# Inhomogeneous Magnetization in Dilute Asymmetric and Symmetric Systems

Ido Kanter

*Joseph Henry Laboratories of Physics, Jadwin Hall, Princeton University, Princeton, New Jersey 08544*

(Received 24 September 1987)

A dilute nonsymmetric ferromagnet and a version of the Hopfield model are solved in the limit where the average number of inputs per spin is finite. The solution is based on the fact that the magnetization of each site is a function of the number and the type of inputs and their connection strengths. Extension of this solution to the symmetric case is discussed. The solution can also be extended to any other distribution of interactions.

PACS numbers: 05.50.+q, 75.10.Nr, 87.10.+e

Spin-glass models for associative memory have been of increasing interest in the last few years.[1-3] As first proposed by Little[1] and Hopfield,[2] these models are based on an Ising Hamiltonian and hence can be treated by equilibrium statistical mechanics.[3] Two assumptions are crucial to allow for an exact solution of the equilibrium properties of the model: The synaptic connections are assumed to be symmetric and each neuron is assumed to be connected to an infinite number of other neurons. However, in biological networks, the synapses are known to be asymmetric and on the average a neuron is connected only to a fraction $\rho \approx 10^{-8}$ of all neurons. Hence, it is important to study the effects of asymmetry and extreme dilution.

The dynamics of nonsymmetric models has recently been investigated in the limit of extreme dilution.[4-7] In such solutions the system is described only by the average magnetization.

In this Letter, I give a solution for the dynamics of a highly dilute nonsymmetric ferromagnet and a nonsymmetric Hopfield model. The solution is based on the fact that the magnetization is inhomogeneous and depends on the detailed geometry of the system. More precisely, the magnetization depends on the number and the type of inputs and on the strength of the connections between each site and its input sites. In particular in neural-network systems, spins with small magnetization indicate the location of errors in the pattern. Therefore, the knowledge of the connection between the structure of the system and the local magnetization *a priori* indicates the most probable location of errors. This theory can be extended to other highly dilute nonsymmetric systems, such as spin-glass models. Finally, I will briefly discuss the extension of this theory to the symmetric case.

Following Refs. 4-6, the model considered here is a system of $N$ Ising spins whose interactions $J_{ij}$ are given by[2]

$$J_{ij} = C_{ij} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \tag{1}$$

where $\xi_i^\mu$ is the value at site $i$ in the pattern $\mu$ and $p$ is the number of stored patterns. The $\xi_i^\mu$ is a quenched random number having the values $\pm 1$ with equal probability. For the ferromagnetic case, $p=1$. The $C_{ij}$ are random independent parameters which represent the dilution and the asymmetry. For each pair $(i,j)$, a value for $C_{ij}$ is chosen at random according to the distribution

$$P(C_{ij}) = (c/N)\delta(C_{ij}-1) + (1-c/N)\delta(C_{ij}). \tag{2}$$

The interactions $J_{ij}$ are not symmetric because for each pair $(i,j)$, $C_{ij}$ and $C_{ji}$ are independent random variables. In the highly dilute nonsymmetric model, $c$ is of the order of unity. For this model the following two dynamics can be considered.[4]

In parallel dynamics, at time $t$ all spins are updated simultaneously in the following way: On each site $i$ which gets input from $l$ other spins and the strength of the connections are $\{J_{ir}\}_{r=1}^{l}$, the local field $h_i(l;J_{i1}, J_{i2}, \ldots, J_{il})$ is computed

$$h_i(l;J_{i1},J_{i2},\ldots,J_{il}) = \sum_{j=1}^{l} J_{ij} S_j(t), \tag{3}$$

and the spins are updated at time $t+\Delta t$ with probability $1+\exp[-2h_i^l(t)/T]^{-1}$ to be $+1$ and with probability $1+\exp[2h_i^l(t)/T]^{-1}$ to be $-1$. For parallel dynamics, the natural time scale is $\Delta t = 1$.

