Journal of Statistical Physics, Vol. 55, Nos. 5/6, 1989

Some Rigorous Results on the Hopfield Neural Network Model

Hans Koch[1] **and Jacques Piasko**[2]

Received July 27, 1988; revision January 10, 1989

We analyze the thermal equilibrium distribution of 2^p mean field variables for the Hopfield model with p stored patterns, in the case where 2^p is small compared to the number of spins. In particular, we give a full description of the free energy density in the thermodynamic limit, and of the so-called "symmetric solutions" for the mean field equations.

KEY WORDS: Neural network; random interaction; mean field theory; critical points.

1. INTRODUCTION AND MAIN RESULTS

We consider Hopfield's model[1-3] of an associative read-only memory with p stored patterns in the case where 2^p is small compared to the number of degrees of freedom (neurons, spins). The time evolution of this model is the Glauber dynamics for a system of N interacting Ising spins S_i with values $+1$ or -1, governed by a Hamiltonian of the form

$$H_N = - \sum_{1 \leqslant i < j \leqslant N} J_{ij} S_i S_j - \sum_{1 \leqslant i \leqslant N} T_i S_i \tag{1.1}$$

Here, the values of the coupling constants J_{ij} and T_i depend on the content of the memory: If $\xi = (\xi^1, \xi^2, ..., \xi^p)$ is an arbitrary but fixed collection of spin configurations, representing the stored patterns, then the following constants are chosen:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \qquad T_i = \eta \xi_i^\nu \tag{1.2}$$

where ν and η are external field parameters which will be specified later.

[1] Department of Mathematics, University of Texas at Austin, Austin, Texas 78712.

[2] Institut für Theoretische Physik, ETH Hönggerberg, CH-8093 Zürich, Switzerland.

0022-4715/89/0600-0903$06.00/0

The generalized Hopfield dynamics[4] describes the retrieval, from a "noisy" memory, of p stored patterns by association with some input pattern. The retrieval process can be viewed as a random walk on the set $\mathscr{S}$ of spin configurations: If ρ is the delta function on $\mathscr{S}$ with peak at the input pattern, or any other probability distribution on $\mathscr{S}$, then the spin distribution after one unit of time is given by $W\rho$,

$$(W\rho)(S') = \sum_{S \in \mathscr{S}} W(S', S)\, \rho(S) \tag{1.3}$$

The transition probabilities $W(S', S)$ depend on the inverse temperature β of the noise, and are defined as follows. If S' can be obtained from S by flipping a single spin, then

$$W(S', S) = \frac{1}{N} \frac{\omega(S')}{\omega(S) + \omega(S')} \tag{1.4}$$

where ω denotes the normalized Gibbs distribution

$$\omega(S) = \left(\sum_{S \in \mathscr{S}} e^{-\beta H_N(\eta, \xi, S)} \right)^{-1} e^{-\beta H_N(\eta, \xi, S)} \tag{1.5}$$

All other off-diagonal entries of W are zero, and the diagonal entries are determined by requiring that the probabilities $W(S', S)$, for any fixed S, add up to one.

From this definition it follows immediately that all matrix elements of W^N are positive, and that W satisfies the detailed balance condition $\omega(S)\, W(S', S) = \omega(S')\, W(S, S')$, for any pair (S', S) of spin configurations. Thus, by the fundamental theorem of Monte Carlo calculus, $W^n \rho$ converges to ω as $n \to \infty$, for every probability distribution ρ on $\mathscr{S}$, in contrast to the situation at zero temperature,[2,5,13,14] where every local minimum of the energy function H_N corresponds to an attractor of the Hopfield dynamics. As long as β is finite, the retrieval of information is a transient process; after a sufficiently long time, the system starts to forget its initial condition, and approaches the thermal equilibrium state ω.

The first part of our analysis deals with the equilibrium properties of the Hopfield model with "unbiased" memories. More precisely, we consider the free energy per spin, averaged over all 2^{pN} possible choices of p patterns,

$$F_N(\beta, \eta) = 2^{-pN} \sum_{\xi} \frac{-1}{\beta N} \ln \left(\sum_{S \in \mathscr{S}} e^{-\beta H_N(\eta, \xi, S)} \right) \tag{1.6}$$

and we assume that $2^p \ll N$. The averaging will be justified later by showing that, outside a negligible set of "biased" patterns ξ, the free energy per spin converges uniformly to the same value as F_N (i.e., it is self-averaging), for large N.

This model, and variations thereon, have been studied in detail in the case of a fixed finite number of patterns.[4,6,8-10] As for the thermodynamic behavior, it is found that a second-order phase transition occurs at $\beta = 1$, from a paramagnetic high-temperature phase ($\beta < 1$) to a ferromagnetic low-temperature phase ($\beta > 1$). The following theorem establishes the existence of this phase transition in a more general situation where the number of patterns is not necessarily finite. A proof is given in Section 2.

Theorem 1.1. Fix $\beta \neq 1$ and $\alpha < 1$. Then the average free energy density $F_N(\beta, \eta)$ converges as $N \to \infty$ for any positive integer v and for any sequence $p = p(N)$ satisfying $v \leqslant p$ and $2^p \leqslant N^\alpha$. The same holds for the magnetization $m_N(\beta, \eta) = (\partial/\partial\eta)\, F_N(\beta, \eta)$ if $\eta \neq 0$. The corresponding limits F_∞ and m_∞ only depend on (β, η), and they satisfy

$$\begin{aligned} \beta F_\infty(\beta, \eta) &= -\ln 2 - \frac{1}{2}\int_0^\beta a_1(t)^2\, dt + \mathcal{O}(\eta) \\ m_\infty(\beta, \eta) &= \operatorname{sgn}(\eta)\, a_1(\beta) + \mathcal{O}(\eta), \qquad \eta \neq 0 \end{aligned} \tag{1.7}$$

where $a_1(\beta)$ is the largest solution of the equation $\tanh(\beta a_1) = a_1$.

Note that if p is constant, then the $2p$ possible choices for $(v, \operatorname{sgn} \eta)$ lead to $2p$ distinct phases at low temperature (by symmetry, the magnetization in the direction of ξ^μ is zero for all $\mu \neq v$). If 2^p grows like N^α, with $\alpha < 1$, then an infinite number of these low-temperature phases are obtained. The case $\alpha = 1$ represents the borderline case for the methods used in our proof of Theorem 1.1, and possibly also for the validity of Eq. (1.7); but the phase portrait is believed to be the same for all α. On the other hand, if p grows at a rate proportional to N, the Hopfield model is expected to exhibit a spin glass phase.[7,11]

We shall now change to a reduced representation (of the Hopfield model), in which the independent degrees of freedom are $d = 2^p$ mean field variables.[9] Let $\{e_1, e_2, ..., e_d\}$ be a fixed, ordered set which contains all vectors in $\mathbb{R}^p$ whose components are either $+1$ or -1. Any choice of p patterns can then be regarded as a map $\xi: i \mapsto \xi_i = (\xi_i^1, \xi_i^2, ..., \xi_i^p)$, which associates to every site i, $1 \leqslant i \leqslant N$, one of the vectors e_k. The map ξ defines a partition $N = L_1 + L_2 + \cdots + L_d$ of N, where $L_k = L_k(\xi)$ denotes the number of sites in $\xi^{-1}(e_k)$. It also determines a partition of the spin configuration space $\mathcal{S}$ into subsets

$$\mathcal{S}(Y) = \left\{ S \in \mathcal{S} : \sum_{i \in \xi^{-1}(e_k)} S_i = Y_k,\ 1 \leqslant k \leqslant d \right\} \tag{1.8}$$

indexed by vectors $Y \in \mathbb{Z}^d$ with components $Y_k \in \{-L_k, -L_k+2, ..., L_k\}$, $1 \leqslant k \leqslant d$. Such vectors will be referred to as mean field configurations.

It is easy to check that the Hamiltonian H_N is constant on each of the sets $\mathscr{S}(Y)$. In addition, the Hopfield dynamics induces a random walk on the set $\mathscr{Y}$ of mean field configurations, with transition probabilities given by

$$\bar{W}(Y', Y) = \frac{1}{|\mathscr{S}(Y)|} \sum_{S \in \mathscr{S}(Y)} \sum_{S' \in \mathscr{S}(Y')} W(S', S) \tag{1.9}$$

To be more specific, we define a linear transformation A which maps functions on $\mathscr{S}$ to functions on $\mathscr{Y}$, by the equation

$$(A\rho)(Y) = \sum_{S \in \mathscr{S}(Y)} \rho(S) \tag{1.10}$$

Proposition 1.2. The matrix $\bar{W}$ is conjugate to W and satisfies detailed balance, i.e., (i) $AW = \bar{W}A$, and (ii) $(A\omega)(Y)\,\bar{W}(Y', Y) = (A\omega)(Y')\,\bar{W}(Y, Y')$ for every $Y, Y' \in \mathscr{Y}$.

