Robust competitive networks

M. Schmutz

Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Nobelstrasse 12, 7000 Stuttgart 80, Germany

W. Banzhaf

Central Research Laboratory, Mitsubishi Electric Corporation, 1-1, Tsukaguchi, Honmachi 8-chome, Amagasaki 661, Japan (Received 22 October 1991)

In the search for networks that implement the winner-take-all function dynamically in a robust way, we study in this paper a class of models that includes short-range diffusive interactions. The analysis is based on a combination of numerical simulations and analytic results obtained by the application of field-theoretical methods. An examination of the ground-state properties reveals that the implementation of robust competition in dimensions higher than one requires a nonstandard diffusive interaction of at least fourth order.

PACS number(s): 87.10. + e

I. INTRODUCTION

Although the physics of large systems and the theory of artificial neural networks both deal with the collective properties of systems made up of many interacting elementary units, they differ in several aspects. In contrast to the building blocks of physical systems, the properties of artificial neurons may be defined freely. More fundamentally, states of networks are considered to code information, and their dynamics is interpreted as information processing. Because of their inherent parallelism and the very different organization mechanism, neural networks are considered to be an interesting alternative to standard computers and their programming.

In view of the virtually inexhaustible richness in the behavior of large systems and in their freedom to code information, the exploration of neural networks as technical devices for information processing seems to be still in its infancy. The merely remote relationship between properties of elementary units and their collective behavior only adds to this impression. In the absence of a general theory of what networks can really do, progress in network design has to rely mainly on a combination of guessing a promising looking network structure, for instance, by borrowing and reinterpreting concepts from physics and analyzing the actual behavior.

In this paper we demonstrate this procedure by taking on the modest task of the construction of a robust competitive dynamics. As a starting point, we choose the competition dynamics developed by Haken [1] for the purpose of pattern recognition. Here the outcome of a competition is a strictly localized state, the only excited neuron representing the winner and all other neurons being quiescent. In the search for a system with robust coding of the information on the winner by distributing it over a tunable number of neurons, we modify the Haken model by adding a diffusive interaction [2]. The motivation for the resulting diffusive Haken model is based on the expectation that the new ground states have to represent a compromise between localizing forces of the original Haken model and the delocalizing diffusive in-

teraction, thus giving rise to bubbles of excitation, the width of which is controllable by the diffusive coupling strength. The main task then is to find our whether or not the model behaves according to this expectation and how.

The definition of a nearest-neighbor coupling introduces concepts of topology and dimension, thereby changing profoundly the properties of the model, with the unpleasant consequence that the diffusive Haken model can no longer be treated analytically in a straightforward manner. The situation becomes much more favorable when we replace the discrete neurons by an excitable continuous medium. As it turns out, the ground-state properties of the resulting field theory can be discussed in great detail, depending on the dimensions of the system, the order of the diffusive interaction, and the coupling strength. In combination with simulations of the discrete version, we obtain a comprehensive picture of ground-state properties of the diffusive Haken model for the field-theoretical as well as the discrete formulation.

The plan of this paper is as follows. In Sec. II we summarize the properties of the original Haken model. In Sec. III we define the diffusive Haken model, including its field-theoretical formulation. In Sec. IV we discuss the ground-state properties using a variational calculation based on an approximation by Gaussians. In Secs. V and VI we adapt the instanton approach to our problem and we give exact solutions for one dimension. In Sec. VII we summarize analytical and numerical results of the field theory in low dimensions and compare them with simulations of the discrete version. Section VIII contains our conclusion.

II. HAKEN MODEL

Motivated by the analogy between pattern formation in synergetic systems and pattern recognition, Haken [1] has formulated a prototype dynamics for the modeling of competition. In this section we summarize its basic properties (for details we refer to [1] and a recent overview [3]).

Driven by a potential $V^{(1)}(q)$, the states $q \in \mathbb{R}^N$ of the system change according to

$$\dot{q} = -\nabla V^{(1)}(q) , \qquad (2.1)$$

where it is assumed

$$V^{(1)}(q) = -\frac{1}{2}D + \frac{1}{2}D^2 - \frac{1}{4}\sum_{i}q_i^4$$
 (2.2)

and

$$D = \sum_{i} q_i^2 \ . \tag{2.3}$$

The resulting equations of motion,

$$\dot{q}_i = (1 - 2D)q_i + q_i^3 , \qquad (2.4)$$

may be interpreted as the dynamics of a network consisting of N cells with activities q_i .

Equations (2.4) are invariant with respect to the replacement of q_i by $-q_i$, and one can easily show that $q_i(t) \ge 0$ for all t > 0 if $q_i(0) \ge 0$. Hence it suffices to investigate the dynamical properties of the system in the invariant positive section $R_+^N = \{q \in R^N | q_i \ge 0, i = 1, ..., N\}$.

Because each degree of freedom is coupled to all others with equal weight, it follows that each trajectory q'(t), resulting from a solution q(t) by an arbitrary permutation of its components, is also a solution.

Being the end points of any motion, the stationary states are of particular interest. Despite the nonlinear nature of the equations of motion [Eq. (2.4)], they may be discussed in complete detail. It turns out that the stationary states falling into R^N_+ may be grouped into m+1 classes. Those forming class $m, 0 \le m \le N$, are the states with m components equal to 1, all others being equal to 0. The stable stationary states are the N vectors of class m=1. All other stationary states are either unstable (m=0) or saddle points (m>1). The potential associated with these states is given by

$$V_m^{(1)} = -\frac{1}{4} \frac{m}{2m-1} \ . \tag{2.5}$$

Because all stable stationary points have the same potential

$$V_1^{(1)} = -\frac{1}{4} \,, \tag{2.6}$$

it follows that they all are global minima of V(q). There are two homogeneous stationary states

$$q_i = \begin{cases} 0, & m = 0 \\ (2N - 1)^{-1/2}, & m = N \end{cases} , \tag{2.7}$$

for all i, which we call the vacuum state and the dissipative state, respectively. For large systems $(N \to \infty)$ the components of the two states become indistinguishable. Nevertheless, the limit states have to be considered as different states because their respective potentials take different values $V_0^{(1)} = 0$ and $V_\infty^{(1)} = -\frac{1}{8}$.

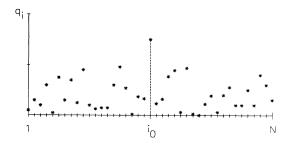
The interpretation of the system (2.1)-(2.3) as a model for competition relies on the conservation of inequalities among the components of a state: If initially $q_i(0) \ge q_i(0)$

for any pair i and j, then $q_i(t) \ge q_j(t)$ for all t > 0. Hence an initial state having typically a unique maximal component $q_{i_0}(0) > q_{i \ne i_0}(0)$ ends up in the stable stationary state with $q_{i_0} = 1$, $q_{i \ne i_0} = 0$. The results of a simulation are shown in Fig. 1. The network thus implements the winner-take-all function [4] in a dynamical way. Such a dynamical competition requires an exchange of information between all degrees of freedom, which is established by the coupling term in the equations of motion [Eq. (2.4)].

