Bistable gradient networks. I. Attractors and pattern retrieval at low loading in the thermodynamic limit

Patrick N. McGraw and Michael Menzinger

Department of Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 3H6
(Received 26 March 2002; published 30 January 2003)

We examine the large-network, low-loading behavior of an attractor neural network, the so-called bistable gradient network (BGN), and compare it with that of the Hopfield network (HN). We use analytical and numerical methods to characterize the attractor states of the network and their basins of attraction. The energy landscape of BGN is more complex than that of the HN and depends on the strength of the coupling among units. At weak coupling, the BGN acts as a highly selective associative memory; the input must be close to the one of the stored patterns in order to be recognized. A category of spurious attractors occurs which is not present in the HN. Stronger coupling results in a transition to a more Hopfield-like regime with large basins of attraction. The basins of attraction for spurious attractors are noticeably suppressed compared to the Hopfield case, even though the Hebbian synaptic structure is the same and there is no stochastic noise.

PACS number(s): 84.35.+i, 07.05.Mh, 87.18.Sn

I. INTRODUCTION

Many neural network models [1,2], in addition to their potential applications to computation, robotics and artificial intelligence, constitute intriguing dynamical systems in their own right, showing unusual manifestations of the statistical mechanics phenomena of order, disorder and frustration. The connection between neural networks and statistical mechanics became especially clear with the introduction of the Hopfield [3,4] model, which furnishes a model of associative memory, or the recall of a memorized pattern from an incomplete stimulus. This model has a well-defined energy function and is closely related to the Sherrington-Kirkpatrick spin glass model [5,6].

In this paper we consider a Hopfield-like network of N bistable elements, the bistable gradient network or BGN, previously introduced in Ref. [7]. A closely related model was also discussed in [8] and suggested as a model for the so-called "bistability of perception" in the interpretation of ambiguous visual stimuli [9]. The network's dynamics consists of a continuous gradient descent described by the coupled differential equations

$$\frac{dx_i}{dt} = -\frac{\partial H}{\partial x_i},\qquad(1)$$

where x_i ($1 \leqslant i \leqslant N$) are continuous-valued real state variables associated with the N nodes of the network and the Hamiltonian or energy function is given by

$$H = H_0 + H_{int} = \sum_{i=1}^{N}\left(\frac{x_i^4}{4} - \frac{x_i^2}{2}\right) - \frac{1}{2}\gamma\sum_{i,j=1}^{N} w_{ij}x_i x_j.\quad(2)$$

The quantities w_{ij} are a symmetric matrix of coupling strengths. γ is a control parameter determining the strength of the internode couplings relative to the local terms. For the purposes of this paper, we do not include any external inputs or linear "bias" terms. The variables x_i can be viewed as components of an N-dimensional state vector $\mathbf{x}$. We define a normalized inner product between two state vectors $\mathbf{x}^1$ and

$\mathbf{x}^2$ by $\mathbf{x}^1 \cdot \mathbf{x}^2 \equiv (1/N)\Sigma_{i=1}^{N} x_i^1 x_i^2$. The first term in the Hamiltonian represents a local double-well potential for each node, making each node individually bistable. This local potential constitutes the main difference between the BGN and the Hopfield model. The classical Hopfield network (HN), which we consider by way of comparison, is described by the Hamiltonian

$$H_{HN} = -\frac{1}{2}\sum_{i,j=1}^{N} w_{ij}x_i x_j,\qquad(3)$$

where the x_i are now *discrete* state variables restricted to the values ± 1. Although continuous versions of the HN have also been studied, these generally lack the bistability property, and their behavior is essentially similar to that of the discrete version [10,11].

The variables x_i can be thought of as the outputs of processing units or neurons. Their dynamical equations can be written as

$$\frac{dx_i}{dt} = x_i - x_i^3 + h_i,\qquad(4)$$

where $h_i \equiv \gamma\Sigma_{j=1}^{N} w_{ij}x_j$ is the input to the neuron from its connections with other nodes. By analogy with Ising spin systems, we also refer to h_i as a "magnetic field." The steady-state output for a given input is a solution of the fixed-point equation

$$x_i - x_i^3 + h_i = 0.\qquad(5)$$

When $h = 0$, there are stable fixed points at $x = \pm 1$ and an unstable fixed point at $x = 0$. An applied field shifts the positions of the fixed points. A saddle-node bifurcation occurs when $|h| = h_c = 2\sqrt{3}/9 \approx 0.385$ so that for larger values of the field there is only one equilibrium, aligned with the field (x and h have the same sign). x is, in principle, unbounded; the output does not truly saturate when the input is large. The double valuedness and the lack of saturation are the principal

differences between the input-output relation for the BGN and that of the Hopfield model, including its continuous versions.

Numerous experimental studies have been made on intriguing chemical analogs of the BGN (see, e.g., Refs. [12,13]). These studies involved networks of bistable chemical reactors coupled either electrically or through mass exchange. Previous computational work on small BGN's [7] suggested that under some conditions the network might permit the storage of a larger number of patterns than in a HN of the same size, without any modification of the basic Hebb learning rule. It was noted, however, that the stability of a particular attractor can depend on the control parameter γ. Some dependence of pattern stability on the coupling strength had also been noted in the experiments on the mass-coupled chemical networks [13].

In this paper we focus on the behavior of the network in the case where the number of nodes is large and the number of memorized patterns is small. Using both analytical techniques and numerical simulations,[1] we examine the retrieval of stored patterns and classify the attractors that occur. We find that there are three types of attractors. In addition to *memory* or *retrieval states*, there are spurious attractors in which no pattern is fully recognized. These include the *mixture* or *spin glass states* familiar from HN studies, as well as an additional category specific to the BGN which we refer to as *uncondensed states*. We examine how the attractors and their basins of attraction change as the control parameter γ is changed. Throughout the paper, we compare our model with the zero-temperature or deterministic discrete Hopfield model. It is hoped that these results can shed more light on the behavior of the BGN and clarify its relation to the HN. The behavior of the BGN under higher memory loading and the question of its maximum storage capacity will be addressed in a companion paper [16].

II. STORAGE AND RETRIEVAL OF BINARY PATTERNS

As in previous work on Hopfield networks [3,14,2], we define the task of associative memory as follows. We are given a set of p distinct N-dimensional vectors or *memory patterns* $\boldsymbol{\xi}^\mu$ ($\mu \in \{1, \ldots, p\}$), which are to be recognized by the network. The patterns should correspond to attractors of the network's dynamics. We will refer to these attractors as *retrieval states*. Input is given by imposing a particular initial condition on the network. If that initial condition is sufficiently close to one of the memorized patterns, then the network's state should converge to the correct nearby attractor,

---

[1]Simulations were conducted by integrating the differential equations numerically using a fourth-order Runge-Kutta algorithm with adaptive step size. The system was judged to have converged to a fixed point if the magnitudes of all derivatives $|\partial H/\partial x_i|$ fell below a convergence criterion that for most examples was taken as 0.001. Our HN simulations, which were used for comparisons, were performed using asynchronous updating in random order. Unless otherwise indicated, all simulations were done on an $N=1000$ network with $p=5$ stored memory patterns.

and we say that the network has recognized or retrieved the pattern. In this paper we follow the HN literature in considering the case where the patterns are random and uncorrelated strings of $+1$'s and -1's. We read the output of the network according to the *signs* of the x_i. Thus we say that the network has recalled pattern 1, for example, if $\text{sgn}(x_i) = \xi_i^1$ for all i. Although variations in the magnitude of x_i can be important to the dynamics, we will for the moment ignore them for the purpose of reading the output. As we will see below, the retrieval states in general do not have $|x_i|=1$ even though the patterns have $|\xi_i|=1$. We focus here on the limiting case $N \to \infty$, $p/N \ll 1$, or large networks with low memory loading. (Strictly speaking, we take N to infinity while p remains finite.) In this case the inner product of a pair of patterns $\boldsymbol{\xi}^\mu \cdot \boldsymbol{\xi}^\nu = (1/N) \Sigma_i \xi_i^\mu \xi_i^\nu$ behaves as a Gaussian random variable with zero mean and variance $1/\sqrt{N}$, so that in the $N \to \infty$ limit the pattern vectors are nearly orthogonal to each other and form a basis for a p-dimensional subspace of the N-dimensional configuration space.