From either of these two properties (whose proof is straightforward) it follows that every probability distribution on $\mathscr{Y}$ converges to $A\omega$ under the mean field dynamics defined by $\bar{W}$. The equilibrium distribution $A\omega$ is the Gibbs distribution for the mean field Hamiltonian $\bar{H}_N$, given by

$$\bar{H}_N(\beta, \eta, L, Y) \equiv H_N(\eta, \xi, S) - \frac{1}{\beta} \ln |\mathscr{S}(Y)|, \qquad S \in \mathscr{S}(Y) \tag{1.11}$$

To simplify our discussion, we assume from now on that $L_k = N/d$, for $1 \leqslant k \leqslant N$. As far as the proof of Theorem 1.1 is concerned, this restriction is justified by the fact that the average (1.6) may be restricted to patterns satisfying $|L_k(\xi) - N/d| < (N/d)^{1/2} \ln N$ for all k, without affecting the limit $N \to \infty$; for details see Section 2. With all L_k set to N/d, and for zero external field, the mean field Hamiltonian becomes

$$\bar{H}_N(\beta, 0, L, Y) = N\beta^{-1} \ln 2 + N\beta^{-1} f_\beta\left(\frac{d}{N} Y\right) + o(N) \tag{1.12}$$

where

$$f_\beta(y) = \frac{1}{d} \sum_{k=1}^{d} \int_0^{y_k} dt \tanh^{-1}(t) - \frac{\beta}{2d} \|Py\|^2 \tag{1.13}$$

Here, $\|\cdot\|$ is the norm defined by the standard inner product on $\mathbb{R}^d$, and P denotes the orthogonal projection in $\mathbb{R}^d$ onto the subspace spanned by the p vectors $e^\mu = (e_1^\mu, e_2^\mu, ..., e_d^\mu)$. Formally, it is now clear that in the limit $N \to \infty$, and for fixed, finite p, the average free energy density is determined by the minimum of f_β on the hypercube $[-1, 1]^d$. It can be shown that $f_\beta(y)$ takes on this minimum value if and only if

$$y = \pm a_1(\beta) e^\mu, \qquad 1 \leqslant \mu \leqslant p \tag{1.14}$$

These $2p$ minimizing vectors (for $\beta > 1$) are commonly referred to as retrieval states, since each of them is associated with exactly one of the stored patterns (e^μ is the mean field analogue of the pattern ξ^μ). Below we will discuss other local minima of f_β, or so-called spurious states, which are associated with several patterns (thus corresponding to a confused memory). In numerical experiments, both types of "states" behave like attractors for the Hopfield dynamics if N is sufficiently large. We expect that any distribution on $\mathscr{Y}$ whose support lies within a distance $o(N)$ of a local minimum Y of $\bar{H}_N$ will evolve first into a distribution which is essentially localized in a ball of radius $\mathcal{O}(\sqrt{N})$ around Y before spreading out significantly. Formal calculations indicate that the time scale for the localization process is of the order of N, while the time needed to reach an approximate thermal equilibrium grows exponentially with N.

In the second part of this paper we consider the set of critical points of the function f_β, or, equivalently, the solutions of the (mean field) equation

$$y_k = \tanh[\beta(Py)_k], \qquad 1 \leqslant k \leqslant d \tag{1.15}$$

Our first result describes the so-called "symmetric solutions" of order n, whose existence, for all $\beta > 1$ and $n \leqslant p$, has been conjectured in ref. 4, based on (numerical calculations and) expansions near $\beta = 1$ and $\beta = \infty$. A symmetric solution of order $n > 0$ (the case $n = 0$ corresponds to the trivial solution $y = 0$) can be obtained by making the ansatz $y = a_n(e^1 + \cdots + e^n) + w$, with $Pw = 0$. As shown in Section 3, this ansatz leads to the following equation for a_n:

$$a_n = 2^{-n+1} \sum_{0 \leqslant m < n/2} \binom{n}{m} \frac{n-2m}{n} \tanh[(n-2m)\beta a_n] \tag{1.16}$$

Theorem 1.3. Given $\beta > 1$ and a positive integer n, Eq. (1.16) has a unique positive solution $a_n = a_n(\beta)$. Furthermore, if $(c_1, c_2, ..., c_p)$ is a vector of length n in $\mathbb{R}^p$ whose components are either 0 or ± 1, and if $y \in \mathbb{R}^d$ is defined by

$$y_k = \tanh\left[\beta a_n(\beta) \sum_{\mu=1}^{p} c_\mu e_k^\mu\right], \qquad 1 \leqslant k \leqslant d \tag{1.17}$$

then the function f_β has a critical point at y.

For $\beta \leqslant 1$, it is easy to see that Eq. (1.15) admits only the trivial solution, and that f_β takes its minimum value for $y = 0$. This minimum turns into a local maximum as β is increased past its critical value $\beta = 1$, and the remaining $3^p - 1$ symmetric solutions bifurcate away from the origin. A qualitative picture of what happens near $\beta = 1$ can be derived from general results in bifurcation theory; see ref. 10. A more direct approach, which also allows for explicit numerical bounds, is presented in Section 3. In particular, we prove the following result.

Theorem 1.4. Let $1 < \beta < 1 + (9d + 500p^8)^{-1}$ and $y \in \mathbb{R}^d$.

(i) If f_β has a critical point at y, then y is a symmetric solution of some order $n \leqslant p$.

(ii) If f_β has a local manimum at y, then y is a symmetric solution of order $n = 1$.

We note that, while some condition on β is necessary in order for the conclusion of Theorem 1.4 to hold, the bound given here is clearly too restrictive. Numerical results[4] indicate that there is an increasing sequence of inverse temperatures β_m, starting with $\beta_1 \approx 2.17$, such that if β is larger (smaller) than β_m, then every symmetric solution of order $2m + 1$ corresponds to a local minimum (saddle point) of f_β. In contrast, the symmetric solutions of even order seem to correspond to saddle points, for all $\beta > 1$.

Our last result concerns the observed qualitative difference between solutions of even and odd order.

Theorem 1.5. Let m be a positive integer not exceeding $p/2$, and assume that

$$\beta \cdot 2^{-2m-1} \binom{2m}{m} > \ln(\beta) > 1 \tag{1.18}$$

Then f_β has a saddle point or local maximum at every symmetric solution of order $2m$, and if $2m < p$, f_β has a local minimum at every symmetric solution of order $2m + 1$.

Further details, including the proofs of Theorems 1.3–1.5, are given in Section 3.

2. THE THERMODYNAMIC LIMIT

In this section we will prove Theorem 1.1. We start by deriving an explicit expression for the mean field Hamiltonian $\bar{H}_N$, as defined in (1.11). Note that if $S \in \mathscr{S}(Y)$, then

$$\sum_{i=1}^{N} \xi_i^\mu S_i = \sum_{k=1}^{d} \sum_{i \in \xi^{-1}(e_k)} \xi_i^\mu S_i = \sum_{k=1}^{d} e_k^\mu \sum_{i \in \xi^{-1}(e_k)} S_i = \sum_{k=1}^{d} e_k^\mu Y_k \equiv \langle e^\mu, Y \rangle \tag{2.1}$$

Using this identity, the mean field Hamiltonian can be written as follows:

$$\begin{aligned}
&\bar{H}_N(\beta, \eta, L, Y) \\
&\quad \equiv -\frac{1}{2N} \sum_{\mu=1}^{p} \left(\sum_{i=1}^{N} \xi_i^\mu S_i \right)^2 - \eta \sum_{i=1}^{N} \xi_i^\nu S_i + \frac{p}{2} - \frac{1}{\beta} \ln |\mathscr{S}(Y)| \\
&\quad = -\frac{1}{2N} \sum_{\mu=1}^{p} \langle e^\mu, Y \rangle^2 - \eta \langle e^\nu, Y \rangle + \frac{p}{2} - \frac{1}{\beta} \sum_{k=1}^{d} \ln \left(\binom{L_k}{\frac{1}{2}(L_k + Y_k)} \right)
\end{aligned} \tag{2.2}$$

The entropy term may be represented more conveniently by using Stirling's formula: There is a function g, satisfying the bound $|g(L_k, Y_k)| \leqslant \ln(L_k) + 1$, such that

$$\ln \left(\binom{L_k}{\frac{1}{2}(L_k + Y_k)} \right) = L_k \ln 2 - L_k \int_0^{Y_k/L_k} dt \tanh^{-1}(t) + g(L_k, Y_k) \tag{2.3}$$

In order to discuss the N dependence of $\bar{H}_N$, let us now change to normalized variables by writing

$$L_k = (1 + \lambda_k) N/d. \qquad Y_k = L_k y_k \tag{2.4}$$

The range of values for (λ, y) is determined from that of the original variables (L, Y). In particular, y takes on values in the set $\mathscr{X} = \{ y \in [-1, 1]^d \colon L_k(1 + y_k)/2 \in \mathbb{N},\ 1 \leqslant k \leqslant d \}$. Denoting by Λ the diagonal $d \times d$ matrix with entries $\Lambda_{kk} = \lambda_k$, we arrive at the following expression for $\bar{H}_N$:

$$\beta \bar{H}_N(\beta, \eta, L, Y) = -N \ln 2 + \frac{N}{d} f(\beta, \eta, \lambda, y) + \frac{p}{2} + \sum_{k=1}^{d} g(L_k, Y_k) \tag{2.5}$$

where

$$\begin{aligned}
f(\beta, \eta, \lambda, y) = &-\frac{\beta}{2} \| P(1 + \Lambda) y \|^2 - \beta \eta \langle e^\nu, (1 + \Lambda) y \rangle \\
&+ \sum_{k=1}^{d} (1 + \lambda_k) \int_0^{y_k} dt \tanh^{-1}(t)
\end{aligned} \tag{2.6}$$

The quantities of interest in this section are the free energy density $F_{N,d}$ and the magnetization $m_{N,d}$,

$$\begin{aligned}
F_{N,d}(\beta, \eta, L) &= -\frac{1}{\beta N} \ln \left\{ \sum_{y \in \mathscr{X}} \exp^{[-\beta \bar{H}_N(\beta, \eta, L, Y)]} \right\} \\
m_{N,d}(\beta, \eta, L) &= \frac{d}{d\eta} F_{N,d}(\beta, \eta, L)
\end{aligned} \tag{2.7}$$

More precisely, if $\phi_{N,d}(L)$ denotes one of the functions defined in (2.7), we would like to compute the average of $\phi_{N,d}(L(\xi))$ over all possible choices of the patterns ξ,

$$2^{-pN} \sum_{\xi} \phi_{N,d}(L(\xi)) = d^{-N} \sum_{L \in \mathscr{P}_{N,d}} \frac{N!}{L_1!\, L_2! \cdots L_d!} \phi_{N,d}(L) \tag{2.8}$$

Here, $\mathscr{P}_{N,d}$ denotes the set of vectors in $\mathbb{R}^d$ whose components are non-negative integers which add up to N. The following proposition (a simple large-deviations estimate) will be used to approximate the sum (2.8) by a sum over "unbiased" patterns, represented by the set $\mathscr{U}_{N,s} = \{L \in \mathscr{P}_{N,d}$: $|\lambda_k| < \delta,\ 1 \leqslant k \leqslant d\}$, for some fixed $\delta > 0$.