III. DIFFUSIVE HAKEN MODEL

The Haken model perfectly performs a dynamical parallel search for the maximum in a set of data. However, the position of the maximum eventually is indicated by a single degree of freedom, i_0 ("cell"), all others being quiescent $(q_{i\neq i_0}=0)$. Therefore, the performance of the system depends on the perfect functioning of each cell, which makes it prone to failure.

It is well known that in large natural systems such as, e.g., the human brain, information typically is spread over many cells in the form of local "bubbles" of excitation or over many synapses [5,6]. This serves as a safeguard against the failure of individual cells. In the following we look for modifications of the Haken model which reflects this behavior in the case of a maximum detector network. A practical motivation for these investigations is that a network with these properties allows, at least in principle, implementation of the learning and recognition algorithm for Kohonen's feature map [5] on a massively parallel architecture. This implementation



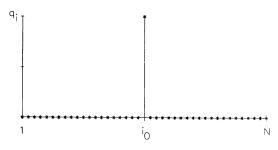


FIG. 1. Example of dynamical competition: (a) initial and (b) final states of a discrete system with components q_i . q_{i_0} is the maximal activity.

avoids the bottleneck of the standard approach [7], which basically relies on a sequential search and thus has a time complexity which increases with the size N of the data.

The potential $V^{(1)}$ of the Haken model drives the system into sharply localized states. Therefore we will add a term $V^{(0)}$ which prefers homogeneous states. As a result of a balance between the two opposing forces, we expect that the ground states (global minima) of the system $V = V^{(0)} + V^{(1)}$ will have a bubble structure of the sort we are looking for. Furthermore, by choosing $V^{(0)}$ symmetrical in the components q_i , we expect that system V again has the properties of a maximum detector, the maximum now being indicated by the center of the bubble. Obvious candidates for $V^{(0)}$ are diffusive potentials of the form

$$V^{(0)}(q) = \frac{1}{2n} \alpha \sum_{\langle i,j \rangle} (q_i - q_j)^{2n} ,$$

$$\alpha \ge 0, \quad n = 1, 2, \dots,$$
 (3.1)

where the bracket indicates summation over pairs of nearest neighbors. Indeed, $V^{(0)}$ takes its minimum equal to zero when all components have the same value.

We thus define the diffusive Haken model by means of the potential

$$V = V^{(0)} + V^{(1)} \,, \tag{3.2}$$

which gives rise to the dynamics

$$\dot{q} = -\nabla V \ . \tag{3.3}$$

The coupling α controls the diffusive interaction, and for $\alpha = 0$ we retain the original Haken model. Whereas in the Haken model the numbering of cells is completely irrelevant, we now assume that the cells are arranged on a d-dimensional square lattice where the notion of nearest neighbors can be defined. To minimize boundary effects in the case of finite systems, we impose periodic-boundary conditions in our simulations.

The diffusive model (3.1)–(3.3) contains three parameters: the power n of the diffusive interaction, its coupling strength α , and the dimension d of the lattice. Except for occasional general considerations, we will concentrate on the standard diffusive interaction $n = 1 \pmod{I}$ and, for reasons which will become obvious later, on the non-standard interaction $n = 2 \pmod{I}$.

The addition of a diffusive interaction turns the Haken model into a difficult analytical problem, and it appears that detailed insight into its properties can only be obtained by means of numerical simulations. In part, the difficulties stem from the discrete nature of the model. Therefore it seems to be helpful to consider a continuous version of the model. In the resulting field theory, the lattice becomes a d-dimensional medium, states are functions $q(x), x \in \mathbb{R}^d$, and the potential reads

$$V[q] = V^{(0)}[q] + V^{(1)}[q], (3.4)$$

where

$$V^{(0)}[q] = \frac{1}{2n} \alpha \int dx \, d(\nabla q)^{2n} , \qquad (3.5)$$

$$V^{(1)}[q] = -\frac{1}{2}D + \frac{1}{2}D^2 - \frac{1}{4}\int dx^d q^4, \qquad (3.6)$$

$$D = D[q] = \int dx \,^d q^2 \ . \tag{3.7}$$

The equation of motion is given as the functional deriva-

$$\dot{q} = -\frac{\delta V[q]}{\delta a} \ . \tag{3.8}$$

As long as the equilibrium states have only a weak spatial dependence, the discrete model and field theory will show very similar behavior.

IV. GROUND STATES: VARIATIONAL CALCULATIONS

With some luck in the choice of a set of test functions, a variational calculation gives valuable information on the ground states of a system. If the functions are simple and, at the same time, catch essential features of the ground states, we obtain with little effort rigorous upper bounds to the ground-state potential close to the exact value and quantitative insight into the properties of the ground states.

To begin with let us consider the constant states $q(x) = q_0$. For the vacuum state we trivially have

$$q(x)=0, D[q]=0, V[q]=0.$$
 (4.1)

For a finite system with periodic-boundary conditions, we find the two optimal constant states $q(x) = \pm q_0$, with

$$q(x)^2 = 1/(2N-1), D[q] = N/(2N-1),$$

 $V[q] = -N/(8N-4),$
(4.2)

where $N = \int dx^d$ is a measure of the size of the system.

Being unaffected by the diffusive interaction, the constant states (4.1) and (4.2) simply are the continuous versions of the homogeneous states (2.7) of the discrete original Haken model. In both models, for $N \to \infty$, the nontrivial state (4.2) becomes pointwise identical with the vacuum state while taking a nonzero norm $D[q] = \frac{1}{2}$ and a negative potential $V = -\frac{1}{8}$ well below that of the vacuum state. This dissipative state can be considered as the limit of a sequence of states, where the excitation dissipates to infinity, while the norm and potential remain nonzero.

Because the dissipative state is the homogeneous state with lowest potential energy, it follows that the ground states of an infinite system are spatially structured if and only if there exist states with $V[q] < -\frac{1}{8}$. This criterion will play a decisive role in our further discussion.

The potential V is invariant against translations and rotations of the coordinates x. In addition, V has the discrete symmetry V[q] = V[-q]. If each of the respective functions $\pm q(x)$ has a definite sign, we will consider only the positive state.

Because of the symmetries of V, it is not unreasonable to guess that a bubblelike ground state has a rotational symmetry around its center. If we choose the center of symmetry to be the origin of coordinates, then an obvious choice for a two-parameter set of test functions is given by the Gaussians

$$q(r) = q_0 e^{-(sr)^2}, \quad r = |x|,$$
 (4.3)

where q_0 is the amplitude and $\sqrt{d}/\sqrt{2}s$ characterizes the width of the bubble.