In random sequential dynamics,[4] at time $t$ one chooses at random a site $i$ among the $N$ sites and updates this site at time $t+\Delta t$ by use of the same rules as in the parallel dynamics. Since at each time step only one spin is updated, one should scale the time with the system size $\Delta t = 1/N$.

In this Letter I obtain exact results for the dynamical properties of this model in the thermodynamic limit $(N \to \infty)$.

The evolution of the configuration $\{S_i(t)\}$ having a macroscopic overlap on one stored pattern and a microscopic overlap on the other $p-1$ random patterns, is described by the average magnetization at time $t$

$$m(t;K;J_1,\ldots,J_K) = [\bar{N}_K P(J_1,\ldots,J_K)]^{-1} \sum_{i=1}^{\bar{N}_K P(J_1,\ldots,J_K)} \xi_i^1 S_i(t), \tag{4}$$

where $P(J_1,\ldots,J_K)$ is the probability that $K$ input connections have the strength $J_1,J_2,\ldots,J_K$. The summation in Eq. (4) is only for sites which receive an input from $K$ other sites and the connection strengths are $J_1,\ldots,J_K$. The density of sites[4-6] which received any input from $K$ other sites is given by

$$N_K \equiv \bar{N}_K/N = e^{-c}c^K/K!. \tag{5}$$

The origin for the inhomogeneous magnetization in the system is the dependence of the local field [Eq. (3)] on the number of inputs, the magnetization of the inputs, and the strength of the connections.

The dynamical evolution of $m(t;K;J_1,\ldots,J_K)$ is given for the parallel dynamics by

$$m(t+1;K;J_1,\ldots,J_K) = f_K\{m(t;l;J_1\ldots J_l)\}, \tag{6}$$

and for random sequential updating by

$$dm(t;K;J_1,\ldots,J_K)/dt = f_K\{m(t;l;J_1,\ldots,J_l)\} - m(t;K;J_1,\ldots,J_K), \tag{7}$$

where $f_K\{m(l;J_1,\ldots,J_l)\}$ is given by

$$f_K\{m_l\{J\}\} = \frac{1}{2^K}\operatorname*{Tr}_{\{K_\rho\}}\sum_{l=0}^{K}\left[\binom{K}{l}\prod_{\rho=1}^{K-l}N_{K_\rho}P_{K_\rho}\{J\}(1+m_{K_\rho}\{J\})\right]\left(\prod_{\rho=l+1}^{K}N_{K_\rho}P_{K_\rho}\{J\}(1-m_{K_\rho}\{J\})\right)$$

$$\times\sum_{n=0}^{K(p-1)}\left\langle\left\langle\delta\left(\sum_{\rho=1}^{K}\sum_{\mu=2}^{p}\xi_i^\mu\xi_\rho^\mu - K(p-1)+2n\right)\prod_{\rho=1}^{K}\delta\left(\sum_\mu \xi_i^\mu\xi_\rho^\mu - J_\rho\right)\right\rangle\right\rangle\tanh\left(\frac{Kp-2l-2n}{T}\right), \tag{8}$$

where $P_{K_\rho}\{J\}\equiv P(J_1,\ldots,J_{K_\rho})$, $m_{K_\rho}\{J\}\equiv m(K_\rho;J_1,\ldots,J_{K_\rho})$, $K_\rho=0,1,2,\ldots$ with equal probability and the notation $\langle\langle\cdots\rangle\rangle$ stands for the average over the distribution of $\{\xi_i^\mu\}$.

The derivation of Eq. (8) is very similar to what has been done in Ref. 4. Briefly, it is derived as follows. First, as long as $c$ is small

$$c \ll \log N; \tag{9}$$

the magnetization in the quenched and annealed models are both given by expression (8).[8-10] This result is due to the fact that the inputs of almost all sites are not correlated at any finite time. Therefore, under the restriction Eq. (9) the calculation can be carried out for the annealed model.