Proposition 2.1. There exists $\delta_0 > 0$ such that if $d/N \leqslant \delta \leqslant \delta_0$, then

$$d^{-N} \sum_{L \in \mathscr{P}_{N,d} \setminus \mathscr{U}_{N,d}} \frac{N!}{L_1!\, L_2! \cdots L_d!} < dN \exp\left(-\frac{N\delta^2}{2d}\right) \tag{2.9}$$

Proof. Assume that $d \leqslant \delta N$, and denote by $\mathscr{B}$ the set of all non-negative integers $n \leqslant N$ which satisfy $|n - N/d| \geqslant (N/d)\,\delta$. Since every vector in $\mathscr{P}_{N,d} \setminus \mathscr{U}_{N,d}$ has at least one of its d components in $\mathscr{B}$, the left-hand side of (2.9) is bounded by

$$\begin{aligned} \sigma &\equiv d \cdot d^{-N} \sum_{L \in \mathscr{P}_{N,d}:\, L_1 \in \mathscr{B}} \frac{N!}{L_1!\, L_2! \cdots L_d!} \\ &= d \cdot d^{-N} \sum_{L_1 \in \mathscr{B}} \left[\binom{N}{L_1} (d-1)^{N-L_1} \right] \end{aligned} \tag{2.10}$$

It is easy to check that the expression in square brackets, when considered as a function of L_1, is increasing for $L_1 < (N-d+1)/d$ and decreasing for $L_1 > (N+1)/d$. Thus, if ε is chosen such that $L_1 = (1+\varepsilon)N/d$ maximizes $[\cdots]$ on $\mathscr{B}$, then $\delta \leqslant |\varepsilon| \leqslant 2\delta$, and

$$\sigma \leqslant dN \cdot d^{-N} \binom{N}{(N/d)(1+\varepsilon)} (d-1)^{N-(1+\varepsilon)N/d} \tag{2.11}$$

By applying Stirling's formula to the combinatorial factor on the right-hand side of inequality (2.11) and then simplifying the result, we obtain

$$\ln(\sigma) \leqslant \ln(dN) - Ng(\varepsilon) - \frac{1}{2} \ln\left[\frac{N}{d}(1+\varepsilon)\right] + \text{const} \tag{2.12}$$

where

$$g(\varepsilon) = \frac{1}{d}(1+\varepsilon)\ln(1+\varepsilon) + \left[1 - \frac{1}{d}(1+\varepsilon)\right] \ln\left(1 - \frac{\varepsilon}{d-1}\right) \tag{2.13}$$

An explicit calculation shows that $g(0) = g'(0) = 0$, and that $g''(\varepsilon) > 1/d$, for ε sufficiently close to zero. Since $\delta \leqslant |\varepsilon| \leqslant 2\delta$, we can now bound σ as follows:

$$\sigma \leqslant dN\, e^{-Ng(\varepsilon)} \leqslant dN \exp\left(-\frac{N\delta^2}{2d}\right) \tag{2.14}$$

provided that δ is sufficiently small, and $d \leqslant \delta N$. This proves the assertion of Proposition 2.1. ▮

In what follows, the number p of patterns is assumed to be sufficiently small, such that $d \equiv 2^p \leqslant N^\alpha$, for some fixed, positive constant $\alpha < 1$. We also choose, once and for all,

$$\delta = (d/N)^{1/2} \ln(N) \tag{2.15}$$

Corollary 2.2 (Self-averaging). Let $(N, d) \mapsto (\phi_{N,d}\colon \mathscr{P}_{N,d} \to \mathbb{R})$ be a two-parameter sequence of functions, and assume that there are constants ϕ_∞ and κ, λ, $M > 0$ such that for $N > M$ and for $d \leqslant N^\alpha$ the following hold

(a) $|\phi_{N,d}(L)| \leqslant N^\lambda$ for all $L \in \mathscr{P}_{N,d}$.

(b) $|\phi_{N,d}(L) - \phi_\infty| \leqslant N^{-\kappa}$ for $L \in \mathscr{U}_{N,d}$.

Then

$$\left| d^{-N} \sum_{L \in \mathscr{P}_{N,d}} \frac{N!}{L_1!\, L_2! \cdots L_d!} \phi_{N,d}(L) - \phi_\infty \right| \leqslant 2N^{-\kappa} \tag{2.16}$$

provided that N is sufficiently large, and $d \leqslant N^\alpha$.

Proof. Using Proposition 2.1 and assuming properties (a) and (b), we can bound the left-hand side of (2.16) as follows:

$$\begin{aligned} &d^{-N} \sum_{L \in \mathscr{P}_{N,d} \setminus \mathscr{U}_{N,d}} \frac{N!}{L_1! \cdots L_d!} |\phi_{N,d}(L) - \phi_\infty| \\ &\qquad + d^{-N} \sum_{L \in \mathscr{U}_{N,d}} \frac{N!}{L_1! \cdots L_d!} |\phi_{N,d}(L) - \phi_\infty| \\ &\qquad \leqslant (N^\lambda + |\phi_\infty|) N^2 e^{-(\ln N)^2/2} + N^{-\kappa} \end{aligned} \tag{2.17}$$

For sufficiently large N, the last expression is bounded by $2N^{-\kappa}$. ▮

Our aim is to apply this corollary to the free energy and to the magnetization, as defined in (2.7). The hypothesis (a) is easy to check in these two cases: $F_{N,d}$ and $m_{N,d}$ are bounded in absolute value by $p/2 + \text{const} = \mathcal{O}(\ln N)$ and 1, respectively. We shall now work toward the proof of property (b).

Proposition 2.3. For $L \in \mathscr{U}_{N,d}$,

$$\left|\beta \bar{H}_N(\beta, \eta, L, Y) + N \ln 2 - \frac{N}{d} f(\beta, \eta, 0, y)\right| \leqslant (1+\beta)(3+|\eta|)\,\delta N \quad (2.18)$$

Proof. By using Eq. (2.5) and the fact that $|y_k| \leqslant 1$ and $|\lambda_k| \leqslant \delta$, we have

$$\begin{aligned}
&\left|\beta \bar{H}_N(\beta, \eta, L, Y) + N \ln 2 - \frac{N}{d} f(\beta, \eta, 0, y)\right| \\
&\quad = \left|\frac{N}{d} f(\beta, \eta, \lambda, y) - \frac{N}{d} f(\beta, \eta, 0, y) + \frac{p}{2} - \sum_{k=1}^{d} g(L_k, Y_k)\right| \\
&\quad \leqslant \frac{N}{d}\left|\frac{\beta}{2}\langle P(2+\Lambda)\, y, \Lambda y\rangle + \beta\eta\langle e^{\nu}, \Lambda y\rangle \right. \\
&\qquad \left. - \sum_{k=1}^{d} \lambda_k \int_0^{y_k} dt \tanh^{-1}(t)\right| + \frac{p}{2} + \sum_{k=1}^{d} |g(L_k, Y_k)| \\
&\quad \leqslant \frac{N}{d}\left[\frac{\beta}{2}(2+\delta)\,\delta d + \beta|\eta|\,\delta d + \delta d \ln 2\right] + 2d \ln N \\
&\quad \leqslant (1+\beta)(3+|\eta|)\,\delta N \quad \blacksquare
\end{aligned} \quad (2.19)$$

Since the dominant contributions to the free energy density and the magnetization come from mean field configurations Y which minimize $\bar{H}_N$, we continue by estimating $f(\beta, \eta, 0, y)$ near its minimum. To do so, let us write

$$f(\beta, \eta, 0, y) = -\sum_{k=1}^{d} h_k(y_k) + \frac{\beta}{2}\|(1-P)\, y\|^2 \quad (2.20)$$

where, for $|s| \leqslant 1$,

$$h_k(s) = \int_0^s dt\, [\beta\eta e_k^{\nu} + \beta t - \tanh^{-1}(t)] \quad (2.21)$$

At this point, it is necessary to distinguish between the high-temperature phase ($\beta < 1$) and the low-temperature phase ($\beta > 1$), and between large and small external fields η. To avoid undue repetition, we will limit our discussion to $\beta > 1$ and to small values of η; the other cases can be treated similarly.

If η is sufficiently small (depending on β, but not on d), then each of the functions h_k has a unique positive maximum on each of the intervals

$(-1, 0)$ and $(0, +1)$. Denote by $v_k(-)$ and $v_k(+)$ the location of these maxima; then $v_k(\pm) = \pm a_1(\beta) + \mathcal{O}(\eta)$, where $a_1(\beta)$ is the positive solution of the equation $a_1 = \tanh(\beta a_1)$.

Definition 2.4. Given a vector $y \in \mathbb{R}^d$, define $v(y)$ to be the vector in $\mathbb{R}^d$ whose kth component is $v_k(\operatorname{sgn} y_k)$ for $1 \leqslant k \leqslant d$. Here, we use, e.g., the convention that sgn 0 is positive. Furthermore, define $u = v(\eta e^\nu)$.