For these functions the calculation of the various terms of V[q] reduces to Gaussian integrals which can be performed easily for arbitrary dimensions d. We obtain

$$D = \int dx^{d}q(x)^{2} = \pi^{d/2}q_{0}^{2}s^{-d}2^{-d/2}, \tag{4.4}$$

$$\int dx^d q(x)^4 = \pi^{d/2} q_0^4 s^{-d} 2^{-d} , \qquad (4.5)$$

$$\int dx^{d} [\nabla q(x)]^{2n} = 2n\mu \pi^{d/2} q_0^{2n} s^{2n-d} 2^{-d} , \qquad (4.6)$$

where

$$\mu = (2/n)^n (2n)^{-d/2} - 2^d \Gamma(d/2 + n) / \Gamma(d/2) . \tag{4.7}$$

Using these expressions, we end up with an explicit function $V(q_0,s)$ for the potential V [Eqs. (3.4)–(3.7)]. The best approximation by a Gaussian is then the minimum of $V(q_0,s)$ and, hence, is a solution to the equations

$$\partial_{q_0} V(q_0, s) = 0, \quad \partial_s V(q_0, s) = 0.$$
 (4.8)

They may be written as

$$4\alpha\mu(2n/d-1+n)q_0^{2(n-2)}s^{2n}=1, \qquad (4.9)$$

$$nd/[4n+2(n-1)d]-q_0^{-2}2^{d/2}+2\pi^{d/2}s^{-d}=1$$
. (4.10)

We restrict a detailed discussion to the standard diffusive interaction $n = 1 \pmod{I}$ and the nonstandard diffusive interaction $n = 2 \pmod{II}$.

A. Model I

In one dimension (d=1) Eqs. (4.9) and (4.10) may for n=1 be written as

$$q_0^2 = 2^{5/2} \alpha s^2 \,, \tag{4.11}$$

$$s^2 - 8\sqrt{\pi}s/3 + \frac{1}{3\alpha} = 0. \tag{4.12}$$

The quadratic equation (4.12) has no real solution below the critical coupling

$$\alpha_c = \frac{3}{16\pi} \ . \tag{4.13}$$

Keeping for $\alpha \ge \alpha_c$ only the solution with the lower potential, we get

$$s = s(\alpha) = \frac{4}{3} \sqrt{\pi} (1 - \sqrt{1 - \alpha_c/\alpha}),$$
 (4.14)

which, in combination with (4.11), characterizes the best Gaussian approximation. Its potential is given as

$$V(\alpha) = \frac{1}{8} \left[\frac{3}{32} q_0^4 - \frac{1}{2\sqrt{2}} q_0^2 - 1 \right]. \tag{4.15}$$

For the critical coupling we have

$$q_{0c}^2 = \frac{2^{5/2}}{3}, \quad s_c = \frac{4\sqrt{\pi}}{3}, \quad V_c = -\frac{1}{6}$$
 (4.16)

For increasing values of α , the Gaussian becomes broader and broader and its amplitude decreases. At the same time its potential increases monotonically. For $\alpha \gg \alpha_c$

we have approximately

$$q_0^2 \approx (2^{7/2}\pi\alpha)^{-1}, \quad s \approx (8\sqrt{\pi}\alpha)^{-1},$$
 (4.17)

$$V \approx -\frac{1}{8} - \frac{1}{256\pi\alpha} \ . \tag{4.18}$$

For $\alpha \to \infty$ the Gaussian tends to the dissipative state and $V = -\frac{1}{8}$.

Because the potential of the Gaussians is always below that of the dissipative state, it follows that in one dimension the exact ground states are spatially structured, at least for all $\alpha \ge \alpha_c$.

In two dimensions (d=2), Eqs. (4.9) and (4.10) have a single solution:

$$q_0^2 = 4(\alpha/\alpha_c - 1), \quad \alpha_c = \frac{1}{8\pi}$$
 (4.19)

$$s = \left[\frac{\alpha - \alpha_c}{2\alpha\alpha_c}\right]^{1/2},\tag{4.20}$$

with

$$V = 2\pi\alpha \left[\frac{\alpha}{2\alpha_c} - 1 \right] . \tag{4.21}$$

Here, too, the Gaussians are real only for couplings beyond a critical value α_n .

According to relation (4.20), the width of the optimal Gaussians decreases for increasing values of the coupling α . This behavior runs completely contrary to our intuition about the effect of the diffusive interaction on the ground-state structure. In fact, we have, according to (4.21), $V(\alpha) > -\frac{1}{8}$, for all $\alpha > \alpha_c$ and $V(\alpha_c) = -\frac{1}{8}$, where $q_0(\alpha_c) = 0$ and $s(\alpha_c) = 0$. Hence the potential of all Gaussians with finite amplitude is above that of the dissipative state. We take this result as an indication of a complete absence of bubblelike ground states for model I in two dimensions. We can extend this result to dimensions $2 \le d \le 4$. By use of the conditions (4.9) and (4.10), it follows that, for n = 1,

$$V = -\frac{1}{8} + \frac{1}{16} [(4-d)d2^{-d-3}q_0^4 + (d-2)2^{-d/2}q_0^2],$$
(4.22)

and hence $V \ge -\frac{1}{8}$ for $2 \le d \le 4$.

B. Model II

For the nonstandard diffusive interaction n=2, condition (4.9) for the optimal Gaussian takes a particularly simple form, and we get

$$q_0^2 = (d+4)2^{d/2-2}[(\alpha/\alpha_c)^{d/4} - 1]^{-1},$$
 (4.23)

$$s = \left[\frac{(d+2)(d+4)}{4} \alpha \right]^{-1/4}, \tag{4.24}$$

$$V = -\frac{1}{8} \left[1 + \frac{1}{(\alpha/\alpha_c)^{d/4} - 1} \right], \tag{4.25}$$

where the critical coupling α_c depends on the dimension d:

$$\alpha_c = 4[2/(d+4)]^{4/d}[\pi^2(d+2)(d+4)]^{-1}$$
. (4.26)

Because $V \le -\frac{1}{8}$ for arbitrary dimensions, the variational calculation proves rigorously that the ground states of model II are always spatially structured, at least for $\alpha \ge \alpha_c$. The corresponding considerations for model I are only conclusive for d=1. The calculated upper bound to the ground-state energy for $2 \le d \le 4$ does not exclude the existence of states very different from Gaussians with a potential lower than $-\frac{1}{8}$. However, the discussion of the exact solutions will show that the simple variational calculation given in this section describes the essential properties of ground states astonishingly well.

We end this section with a brief comment on the dissipative state. We can prove its existence directly for an infinite system in the following way: Consider, again, a Gaussian of the form (4.3) and impose the condition $D = \frac{1}{2}$. According to (4.4), this establishes a relation between the amplitude and width of the Gaussian:

$$q_0^2 = 2^{d/2 - 1} \pi^{-d/2} s^d . (4.27)$$

In the limit $q_0 \rightarrow 0$, which implies $s \rightarrow 0$, it follows that $q(x) \rightarrow 0$ and $V[q] \rightarrow -\frac{1}{8}$ for all n and d.

V. INSTANTON APPROACH

In the case of model I, the localized states we are looking for are closely related to the "instantons" of quantum field theory [8-10]. Except for slight modifications resulting from the presence of a second-order volume integral in the potential V, we may directly adopt the methods developed in field theory to our problem. In this section we show that the basic ideas of the instanton approach also form a sufficiently general framework for a discussion of models with nonstandard diffusive interactions.