The term $N_K = e^{-c}c^K/K!$ in Eq. (8) is the density of sites having $K$ nonzero inputs. The probability that $K-l$ inputs are parallel to the pattern and $l$ are antiparallel to the pattern, and each input site receives by itself $K_\rho$ inputs whose connection strengths $J_1,\ldots,J_{K_\rho}$ are given by the first two square brackets in Eq. (8). The last term is the probability contribution of the remaining $p-1$ patterns to the local field to be parallel to the first pattern. This calculation must be done under the constraints $\pi\delta(\sum\xi_i^\mu\xi_\rho^\mu - J_\rho)$. The local field consists of two parts; the contribution $k-2l$ from the macroscopic overlap and the contribution $k(p-1)-2n$ from the remaining $p-1$ patterns.

When one averages over all the shapes of highly dilute random graphs, all sites which receive inputs from any $K$ input sites with the connection strengths $J_1,\ldots,J_K$ have the same average $m(K;J_1,\ldots,J_K)$. Therefore, one finds

$$m(t+\Delta t;K;J_1,\ldots,J_K) = f\{t;m;J_1,\ldots,J_K\},$$

where $m$ is the average magnetization.[4] It is important to note that $m(K=0)$ is equal to zero by definition.

I now concentrate my discussion primarily on the zero-temperature limit.

In the *ferromagnetic case* all the remaining bonds are equal to $J$. Therefore, one can obtain from Eq. (8) the following results:

$$m(2K) = K\binom{2K}{K}\int_{(1-m)/2}^{(1+m)/2}t^{K+1}(1-t)^K dt, \tag{10}$$

$$m(2K+1) = (2K+1)\binom{2K}{K}\int_{(1-m)/2}^{(1+m)/2}[t(1-t)]^K dt, \tag{11}$$

where $m(K)$ is the average magnetization of sites which received any type of $K$ inputs and $m$ is the average magnetization $m\equiv\sum_K N_K m(K)$. The magnetization is inhomogeneous in the system; for example, $m(1)=m$, $m(2)=m$, and $m(3)=(m/2)(3-m^2)$. The result $m(1)=m$ is trivial because the spin receives only one input and therefore $m(1)=\sum_K N_K m(K)\equiv m$. It is clear that the $m(K)$ depends only on odd powers of $m$ because of the inversion symmetry in the system. From Eqs. (10) and (11), one obtains

$$m(2K+1)-m(2K) = mK\binom{2K}{K}\int_0^{(1-m)/2}[t(1-t)]^{K-1}(1-2t)dt. \tag{12}$$

These results indicate that for $K \geq 2$, the magnetization $m(K)$ is an increasing function of $K$, and $m(K \to \infty) \to 1$. This result is due to the fact that in the ferromagnetic case, the magnetization of each spin is always greater than or equal to zero. It is also important to note that $m(K)$ is a function of $c$ only through the average magnetization $m$.

As a function of $c$, the system undergoes a second-order phase transition. By expanding Eq. (8) for small $m(K)$, one obtains that the equation for the critical $c$ is given by $1 = ce^{-c}[I_0(c) + I_1(c)]$, which indicates that the transition occurs at $c \simeq 1.85$. This result was confirmed in simulations on systems containing from 1000 up to 5000 spins. It is important to note that $m(K) \propto m$ for all $K$, which implies that there is only one transition for all the $m(K)$'s.

In the *Hopfield case,* the $J_{ij}$ are given by Eq. (1) where $p$ is the number of stored patterns. Using Eq. (8), one finds for odd $p$,

$$m(K) = \sum_{l=0}^{K} \binom{K}{l} \left(\frac{1+m}{2}\right)^{K-l} \left(\frac{1-m}{2}\right)^{l} \langle\langle \operatorname{sgn}(Kp - 2l - 2x) P[K(p-1) - 2x, K(p-1)] \rangle\rangle_x, \tag{13}$$

where $P(x,y)$ is the probability of the equal-step random walk being at a distance $x$ after $y$ steps. The symbol $\langle\langle \cdots \rangle\rangle_x$ indicates the average over $x$. From Eq. (13), one should be able to find the $m(K)$ for any $K$, for example,

$$m(1) = mP(0, p-1), \quad m(2) = m[P(0, 2(p-1)) + P(2, 2(p-1))],$$

and

$$m(3) = (m/2)(3 - m^2)P(0, 3(p-1)) + m(3/4)P(2, 3(p-1)).$$

Here again, $m(K)$ is a function of $c$ only through the average magnetization $m$. The average $m(K)$ is a non-decreasing function of $K$, and $m(K \to \infty) \to 1$. This result is due to the fact that each input spin makes on the average a positive contribution to the local field being parallel to the condensed pattern.