Proposition 2.5. Let $\beta > 1$. Then there are two positive constants $c_l < c_u$ such that for any vector y in $[-1, 1]^d$ and for $|\eta|$ sufficiently small,

$$\begin{aligned} f(\beta, \eta, 0, y) = f(\beta, \eta, 0, u) &+ \beta\, |\langle u - v(y), le^\nu \rangle| \\ &+ c(y)\, \|y - v(y)\|^2 + \tfrac{1}{2}\beta\, \|(1-P)\, y\|^2 \end{aligned} \tag{2.22}$$

where $c(y)$ and l are constants (depending on η) which satisfy the bounds $c_l \leqslant c(y) \leqslant c_u$ and $|l - \eta| = \mathcal{O}(\eta^3)$, respectively.

Proof. l is defined by writing the local minima of h_k in the form $h_k(v_k(\pm)) = E_k + \beta l e_k^\nu v_k(\pm)$. By computing the difference of these two values, we obtain

$$\begin{aligned} le_k^\nu &= \eta e_k^\nu + \frac{\beta^{-1}}{|v_k(+)| + |v_k(-)|} \int_{|v_k(-)|}^{|v_k(+)|} dt\, [\beta t - \tanh^{-1}(t)] \\ &= [\eta + \mathcal{O}(\eta^3)]\, e_k^\nu \end{aligned} \tag{2.23}$$

Given $y \in [-1, 1]^d$, define $x_k = \operatorname{sgn}(y_k)$. Then we can write $f(\beta, \eta, 0, y)$ as follows:

$$\begin{aligned} f(\beta, \eta, 0, y) = {} & -\beta \langle v(y), le^\nu \rangle + \sum_{k=1}^{d} [h_k(v_k(x_k)) - h_k(y_k) - E_k] \\ & + \frac{\beta}{2}\, \|(1-P)\, y\|^2 \end{aligned} \tag{2.24}$$

Since, for the values of β and η considered, h_k has a quadratic maximum at $v_k(x_k)$ which is unique in the interval bounded by 0 and x_k, we have

$$c_l \leqslant \frac{h_k(x_k) - h_k(y_k)}{|v_k(x_k) - y_k|^2} \leqslant c_u \tag{2.25}$$

with bounds c_l, c_u that are independent of k and y. As a consequence,

$$\begin{aligned} f(\beta, \eta, 0, y) = {} & -\beta \langle v(y), le^\nu \rangle + c(y)\, \|y - v(y)\|^2 \\ & - \sum_{k=1}^{d} E_k + \frac{\beta}{2}\, \|(1-P)\, y\|^2 \end{aligned} \tag{2.26}$$

for some constant $c(y)$ contained in the interval $[c_l, c_u]$.

Equation (2.22) is now obtained by using that the two norms in (2.26) vanish for $y=u$, and that $\langle v(y), le^{\nu}\rangle$ is maximized by $y=u$ [note also that $v(u)=u$]. Both of these properties follow from the fact that for $1\leqslant k\leqslant d$,

$$\operatorname{sgn}(u_k)=\operatorname{sgn}(\eta e_k^{\nu})=\operatorname{sgn}(le_k^{\nu}), \qquad |u_k|=\max\{|v_k(-)|, |v_k(+)|\}=t \tag{2.27}$$

where t is the largest solution of the equation $\beta|\eta|+\beta t-\tanh^{-1}(t)=0$. ▌

The following two propositions, together with Corollary 2.2, prove the assertions of Theorem 1.1.

Proposition 2.6. There is a function $F_{\infty}(\beta, \eta)$ with the following properties. If $\beta>1$ and if $|\eta|$ is sufficiently small, then

$$|F_{N,d}(\beta, \eta, L)-F_{\infty}(\beta, \eta)|\leqslant(d/N)^{1/3} \tag{2.28}$$

for all $L\in\mathscr{U}_{N,d}$ and for N sufficiently large. Furthermore,

$$F_{\infty}(\beta, \eta)=-\frac{\ln 2}{\beta}-\frac{1}{2\beta}\int_1^{\beta} dt\, a_1(t)^2+\mathcal{O}(\eta) \tag{2.29}$$

Proof. From the definition (2.7) and from Proposition 2.3, it follows that

$$\begin{aligned} F_{N,d}(\beta, \eta, L)=&-\frac{\ln 2}{\beta}+\frac{1}{\beta d}f(\beta, \eta, 0, u)+\mathcal{O}(\delta)\\ &-\frac{1}{\beta N}\ln\left(\sum_{y\in\mathscr{X}} e^{-[f(\beta,\eta,0,y)-f(\beta,\eta,0,u)]N/d}\right) \end{aligned} \tag{2.30}$$

for all $L\in\mathscr{U}_{N,d}$. To get an upper bound on the sum over y, we will use that, for large N, the number of elements in $\mathscr{X}$ is bounded by

$$|\mathscr{X}|=\prod_{k=1}^{d}(L_k+1)\leqslant(2N/d)^d\leqslant e^{\delta N} \tag{2.31}$$

To get the corresponding lower bound, we note that there is a vector $y'\in\mathscr{X}$, which is sufficiently close to u, such that $\operatorname{sgn}(y_k')=\operatorname{sgn}(u_k)$, and $|y_k'-u_k|\leqslant 3d/N$, for $1\leqslant k\leqslant d$. By Proposition 2.5,

$$\begin{aligned} 0&\leqslant|f(\beta, \eta, 0, y')-f(\beta, \eta, 0, u)|\\ &\leqslant\left(c_u+\frac{\beta}{2}\right)\|y'-u\|^2\leqslant 9\left(c_u+\frac{\beta}{2}\right)\frac{d^3}{N^2}\leqslant\delta d \end{aligned} \tag{2.32}$$

and therefore

$$|\mathscr{X}| \geqslant \sum_{y \in \mathscr{X}} e^{-[f(\beta,\eta,0,y)-f(\beta,\eta,0,u)]N/d} \geqslant e^{-\delta N} \tag{2.33}$$

This inequality, together with (2.31), shows that the last term in (2.30) is bounded by $\pm\delta$. The bound (2.28) follows if we define $F_\infty(\beta, \eta) \equiv [-\ln 2 + f(\beta, \eta, 0, u)/d]/\beta$.

In order to prove (2.29), we need only consider the case $\eta = 0$, since

$$|f(\beta, \eta, 0, u) - f(\beta, 0, 0, u)| = |\beta\eta\langle e^\nu, u\rangle| \leqslant \beta d\,|\eta| \tag{2.34}$$

From Eq. (2.6), using the fact that $y = u(\beta) = \operatorname{sgn}(\eta)\, a_1(\beta) e^\nu$ minimizes $f(\beta, 0, 0, y)$, we obtain

$$\begin{aligned} \frac{d}{d\beta} f(\beta, 0, 0, u(\beta)) &= \left(\frac{\partial}{\partial\beta} f\right)(\beta, 0, 0, u(\beta)) \\ &\quad + \left\langle \left(\frac{\delta}{\delta y} f\right)(\beta, 0, 0, u(\beta)), \frac{d}{d\beta} u(\beta)\right\rangle \\ &= \frac{1}{2}\,\|Pu(\beta)\|^2 + 0 = -\frac{1}{2}\, a_1(\beta)^2\, d \end{aligned} \tag{2.35}$$

The assertion now follows since $f(1, 0, 0, u(1)) = f(1, 0, 0, 0) = 0$. ∎

Proposition 2.7. There is a function $m_\infty(\beta, \eta)$ with the following properties. If $\beta > 1$ and if $|\eta| > 0$ is sufficiently small, then

$$|m_{N,d}(\beta, \eta, L) - m_\infty(\beta, \eta)| \leqslant 3(d/N)^{1/5} \tag{2.36}$$

for all $L \in \mathscr{U}_{N,d}$ and for N sufficiently large. Furthermore,

$$m_\infty(\beta, \eta) = \operatorname{sgn}(\eta)\, a_1(\beta) + \mathscr{O}(\eta) \tag{2.37}$$

Proof. The magnetization $m_{N,d}$ is given by the following expression:

$$\begin{aligned} m_{N,d}(\beta, \eta, L) &= \left(\sum_{y \in \mathscr{X}} e^{-\beta \bar{H}_N(\beta,\eta,L,Y)}\right)^{-1} \\ &\quad \times \sum_{y \in \mathscr{X}} \frac{1}{d} \langle e^\nu, (1 + A)\, y\rangle e^{-\beta \bar{H}_N(\beta,\eta,L,Y)} \end{aligned} \tag{2.38}$$

By writing the inner product in Eq. (2.38) as the sum

$$\frac{1}{d}\langle e^\nu, (1 + A)\, y\rangle = \frac{1}{d}\langle e^\nu, u\rangle + \frac{1}{d}\langle e^\nu, (1 + A)\, y - u\rangle \tag{2.39}$$

we split the magnetization into a leading term $m_\infty(\beta, \eta) \equiv (1/d)\langle e^v, u\rangle$ and a remainder. The sum over y in the remainder is now estimated separately on the set $R = \{y \in \mathscr{X}: |\langle y-u, e^v\rangle| \leqslant \varepsilon d\}$ and on its complement, where $\varepsilon = (d/N)^{1/5}$. For $y \in R$, we have

$$\frac{1}{d}|\langle e^v, (1+\Lambda)\,y-u\rangle| \leqslant \frac{1}{d}|\langle e^v, \Lambda y\rangle| + \frac{1}{d}|\langle e^v, y-u\rangle| \leqslant \delta + \varepsilon \tag{2.40}$$