The exact ground states are the global minima of the potential V[q(x)]. By assuming rotational symmetry q(x) = q(r), r = |x|, the search for these states is reduced for all dimensions d to a one-dimensional problem,

$$V[q(r)] = \gamma_d \int_0^\infty dr \, r^{d-1} \left[\frac{\alpha}{2n} q'^{2n} - \frac{1}{2} q^2 - \frac{1}{4} q^4 \right] + \frac{1}{2} \left[\gamma_d \int_0^\infty dr \, r^{d-1} q^2 \right]^2, \tag{5.1}$$

where q' is the derivative of q(r) and

$$\gamma_d = d\pi^{d/2} / \Gamma(d/2 + 1) \tag{5.2}$$

is the surface of a unit sphere in d dimensions.

The minima of V[q(r)] are solutions to the Euler-Lagrange equation $\delta V/\delta q(r)=0$, which takes the form

$$\alpha q'^{2(n-1)} \left[(2n-1)q'' + \frac{d-1}{r} q' \right] = -\frac{d}{dq} U(q, A) ,$$
(5.3)

where

$$U(q, A) = -\frac{1}{2}Aq^2 + \frac{1}{4}q^4, \qquad (5.4)$$

$$A = A[q(r)] = 2D[q(r)] - 1$$
, (5.5)

and

$$D[q(r)] = \gamma_d \int_0^\infty dr \, r^{d-1} q^2 \,. \tag{5.6}$$

The quantity A which appears in the auxiliary potential U(q,A) is, according to (5.5) and (5.6), a function of the solution q(r). As a result, the Euler-Lagrange equation represents an integro-differential equation. For its solution we proceed in two steps. In the first step we treat A as a parameter, and we thus consider Eq. (5.3) as an ordinary differential equation. Second, from the entire set of its solutions q(r,A), we then pick candidates for the ground state by imposing the self-consistency (SC) condition

$$A = A[q(r, A)], (5.7)$$

given by (5.5) and (5.6).

Because we already know from our variational calculation that the potential of ground states does not exceed the value $-\frac{1}{8}$, we may ignore all solutions q(r,A) with higher potential. In particular, we may, of course, ignore all solutions with a diverging potential $V \to \infty$. A very useful criterion in this respect is the following: Assume that |q(r)| is bounded and suppose that |q(r)| remains nonzero or tends to zero so slowly for $r \to \infty$ that the square of its norm D diverges. Then the potential V[q] is dominated by the second-order volume integral $\frac{1}{2}D^2$ [see (3.4)-(3.7)] and hence $V[q] \to \infty$. As a result, we may ignore all non-normalizable states.

Turning to the first step, the starting point for the instanton approach is the interpretation of q as a particle position and of r as a time. With this interpretation Eq. (5.3) is the equation of motion of a particle moving in one dimension and subject for n > 1 to a somewhat peculiar dynamics. The trajectories q(r, A) obey a least-action principle, where the action may be written as

$$S[q] = \gamma_d \int_0^\infty dr \, r^{d-1} L(q, q')$$
, (5.8)

$$L[q,q'] = \frac{\alpha}{2n} q'^{2n} - U(q,A) . {(5.9)}$$

For a Hamiltonian formulation we refer to Appendix A. There the energy of the particle is defined as

$$E = \alpha \frac{2n-1}{2n} q^{2n} + U(q, A) , \qquad (5.10)$$

which changes along a trajectory according to

$$E' = -\alpha \frac{d-1}{r} q'^{2n} \begin{cases} = 0 & \text{for } d = 1\\ \le 0 & \text{for } d = 2 \end{cases}$$
 (5.11)

The one-dimensional case d=1 is unique in that for arbitrary diffusive interactions the energy is a constant of motion. As a consequence, the particle performs an undamped motion and solutions may be grouped into the following classes. First, all solutions with $E\neq 0$ are nonnormalizable periodic functions of r [see Figs. 2(a) and 2(b)]. Second, for E=0 we have for all values of A the trivial solution q(r)=0, which has the two interpretations as the vacuum state and the dissipative state. Finally, for E=0 and A>0 there exists the unique instanton

solution where the particle starts for r=0 at $q_0=(2A)^{1/2}$, accelerates downhill the auxiliary potential U(q,a), moves uphill again, and comes to a rest at the local hill q=0 [see Fig. 2(b)].

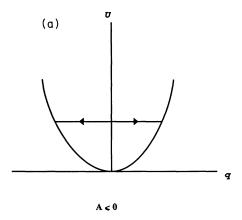
Because we may ignore non-normalizable solutions, for d=1 there are only two candidates for the ground state: the dissipative state with $V=-\frac{1}{8}$ and the self-consistent instanton state. The latter follows by integrating the energy-conservation law (5.10) specialized to E=0,

$$q' = -\left[-\frac{2n}{\alpha(2n-1)}U(q,A)\right]^{1/(2n)},$$
 (5.12)

and then imposing the SC condition (5.7).

For dimensions $d \ge 2$ the energy E is a monotonically decreasing Liapunov function. The particle undergoes a damped motion caused by the second term on the left-hand side (LHS) of the equation of motion [Eq. (5.3)], and each solution eventually settles down for $r \to \infty$ at a stationary point of U(q, A).

For $A \le 0$ all solutions tend, for $r \to \infty$, to the single minimum q = 0. During the approach, the particle performs damped oscillations where the damping rate tends to zero. By treating the damping in this limit as an adiabatic perturbation, we may determine the asymptotic behavior of q(r, A). In Appendix B we show that in two



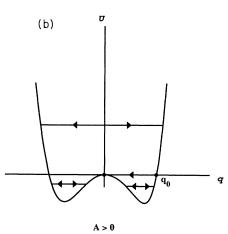


FIG. 2. Undamped motion of the particle in the auxiliary potential U(q, A) for (a) $A \le 0$ and (b) A > 0. $q_0 = \sqrt{2A}$.

and three dimensions all solutions of models I and II (n=1,2) different from q(r)=0 have a divergent norm and may therefore be ignored.

For A>0 the potential U(q,A) has two minima and one local maximum. As a consequence, two classes of solutions exist. Those tending to the minima trivially are non-normalizable because they approach a nonzero constant for $r\to\infty$. In addition, there exists a stationary solution q(r)=0 and a denumerable set of instanton solutions which either directly tend to q=0 or approach this value after one or several bounces at the auxiliary potential U(q,A) (see Fig. 3). Because the lowest-order instanton has no node, it seems reasonable to guess that its potential V is lower than that of all other instantons.

In summary, the search for the ground state in the cases of interest has been reduced to the question of whether or not the unique or lowest-order self-consistent instanton solution has a potential V lower than the dissipative state. The results of our variational calculation show that in case of model II for all dimensions d the answer is in favor of instantons. A similar conclusive answer follows in model I only for d=1.