As in the HN, we construct the coupling matrix from the stored patterns according to the Hebb [15] learning rule,

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu - \frac{p}{N} \delta_{ij}. \tag{6}$$

The term $-(p/N)\delta_{ij}$ is included to make all diagonal elements of the coupling matrix zero. Nonzero diagonal entries would have the effect of adding an additional quadratic self-interaction term. Following the usual practice we omit them here so that the quadratic term is contained only in the local potential. For the case $p/N \to 0$, however, the effect of the diagonal elements is negligible and we can substitute the simpler learning rule

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu. \tag{7}$$

A useful set of order parameters are the overlaps m_μ, which are inner products of the network's state with each of the stored patterns,

$$m_\mu \equiv \boldsymbol{\xi}^\mu \cdot \mathbf{x} = \frac{1}{N} \sum_{i=0}^{N} \xi_i^\mu x_i. \tag{8}$$

For the discrete HN, these variables take values $-1 \leq m_\mu \leq 1$, while for the BGN any real values are possible. It will be useful to define another set of variables, which we will call "bit overlaps," by

$$b_\mu \equiv \frac{1}{N} \sum_{i=0}^{N} \xi_i^\mu \, \text{sgn}(x_i). \tag{9}$$

The bit overlap is simply related to the Hamming distance by $b_\mu = (1/2N)[N - 2d(\mathbf{x}, \boldsymbol{\xi}^\mu)]$, where the Hamming distance $d(\mathbf{x}, \mathbf{y})$ between two vectors is defined as the number of elements for which their signs differ, or the number of positions i such that $x_i y_i < 0$. Unlike m_μ, the bit overlaps always

obey $-1 \leqslant b_\mu \leqslant 1$. They encode information about sign agreements but not about magnitudes of the outputs x_i.

The definitions of the overlap variables allow us to rewrite the Hamiltonian and the dynamical equations in useful forms [2]. In particular, if the synaptic matrix **w** is given by the simplified Hebb rule (7), then the interaction term of the Hamiltonian can be rewritten in terms of m_μ as follows:

$$H_{int} = -\frac{\gamma}{2} \sum_{i,j=1}^{N} w_{ij} x_i x_j = -\frac{\gamma}{2N} \sum_{i,j=1}^{N} \sum_{\mu=1}^{p} x_i \xi_i^\mu \xi_j^\mu x_j$$

$$= -N\frac{\gamma}{2} \sum_{\mu=1}^{p} (m_\mu)^2, \tag{10}$$

and the net input to a given node from the other nodes is given by

$$h_i \equiv -\frac{\partial H_{int}}{\partial x_i} = \gamma \sum_{j=1}^{N} w_{ij} x_j = \frac{\gamma}{N} \sum_{j=1}^{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu x_j$$

$$= \gamma \sum_{\mu=1}^{p} \xi_i^\mu m_\mu. \tag{11}$$

A. Retrieval states at low memory loading

To show that the network functions properly as an associative memory, we exhibit the attractor states corresponding to the stored patterns, demonstrate their stability, and show that a pattern can be retrieved from an initial condition which lies close to the pattern but differs from it by one or more incorrect signs.

Consider the state $\mathbf{x} = M\boldsymbol{\xi}^\nu$, where M is a scalar and $\boldsymbol{\xi}^\nu$ is a particular one of the stored patterns. We will show that for a suitable value of M this state represents a stable fixed point of the dynamics and is therefore the retrieval state we seek. In this state, $m_\nu = M$, $b_\nu = 1$, and all other overlaps are small. The field acting on the ith node can be written as follows:

$$h_i = \gamma \xi_i^\nu M + \gamma \sum_{\mu \neq \nu} \xi_i^\mu m^\mu. \tag{12}$$

The sum over patterns $\mu \neq \nu$ is called the crosstalk term. For the overlaps with these other patterns we have

$$m^\mu = \frac{1}{N} \sum_i x_i \xi_i^\mu = M \frac{1}{N} \sum_i \xi_i^\nu \xi_i^\mu \quad (\mu \neq \nu). \tag{13}$$

Since the patterns are random, each of these overlaps is of order $M/\sqrt{N}$. The number of patterns remains finite as $N \rightarrow \infty$, so the sum in the crosstalk term vanishes in this limit and $h_i \approx \gamma \xi_i^1 M$. A stationary state must satisfy the fixed-point condition (5) for each node, which leads to a self-consistency condition on M,

$$0 = x_i^3 - x_i - h_i = (M\xi_i^\nu)^3 - (1+\gamma)(M\xi_i^\nu),$$

$$0 = M^3 - (1+\gamma)M. \tag{14}$$

In the last step, we have used the fact that $\xi_i^3 = \xi_i$ when $\xi_i = \pm 1$ and then divided out the common factor of ξ_i. Solutions to this condition are an unstable equilibrium $M = 0$ and two stable equilibria

$$M = m_\nu = \pm \sqrt{1+\gamma}. \tag{15}$$

The two stable solutions represent perfect retrieval of pattern ν and its mirror state, respectively. The doubled state is a consequence of the Z_2 symmetry in the Hamiltonian. Since the overlap m_ν is equal to M and all other overlaps vanish in the thermodynamic limit, the energy of this retrieval state is easily calculated using expression (10) for the energy in terms of the overlap variables, giving

$$E = \sum_i \left(\frac{x_i^4}{4} - \frac{x_i^2}{2} \right) - \frac{\gamma}{2} \sum_\mu m_\mu^2 = N\left(\frac{M^4}{4} - \frac{M^2}{2} \right) - \frac{\gamma}{2} M^2$$

$$= N\left(-\frac{1}{4} - \frac{\gamma}{4} - \frac{\gamma^2}{4} \right). \tag{16}$$

Note that this energy expression is extensive (proportional to N) and a monotonically decreasing function of γ.

Having identified the state

$$\mathbf{x}^\nu = \sqrt{1+\gamma}\boldsymbol{\xi}^\nu \tag{17}$$

as an equilibrium state, we now demonstrate its stability using a linear stability analysis of the dynamical equations

$$\frac{dx_i}{dt} = y_i \equiv x_i - x_i^3 + \gamma \sum_j w_{ij} x_j. \tag{18}$$

Evaluating the Jacobian $\partial y_i / \partial x_j$ at the fixed point (17) we obtain

$$\frac{\partial y_i}{\partial x_j} = (1 - 3x_i^2)\delta_{ij} + \gamma w_{ij} = [1 - 3(\gamma+1)]\delta_{ij} + \gamma w_{ij}$$

$$= (-2 - 3\gamma)\delta_{ij} + \gamma w_{ij}, \tag{19}$$

where δ_{ij} is the Kronecker delta. The fixed point is linearly stable if and only if the Jacobian has no eigenvalues with positive real part. This depends in turn on the eigenvalues of the synaptic connection matrix **w**. But in the limit where all of the stored patterns $\boldsymbol{\xi}^\mu$ are mutually orthogonal, the stored patterns are themselves eigenvectors spanning a degenerate subspace with eigenvalue 1, while the complement of this subspace has eigenvalue 0. [The Hebb rule (7) itself gives a spectral decomposition of **w**.] Since the maximum eigenvalue of **w** is 1, we see that the Jacobian at the retrieval fixed point has no positive eigenvalues and so the retrieval state is linearly stable for any value of γ. In fact, all eigenvalues become more negative as γ increases. We reiterate that this result is valid in the ideal limit of large N and low loading where the stored patterns are orthogonal. For finite-sized networks with finite overlaps among the patterns, it is possible for the memory states to be destabilized by the crosstalk terms. This issue will be examined in the companion paper [16].

Numerical results for a network with $N = 1000$ nodes and $p = 5$ random patterns agree excellently with the above description. To study a retrieval state numerically, we initialized the network to the state $\mathbf{x} = \boldsymbol{\xi}^1$ (arbitrarily choosing the first pattern). Starting at $\gamma = 0$, we increased γ by small steps to $\gamma = 6$. At each step, we integrated the dynamical equations until they converged. This procedure allows us to examine the evolution of a state under quasistatic changes in the control parameter γ. We verified that b_1 remained equal to 1 over the whole range $0 \le \gamma \le 6$, indicating that the retrieval state is stable. The measured values of m_1 and E were within 1% of the theoretical expressions (15) and (16), respectively.

B. Error correction and basins of attraction

Linear stability analysis has shown that the retrieval states are stable against *infinitesimal* perturbations for any value of γ, but this does not guarantee their stability against the flipping of signs of one or more nodes. In order to function as an associative memory, a network must be capable of dynamically correcting sign errors. When presented with an input at a small, nonzero Hamming distance from one of the stored patterns (i.e., differing from it by a few reversed signs) it must be able to flip the reversed signs and restore the correct pattern. We will now show that there is a critical value $\gamma_c = \frac{1}{3}$ above which the correction of sign-flip errors can occur. For smaller values of γ the BGN does not correct sign errors and thus does not truly function as an associative memory, but as γ increases above $\frac{1}{3}$ the retrieval states develop increasingly large basins of attraction.

