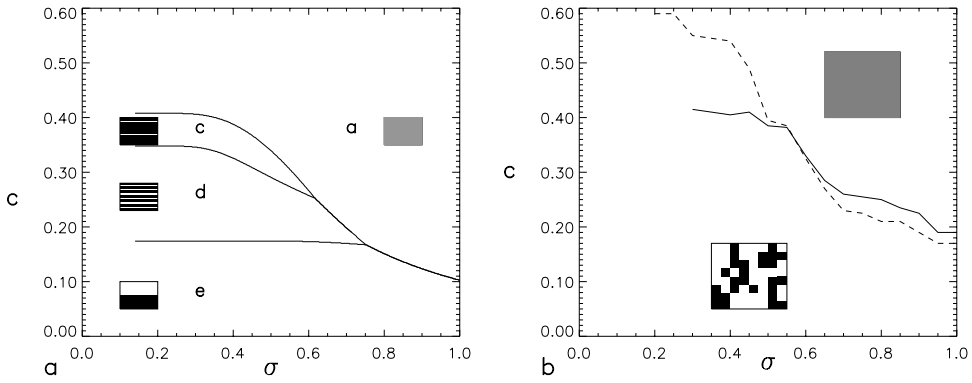


**Figure 4**: Phase diagrams for a SOM-model for the development of ocular dominance maps. **a:** analytical solution. **b:** numerical solution. Solid line: multiplicative normalization Eq. (1), dashed line: subtractive normalization.

### 3.3 Development of oriented receptive fields from non-oriented stimuli

Finally we briefly describe results of a recent project involving an SOM model for the development of orientation maps based on a competition of circular symmetric on-center and off-center cell inputs. The ability or non-ability of SOM-based models to generate a) the receptive fields of individual neurons in addition to the map layout, and b) oriented receptive fields from non-oriented stimuli [3] has been at the center of a recent debate about the merits and disadvantages of different map formation models. A formation of receptive field shapes without prior parametrization of receptive field properties cannot be achieved in feature map approximation but only in the high-dimensional SOM variant. We here report that our above-described method allowed us to calculate the regime of stimulus and map parameters for the On-center- Off-center-cell competition problem. In subsequent simulations we then obtained maps with oriented receptive fields [14]. With regard to the above-mentioned debate we can now state that SOM models exhibit on a qualitative level the same pattern formation properties as a the competing correlation-based models, plus the effect of an increase in ocular dominance stripe width upon less correlated stimulation as observed in a recent experiment. An assessment of more quantitative properties of high-dimensional SOM maps, as well as investigation of a combined orientation and ocular dominance map formation will be adressed in future work. We expect that our new method of analysis will provide a valuable guiding line in these investigations.

## References


1. T. Kohonen, *Self-Organizing Maps*, Springer, Berlin (1995).
2. C. von der Malsburg, Kybernetik **14**, 85 (1973); Biol. Cyb. **32**, 49 (1979).
3. K. D. Miller, J. Neurosci. **14**, 409 (1994).
4. K. Obermayer, H. Ritter, K. Schulten, Proc. Nat. Acad. Sci. USA **87**, 8345 (1990).
5. H.-U. Bauer, M. Riesenhuber, T. Geisel, submitted to Phys. Rev. Letters (1995).
6. M. Riesenhuber, H.-U. Bauer, T. Geisel, submitted to Biol. Cyb. (1995).
7. E. Erwin, K. Obermayer, K. Schulten, Biol. Cyb. **67**, 47 (1992).
8. K. Obermayer, *Adaptive Neuronale Netze und ihre Anwendung als Modelle der Entwicklung kortikaler Karten*, infix Verlag, Sankt Augustin (1993).
9. K. Obermayer, G. G. Blasdel, K. Schulten, Phys. Rev. A **45**, 7568 (1992).
10. K. Pawelzik, H.-U. Bauer, F. Wolf, T. Geisel, Proc. ICANN 96, this volume (1996).
11. G. J. Goodhill, Biol. Cyb. **69**, 109 (1993).
12. S. Löwel, J. Neurosci. **14**, 7451 (1994).
13. G. J. Goodhill, D. J. Willshaw, Network **1**, 41 (1990); P. Dayan, Neur. Comp. **5**, 392 (1993).
14. M. Riesenhuber, H.-U. Bauer, T. Geisel, submitted to CNS 96, Boston (1996).


This article was processed using the LaTeX macro package with LLNCS style