VI. INSTANTONS IN ONE DIMENSION

In one dimension the instantons follow by integration from (5.1). When setting

$$q = \sqrt{2A}Q , \qquad (6.1)$$

$$a = a(A, n, \alpha) = \left[\alpha 2 \frac{n-1}{2n} 2^n A^{n-2} \right]^{1/2n}, \quad (6.2)$$

the differential equation (5.12) acquires the form

$$Q' = -\frac{1}{a(A)}Q^{1/n}(1-Q^2)^{1/2n} . (6.3)$$

The inverse of the instanton solution r = r(Q, A) then follows as

$$r = -a(A) \int_{1}^{Q} dq \ q^{-1/n} (1 - q^{2})^{-1/2n} \ . \tag{6.4}$$

Relation (6.3) gives Q' as a function of Q. This allows us

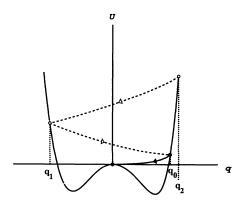


FIG. 3. Instantons of different order in the presence of damping. q_0 , q_1 , q_2 are the starting points.

to express any time integral involving the solution Q(r, A) and its derivative as an explicit integral:

$$\int_0^\infty dr \, F(Q, Q') = -\int_0^1 dQ \frac{F(Q, Q'(Q))}{Q'(Q)} \ . \tag{6.5}$$

Making use of this relation, the SC condition (5.7) may be formulated as

$$A = 8 Ad(A) - 1 , (6.6)$$

where

$$d(A) = a(A) \int_0^1 dq \ q^{2-1/n} (1-q^2)^{-1/2n} \ . \tag{6.7}$$

The complete determination of the self-consistent instantons requires the calculations of the two integrals on the RHS of (6.4) and (6.7) and the solution of SC condition (6.6). Both integrals involve binomial expressions, and because of the integer nature of n, they can be expressed in terms of elementary functions only for n = 1 and 2.

A. Model I

Taking into account (6.1) and (6.2), we readily obtain the instanton solution for n = 1:

$$q(r) = q_0 \left[\cosh \frac{q_0}{\sqrt{2\alpha}} r \right]^{-1}, \tag{6.8}$$

where the amplitude q_0 is related to the parameter A by

$$q_0 = \sqrt{2A} \quad . \tag{6.9}$$

Making use of this relation, the SC condition becomes a quadratic equation for q_0 . Keeping only the solution with the lowest potential, we obtain the amplitude as a function of the coupling

$$q_0 = q_0(\alpha) = \sqrt{2\alpha/\alpha_c} (1 - \sqrt{1 - \alpha_c/\alpha})$$
, (6.10)

which is real only for couplings larger than the critical value

$$\alpha_c = \frac{1}{16} . \tag{6.11}$$

Having arrived at an analytic expression for the instanton solution, all quantities of interest can be calculated explicitly. For the norm and potential we obtain

$$D[q] = (8\alpha)^{1/2}q_0, (6.12)$$

$$V[q] = -(2\alpha)^{1/2}q_0 + 4\alpha q_0^2 - \frac{(2\alpha)^{1/2}}{6}q_0^3.$$
 (6.13)

The geometrical structure of the instanton may be roughly characterized by its amplitude $q_0 = q_0(\alpha)$ given by (6.10) and by its half-width

$$\Delta x^{2} = \left[\int_{0}^{\infty} dr \, r^{2} q(r) \right] / \int_{0}^{\infty} dr \, q(r) , \qquad (6.14)$$

for which we obtain

$$\Delta x = \frac{\pi}{q_0} \left[\frac{\alpha}{2} \right]^{1/2} . \tag{6.15}$$

At the critical coupling α_c , the various quantities take

the values

$$q_{0c}^2 = 2$$
, $\Delta x_c = \pi/8$, (6.16)

$$D_c = 1, \ V_c = -\frac{1}{6} \ .$$
 (6.17)

For $\alpha \gg \alpha_c$ the α dependence may be approximated by

$$q_0^2 \approx \frac{1}{32\alpha}, \quad \Delta x \approx 4\pi\alpha$$
, (6.18)

$$V \approx -\frac{1}{8} - \frac{1}{768\alpha} \ . \tag{6.20}$$

B. Model II

For n=2 the algebraic integral in (6.4) may be turned into an integral of a rational function by use of the substitution

$$t = t(q) = \left[\frac{1}{q^2} - 1\right]^{1/4}.$$
 (6.21)

The inverse of the instanton solution Q = Q(r), which here is independent of A, then follows as

$$r = \frac{(3\alpha)^{1/4}}{2^{3/2}} \left[\ln \frac{t^2 - \sqrt{2}t + 1}{t^2 + \sqrt{2}t + 1} + 2 \arctan \frac{\sqrt{2}t}{1 - t^2} + 2\pi\Theta(t - 1) \right], \quad (6.22)$$

where t = t(Q) and $\Theta(t)$ is the step function.

We observe that Q(r) vanishes for $r \ge r_0$, where

$$r_0 = \frac{(3\alpha)^{1/4}}{\sqrt{2}}\pi \ . \tag{6.23}$$

For values $r \lesssim r_0$ the solution Q(r) may be approximated by

$$Q(r) \approx \frac{1}{4\sqrt{3}\alpha} (r - r_0)^2$$
 (6.24)

In the language adopted in the last section, the existence of a finite value r_0 means that the "particle" arrives at the local hill of U(q, A) at q = 0 in a finite "time."

From (6.6) and (6.7) we obtain the self-consistent value of A

$$A = \frac{1}{2r_0 - 1} , (6.25)$$

and hence

$$q_0^2 = 2A = \frac{2}{\sqrt{2}\pi(3\alpha)^{1/4} - 1} . \tag{6.26}$$

The amplitude q_0 has to be real, and hence the instantons exist only for $\alpha > \alpha_c$, where

$$\alpha_c = \frac{1}{12\pi^4} \ . \tag{6.27}$$

Because of the complicated form of the instanton solution, we approximate its half-width simply by $r_0/2$; i.e., we set

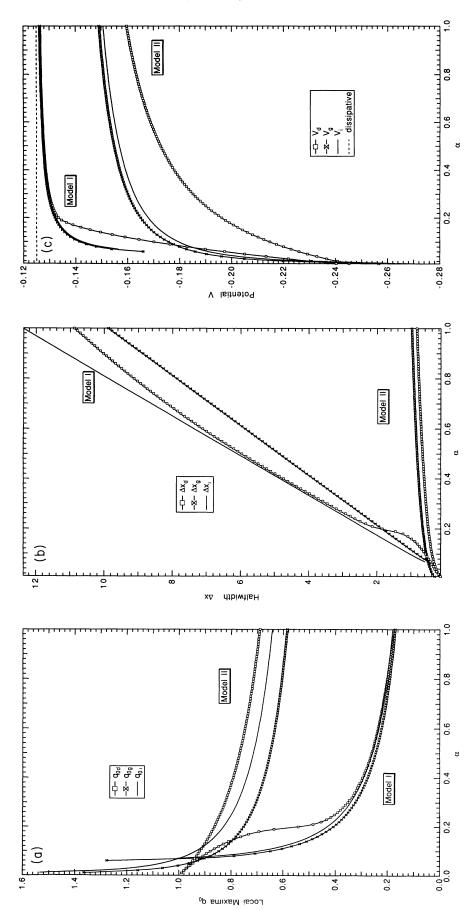


FIG. 4. Characteristic quantities of ground states of models I and II in one dimension as a function of the diffusive coupling: (a) amplitude, (b) half-width, and (c) potential V. As indicated by indices in the legend, the different types of curves refer to the discrete model (d), the Gaussian approximation (g), and the instanton solution (i). Number of neurons,

$$\Delta x \approx \frac{(3\alpha)^{1/4}}{2\sqrt{2}}\pi \ . \tag{6.28}$$

In contrast, the various terms of the potential V may easily be calculated by use of relation (6.5), and we arrive at the simple expression

$$V[q] = -\frac{1}{8} - \frac{1}{16}q_0^2 . {(6.29)}$$

From (6.13) and (6.29) we see explicitly that in one dimension the potential of the self-consistent instantons of models I and II is lower than that of the dissipative state $(V[q] < -\frac{1}{R})$.