Consider a state of the network which is a slightly corrupted retrieval state: all node variables have the values $x_i = M \xi_i^\nu = \sqrt{1 + \gamma} \xi_i^\nu$ with the exception of one or possibly some number $\ll N$ of nodes which may be misaligned. In such a state the few misaligned nodes make only a small contribution to the overlap sums, so we have $m_\nu \approx M$ and $m_\mu \approx 0 (\mu \ne \nu)$. The field acting on each node is therefore $h_i \approx \gamma M \xi_i^\nu$. The misaligned bits experience a field opposite to their signs. If the field becomes larger than the critical value $2\sqrt{3}/9$, then there is only one stable equilibrium for each node, and the misaligned nodes will flip to conform with the stored pattern. Error correction therefore occurs if

$$|h_i| \approx \gamma M \approx \gamma \sqrt{1 + \gamma} > \frac{2\sqrt{3}}{9} \approx 0.385. \qquad (20)$$

The critical value occurs when the equality $\gamma \sqrt{1 + \gamma} = 2\sqrt{3}/9$ holds, or at $\gamma_c = \frac{1}{3}$.

If the pattern is more strongly corrupted (a significant number of bits misaligned) then the situation is more complicated, because the presence of a larger number of misaligned bits may reduce the value of m_ν and thus the magnitude of the field. The misaligned bits have a significant back reaction on the ones with the correct sign. The correction of larger numbers of sign errors requires higher values of γ. We will return to this point later; the basic result is that when γ is only slightly above the threshold of $\frac{1}{3}$, the memory states have rather small basins of attraction, but these basins grow as γ increases.

III. SPURIOUS ATTRACTORS: SPIN GLASS STATES

In the case of the HN, the Hebb learning rule results in a large number of "spurious" attractors in addition to the retrieval states. The energy function defines a rugged landscape, and a trajectory that does not start sufficiently close to one of the stored patterns may become trapped in one of the spurious local minima instead of one corresponding to a recalled pattern. It is possible to suppress the spurious minima by introducing thermal noise that allows trajectories to jump out of the shallower basins of attraction into deeper ones.

At low levels of loading, the HN possesses spurious attractors that are nonlinear combinations of the stored patterns. There is a hierarchy of symmetric *mixture states* of the form [14,2]

$$x_i = \text{sgn}(\xi_i^{\mu_1} \pm \xi_i^{\mu_2} \pm \cdots \pm \xi_i^{\mu_n}). \qquad (21)$$

These states overlap equally with n different patterns: $m_{\mu_1}^2 = m_{\mu_2}^2 = \ldots m_{\mu_n}^2 < 1$. For the HN, only the mixtures with odd n are stable. The $n = 3$ mixtures have the lowest energy in this category, and the energies increase with n, asymptotically approaching $-1/\pi$. As the number of stored patterns p increases, these spurious states proliferate exponentially; their number is of order 3^p. There are also nonsymmetric mixtures. The proliferation of spurious states is associated with spin glass type behavior in the HN. Accordingly, we also refer to these mixed states somewhat loosely as *spin glass states*.[2]

We will show here that the BGN possesses mixture states analogous to those of the HN, but their structure is slightly more complex. Let us focus on the $n = 3$ symmetric mixture state with positive signs. For the HN, this state is given by

$$x_i = \xi_i^S \equiv \text{sgn}(\xi_i^{\mu_1} + \xi_i^{\mu_2} + \xi_i^{\mu_3}). \qquad (22)$$

This state is stable against individual sign flips because each node is subject to a nonzero magnetic field that maintains its alignment. To see this, note that there are two possibilities for each bit. Either all three patterns agree at that particular site ($\xi_i^{\mu_1} = \xi_i^{\mu_2} = \xi_i^{\mu_3}$) or one of the patterns has the opposite sign from the other two, for example, $\xi_i^{\mu_1} = \xi_i^{\mu_2} = -\xi_i^{\mu_3}$. When all three agree, we say that the ith bit is a "unanimous" bit. If the patterns are random, then each ξ_i^μ is ± 1 with equal probability, giving a probability of $\frac{1}{4}$ that a given bit is unanimous. The mixture state has equal overlaps with all three of the patterns. Since for a given i there is a prob-

---

[2]More precisely, the mixture states "melt" into spin glass states when the number of memorized patterns becomes higher. With higher numbers of patterns, there is a distinction between mixture states that overlap with a small set of the memory patterns and spin glass states that overlap with nearly all, but that distinction is not salient for the cases we consider with very few patterns (see Ref. [14].)

ability of $\frac{3}{4}$ that $\xi_i^S = \xi_i^{\mu_1}$, we have $m_{\mu_1} = (1/N)\Sigma_{i=1}^N \xi_i^S \xi_i^{\mu_1}$ $= \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$, and likewise $m_{\mu_2} = m_{\mu_3} = \frac{1}{2}$. The field acting at the ith site is

$$h_i = \sum_\mu m_\mu \xi_i^\mu = \frac{1}{2}(\xi_i^{\mu_1} + \xi_i^{\mu_2} + \xi_i^{\mu_3}). \tag{23}$$

(Consistent with the low-loading, large-N limit, we ignore all other overlaps that are of order $1/\sqrt{N}$.) This gives $h_i = \frac{3}{2}\xi_i^S$ for unanimous bits and $h_i = \frac{1}{2}\xi_i^S$ for the others. In any case, each node of the HN experiences a field that stabilizes its alignment.

We will now show that the BGN, like the HN, possesses a mixture state in which the sign of x_i is given by the majority vote of three of the stored patterns,

$$\text{sgn}\, x_i^S = \text{sgn}(\xi_i^{\mu_1} + \xi_i^{\mu_2} + \xi_i^{\mu_3}). \tag{24}$$

This state has a more complicated structure, however, because the magnitude of x_i at a given node depends significantly on the local field at that node. Since the field at a unanimous bit is stronger than the field at a nonunanimous bit, we expect the magnitude of x_i to be larger for a unanimous bit. Therefore, we make the *ansatz*

$$x_i^S = \begin{cases} A\,\text{sgn}(\xi_i^{\mu_1} + \xi_i^{\mu_2} + \xi_i^{\mu_3}) & \text{if } \xi_i^{\mu_1} = \xi_i^{\mu_2} = \xi_i^{\mu_3} \\ D\,\text{sgn}(\xi_i^{\mu_1} + \xi_i^{\mu_2} + \xi_i^{\mu_3}) & \text{otherwise}, \end{cases} \tag{25}$$

where A and D are real numbers. The dynamical equations for the network give a pair of self-consistency equations that can be solved numerically for A and D. First, we need an expression for the overlap of the mixture state with one of the three patterns, say, pattern μ_1. If the ith bit is a unanimous bit, then x_i^S has magnitude A and agrees in sign with $\xi_i^{\mu_1}$. On the other hand, if it is a nonunanimous bit, then x_i^S has magnitude D and has a 2/3 probablity of agreeing in sign with $\xi_i^{\mu_1}$. The result is that

$$m_{\mu_1} = \frac{1}{N}\sum_i x_i \xi_i^{\mu_1} = \frac{1}{4}A + \frac{1}{2}D + \frac{1}{4}(-D) = \frac{1}{4}(A+D). \tag{26}$$

Note that in the special case $A = D = 1$, the above expression reduces to the Hopfield value $1/2$, as it should. Again, all three overlaps have the same size: $m_{\mu_1} = m_{\mu_2} = m_{\mu_3}$. The total energy of the network in this state is given by

$$E(A,D) = \sum_i \left(\frac{x_i^4}{4} - \frac{x_i^2}{2}\right) - \frac{N}{2}\gamma \sum_\mu m_\mu^2$$

$$= \frac{N}{4}\left(\frac{A^4}{4} - \frac{A^2}{2}\right) + \frac{3N}{4}\left(\frac{D^4}{4} - \frac{D^2}{2}\right)$$

$$- \frac{3N}{2}\gamma\left(\frac{1}{4}(A+D)\right)^2$$

$$= N\left[\frac{A^4}{16} + \frac{3D^4}{16} - \frac{A^2}{8} - \frac{3D^2}{8}\right.$$

$$\left. - \gamma\left(\frac{3A^2}{32} + \frac{3D^2}{32} + \frac{3AD}{16}\right)\right]. \tag{27}$$

We can think of this as an an energy function on the restricted family of states parametrized by Eq. (25).