In contrast to the ferromagnetic case, in the Hopfield case there are spins with negative average magnetization. This is due to the fact that the magnetization is also a function of the connection strengths [see Eq. (8)]. For example, the fraction $[1 - P(0, p-1)]/2$ of $m(1)$ has a negative magnetization $-m$. It is clear from Eq. (13) that the fraction of spins which receive $K$ inputs and have a negative magnetization is a decreasing function of $K$ and decreases to zero as $K \to \infty$. In principle, from Eq. (8) one can calculate $m(K, J_1, \ldots, J_K)$ for any $K$. More detailed results on the Hopfield model will be given elsewhere.

One can find the distribution of higher moments following the procedure used to find the distribution of the magnetization Eq. (8) (the calculations of higher moments in the homogeneous case are presented in Ref. 7). For example, the compatible Edwards-Anderson[11] order parameter for any distribution of bonds is given by

$$q_{\mathrm{EA}}(K; J_1, \ldots, J_K) = \frac{1}{2^{2K}} \operatorname*{Tr}_{\{S_l^1, S_l^2, K_\rho\}} \left( \prod_{l=1}^{K} N_{K_\rho}[1 + m(K_\rho) S_l^1] \right) \times \left( \prod_{l=1}^{K} N_{K_\rho}[1 + m(K_\rho) S_l^2] \right) \operatorname{sgn}\left[ \left( \sum_{j=1}^{K} J_j S_j^1 \right) \left( \sum_{j=1}^{K} J_j S_j^2 \right) \right], \tag{14}$$

where $K_\rho = 0, 1, \ldots$ with equal probability. For the ferromagnetic case, one can show, for example, $q_{\mathrm{EA}}(1) = \sum_K N_K m^2(K) \equiv \overline{m^2(K)}$ and $q_{\mathrm{EA}}(2) = \overline{m(K)}^2$. It is important to note that

$$\sum_K N_K q_{\mathrm{EA}}(K) = \sum_K [\langle S(K; J_1, \ldots, J_K) \rangle^2] \neq N^{-1} \sum_i m_i^2.$$

The symbol $\langle \cdots \rangle$ indicates the trace over the input spins, and the symbol $[\cdots]$ indicates an average over the distribution of the connection strengths.

One can easily generalize this theory and find the magnetization

$$m(K; J_1, \ldots, J_K, m(l_1, J_1^1, \ldots, J_{l_1}^1), \ldots, m(l_K; J_1^K, \ldots, J_{l_K}^K)):$$

The magnetization of spins which receive $K$ inputs whose connection strengths are equal to $\{J_i\}$, respectively, and the input spins themselves received $\{l_i\}$ inputs and their connection strengths are $\{l_i; J_1^i, \ldots, J_{l_i}^i\}$. These kinds of magnetizations can easily be found from Eq. (8). One can generalize this picture further and claim that the distribution of the magnetization depends not only on the nature of the input sites, but also on the nature of input to the input sites, and so on. As long as the input depends only on a finite number (compared to $\log N$) of "levels" of the input tree, all the spins which fixed each site are uncorrelated. One can find the distribution of the magnetization in such a case, in a way similar to the derivation of Eq. (8). Such a generalization is required to calculate the freezing in the system.

In the *symmetric case* the distribution of the bonds is again given by Eqs. (1) and (2), but with the constraint that $C_{ij}=C_{ji}$. The full description of the ordered phase requires knowledge of the entire distribution of the local magnetization $m_i$. Here again, I assume that the magnetization of each spin is a function of the type and the number of the coupled spins and their connection strengths. This is in contrast to the infinite-range case in which the local field is a Gaussian variable.