VII. GROUND STATES OF THE DIFFUSIVE HAKEN MODEL IN ONE AND TWO DIMENSIONS

Because of damping, an investigation of instantons in dimensions d > 1 requires a numerical analysis. It is greatly simplified by use of scaling laws which relate selfconsistent instantons and other quantities of interest for arbitrary couplings α to corresponding quantities for, say, $\alpha = 1$ (for details see Appendix C). Combining the calculations for d=2 with the analytical results derived so far and supplementing them with simulations of discrete versions of models I and II for dimensions d = 1and 2, we finally are in the position to discuss in detail the ground-state properties of these models in low dimensions. The results give insight into the similarities and dissimilarities of both models; they allow comparison of discrete and continuous versions and, in latter cases, a judgment of the accuracy of variational calculations based on a Gaussian approximation.

Figures 4(a)-4(c) show characteristic quantities of ground states (amplitude, half-width, potential) of both models for d=1 as functions of the coupling α . We observe that the discrete models show at $\alpha=0$ a continuous transition from strictly localized to delocalized ground states, which for $\alpha\to\infty$ seemingly tend to the dissipative state. This approach is much more rapid for model I than for model II. This behavior suggests the interpretation that for given α value the nonstandard diffusive interaction has a much less delocalizing effect than the standard one.

The same behavior is reflected by the ground states of the field theories. The differences in the properties of the continuous and discrete versions present for weak couplings diminish with increasing coupling. This is in complete agreement with our expectations because for large values of α the ground states show only weak spatial variations.

The most striking result is that, except for very weak couplings, the properties of the simple Gaussian approximation not only agree astonishingly well with those of the exact ground states (instantons) of the field theories, but also with those of the discrete models.

Turning to two dimensions, we recall that the only available analytical result in the form of a Gaussian approximation did not definitely decide on the nature of the ground state of continuous model I. Figure 5 shows that the potential of the lowest-order instanton stays, in fact, always above that of the dissipative state. Since higher-order instantons have even higher potentials, we conclude that the dissipative state actually is the ground state for all values of $\alpha > 0$. Simulations of the discrete version of model I lead to the same conclusion. Hence model I undergoes for d=2 a singular transition at $\alpha=0$ in the sense that the presence of an arbitrarily weak standard diffusive interaction delocalizes the ground states of the Haken model completely. Taking the analytical results of the variational calculation as sufficient evidence, we may extend this behavior, at least to dimensions d=4.

In contrast, Figs. 6(a)-6(c) demonstrate that ground states of model II behave in two dimensions very similar to those of models I and II in one dimension. The results of the Gaussian approximation suggests that the non-standard diffusive interaction gives rise to a continuous transition to delocalized states in all dimensions.

Turning to the dynamical behavior of the discrete models I and II, we present results of simulations in Figs. 7 and 8. They show the final states evolving from given one- and two-dimensional initial states. In agreement

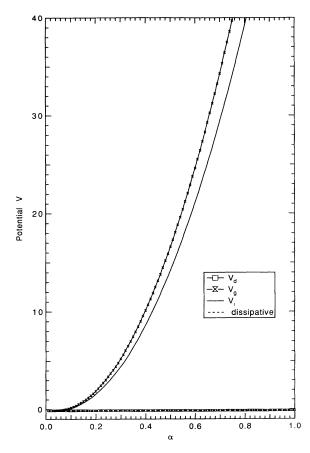


FIG. 5. Potential V of the ground states of model I in two dimensions as a function of the diffusive interaction: A comparison of simulations of the discrete model (d) with the Gaussian approximation (g), the lowest-order instanton (i), and the dissipative state.

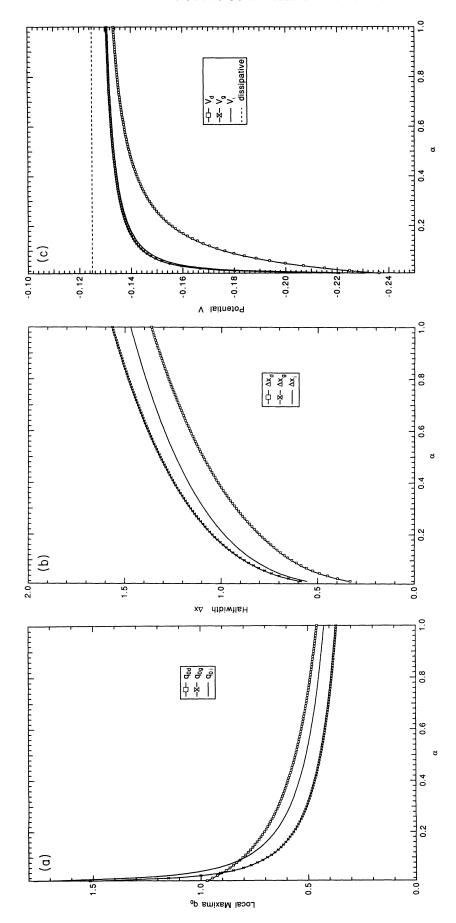


FIG. 6. (a) Amplitude, (b) half-width, and (c) potential V of the ground states of model II in two dimensions as a function of the diffusive coupling: Comparison of the discrete model (d), Gaussian approximation (g), and lowest-order instantion (i). $N=40\times40$.

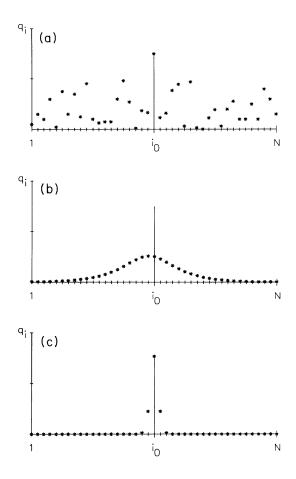
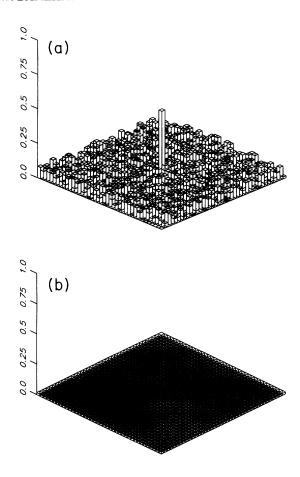


FIG. 7. Robust dynamical competition in one dimension: Neuron activity for (a) initial state, (b) final state of model I, and (c) final state of model II. N = 40.

with the foregoing discussion, the final state of model I for d=2 is the flat dissipative state [Fig. 8(b)]. In all other cases the final states are bubblelike excitations (ground states) usually centered at the position of the maximal initial excitation, thus demonstrating that here the models implement the winner-take-all function in a robust way.