A necessary condition for the mixture state (25) to be a fixed point is that $\partial E/\partial A = \partial E/\partial D = 0$. This gives us two self-consistency conditions for the parameters A and D,

$$\frac{A^3}{4} - \frac{A}{4} - \frac{3\gamma}{16}(A+D) = 0,$$

$$\frac{3D^3}{4} - \frac{3D}{4} - \frac{3\gamma}{16}(A+D) = 0. \tag{28}$$

These are the equations of the nullclines of the energy function (27). Alternatively, the above equations could be derived directly from the dynamical equations for each node and the expressions for h_i instead of using the energy function. For a *stable* fixed point, (A,D) must be a local minimum of the energy function (27). The graphs of the two cubic equations 28 are plotted for $\gamma = 3$ in Fig. 1(a). Solutions are points where the two nullclines intersect. Note that the slope, dD/dA, of the first curve at the origin is $-(1+4/3\gamma)$ < -1, and for the second curve $dD/dA = -1/(1+4/\gamma$ $> -1)$. These two inequalities satisfied by the slopes ensure that the curves always intersect in at least five points. The five solutions can be classified by looking at the energy function and its gradient. We see that the solution $A = D = 0$ is an unstable node (maximum of the energy), the two solutions in the second and fourth quadrants are saddle points, and the two solutions in the first and third quadrants are the stable solutions we seek. (There are two because of the Z_2 sign reversal symmetry—one is a mirror state of the other.)

The self-consistency equations for A and D were solved numerically for a range of γ values using a gradient descent algorithm. These values were, in turn, substituted into Eq. (27) to find the energy as a function of γ. The results are plotted in Fig. 2 where they are also compared with numerical results from dynamical simulations of a BGN with $N = 1000$. We studied the mixture state numerically by initializing a network to the state $x_i = \text{sgn}(\xi_i^1 + \xi_i^2 + \xi_i^3)$ and incrementing γ beginning at 0, much as was done for the retrieval states. Figure 2(a) shows the magnitudes of unanimous and nonunanimous bits in the mixture state as functions of γ. The solid lines show the solutions of the self-consistency equations for A and D. The symbols show the observed mag-
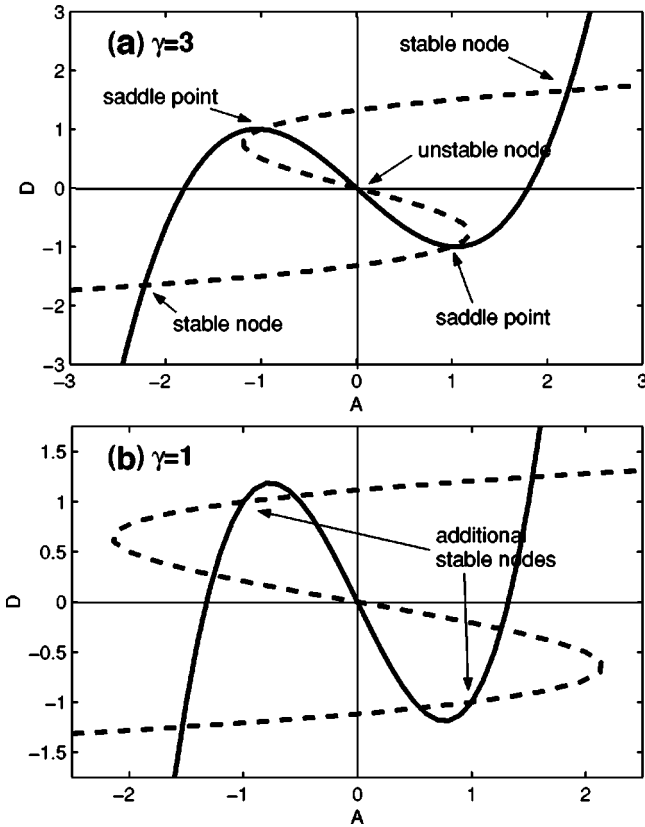
FIG. 1. Self-consistency equations for the $n=3$ mixture state. (a) Stable solutions occur in the first and third quadrants. (b) When $\gamma < 2$ two additional stable solutions appear in the second and fourth quadrants.

nitudes $|x_i|$ for $1 \leq i \leq 6$ in the simulated network state. Two of these first six bits are unanimous while the other four are not. There is good agreement between the observed values of $|x_i|$ and the values obtained from the self-consistency equations. For comparison, $\sqrt{1+\gamma}$ is plotted as a dotted line on the same axes. (Recall that this is the value of all $|x_i|$ in a pure retrieval state.) Figure 2(b) shows a corresponding comparison of the observed and theoretical energies. Finally, Fig. 2(c) compares the $n=3$ mixture state with the retrieval state by plotting the ratios $A/\sqrt{1+\gamma}$ and $D/\sqrt{1+\gamma}$ as well as the ratio of the mixture state energy E_{mix} to that of a retrieval state E_{ret}. All three of these ratios appear to approach constant asymptotic values as γ increases. Asymptotically, $E_{mix}/E_{ret} \approx 0.7$, while for the HN the corresponding ratio is 0.75. The strength of the field acting on each unanimous bit, h_A, and that acting on the nonunanimous bits, h_D, both increase as γ increases. The mixture state is stable against single sign flips of the unanimous bits when $h_A > 2\sqrt{3}/9$ and stable against any single sign flip when $h_D > 2\sqrt{3}/9$. Thus as γ increases, the mixture state begins to develop a nontrivial basin of attraction of its own.

It is interesting to note that at $\gamma = 2$ a saddle-node bifurcation occurs and for $\gamma < 2$ two additional stable solutions to the self-consistency equations (28) appear in the second and fourth quadrants, at $(A,D) = (1,-1)$ and $(-1,1)$ [see Fig. 1(b)]. These are states with $E = -\frac{1}{4}N$ in which $m_{\mu_1} = m_{\mu_2}$

$= m_{\mu_3} = 0$ and the net field acting on each node exactly cancels. They are not stable against sign flips, and become completely destabilized when $\gamma > 2$. As we will see in the following section, they do not properly belong to the category of mixture or spin glass states, but rather to another class of spurious attractors present only in the BGN at low values of γ.

Here we examined only the $n=3$ mixture state, but similar methods may be used to characterize higher-order mixtures. In general, they are more complex as there are more possibilities for the size of the majority by which the sign is determined. The magnitudes of x_i then take a greater number of distinct values.

IV. UNCONDENSED STATES AND THEIR COLLAPSE TO THE PATTERN SUBSPACE

In Sec. II B, we noted that for values of γ not far above $\frac{1}{3}$, it is possible that a state may have a significant overlap with one of the stored patterns but that the field acting on the nodes may, nonetheless, not be strong enough to overcome the potential barrier and correct the sign errors. Indeed if γ is below $\frac{1}{3}$, then even a single sign error may go uncorrected. This consequence of the bistability of the BGN units contrasts with the behavior of the HN.

Consider first the case of the HN. A typical random initial state has small but nonzero overlaps with the memorized patterns, $m_\mu \sim O(1/\sqrt{N})$, resulting in fields $h_i = \Sigma \xi_i^\mu m_\mu$ that are random with zero mean and variance of order $1/\sqrt{N}$. Typically, for approximately half of the nodes x_i and h_i will initially have opposite signs. Since there is no potential barrier against sign flips, those nodes will change their signs, and the sign flips will continue until the field experienced by every node is aligned with x_i. Every sign flip will increase the magnitude of one or more of the overlap variables. If, for example, one overlap m_ν is larger than all of the others, then most nodes will experience fields that tend to align them with pattern ν. Every sign flip further increases the value of m_ν, and eventually m_ν will be fully retrieved even though the initial overlap may have been quite small. However, if one overlap does not clearly dominate the others, then the trajectory may arrive at a spurious attractor that has roughly equal overlaps with several patterns, instead of at a single one of the memory states. Even a state that is initially orthogonal to all of the memorized patterns can be rendered unstable by changing a single sign: even a single sign flip will create a small but nonzero field affecting the other nodes, resulting in further sign flips, and so on. In summary, for the HN essentially any initial condition converges under the dynamics to an attractor lying in or close to the p-dimensional subspace of the patterns.