Using Proposition 2.3 and Proposition 2.6, we arrive at the bound

$$\begin{aligned}
&\left| m_{N,d}(\beta, \eta, L) - \frac{1}{d}\langle e^v, u\rangle \right| \\
&\quad \leqslant \delta + \varepsilon + \left(\sum_{y \in \mathscr{X}} e^{-\beta \bar{H}_N(\beta, \eta, L, Y)} \right)^{-1} \\
&\qquad \times \sum_{y \in \mathscr{X}\setminus R} \frac{1}{d}|\langle e^v, (1+\Lambda)\,y-u\rangle|\, e^{-\beta \bar{H}_N(\beta, \eta, L, Y)} \\
&\quad \leqslant 2\varepsilon + e^{\mathcal{O}(\delta N)} \sum_{y \in \mathscr{X}\setminus R} e^{-[f(\beta, \eta, 0, y) - f(\beta, \eta, 0, u)]N/d}
\end{aligned} \tag{2.41}$$

In order to estimate the last term in (2.41), we use the fact that for $y \in \mathscr{X}\setminus R$, $f(\beta, \eta, 0, y)$ cannot be very close to its minimum value. More precisely, if y lies in $\mathscr{X}\setminus R$, then either $|\langle e^v, u - v(y)\rangle| > \varepsilon d/2$ holds, or $|\langle e^v, y - v(y)\rangle| > \varepsilon d/2$. In the first case, it follows from Proposition 2.5 that, for small $|\eta|$,

$$f(\beta, \eta, 0, y) - f(\beta, \eta, 0, u) \geqslant \tfrac{1}{3}\beta|\eta|\,\varepsilon d \tag{2.42}$$

In the second case, we combine Proposition 2.5 with the inequality

$$\|y - v(y)\|^2 \geqslant |\langle e^v, y - v(y)\rangle|^2\, \|e^v\|^{-2} \geqslant \varepsilon^2 d/4 \tag{2.43}$$

to obtain a similar result:

$$f(\beta, \eta, 0, y) - f(\beta, \eta, 0, u) \geqslant (c_l/4)\,\varepsilon^2 d \tag{2.44}$$

By substituting these two bounds into (2.41), we find that, for $\varepsilon < |\eta|$,

$$\left| m_{N,d}(\beta, \eta, L) - \frac{1}{d}\langle e^v, u\rangle \right| \leqslant 2\varepsilon + e^{\mathcal{O}(\delta N)} \sum_{y \in \mathscr{X}\setminus R} e^{-\kappa\varepsilon^2 N} \tag{2.45}$$

where κ is some positive constant (depending on β and η). The number of terms in the sum over y is bounded by $\exp[\mathcal{O}(\delta N)]$, as in (2.31). The assertion (2.36) now follows since $\delta/\varepsilon^2 \to 0$ as N tends to infinity, while (2.37) follows from the fact that $u = \mathrm{sgn}(\eta)\, a_1(\beta) e^v + \mathcal{O}(\eta)$. ∎

3. THE SYMMETRIC SOLUTIONS

In this section we describe in more detail the set of critical points of the function f_β,

$$f_\beta(y) \equiv \frac{1}{d} f(\beta, 0, 0, y) = \frac{1}{d} \sum_{k=1}^{d} \int_0^{y_k} dt \tanh^{-1}(t) - \frac{\beta}{2d} \|Py\|^2 \tag{3.1}$$

defined for $y \in (-1, 1)^d$. The number of patterns p is assumed to be fixed, but arbitrary, and $d = 2^p$. As mentioned in the introduction, the local minima of f_β are expected to play an important role for the dynamics of the Hopfield model.

A well-known procedure for finding (e.g., numerically) the local minima of a function g is the method of steepest descent, which (in its simplest form) consists in iterating a map $\Omega: y \mapsto y - \lambda \nabla g(y)$. If $\lambda > 0$ is chosen sufficiently small (such that the Hessian of λg has only eigenvalues smaller than 2), then the stable fixed points of Ω are precisely the local minima of g. In the following, a map of this type will be used in order to distinguish local minima of f_β from other critical points; this map is also closely related to the one used in refs. 8 and 10, and somewhat similar to the learning algorithm of ref. 15.

Before applying the method of steepest descent, we may of course perform a change of variables $z \mapsto y$ such as the one defined by the equation

$$y = \operatorname{Tanh}(\beta z) \equiv (\tanh(\beta z_1), \tanh(\beta z_2), ..., \tanh(\beta z_d)) \tag{3.2}$$

Proposition 3.1. If y is a local minimum of f_β in the hypercube $(-1, 1)^d$, then $z = Py$ is a stable fixed point of the map

$$\Omega_\beta: \quad z \mapsto P \operatorname{Tanh}(\beta P z), \qquad z \in \mathbb{R}^d \tag{3.3}$$

Conversely, if z is a stable fixed point of Ω_β, then $y = \operatorname{Tanh}(\beta z)$ is a local minimum of f_β.

Proof. The derivative of the function $g_\beta = f_\beta(\operatorname{Tanh}(\beta \cdot))$ can be written as follows:

$$Dg_\beta(z; u) = \frac{\beta^2}{d} \langle z - P \operatorname{Tanh}(\beta z), \operatorname{Tanh}'(\beta z) \bullet u \rangle \tag{3.4}$$

where "$v \bullet$" denotes the diagonal matrix associated with a vector v, i.e., $(v \bullet u)_k = v_k u_k$. Since $\tanh'(\beta z) > 0$, we see that the critical points of g_β coincide with the fixed points of Ω_β. Assume now that $\Omega_\beta(z) = z$. Then the second derivative of g_β at z is given by

$$D^2 g_\beta(z; u, v) = \frac{\beta^2}{d} \langle [\mathrm{Id} - \beta P(\operatorname{Tanh}'(\beta z) \bullet)] v, \operatorname{Tanh}'(\beta z) \bullet u \rangle \tag{3.5}$$

Note that the matrix $[\cdots]$ in (3.5) is self-adjoint with respect to the inner product $(v, u) = \langle v, \mathrm{Tanh}'(\beta z) \bullet u \rangle$. Thus, we have

$$\inf_{(v,v)=1} D^2 g(z; v, v) = \frac{\beta^2}{d}(1-\lambda) \tag{3.6}$$

where λ is the largest eigenvalue of $\beta P(\mathrm{Tanh}(\beta z) \bullet)$, or, equivalently (if $\lambda \neq 0$), the largest eigenvalue of the tangent map $D\Omega_\beta(z) = \beta P(\mathrm{Tanh}'(\beta P z) \bullet) P$ of Ω_β at the fixed point z. ▌

At high temperatures ($\beta < 1$), the map Ω_β is easily seen to be a contraction, with fixed point $z = 0$. For $\beta \geqslant 1$ the situation is more complicated; but fortunately, the Hopfield model with orthogonal patterns (i.e., when $L_k = N/d$ for all k) has many symmetries. In order to describe these symmetries, let us denote by C_p the set of corners of the hypercube $[-1, 1]^p$, and by E the map $k \mapsto e_k$ which was introduced earlier for the purpose of (arbitrarily) enumerating the elements of C_p. If ψ is a permutation of the set C_p, we associate with ψ a linear transformation Ψ on $\mathbb{R}^d$ by defining

$$(\Psi y)_k = y_j, \qquad j = E^{-1}(\psi^{-1}(E(k))) \tag{3.7}$$

for all $y \in \mathbb{R}^d$, and for $1 \leqslant k \leqslant d$. Note that Ψ is orthogonal with respect to the standard inner product in $\mathbb{R}^d$. The following permutations are of particular interest; see also ref. 12. For $1 \leqslant \nu, \kappa, \lambda \leqslant p$ we define ψ_ν and $\psi_{\kappa\lambda}$ by setting

$$\begin{aligned} (\psi_\nu(c))^\mu &= \begin{cases} -c^\mu & \text{if } \mu = \nu \\ +c^\mu & \text{if } \mu \neq \nu \end{cases} \\ (\psi_{\kappa\lambda}(c))^\mu &= \begin{cases} c^\kappa & \text{if } \mu = \lambda \\ c^\lambda & \text{if } \mu = \kappa \\ c^\mu & \text{if } \mu \notin \{\kappa, \lambda\} \end{cases} \end{aligned} \tag{3.8}$$

for all c in C_p. If ψ is any of these permutations, then $\Psi^2 = \mathrm{Id}$, and thus Ψ is symmetric. Furthermore, the identity $(\Psi e^\mu)_k = (\psi^{-1}(e_k))^\mu$, which follows directly from (3.7), shows that Ψ_ν acts on the set $S = \{e^1, e^2, ..., e^p\}$ by multiplying the vector e^ν by -1, and $\Psi_{\kappa\lambda}$ acts on S by exchanging e^κ with e^λ. As a consequence, all of these transformations commute with P, the orthogonal projection onto the span of S. This proves the following proposition.

Proposition 3.2. Let ψ be one of the permutations defined in (3.8). Then $\Omega_\beta \circ \Psi = \Psi \circ \Omega_\beta$ for all β.

As another immediate consequence we have the following orthogonality property. Let $I = \{1, 2,..., p\}$, and for every subset $J \subset I$ let

$$e_k^{(J)} = \prod_{\mu \in J} e_k^\mu, \qquad 1 \leqslant k \leqslant d \tag{3.9}$$

where the value of an empty product is defined to be 1. It is easy to see that $e^{(J)}$ is an eigenvector of Ψ_ν for every $J \subset I$ and every $\nu \in I$; the corresponding eigenvalue is -1 if $\nu \in J$ and 1 if $\nu \notin J$. Since the operators Ψ_ν are symmetric and commute with each other, the set $\{e^{(J)}: J \subset I\}$ is an orthogonal basis for $\mathbb{R}^d$.