In general, each spin can be coupled to $K$ other spins $S_1,\ldots,S_K$ with a probability $e^{-c}c^K/K!$. Because each of the $K$ spins is distant from the others with probability 1, one can show that the $K$ coupled spins are uncorrelated (for more details, see Mezard and Parisi[12]). Within the hypothesis that there is only one pure state, self-consistent equations can be written for the magnetizations $m(K)$

$$m(K)=\left\langle \tanh\left[\sum_{l=1}^{K}\tanh^{-1}[\tanh(\beta J_l)m(K_l)]\right]\right\rangle, \qquad K=0,1,\ldots, \tag{15}$$

where $K$ is the number of coupled spins and the symbol $\langle\cdots\rangle$ indicates the average over the Poisson distribution of the $\{m(K_l)\}$ and over the distribution of the bonds $\{J_l\}$.

In the ferromagnetic case, all the nonzero bonds are equal to 1. At zero temperature, one can verify from Eq. (15) that the system undergoes a second-order phase transition at $c=1$. Solving Eq. (15) numerically at $c=1.5$, for example, one finds that the $m(K)$ are given by 0.59, 0.84, 0.93, 0.98, ... for $K=1,2,\ldots$, respectively. One can prove that $m(K)$ is an increasing function of $K$, and $m(K\to\infty)\to 1$. The mass of freezing spins is given by $m\equiv\sum e^{-c}c^K/K!m(K)\simeq 0.59$, which is equal to the solution of the homogeneous case.[12,13] The magnetization is inhomogeneous even at zero temperature due to the fact that the distribution of spins which couple to $K$ other spins is *different* for the infinite cluster and for the finite clusters. At finite temperature, $T\simeq 0.59$ for example, the $m(K)$ are given by 0.12, 0.246, 0.357 for $K=1,2,3$, respectively. The system undergoes a second-order phase transition at $T_c\simeq 0.61$. From these results and from Eq. (15) one can verify that near the transition temperature $m(K)\simeq K\tanh(\beta_c m)$, where $m$ is the average magnetization. Here again, one can extend this picture to a theory in which the magnetization is a function of the type of the trees coupled to each spin in the system.

In summary, I have shown that the magnetization of each spin is a function of the type and the number of inputs and the connection strengths. As a first approximation all spins with the same local structure have the same average magnetization. I believe that this idea has an important application in neural networks. The geometric structure *a priori* indicates that the most probable location of errors are spins with low connectivity. Therefore, it seems plausible to embed the most important information of each pattern (classification, category, etc.) in spins with high connectivity.

This theory also indicates that it is preferable to use an inhomogeneous temperature in the system, and also to update the spins in order which depends on the local geometry structure. For example, when the system has a macroscopic overlap with one of the pattern, the compatible temperature for each spin may decrease as a function of the connectivity. This procedure prevents a disorder in spins with large magnetization. For the same reason it is plausible to update first spins with low connectivity.

I would like to thank Professor D. J. Amit, Professor D. S. Fisher, and Professor H. Sompolinsky for helpful conversations. This work supported by Weizmann Fellowship and also supported in part by the National Science Foundation Grant No. DMR 8518163.

[1]W. A. Little, Math. Biosci. **19**, 101 (1974).

[2]J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982), and **81**, 3088 (1984).

[3]D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985).

[4]B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).

[5]B. Derrida, J. Phys. A **20**, L721 (1987).

[6]R. Kree and A. Zippelius, to be published.

[7]E. Barkai, I. Kanter, and H. Sompolinsky, to be published.

[8]B. Derrida and G. Weisbuch, J. Phys. (Paris) **47**, 1297 (1986).

[9]B. Derrida and Y. Pomeau, Europhys. Lett. **1**, 45 (1986).

[10]S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969), and Physica (Amsterdam) **10D**, 145 (1984).

[11]S. Edwards and P. W. Anderson, J. Phys. F **5**, 965 (1975).

[12]M. Mezard and G. Parisi, Europhys. Lett. **3**, 1067 (1987).

[13]I. Kanter and H. Sompolinsky, Phys. Rev. Lett. **58**, 164 (1987).