For the BGN, on the other hand, the situation is different due to the presence of potential barriers. Just as with the HN, given any state **x** every node i experiences a field h_i, which may be aligned with or opposed to x_i. However, the antiparallel local fields may not be strong enough to flip their nodes into the parallel direction. If most h_i are well below the threshold $2\sqrt{3}/9$ then the flipping of one or a few nodes will
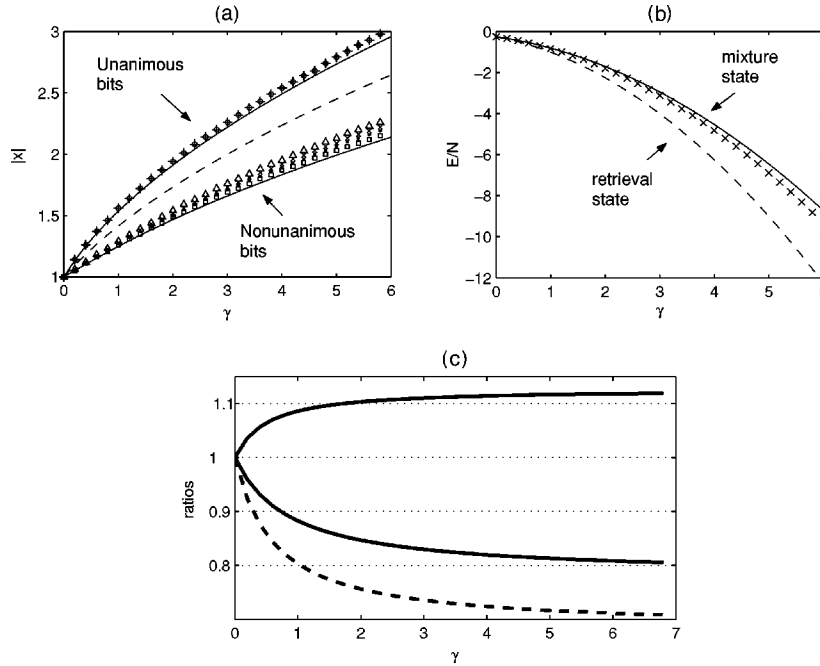
FIG. 2. $n=3$ mixture state compared with retrieval state. (a) The two solid lines represent A and D, the values of $|x|$ for unanimous and nonunanimous bits, respectively, obtained by numerical solution of the self-consistency equations. The symbols show the *observed* values $|x_i|$ ($1 \leqslant i \leqslant 6$) for a mixture state of the dynamically simulated network. The dotted line is $\sqrt{1+\gamma}$, which is the theoretical value of all $|x_i|$ for a retrieval state. (b) Energy E_{mix} for the mixture state. Solid line: solution of self-consistency equations. Symbols: observed energy of simulated mixture state. Dotted line: energy E_{ret} of retrieval state from Eq. (16). (c) The ratios $A/\sqrt{1+\gamma}$ and $D/\sqrt{1+\gamma}$ (upper and lower solid lines) and E_{mix}/E_{ret} (dotted line) approach asymptotically constant values.

not change the field enough to cause any further flips. Thus there might be a large number of initial conditions that remain stuck with low overlaps, far away from any of the patterns. We refer to such states with subthreshold fields as "uncondensed" states, because in those states none of the order parameters m_μ are condensed. However, we will show below that states with low overlaps cannot remain stable for $\gamma > 2$, and thus for higher values of γ the behavior is Hopfield-like, with all trajectories collapsing toward the pattern subspace.

Consider a hypothetical state that is strictly orthogonal to all memory patterns, so that $m_\mu = 0$ for all μ. (The extra solutions appearing in the self-consistency equations for the mixture state when $\gamma < 2$ are examples of such states.) If $m_\mu = 0$ for all μ, then $h_i = 0$ for all i. In this case, the steady state of each node is $x_i = \pm 1$. Proceeding with linear stability analysis as above, we find the relevant Jacobian

$$\frac{\partial y_i}{\partial x_j} = (1 - 3x_i^2)\delta_{ij} + \gamma w_{ij} = -2\delta_{ij} + \gamma w_{ij}. \qquad (29)$$

The equilibrium is unstable if $\gamma w_{max} > 2$, where w_{max} is the largest eigenvalue of the coupling matrix. w_{max} is at least 1. Therefore, if $\gamma > 2$, states orthogonal to the stored patterns are all unstable. If all of the p stored patterns are mutually orthogonal (which is approximately true in the limit $N \rightarrow \infty$) then for $\gamma > 2$ there are p unstable eigenvalues. Each is associated with one of the stored patterns, to which the uncondensed state can evolve if it is perturbed.

Thus, we see that there is an absolute upper limit for the existence of stable uncondensed states. In fact, $\gamma = 2$ turns out to be a high upper bound. The example of a state with all m_μ equal to zero is a sort of "worst-case scenario." For a finite-sized network the typical random initial condition has small but nonzero overlaps. In addition, if the patterns are truly random then they will not be exactly orthogonal but
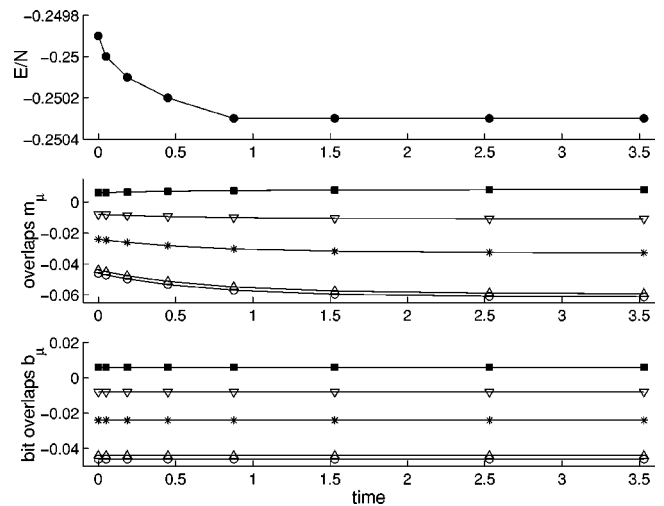

FIG. 3. Trajectory of a random initial condition for $\gamma = 0.5$. Note: in the lower two plots, five curves marked with different symbols show overlaps with each of the five stored patterns. The unequal time steps result from the adaptive step size control in our integration algorithm.
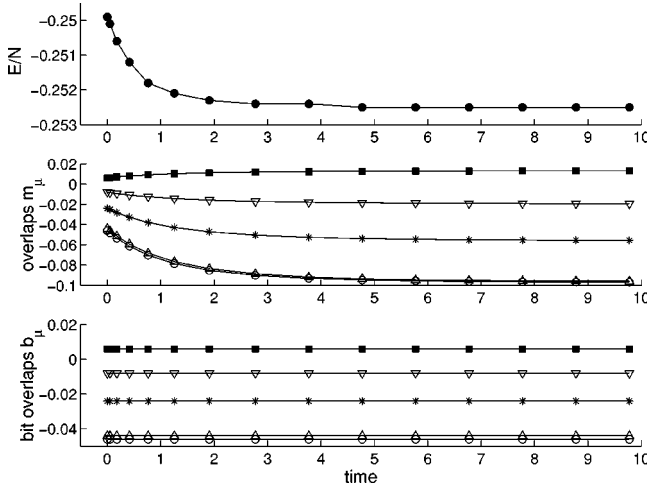
FIG. 4. Trajectory of the same initial condition, $\gamma = 1.0$.

FIG. 5. Same initial condition, $\gamma = 1.2$.

have small overlaps and so the largest eigenvalue of the synaptic matrix will be slightly larger than unity. Because of these factors the typical uncondensed state becomes unstable at values of γ lower than 2. In numerical simulations we found that for the case $N = 1000$, $p = 5$ most uncondensed states become unstable between $\gamma = 1$ and 1.5.

Figures 3–6 show numerical results for the fate of a typical random initial condition of the BGN with $N = 1000$, $p = 5$, demonstrating the trapping of a typical trajectory in an uncondensed state and the destabilization of that state at higher values of γ. Figures 3–5 show the dynamical evolution of the same initial condition, at different values of γ. The initial condition was a random string of ± 1 values. We plot the energy per node, all five overlap variables m_μ, and all five bit overlap variables b_μ as functions of time. Recall that b_μ contain information about sign agreements only. For $\gamma = 0.5$ (Fig. 3) the state changes very little before convergence occurs. The energy per node remains very close to -0.25, as expected for a state with no large overlaps. The overlaps m_μ increase slightly in magnitude,[3] but the bit overlaps do not change at all, indicating that no sign flips occur. The evolution is thus trapped in an uncondensed state. For $\gamma = 1$ (Fig. 4) the trajectory is similar, except that the small initial overlaps are amplified to a greater extent (we will explain this effect below.) The bit overlaps still do not change. When $\gamma = 1.2$, however, the trajectory changes qualitatively (Fig. 5) and does not remain in an uncondensed state. The magnitudes of the overlaps m_μ grow slowly until at $t \sim 7$ the resulting field becomes strong enough to begin flipping some signs. At this point the bit overlaps begin to change, the energy drops significantly and the trajectory moves close to the pattern subspace. After some further evolution, the system converges to a mixture state that overlaps with several patterns. A different random initial condition, followed again at $\gamma = 1.2$, leads instead to a memory state (Fig. 6). Here one of the five overlaps becomes dominant and
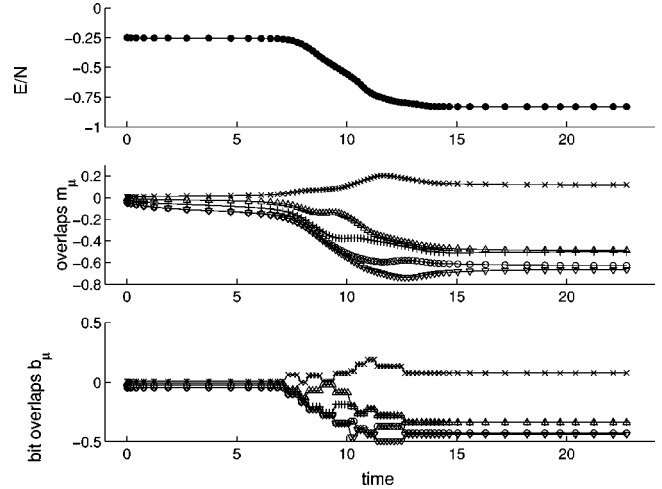
the others shrink away. In this case the mirror state of one of the five patterns is retrieved. These trajectories are typical examples representing descent on a rugged energy landscape. Different initial conditions lead to different attractors, of which some are memory states and some are mixtures. Frequently the trajectory lingers at one or several states before settling at its asymptotic attractor.