The next two propositions establish, for $\beta > 1$, the existence of 3^p "symmetric" fixed points for Ω_β. Each of these fixed points is associated with a nonnegative integer $n < p$ and with one of the following $2^n\binom{p}{n}$ vectors $v \in \mathbb{R}^d$:

$$v = \sum_{\mu=1}^{p} c_\mu e^\mu, \qquad c_\mu \in \{-1, 0, 1\}, \qquad \sum_{\mu=1}^{p} c_\mu^2 = n \tag{3.10}$$

Proposition 3.3. Let $1 \leqslant n \leqslant p$. If v satisfies (3.10), and if a is any real number, then $\Omega_\beta(av) = \gamma_n(\beta a) v$, where γ_n is the function defined by the following equation:

$$\gamma_n(x) = 2^{-n+1} \sum_{0 \leqslant m < n/2} \binom{n}{m} \frac{n-2m}{n} \tanh[(n-2m)x] \tag{3.11}$$

Proof. Given $n > 0$ and a vector v as in (3.10), denote by S the set of linear transformations Ψ which contains Ψ_ν if and only if $c_\nu = 0$, $\Psi_{\kappa\lambda}$ if and only if $c_\kappa = c_\lambda \neq 0$, $\Psi_\lambda \Psi_{\kappa\lambda} \Psi_\lambda$ if and only if $-c_\kappa = c_\lambda \neq 0$, and no other elements. It is easy to check that the only vectors $z \in P\mathbb{R}^d$ which satisfy $\Psi z = z$ for all Ψ in S are the multiples of v. Since by Proposition 3.2 we have $\Psi\Omega_\beta(av) = \Omega_\beta(a\Psi v) = \Omega_\beta(av)$ for all Ψ in S, it follows that $\Omega_\beta(av) = a'v$ for some real number a'.

In order to see that $a' = \gamma_n(\beta a)$, it is useful to write the components of v in the form

$$v_k = \sum_{\mu=1}^{p} c_\mu e_k^\mu = m \cdot 1 + (n-m)\cdot(-1) + (p-n)\cdot 0 \tag{3.12}$$

where $m = m(k)$ is the number of elements μ in $\{1, 2,..., p\}$ for which $c_\mu e_k^\mu$ is equal to 1. A moment's reflection shows that, given j, there are exactly

$$\sum_{m=0}^{n} 2^{p-n} \binom{n}{m} \delta(2m-n-j) \tag{3.13}$$

values of k for which v_k is equal to j. Thus, we have

$$\begin{aligned}
a' &= \|v\|^{-2} \langle v, \Omega_\beta(av)\rangle \\
&= \frac{1}{nd} \sum_{k=1}^{d} v_k \tanh(\beta a v_k) \\
&= \frac{1}{nd} 2^{p-n} \sum_{j=0}^{n} \sum_{m=0}^{n} \binom{n}{m} \delta(2m-n-j)\, j \tanh(\beta a j) \\
&= 2^{-n} \sum_{m=0}^{n} \binom{n}{m} \frac{2m-n}{n} \tanh[(2m-n)\beta a] \\
&= 2^{-n+1} \sum_{0 \leqslant m < n/2} \binom{n}{m} \frac{n-2m}{n} \tanh[(n-2m)\beta a] \quad \blacksquare
\end{aligned} \tag{3.14}$$

Proposition 3.4. For $\beta > 1$, the equation $\gamma_n(\beta a) = a$ has a unique positive solution $a = a_n(\beta)$. Moreover,

(a) $a_n(\beta)$ is an increasing function of β

(b) $a_n(\beta) \to 2^{-n+1}\binom{n-1}{\lfloor (n-1)/2 \rfloor}$ as $\beta \to \infty$
where $\lfloor r \rfloor$ denotes the integer part of r

(c) $a_n(\beta)^2 = (3/(3n-2))(\beta - 1) + \mathcal{O}((\beta-1)^2)$ as $\beta \downarrow 0$

The proof of these statements is straightforward, given the following properties of γ_n.

Proposition 3.5. The functions γ_n are odd and satisfy

(a) $\gamma_n(x) > 0, \quad \gamma_n'(x) > 0, \quad \gamma_n''(x) < 0$ for all $x > 0$

(b) $\gamma_n(x) \to 2^{-n+1}\binom{n-1}{\lfloor (n-1)/2 \rfloor}$ as $x \to \infty$

(c) $\gamma_n'(0) = 1, \quad \gamma_n'''(0) = -2(3n-2)$

Proof. The inequalities (a) are obtained from the corresponding inequalities for the function tanh. Property (b) follows from (3.11): If we define $\binom{n-1}{m} = 0$ for $m < 0$, then

$$\begin{aligned}
\lim_{x \to \infty} \gamma_n(x) &= 2^{-n+1} \sum_{0 \leqslant m < n/2} \binom{n}{m} \frac{n-2m}{n} \\
&= 2^{-n+1} \sum_{m < n/2} \left(\binom{n-1}{m} - \binom{n-1}{m-1} \right) \\
&= 2^{-n+1} \binom{n-1}{\lfloor (n-1)/2 \rfloor}
\end{aligned} \tag{3.15}$$

To prove (c), we use the representation $\gamma_n(x) = (1/nd)\langle v, \mathrm{Tanh}(xv)\rangle$, with $v = \sum_{\mu=1}^n e^\mu$. The mth derivative of γ_n at the origin is then given by the following equation:

$$\begin{aligned}\gamma_n^{(m)}(0) &= \frac{1}{nd}\tanh^{(m)}(0)\sum_{k=1}^{d} v_k^{m+1}\\ &= \tanh^{(m)}(0)\frac{1}{n}\sum_M\left[\frac{1}{d}\sum_{k=1}^{d} e_k^{\mu_1}e_k^{\mu_2}\cdots e_k^{\mu_{m+1}}\right]\end{aligned} \tag{3.16}$$

where $\sum_M$ denotes the sum over all ordered sets $M = (\mu_1, \mu_2,..., \mu_{m+1})$ with $1 \leqslant \mu_j \leqslant n$. Because of the orthogonality of the vectors (3.9), the expression $[\cdots]$ in (3.16) vanishes unless every element of M occurs an even number of times in M. If it does not vanish, then $[\cdots] = 1$. For $m = 1$ there are n sets left which contribute to $\sum_M$, and thus $\gamma_n'(0) = 1$. If $m = 3$, then there are $n(3n-2)$ such sets, and $\gamma_n'''(0) = -2(3n-2)$ follows since $\tanh'''(0) = -2$. ∎

In the remaining part of this section, we discuss the stability of the symmetric fixed points for β in the interval

$$1 < \beta < 1 + (9d + 500p^8)^{-1} \tag{3.17}$$

as well as for large values of β. In addition, we show that the symmetric fixed points are unique if β satisfies (3.17); the following estimate is the first step of the proof.

Proposition 3.6. For $0 < \beta < 1 + (1/9d)$, every nonzero fixed point z of Ω_β satisfies

$$\|z\|^2 < 23d(\beta - 1) \tag{3.18}$$

Proof. Assume that $\Omega_\beta(z) = z$, and let $y = \mathrm{Tanh}(\beta z)$. By using that $Py = z$ and that $z_k y_k = z_k \tanh(\beta z_k) \geqslant 0$ for all k, we obtain the following identities:

$$\begin{aligned}(\beta-1)\|z\|^2 &= (\beta-1)\|z\|^2 + \sum_{k=1}^{d} y_k[y_k - \tanh(\beta z_k)]\\ &= \|y-z\|^2 + \sum_{k=1}^{d} y_k[\beta z_k - \tanh(\beta z_k)]\\ &= \|y-z\|^2 + \sum_{k=1}^{d} |y_k|\,[\beta|z_k| - \tanh(\beta|z_k|)]\end{aligned} \tag{3.19}$$

Since none of the terms in the last sum is negative, it follows that either $z=0$ or $\beta>1$.

Let us now assume that $1<\beta<1+(1/9d)$. Then (3.19) implies that $\|y-z\|^2$ has to be small; in particular, it follows that

$$|\beta z_k| = \beta(|y_k| + |z_k - y_k|) \leqslant \beta[1 + (\beta-1)^{1/2}\|z\|]$$
$$< \left(1+\frac{1}{9d}\right)\left(1+\frac{1}{3\sqrt{d}}\cdot\sqrt{d}\right) < \sqrt{2} \tag{3.20}$$

The last term in (3.19) can now be bounded from below, by using (3.20) and the inequality $\tanh'''(x) \leqslant -2+8x^2$.

$$\sum_{k=1}^{d} |y_k| (\beta|z_k| - \tanh(\beta|z_k|))$$
$$\geqslant \sum_{k=1}^{d} |y_k| \left(\frac{1}{3}(\beta|z_k|)^3 - \frac{1}{15}(\beta|z_k|)^5\right)$$
$$\geqslant \sum_{k=1}^{d} |y_k| \frac{1}{15}(\beta|z_k|)^3$$
$$\geqslant \frac{\beta^3}{15}\left[\sum_{k=1}^{d} z_k^4 - \sum_{k=1}^{d} |y_k - z_k| \cdot |z_k|^3\right] \tag{3.21}$$

The two sums in the square bracket can be compared by using Eq. (3.19) again, together with the Schwartz inequality:

$$\sum_{k=1}^{d} |y_k - z_k| \cdot |z_k|^3$$
$$\leqslant \|y-z\| \cdot \|z\| \left(\sum_{k=1}^{d} z_k^4\right)^{1/2}$$
$$\leqslant (\beta-1)^{1/2}\|z\|^2 \left(\sum_{k=1}^{d} z_k^4\right)^{1/2} \leqslant \frac{1}{3}\sum_{k=1}^{3} z_k^4 \tag{3.22}$$

As a consequence of (3.19), (3.21), and (3.22), we have

$$(\beta-1)\|z\|^2 \geqslant \frac{\beta^3}{15}\frac{2}{3}\sum_{k=1}^{d} z_k^4 \geqslant \frac{2\beta^3}{45d}\|z\|^4 \tag{3.23}$$

and the bound (3.18) follows. ∎

After having localized the fixed points in a small ball around $z=0$, we can now use perturbation theory to rule out the existence of nonsymmetric fixed points of Ω_β for β in the interval (3.17).