VIII. CONCLUSION

In the search for a robust competitive network, we have studied generalized versions of the Haken model which result by adding a diffusive interaction of variable order to its defining potential. The inclusion of such a short-range interaction brings into play the aspects of topology and dimension and thus enriches the structure of the model. We have obtained a comprehensive picture of its ground-state properties and their dependence on the various parameters by combining the results of numerical simulations and of analytic investigations based on methods borrowed from field theory. In particular, it turned out that the ground-state properties can be deter-



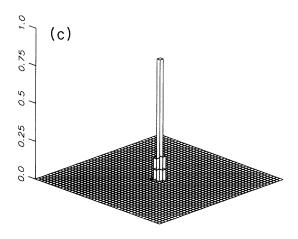


FIG. 8. Simulation of the dynamics in two dimensions: $N = 40 \times 40$. Neuron activity q_{ij} , $1 \le i, j \le 40$, for (a) initial state, (b) final state of model I, and (c) final state of model II.

mined astonishingly well by a simple variational calculation based on Gaussian test functions. As a major result, we found that models which implement robust competition in varying dimensions can be constructed along the lines pursued in this paper only by use of nonstandard diffusive interactions.

ACKNOWLEDGMENTS

This work was made possible by a grant from the Volkswagen-Stiftung to M.S. and by an invitation of W.B. to Mitsubishi Electric's Central Research Laboratory. We thank Dr. J. Bell for assistance in editing the manuscript.

APPENDIX A: HAMILTONIAN FORMULATION

According to (5.8) and (5.9), the action is of the form

$$S[q] = \gamma_d \int_0^\infty dr \, r^{d-1} L(q, q') , \qquad (A1)$$

$$L(q,q') = \frac{\alpha}{2n} q'^{2n} - U(q,A)$$
 (A2)

Note that (A1) differs from the standard form used in mechanics by the "time" function r^{d-1} . We could avoid this deviation by incorporating this factor in a definition of the Lagrangian, which then would depend explicitly on time for d>1. However, such a formulation would give rise to rather clumsy expressions for conjugate momenta. We prefer the form (A1), which explicitly keeps track of the damping in the basic equations.

The Euler-Lagrange equation for (A1) reads

$$\frac{d}{dr}(\partial_{q'}L) + \frac{d-1}{r}\partial_{q'}L = \partial_{q}L .$$
(A3)

When defining conjugate momentum and the Hamiltonian in the usual way,

$$p = \partial_{a'} L , \qquad (A4)$$

$$H(q,p) = pq' - L(q,q'), \qquad (A5)$$

where q' = q'(q,p), the modified Hamilton equations of motion follow as

$$q' = \partial_n H$$
, (A6)

$$p' + \frac{d-1}{r}p = -\partial_q H . (A7)$$

For d = 1 the damping term on the LHS of (A7) is absent and the energy

$$E = H(q(r), p(r)) \tag{A8}$$

is a constant of motion.

For d > 1 the energy changes along a trajectory according to

$$E'(r) = -\frac{d-1}{r}pq' . (A9)$$

For the Lagrangian (A2) the Euler-Lagrange equation reads

$$\alpha q^{\prime(2n-1)}\left[(2n-1)q^{\prime\prime}+\frac{d-1}{r}q^{\prime}\right]=-\frac{d}{dq}U(q,A).$$

(A10)

The corresponding equations of motion in phase space take the form

$$q' = (p/\alpha)^{1/(2n-1)}$$
, (A11)

$$p' + \frac{d-1}{r}p = -\frac{d}{dq}U(q, A)$$
 (A12)

The energy may be written as

$$E(r) = \alpha \frac{2n-1}{2n} q^{2n} + U(q, A)$$
 (A13)

and

$$E'(r) = -\alpha \frac{d-1}{r} q'^{2n} \begin{cases} = 0 & \text{for } d = 1\\ \le 0 & \text{for } d \ge 2 \end{cases}.$$
 (A14)

In agreement with our intuition about the effects of damping, the energy is a monotonically decreasing function of the time for $d \ge 2$. We observe, however, that for large times the damping effects become arbitrarily small and can be considered as a perturbation to a conservative system.

We add the following remarks. A partial initial condition for an instanton solution to (A11) and (A12) consists in setting p(0)=0. Then, for r=0, both, the numerator and denominator in the second term on the LHS of (A12) vanish and it is convenient to switch for small values of r, say, $0 \le r < 1$, to a new time variable $u = \ln r$. When setting

$$q(r)=f(u),$$

$$p(r)=g(u),$$
(A15)

we obtain the nonsingular differential equations

$$f' = e^{u} (g/\alpha)^{1/(2n-1)},$$

$$g' + (d-1)g = -e^{u} \frac{d}{df} U(f, A).$$
(A16)

They form a convenient starting point for numerical integration, e.g., by use of a Runge-Kutta method.

APPENDIX B: ASYMPTOTIC BEHAVIOR

The Euler-Lagrange equation (5.3) cannot be solved in closed form for dimensions d > 1. Nevertheless, we may derive the correct asymptotic behavior of the solutions by treating the damping for large values of r as an adiabatic perturbation. The resulting information is sufficient to decide on the normalizability of the solutions. In particular, we are interested in the norm of the solutions for $A \le 0$. By use of a scaling argument, it suffices to confine our considerations to the two cases A = 0 and -1 for a typical coupling, say, $\alpha = 1$.

The starting point of the perturbation method consists of the periodic solutions of the undamped system for d=1. They obey the differential equation

$$\frac{2n-1}{2n}q'^{2n}+U(q,A)=E,$$
 (B1)

where

$$U(q,0) = \frac{1}{4}q^4$$
 (B2)

For sufficiently small amplitudes we may approximate U(q,-1) by the quadratic term

$$U(q,-1) \approx \frac{1}{2}q^2 . \tag{B3}$$

The value of the energy E fixes the amplitude of a specific solution q(r,E), which is unique up to the shift $r \rightarrow r + r_0$ and up to the symmetry $r \rightarrow -r$. By use of a scaling ansatz, we may obtain from (B1) the E dependence of q(r,E). For A=0 and -1 we find

$$q(r,E) = \begin{cases} E^{1/4}q_1(E^{(2-n)/4n}r), & A = 0 \\ E^{1/2}q_1(E^{(1-n)/2n}r), & A = -1 \end{cases},$$
 (B4)

where $q_1(r) = q(r, E = 1)$ are the respective solutions for the specific energy E = 1.

For systems with dimensions $d \ge 2$, the equation of motion contains a damping term, which becomes arbitrarily small for large times r, and the energy will be a slowly decreasing function tending to zero for $r \to \infty$. In this limit we approximate the damped solutions by expressions (B4) and (B5), respectively, where now E = E(r).