Figures 3 and 4 illustrated that for sufficiently small values of γ, the dynamics amplifies small initial overlaps without flipping the sign of any node. For more insight into this phenomenon, consider an initial state in which all x_i are ± 1 but somewhat more nodes are aligned parallel with one particular pattern ξ_i^ν than are antiparallel. In other words, b_ν is nonzero but less than unity. For simplicity let us neglect all other overlaps. Initially, each node experiences a small field given by $h_i = \gamma m_\nu \xi_i^\nu = \gamma b_\nu \xi_i^\nu$. This field will push x_i to larger magnitudes (>1) for those nodes that are aligned with pattern ξ_i^ν, and it will push the others to smaller magnitudes
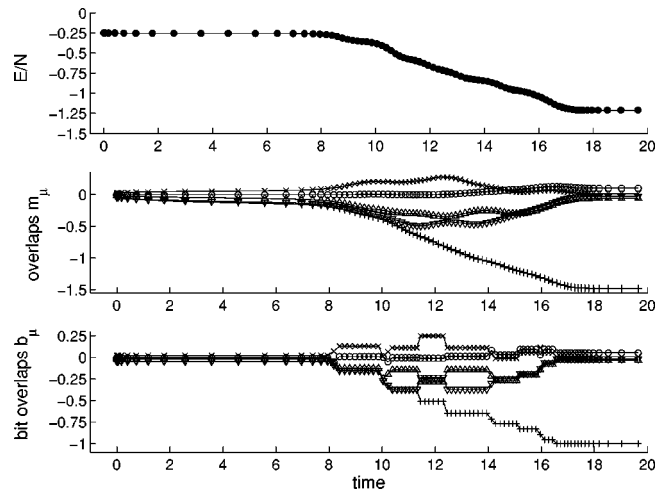

FIG. 6. A $\gamma = 1.2$ trajectory with a different random initial condition. In this case, one of the bit overlaps reaches -1, while the others become small. This indicates that the mirror state of one of the stored patterns has been retrieved.

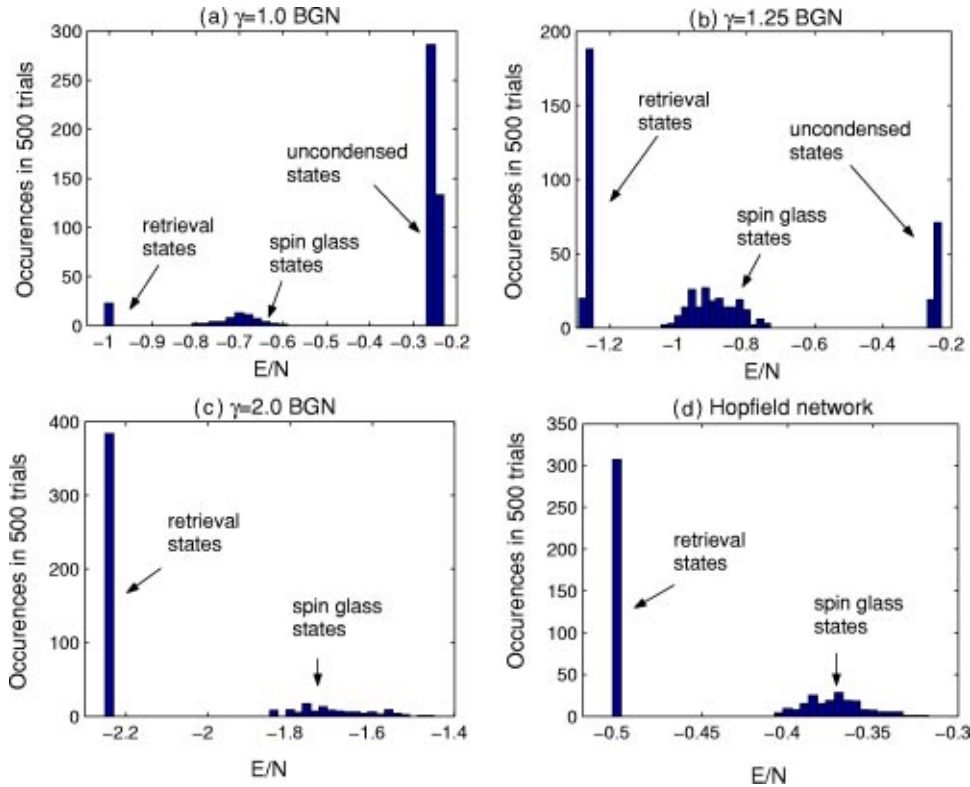[3]Note that negative overlaps with a pattern can be viewed as positive overlaps with the corresponding mirror state.

FIG. 7. Energy per node for attractors reached from random initial conditions, showing a clear separation among different types. (a) For $\gamma=1$, three types of attractors are clearly present. Uncondensed states show up as a peak near $E=-0.25N$. (b) For $\gamma=1.25$, the uncondensed state peak is smaller but occurs at nearly the same energy, whereas the other two peaks are at different energies. (c) For $\gamma=2$, only the retrieval and spin glass states are obtained. (d) HN behavior is qualitatively similar to the BGN with $\gamma=2$.

$x_i<1$. This adjustment in turn increases the value of m_ν, until an equilibrium is reached with $m_\nu>b_\nu$. We might think of this as a kind of "subliminal" recognition of the pattern. The effect becomes stronger as γ increases. Clearly, it has a nonlinear dependence on both γ and b_ν. When the field becomes large enough it will exceed the threshold for sign flips and the state will be attracted toward the pattern retrieval state. The larger γ is, the smaller the initial b_ν that is necessary to fully retrieve the pattern ξ_i^ν. In other words, the basins of attraction of the patterns expand as γ increases.

V. BASINS OF ATTRACTION AND THE ENERGY LANDSCAPE

In this section, we provide numerical support for the three-way classification of attractors into retrieval, spin glass, and uncondensed states and we show how the respective attractor basins change with the control parameter γ. We observe an interpolation between two different regimes. As we showed above, for $\gamma \gtrsim 2$ there are no stable uncondensed states. For lower values of γ, on the other hand, uncondensed states are numerous. Recall that uncondensed states are characterized by local fields too weak to overcome the potential barriers against sign flips, and so their dynamics is dominated by the local potential. In the extreme case $\gamma=0$, there are of course no magnetic fields at all and only the local potential is present.

A. Statistics of attractors reached from random initial conditions

The classification of attractors is very clearly reflected in the energy spectrum. To explore attractors and their basins,

we "seeded" the network with 500 random initial conditions (taken with several different realizations of the five random patterns), integrated the dynamical equations until they converged, and constructed a histogram of the final energies (Fig. 7). For the case $\gamma=1$ [Fig. 7(a)], there are three clearly separated clusters of attractors. Those with the lowest energies are retrieval states, while the states clustered at $E/N=-0.25$ are the uncondensed states, and those in the intermediate range are the glassy states. The picture is qualitatively similar at the slightly larger value $\gamma=1.25$ [Fig. 7(b)], but the peak at $E/N=-0.25$ has shrunk relative to the other two. Note that the energies of the retrieval and spin glass states change with γ, while the uncondensed states remain at nearly the same energy because their dynamics is dominated by the local potential. For $\gamma=2$ [Fig. 7(c)], on the other hand, the cluster at $E/N=-0.25$ is absent as there are no stable uncondensed states. The histogram for a HN [Fig. 7(d)] resembles that for the BGN with $\gamma=2$. One quantitative difference is that the retrieval state peak is slightly higher for the BGN with $\gamma=2$, while the glassy states are comparatively suppressed.