Proof of Theorem 1.4. By Proposition 3.1, we are led to consider the system

$$\frac{1}{d}\left\langle e^{\nu}, \operatorname{Tanh}\left(\beta \sum_{\mu=1}^{p} b_{\mu} e^{\mu}\right)\right\rangle = b_{\nu}, \qquad 1 \leqslant \nu \leqslant p \tag{3.24}$$

which is equivalent to the fixed-point equation $\Omega_{\beta}(z) = z$ if we set $b_{\nu} = (1/d)\langle e^{\nu}, z\rangle$. Note that, since the map Tanh commutes with all transformations Ψ_{μ}, the inner product in (3.24) is an odd function of b_{ν} and even in all the other coefficients b_{μ}. Thus, we may define

$$g^{\nu}(b_1^2, b_2^2,..., b_p^2) \equiv \frac{1}{db_{\nu}} \sum_{k=1}^{d} e_k^{\nu} \tanh\left(\beta \sum_{\mu=1}^{p} b_{\mu} e_k^{\mu}\right), \qquad 1 \leqslant \nu \leqslant p \tag{3.25}$$

Furthermore, since the hyperbolic tangent is analytic on the disc $|\zeta| < \pi/2$, the functions $t \mapsto g^{\nu}(t_1,..., t_p)$ can be continued analytically to the polydisk

$$|t_{\mu}| < (\pi/2p)^2, \qquad 1 \leqslant \mu \leqslant p \tag{3.26}$$

The same holds for the functions g_5^{ν}, which we define as in (3.25), but with tanh replaced by $\tanh_5$, where

$$\tanh_5(\zeta) = \tanh(\zeta) - \zeta - \tfrac{1}{3}\zeta^3, \qquad \zeta \in \mathbb{R} \tag{3.27}$$

In addition, g_5^{ν} vanishes at the origin, together with all its first partial derivatives, for $1 \leqslant \nu \leqslant p$. The second derivatives can be bounded by using the maximum principle for analytic functions and Cauchy's formula with circular integration contours of radius $2p^{-2}$. Since $|\tanh_5(\zeta)| < 4$, for $|\zeta| < \pi/2$, we have

$$\left|\frac{\partial^2}{\partial t_{\lambda}\, \partial t_{\mu}} g_5^{\nu}(t_1,..., t_p)\right| < \frac{2!}{(2p^{-2})^2} \cdot \frac{4}{3/(2p)} < \frac{4}{3} p^5 \tag{3.28}$$

if $|t_{\mu}| \leqslant [1/(2p)]^2$ for all μ. This bound will be used below, together with Taylor's formula, in order to estimate g_5 near the origin.

The difference $(g^{\nu} - g_5^{\nu})$ can be computed explicitly by using the orthogonality of the vectors defined in (3.9); see also the discussion of (3.16). We obtain

$$\begin{aligned}(g^{\nu} - g_5^{\nu})(b_1^2,..., b_p^2) &= \beta + \frac{1}{db_{\nu}} \sum_{k=1}^{d} e_k^{\nu}\left(-\frac{1}{3}\right)\left(\beta \sum_{\mu=1}^{p} b_{\mu} e^{\mu}\right)^3 \\ &= \beta - \frac{\beta^3}{3b_{\nu}} \sum_{\mu,\kappa,\lambda=1}^{p} \left(b_{\mu} b_{\kappa} b_{\lambda} \frac{1}{d} \sum_{k=1}^{d} e_k^{\nu} e_k^{\mu} e_k^{\kappa} e_k^{\lambda}\right) \\ &= \beta - \beta^3 \sum_{\mu=1}^{p} b_{\mu}^2 + \frac{2}{3}\beta^3 b_{\nu}^2\end{aligned} \tag{3.29}$$

This allows us to rewrite Eq. (3.24) as follows:

$$b_\nu\left[(\beta-1)-\beta^3\sum_{\mu=1}^{p} b_\mu^2+\frac{2}{3}\beta^3 b_\nu^2+g_5^\nu(b_1^2,..., b_p^2)\right]=0, \qquad 1\leqslant\nu\leqslant p \tag{3.30}$$

Given a solution of these equations, let us define $V=\{\nu\colon 1\leqslant\nu\leqslant p,\ b_\nu\neq 0\}$ and $n=|V|$. By summing the expression $[\cdots]$ in (3.30) over all $\nu\in V$, it is possible to write the sum $b_1^2+\cdots+b_p^2$ in terms of β, n, and g_5. After substituting the result black into (3.30), we obtain the following equation for the variables $t_\nu=b_\nu^2$.

$$t_\nu=\frac{3}{(3n-2)\beta^3}(\beta-1)-\frac{3}{2\beta^3}g_5^\nu(t_1,..., t_p)+\frac{9}{(6n-4)\beta^3}\sum_{\mu\in V} g_5^\mu(t_1,..., t_p) \tag{3.31}$$

if $\nu\in V$, and $t_\nu=0$ otherwise. In order to establish the existence of a unique solution to (3.31), we will use the contraction mapping principle in a ball $\mathscr{B}(\rho)=\{t\in\mathbb{R}^V\colon |t_\nu-\theta_\nu|\leqslant\rho,\ \nu\in V\}$, centered at the vector θ,

$$\theta_\nu=\frac{3}{(3n-2)\beta^3}(\beta-1), \qquad \nu\in V \tag{3.32}$$

For ρ we choose the value $26(\beta-1)$, so that if $z=\sum_{\nu\in V} b_\nu e^\nu$ satisfies the bound (3.18), then the vector $t=(b_\nu^2)_{\nu\in V}$ lies in $\mathscr{B}(\rho)$. This guarantees that all (real) fixed points of Ω_β will be obtained. Furthermore, as is easy to check, the restriction (3.17) on β ensures that $|t_\nu|\leqslant 29(\beta-1)\leqslant[1/(2p)]^2$ for all $t\in\mathscr{B}(\rho)$, so that the bound (3.28) may be used.

Denote by $M_\nu(t)$ the right-hand side of (3.31). Since $|\theta_\nu|\leqslant 3(\beta-1)$, we obtain

$$\begin{aligned}|M_\nu(\theta)-\theta_\nu| &\leqslant\left(\frac{3}{2\beta^3}+\frac{9n}{(6n-4)\beta^3}\right)\max_\mu |g_5^\mu(\theta)|\\ &\leqslant 6\cdot\frac{4}{3}p^7\cdot\frac{1}{2}\theta_\nu^2\leqslant\frac{1}{40p}\theta_\nu\end{aligned} \tag{3.33}$$

if β satisfies (3.17). This shows that the transformation M defined by $(M(t))_\nu=M_\nu(t)$ maps the center of $\mathscr{B}(\rho)$ into $\mathscr{B}(\rho/2)$. Thus, in order for M to have a unique fixed point in $\mathscr{B}(\rho)$, it is sufficient that M contracts distances (in the norm $\max_\mu |t_\mu|$) by a factor of 2 or more. The following bound shows that this is indeed the case. By (3.31), (3.28), and (3.17), we have

$$\max_{\nu \in V} \sum_{\lambda \in V} \left| \frac{\partial}{\partial t_\lambda} M_\nu(t) \right| \leqslant n \max_{\lambda, \nu \in V} \left| \frac{\partial}{\partial t_\lambda} M_\nu(t) \right|$$

$$\leqslant p \left(\frac{3}{2\beta^3} + \frac{9n}{(6n-4)\beta^3} \left| \frac{\partial}{\partial t_\lambda} g_5^\mu(t) \right| \right.$$

$$\leqslant p \cdot 6 \cdot \frac{4}{3} p^6 \cdot 29(\beta - 1) < \frac{1}{2} \tag{3.34}$$

As a consequence, the only solution of Eq. (3.31) is the symmetric solution, i.e., $t_\nu = a_n(\beta)^2$ for all $\nu \in V$. The assertion now follows from Proposition 3.7 below. ∎

For the following discussion of stability, let $n \geqslant 2$, and assume that z is a symmetric fixed point of order n for Ω_β, i.e., that $z = a_n(\beta) v$, with v satisfying Eq. (3.10). Denote by P_1 and P_3 the orthogonal projections in $\mathbb{R}^d$ onto the subspaces span$\{v\}$ and $\{w \in P\mathbb{R}^d \colon \langle w, c_\mu e^\mu \rangle = 0,\ 1 \leqslant \mu \leqslant p\}$, respectively, and let $P_2 = P - P_1 - P_3$. By using Proposition 3.2, it can be shown that the linearization of Ω_β at the point z has the foloowing spectral representation:

$$D\Omega_\beta(z) = sP_3 + (s-r) P_2 + [s + (n-1) r] P_1 \tag{3.35}$$

where s and r are given by the equations

$$\begin{aligned} s &= \beta - \frac{\beta}{d} \sum_{k=1}^{d} \tanh^2(\beta z_k) \\ r &= -\frac{\beta}{d} c_\mu c_\nu \sum_{k=1}^{d} \tanh^2(\beta z_k) e_k^\mu e_k^\nu, \qquad \mu \neq \nu, \quad c_\mu c_\nu \neq 0 \end{aligned} \tag{3.36}$$

We note that the eigenvalue s of $D\Omega_\beta(z)$ is always bounded on one side by $(s-r)$ and on the other side by $[s+(n-1)r]$. Furthermore, as a consequence of Proposition 3.4, the eigenvalue $[s+(n-1)r]$ lies between 0 and 1 for all $\beta > 1$; it is the slope of the function $a \mapsto \gamma_n(\beta a)$ at the fixed point $a = a_n(\beta)$. Thus, the eigenvalue $(s-r)$ alone determines whether or not the fixed point z is stable.