The major task then consists of a determination of the time dependence of the energy E. For this purpose we consider its decay rate

$$E'(r) = -\frac{d-1}{r} q'^{2n} , (B6)$$

which will show a long-time trend, onto which is superimposed a small modulation fluctuating at the time scale of a local period. The basic idea for an adiabatic approximation consists in replacing E'(r) by its average value during a local period R(E),

$$E' \approx -\frac{d-1}{r} \frac{e(E)}{R(E)}$$
, (B7)

where

$$e(E) = \int_{r}^{r+R(E)} d\rho [q'(\rho, E)]^{2n},$$
 (B8)

and where the undamped solutions q(r, E) are used for the calculation of the quantities e(E) and R(E).

In both cases A = 0 and -1, we end up with a differential equation for E(r):

$$E' = -\frac{\kappa}{r}E , \qquad (B9)$$

where

$$\kappa = \begin{cases} \frac{(d-1)4n}{5n-2} & \text{for } A = 0\\ \frac{(d-1)2n}{3n-1} & \text{for } A = -1 \end{cases}$$
 (B10)

Equation (B9) is readily integrated:

$$E(r) \sim r^{-\kappa}$$
 (B11)

Inserting this expression in (B4) and (B5), respectively, we arrive at the asymptotic form of the damped solutions q(r) = q(r, E(r)) for large values of r. The contribution of the asymptotic part to the norm D may be cast into the form

$$\frac{\gamma_d}{B} \int_0^\infty du \ u^{\lambda} q_1(u)^2 \ , \tag{B12}$$

where

$$\beta = \begin{cases} 1 - \frac{(d-1)(2-n)}{5n-2} & \text{for } A = 0\\ 1 - \frac{(d-1)(1-n)}{3n-1} & \text{for } A = -1 \end{cases},$$
 (B13)

$$\lambda = \begin{cases} \frac{(d-1)2n}{\beta(5n-2)} & \text{for } A = 0\\ 0 & \text{for } A = -1 \end{cases},$$
 (B14)

and where $q_1(r)$ are the respective undamped periodic solutions for E=1. The norm D diverges whenever $\lambda > -1$. For A=-1 this is the case for all n and all $d \ge 2$. For A=0 we find $\lambda \ge 0$ for n=1 and $2 \le d \le 4$ as well as for n=2 and $d \ge 2$.

APPENDIX C: SCALING LAWS

The Euler-Lagrange equations (5.3) and (5.4) contain two parameters α and A, where the latter is to be determined from the SC conditions (5.5) and (5.6). By changing the scale of q and r, we may express the solutions for arbitrary couplings by those of a specific value α , say, $\alpha=1$. Denoting all quantities referring to $\alpha=1$ by an index, we can write

$$q(r) = \lambda q_1(\mu r) , \qquad (C1)$$

where $q_1(r)$ is a self-consistent solution for $\alpha = 1$. From this ansatz follow the relations

$$\alpha \lambda^{2n-4} \mu^{2n} = 1 . \tag{C2}$$

$$A = \lambda^2 A_1 . (C3)$$

Taking into account the SC condition, we obtain

$$\lambda^2 = [2D_1(\mu^{-d} - 1) + 1]^{-1}, \tag{C4}$$

where we have used that D scales according to

$$D = \lambda^2 \mu^{-d} D_1 . (C5)$$

When writing the potential in the form

$$V = K - \frac{1}{2}D + \frac{1}{2}D^2 , \qquad (C6)$$

where

$$K = \int dx^{d} \left[\alpha \frac{1}{2n} (\nabla q)^{2n} - \frac{1}{4} q^{4} \right] , \qquad (C7)$$

it follows by use of (C2) that

$$K = \lambda^4 \mu^{-d} K_1 . (C8)$$

Furthermore, the half-width of a solution scales according to

$$\Delta x = \mu^{-1} (\Delta x)_1 , \qquad (C9)$$

and, by definition, we have for the amplitude

$$q_0 = \lambda q_{1_0} . \tag{C10}$$

Given the quantities D_1 , K_1 , and $(\Delta x)_1$ for $\alpha = 1$, we may calculate all quantities of interest for other values of the coupling α . The scale parameters λ, μ follow from Eqs. (C2) and (C4).

We note that (C2) takes a particularly simple form for n = 2 (model II). Here we find

$$\mu = \alpha^{-1/4} , \qquad (C11)$$

$$\lambda^2 = [2(\alpha^{d/4} - 1)D_1 + 1]^{-1}. \tag{C12}$$

- [1] H. Haken, in Computational Systems, Natural and Artificial, Proceedings of the Elmau International Symposium on Synergetics 1987, edited by H. Haken (Springer, Berlin, 1987), p. 2.
- [2] Another way to introduce robustness is to interpret the competition as the effective dynamics of entire populations of neurons; see W. Banzhaf and M. Schmutz, Int. J. Neural Syst. (to be published).
- [3] H. Haken, Synergetic Computers and Cognition (Springer, Berlin, 1991).
- [4] J. A. Feldman and D. H. Ballard, Cognit. Sci. 6, 205 (1982).
- [5] T. Kohonen, Self-Organization and Associative Memory, 2nd ed. (Springer, Berlin, 1987).

- [6] G. Palm, Neural Assemblies (Springer, Berlin, 1988).
- [7] K. M. Marks and K. F. Goser, in Neuro-Nimes '88, Proceedings and Exhibition Catalog of the International Workshop on Neural Networks and their Applications (EC2 and Chambre de Commerce et d'Industrie de Nimes, Nimes, 1988), p. 239.
- [8] S. Coleman, in *The Whys of Subnuclear Physics*, edited by A. Zichichi (Plenum, New York, 1979).
- [9] B. Feldsager, *Geometry, Particles and Fields* (Odense University Press, Odense, 1983).
- [10] J. Zinn-Justin, in Recent Advances in Field Theory and Statistical Mechanics, edited by J. B. Zuber and R. Stora (North-Holland, Amsterdam, 1984), p. 39.

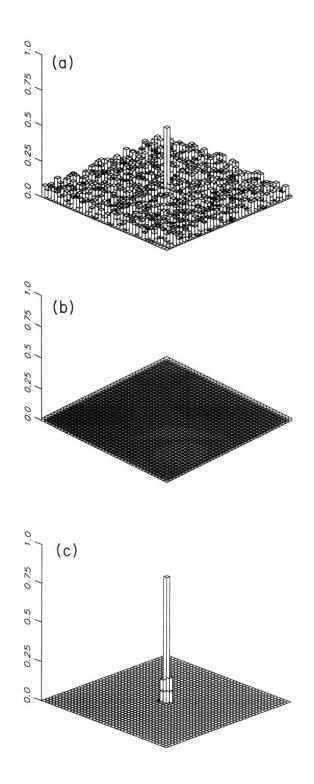


FIG. 8. Simulation of the dynamics in two dimensions: $N=40\times40$. Neuron activity q_{ij} , $1\leq i,j\leq40$, for (a) initial state, (b) final state of model I, and (c) final state of model II.