We performed this experiment at a range of values of γ. In all cases the classification of states was clear from the energy spectrum and was verified by examining the final values of b_μ. Figure 8 shows the probabilities of convergence to each of the three types of attractors from a random initial condition as functions of γ. At $\gamma=0.5$ the landscape is dominated by the uncondensed states. Even though $\gamma=0.5$ lies above the threshold of $\frac{1}{3}$ and the retrieval states have nontrivial basins of attraction, these basins still occupy a very small fraction of the total configuration space volume. The patterns can be retrieved only if the initial overlaps are
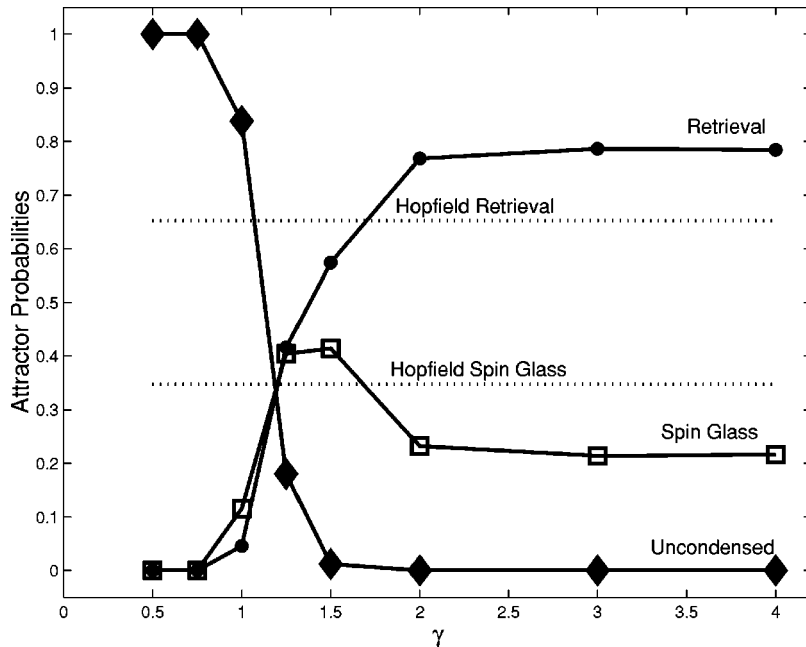
FIG. 8. Probability of convergence of a random initial condition to each of the three types of attractors, plotted as functions of γ. Hopfield values are shown for comparison.

relatively high, and the probability of a *random* initial condition being sufficiently close is very low. The retrieval probability becomes significant only as γ approaches 1. As γ increases from 1 to 1.5, basins for the memory and spin glass states grow at the expense of of the uncondensed states until the latter disappear. The retrieval state basins grow faster than those of the spin glass states. Beyond $\gamma = 1.5$, the probability of retrieving a memory state saturates at approximately a 10% higher value than in the Hopfield case, and the probability of falling into a spin glass state is correspondingly lower.

B. Mapping the boundaries of basins of attraction

In an attempt to map the attractor basins in more detail, we generated configurations at specified initial Hamming distances from particular memory patterns. This was done by starting with a pattern ξ^{ν} and flipping the signs of a specified number of randomly chosen bits. Using an ensemble of such initial conditions, we measured the probability of retrieval of the target pattern ξ^{ν} as a function of the initial distance from it. As a rule, the probability of recognizing the pattern is high if only a few signs are flipped, but drops sharply if a certain threshold Hamming distance is exceeded. We are interested in learning where this threshold lies, and thus answering the question of how close an initial condition must be to a pattern in order to be attracted to it. We are also interested in the fate of states lying just outside the boundaries of a basin of attraction. In other words, does the basin share a boundary with the basins of other patterns, or only with spurious attractors? The results are presented in Fig. 9 for a BGN with $N = 1000$ and $p = 5$, for the three values $\gamma = 0.5, 1.0,$ and 2.0, and also for the HN. In each of these cases, we generated an ensemble of initial conditions at a particular initial value of b_{ν} for some pattern. Each initial condition was allowed to evolve under the dynamics and the resulting attractor was classified as: (a) the target pattern ξ^{ν}, (b) one of the other

patterns ξ^{μ} ($\mu \neq \nu$), (c) a spin glass spurious state, or (d) an uncondensed state. In this case, we classified a state as uncondensed if no sign flips occured during the dynamical evolution. The probabilities of each of these four outcomes were averaged over several realizations of the random patterns and plotted as functions of the initial bit overlap b_{ν}.

Consider first the HN data from Fig. 9(d). A pattern can evidently be retrieved even if the initial overlap is fairly small; the probability is close to unity if $b_{init} \gtrsim 0.1$. If the target pattern is not retrieved, then either a spurious attractor or one of the other patterns may be retrieved. There is a range of b_{init} over which all three probabilities are significant, indicating that the basins for the memory states border on each other as well as those of spurious states. For an $N = 1000$ network, the expected magnitude of the overlap of a random state with any given one of the stored patterns is $1/\sqrt{N} \approx 0.03$, which is not much smaller than the apparent threshold of $b_{init} \approx 0.1$. This is consistent with the view that for the HN, a pattern is likely to be retrieved as long as the initial overlap with that pattern is significantly larger than all of the other overlaps. The $\gamma = 2$ BGN [Fig. 9(c)] shares the qualitative features of the HN. Note, however, that the probability of becoming trapped in a spurious state is smaller for the BGN, consistent with the results in Fig. 8.

A contrasting case is the $\gamma = 0.5$ BGN [Fig. 9(a)]. In this case retrieval of the target pattern requires an initial overlap of more than 0.5. Although this represents a significant basin, it is highly unlikely that a *random* initial condition will have such a large overlap, thus explaining why random initial conditions almost never flow to a memory state. The basins of the memory states are bordered only by spurious states, not by other memory states. Interestingly, the states that lie adjacent to the basin of a memory pattern are not all uncondensed—some are spurious states of the mixture or spin glass type. Examination of the states retrieved near the boundary shows that these are typically *asymmetric* mixture
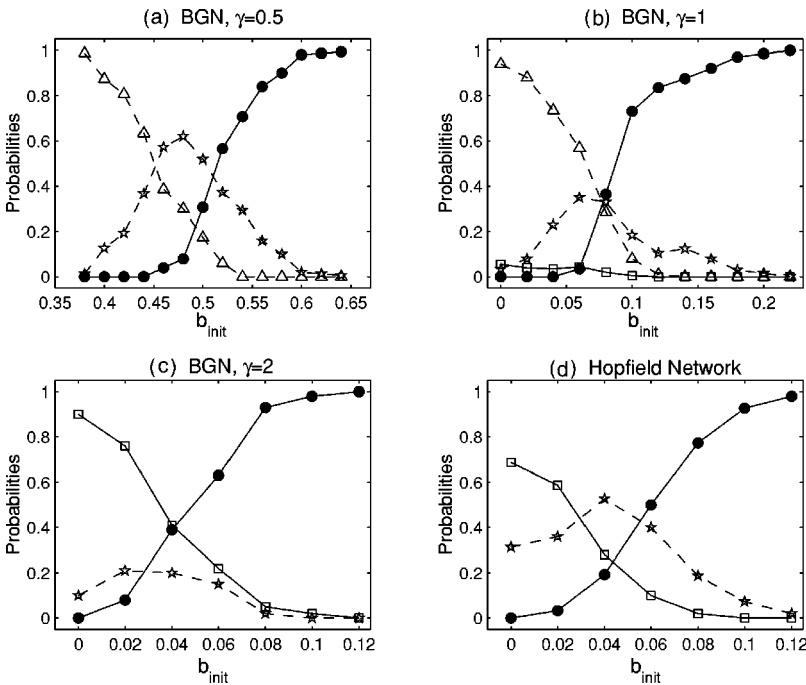
FIG. 9. Attractors retrieved from states with a specified initial overlap with a target pattern. Plots show probability of retrieving the target pattern (solid circles), one of the other memory patterns (squares), an uncondensed state (triangles), or a spin glass spurious state (stars).

states with one large overlap and two or more smaller (but greater than random) overlaps. Finally, for $\gamma = 1$ [Fig. 9(b)], the basins of the memory state are almost as large as in the HN case, and near the boundaries there is a small but non-zero probability of retrieving one of the other memory patterns, indicating that the basins of different memory states almost touch each other.