Proposition 3.7. The symmetric fixed points of order $n = 1$ are stable for all $\beta > 1$. For β in the interval (3.17), all other fixed points of Ω_β are unstable.

Proof. Let $z = a_n(\beta) v$ be a symmetric fixed point of order n. If $n = 0$, then z is clearly unstable for all $\beta > 1$, since $z = 0$ and $D\Omega_\beta(0) = \beta P$. If

$n=1$, we obtain $D\Omega_\beta(z)=sP$, with s given by Eq. (3.36), and it is easy to check that $s<1$ for all $\beta>1$.

Consider now $n>2$, and β in the interval (3.17). By repeating the discussion of Eq. (3.31), but this time for the ball $\mathscr{B}((1/4p)\theta_\nu)$, we get the following bound on $a_n(\beta)$:

$$a_n(\beta)^2<\left(1+\frac{1}{4p}\right)\theta_\nu<4(\beta-1) \tag{3.37}$$

where θ_ν is given by (3.32). In particular, since $|\tanh'''(x)|\leqslant 2$ and $|v_k|\leqslant p$, we have

$$\tanh(|\beta z_k|)\leqslant(1+\tfrac{1}{3}|\beta z_k|^2)\,|\beta z_k|\leqslant\left(1+\frac{1}{4p}\right)|\beta z_k| \tag{3.38}$$

The eigenvalue $(s-r)$ of $D\Omega_\beta(z)$ can now be estimated as follows:

$$\begin{aligned}
s-r&=\beta-\sum_{k=1}^{d}\tanh^2(\beta z_k)(1-c_\mu c_\nu e_k^\mu e_k^\nu)\\
&\geqslant\beta-\frac{\beta}{d}\left(1+\frac{1}{4p}\right)^2\sum_{k=1}^{d}(\beta z_k)^2\,(1-c_\mu c_\nu e_k^\mu e_k^\nu)\\
&\geqslant\beta-\beta^3\left(1+\frac{1}{4p}\right)^3\theta_\nu\left[\frac{1}{d}\sum_{k=1}^{d}v_k^2(1-c_\mu c_\nu e_k^\mu e_k^\nu)\right]\\
&\geqslant\beta-\beta^3\left(1+\frac{1}{p}\right)\theta_\nu(n-2)=1+\frac{4-3(n-2)/p}{3n-2}(\beta-1)>1
\end{aligned} \tag{3.39}$$

This shows that the fixed point z is unstable. For details on how to evaluate the expression $[\cdots]$ in (3.39) we refer to the discussion following Eq. (3.16). ∎

Remark. For $\beta>1$, the symmetric fixed points of order $n=1$ are not only stable, but they also minimize f_β. This follows from Proposition 2.5.

Proposition 3.8. Let n be an even positive integer, and assume that β satisfies

$$\beta\cdot 2^{-n}\binom{n}{n/2}>1 \tag{3.40}$$

Then all symmetric fixed points of order n are unstable.

Proof. Let $z=a_n(\beta)v$ be a symmetric fixed point of even order $n>0$, and define $Z=\{k\colon 1\leqslant k\leqslant d,\ v_k=0\}$. It is easy to see that Z contains exactly $2^{p-n}\binom{n}{n/2}$ elements. Thus, if β satisfies (3.40), the eigenvalue s of $D\Omega_\beta(z)$ can be bounded as follows:

$$s=\frac{\beta}{d}\sum_{k=1}^{d}\left[1-\tanh^2(\beta z_k)\right]\geqslant\frac{\beta}{d}2^{p-n}\binom{n}{n/2}=\beta\cdot 2^{-n}\binom{n}{n/2}>1 \quad \blacksquare \tag{3.41}$$

Proposition 3.9. Let n be an odd integer larger than 1, and assume that β satisfies

$$\beta\cdot 2^{-n}\binom{n-1}{(n-1)/2}>\ln\beta>1 \tag{3.42}$$

Then all symmetric fixed points of order n are stable.

Proof. Let $z=a_n(\beta)v$ be a symmetric fixed point of order $n>1$. If we assume that n is odd, then $|v_k|\geqslant 1$ for all k, and the eigenvalue $(s-r)$ of $D\Omega_\beta(z)$ can be bounded as follows: If $\mu\neq\nu$ and $c_\mu c_\nu\neq 0$, then

$$\begin{aligned} s-r&=\frac{\beta}{d}\sum_{k=1}^{d}\{1-\tanh^2[\beta a_n(\beta)v_k]\}(1-c_\mu c_\nu e_k^\mu e_k^\nu)\\ &\leqslant\beta\{1-\tanh^2[\beta a_n(\beta)]\}<4\beta e^{-2\beta a_n(\beta)} \end{aligned} \tag{3.43}$$

In the first inequality we have used that $1-c_\mu c_\nu e_k^\mu e_k^\nu=2$ for half of the values of k, and $1-c_\mu c_\nu e_k^\mu e_k^\nu=0$ for the other half. Since $a_n(\beta)$ converges to a positive value as $\beta\to\infty$, it is clear that z is stable for large β.

In order to estimate $\beta a_n(\beta)$, we bound the factors $\tanh[(n-2m)x]$ in (3.11) from below by $\tanh(x)$ and then sum as in (3.15):

$$\begin{aligned} a_n(\beta)&=\gamma_n(\beta a_n(\beta))\\ &\geqslant\sum_{0\leqslant m\leqslant n/2}\binom{n}{m}\frac{n-2m}{n}\tanh[\beta a_n(\beta)]\\ &=2^{-n+1}\binom{n-1}{(n-1)/2}\tanh[\beta a_n(\beta)] \end{aligned} \tag{3.44}$$

By using (3.42) and the fact that $\tanh''(x)>-2x$, for $x>0$, we obtain

$$\beta a_n(\beta)>2\tanh[\beta a_n(\beta)]\geqslant 2\beta a_n(\beta)-\frac{2}{3}[\beta a_n(\beta)]^3 \tag{3.45}$$

which implies that $\beta a_n(\beta)>1$. This bound can be improved by applying (3.44) and (3.42) again: Since $\tanh(1)>1/2$, we have

$$\beta a_n(\beta)>\beta\cdot 2^{-n}\binom{n-1}{(n-1)/2}>\ln\beta \tag{3.46}$$

Substituting this into (3.43) and using the fact that $\beta > 4$ by assumption (3.42), we see that the eigenvalue $(s-r)$ is smaller than 1. This completes the proof of Proposition 3.9. ▌

ACKNOWLEDGMENTS

H. K. would like to thank Prof. K. Hepp for his hospitality at the ETH in Zürich, where part of this work has been carried out. The work of H. K. was supported by NSF grant DMS-8802540.

REFERENCES

1. W. A. Little, The existence of persistent states in the brain, *Math. Biosci* **19**:101 (1974).
2. J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA* **79**:2554 (1982).
3. J. P. Provost and G. Vallée, Ergodicity of the coupling constants and the symmetric *n*-replicas trick for a class of mean-field spin-glass models, *Phys. Rev. Lett.* **50**:598 (1983).
4. D. J. Amit, H. Gutfreund, and H. Sompolinsky, Spin-glass models of neural networks, *Phys. Rev. A* **32**:1007 (1985).
5. R. J. McEliece, E. C. Posner, E. R. Rodemich, and S. S. Vankatesh, The capacity of the Hopfield associative memory, *IEEE Trans. Information Theory* **33**:461 (1986).
6. D. J. Amit, The properties of models of simple neural networks, in *Heidelberg Colloquium on Glassy Dynamics (1986)*, J. L. van Hemmen and I. Morgenstern, eds. (Lecture Notes in Physics, No. **275**, 1987), p. 430.
7. D. J. Amit, H. Gutfreund, and H. Sompolinsky, Statistical mechanics of neural networks near saturation, *Ann. Phys.* **173**:30 (1987).
8. J. L. van Hemman, Spin-glass model of a neural network, *Phys. Rev. A* **34**:3435 (1986).
9. D. Grensing and R. Kühn, Random-site spin-glass models, *J. Phys. A* **19**:L1153 (1986).
10. D. Grensing, J. L. van Hemmen, A. Huber, and R. Kühn, Nonlinear neural networks: I. General theory, *J. Stat. Phys.* **50**:231 (1987); II. Information processing, *J. Stat. Phys.* **50**:259 (1987).
11. J. L. van Hemmen and V. A. Zagrebnov, Storing extensively many weighted patterns in a saturated neural network, *J. Phys. A* **20**:3989 (1987).
12. P. Baldi, Symmetries and learning in neural network models, *Phys. Rev. Lett.* **59**:1976 (1987).
13. C. M. Newman, Memory capacity in neural network models: Rigorous lower bounds, *Neural Networks* **1** (1988).
14. J. Komlós and R. Paturi, Convergence results in the Hopfield model, Preprint, UC San Diego (1987).
15. C. Peterson and J. R. Anderson, A mean field theory learning algorithm for neural networks, Preprint MCC-EI-259-87, MCC Austin (1987).