C. Qualitative picture of the energy landscape

Taken together, the above results suggest a qualitative, schematic picture of the energy landscape illustrated in Figs. 10–12. The representation of the configuration space by two dimensions is not to be taken literally, since it is of course N dimensional. At low values of γ, such as 0.5, the energy landscape is dominated by uncondensed states, which form a series of shallow basins, each limited to roughly a single

orthant of N-dimensional space. These are represented in the diagram by a series of shallow pits. The basins of attraction for the retrieval and mixture states form isolated depressions in this pitted plateau. They occupy nontrivial volumes but do not lie adjacent to each other (with the exception of certain spin glass states that lie near the retrieval states). At intermediate values $\gamma \approx 1$, the basins of attraction of the retrieval states are much larger and in some places almost touch each other, but significant islands of uncondensed states remain. By $\gamma = 2$, however, the uncondensed states have disappeared and the basins of attraction for the other two types of states occupy the entire energy landscape and share boundaries with each other.

VI. CONCLUSIONS

We have studied the behavior of the bistable gradient network in the thermodynamic low-loading limits $N \rightarrow \infty$, $p/N \ll 1$. We described and classified the attractors of the dynam-
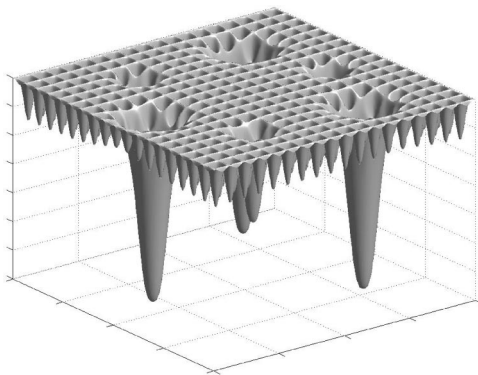

FIG. 10. Schematic illustration showing qualitative features of the energy landscape for low values of γ. The numerous small depressions represent uncondensed states. The retrieval states and mixture states occupy isolated valleys. Retrieval states are represented by the deeper valleys.
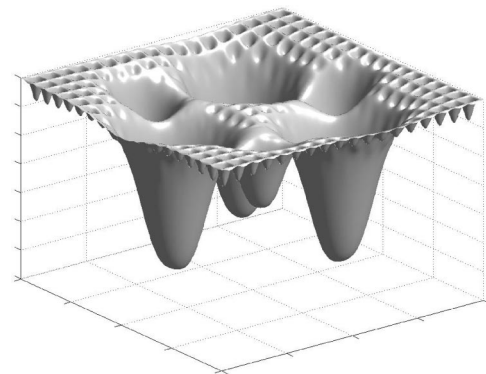

FIG. 11. Energy landscape for intermediate values of $\gamma \sim 1$. The retrieval states and mixture states have large basins of attraction that almost touch, but significant islands of uncondensed states remain.
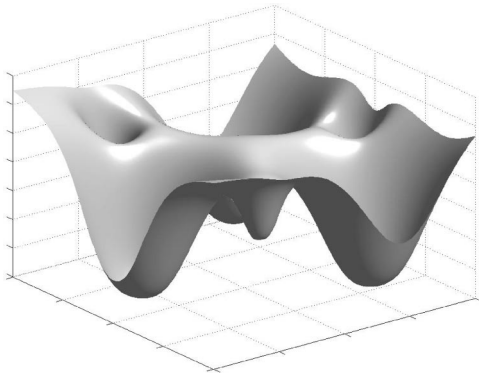
FIG. 12. Energy landscape for large γ. There are no uncondensed states, and large basins of attraction occupy the whole landscape.

ics and also observed the effectiveness of pattern retrieval as a function of the coupling parameter γ. We found that states corresponding to perfect retrieval of the stored patterns are linearly stable at all values of γ, and have an energy that decreases monotonically with increase in γ. Above the threshold $\gamma = \frac{1}{3}$ the retrieval states become stable against sign flips of one or more nodes, and the network begins to function as an associative memory. If γ is not far above this threshold, then the basins of attraction of the retrieval states are small and input must be very close in Hamming distance to a pattern for recognition to occur. The basins of attraction of the retrieval states grow as γ increases.

There are two regimes of behavior, distinguished by the types of attractors that occur. At low γ the configuration space is dominated by the uncondensed states, or states in which no node experiences a field strong enough to overcome its potential barrier. In these states, $|x_i|$ remains close to 1 for all nodes, and the energy remains close to -0.25. Each of these states occupies a basin of attraction confined to approximately a single orthant. In the limit $\gamma = 0$ there are 2^N such states, all degenerate in energy. As γ increases above the threshold $\gamma_c = \frac{1}{3}$, the retrieval states and the mixture or spin glass states at first occupy small isolated basins among the many uncondensed states. However, as γ increases further, these basins grow until they lie adjacent to each other. At some value of γ (observed to lie between 1 and 1.5), the uncondensed states disappear and there is a transition to a Hopfield-like regime where the basins of attraction for retrieval and spin glass states cover the whole configuration space. As γ increases still further, the retrieval basins grow at the expense of the spin glass states, so that the latter can be noticeably suppressed compared to the deterministic Hopfield case. This suppression of the spurious states occurs without thermal noise or a modification of the Hebb learning rule.

The uncondensed states represent a phase that is neither "ferromagnetic" (i.e., strongly ordered and correlated with one pattern) nor "glassy" in the sense that frustration is an important effect, yet they cannot properly be described as "paramagnetic," as paramagnetism is characterized by spins that are able to flip freely from one orientation to the other.

A few words on the application of such networks to practical problems of associative memory are in order. The goal of associative memory is to reconstruct a pattern from a more or less corrupted version or from a fragment of the pattern, without becoming trapped in a spurious local minimum. From this point of view, it appears that increasing γ improves the performance of the network—expanding the basins of attraction for the retrieval states and suppressing the spurious states. The low-γ regime, on the other hand, may be suited to applications where the goal is a *selective* associative memory, one which only recognizes a pattern from a fairly close approximation and thus avoids false recognition. In the low-γ regime, if the input is not close to one of the stored patterns, then the network is likely to remain in an uncondensed state. These can, in general, be distinguished clearly from other states (especially retrieval states) by their relatively high energy ($E/N \approx -0.25$) or by the fact that the magnitudes of all $|x_i|$ remain close to 1. The magnitudes of the outputs can therefore be read as a signal of whether recognition has occurred. Persistence in an uncondensed state corresponds to an "I do not know" or nonrecognition response.

In a subsequent publication, we will examine the behavior of the BGN when the loading level p/N is of order unity, and we will demonstrate another performance trade off. Specifically, we will show that the maximum storage capacity of the network decreases as γ increases. For a low-γ regime, it is possible to stabilize more memorized patterns than in the Hopfield case, while at higher γ, even though the low-loading fault tolerance is increased, the storage capacity decreases.

[1] S. Haykin, *Neural Networks: A Comprehensive Foundation* (Prentice-Hall, Upper Saddle River, NJ, 1999).

[2] Daniel J. Amit, *Modeling Brain Function: The World Of Attractor Neural Networks* (Cambridge University Press, Cambridge, 1989).

[3] J.J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).

[4] W.A. Little, Math. Biosci. **19**, 101 (1974); W.A. Little and G.L. Shaw, *ibid.* **39**, 281 (1978).

[5] S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978).

[6] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

[7] V. Chinarov and M. Menzinger, Biosystems **55**, 137 (2000).

[8] F.C. Hoppensteadt and E.M. Izhikievich, *Weakly Connected Neural Networks* (Springer-Verlag, New York, 1997).

[9] T. Poston and I. Stewart, Behav. Sci. **23**, 318 (1978); I.N.

Stewart and P.L. Peregoy, Psychol. Bull. **94**, 336 (1983).

[10] J.J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **81**, 3088 (1984).

[11] R. Kühn, S. Bös, and J.L. van Hemmen, Phys. Rev. A **43**, 2084 (1991); R. Kühn and S. Bös, J. Phys. A **26**, 831 (1993). These authors develop a formalism that may be applicable to networks with bistable potentials and provide an alternative approach, but the examples they treat do not have the bistability property.

[12] W. Hohmann, M. Krauss, and F.W. Schneider, J. Phys. Chem. A **103**, 7606 (1999); **102**, 3103 (1998); **101**, 7364 (1997); G. Dechert, K.-P. Zeyer, D. Lebender, and F.W. Schneider, *ibid.* **100**, 19 043 (1996).

[13] J.-P. Laplante, M. Pemberton, A. Hjelmfelt, and J. Ross, J. Phys. Chem. **99**, 10 063 (1995); V. Booth, T. Erneux, and J.-P. Laplante, *ibid.* **98**, 6537 (1994); A. Hjelmfelt and J. Ross, *ibid.* **97**, 7998 (1993).

[14] D.J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985); Phys. Rev. Lett. **55**, 1530 (1985); Ann. Phys. (Paris) **173**, 30 (1987); A. Chrisanti, D.J. Amit, and H. Gutfreund, Europhys. Lett. **2**, 337 (1986).

[15] D.O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949).

[16] P. McGraw and M. Menzinger, Phys. Rev. E **67**, 016119 (2003).