Fixed points in a Hopfield model with random asymmetric interactions

Manoranjan P. Singh
Laser Programme, Centre for Advanced Technology, Indore 452013, India

Zhang Chengxiang* and Chandan Dasgupta†
Department of Physics, Indian Institute of Science, Bangalore 560012, India
(Received 3 October 1994)

We calculate analytically the average number of fixed points in the Hopfield model of associative memory when a random antisymmetric part is added to the otherwise symmetric synaptic matrix. Addition of the antisymmetric part causes an exponential decrease in the total number of fixed points. If the relative strength of the antisymmetric component is small, then its presence does not cause any substantial degradation of the quality of retrieval when the memory loading level is low. We also present results of numerical simulations which provide qualitative (as well as quantitative for some aspects) confirmation of the predictions of the analytic study. Our numerical results suggest that the analytic calculation of the average number of fixed points yields the correct value for the typical number of fixed points.

PACS number(s): 87.10.+e, 64.60.Cn, 75.50.Lk, 05.50.+q

I. INTRODUCTION

In the Hopfield model [1,2] of associative memory, a chosen set of random binary patterns ("memories") are stored by using the modified Hebb rule which leads to a symmetric synaptic connection matrix. The symmetry of the synaptic connection matrix has been crucial in the analysis of the behavior of this model by using the techniques of equilibrium statistical mechanics [3,4]. There are, however, many reasons for considering the behavior of Hopfield-like models in which the synaptic matrix is not symmetric. The constraint of symmetry on the synaptic matrix is known to be biologically unrealistic. The synaptic connection between two neurons in biological networks is often found [5] to be only one-way. It is also believed [5] that biological neural networks generally obey Dale's law according to which a neuron can have only one kind of outgoing synapses, either excitatory or inhibitory. Clearly, a network obeying this law cannot have a symmetric synaptic connection matrix—neurons must receive synaptic inputs of both positive and negative sign and send output of single sign.

There are reasons to believe that the presence of asymmetry in the synaptic matrix may improve the performance of the network as an associative memory. It is well known [4] that the Hopfield model possesses a large number of spurious fixed-point attractors which are locally stable under the usual single spin-flip dynamics. Gardner [6] has shown that in the limit of large number of neurons, the spurious fixed points form a much broader band compared to the narrow band of retrieval fixed points found below the critical level of memory loading. The band of spurious attractors is centered around those with no macroscopic overlap with any stored pattern and extends up to attractors having finite correlations with one or more memory states. The presence of these spurious attractors adversely affects the performance of the network as an associative memory. Initial states which are not very close to stored memories are "captured" by one of these spurious attractors, thereby reducing the basins of attraction of the stored memory states. In order to alleviate this problem, one may use stochastic dynamics [4] which corresponds to a finite temperature. As the temperature is increased, more and more spurious attractors are destabilized [3,4]. However, the retrieval quality also decreases monotonically with increase in temperature. An alternative, which has been proposed by many researchers following an early suggestion of Hopfield [1], involves the use of asymmetry in the synaptic connections. Hertz, Grinstein, and Solla [7] studied the dynamics of an analog version of a Hopfield network with random asymmetry. On the basis of an analytic calculation which is strictly valid in the $n \to \infty$ limit of an n-compound spin model, they argued that the presence of any amount of asymmetry in the synaptic connections destabilizes all "spin-glass-like" spurious attractors. Similar conclusions have been obtained by Crisanti and Sompolinsky [8] from an approximate analytic treatment (using the spherical approximation) of a Hopfield model in which a random antisymmetric part is added to the symmetric Hebbian synaptic matrix. On the other hand, these and other [9–11] studies show that retrieval states, which are highly correlated with the stored memories, remain stable in the presence of moderate amounts of asymmetry. Similar conclusions about a de-

*Permanent address: Physics Department, Jilin University, Changchun 130023, People's Republic of China.

†Also at Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560012, India.

struction of the spin-glass state in asymmetric versions of the Sherrington-Kirkpatrick model [14] of Ising spin glass have been obtained by several authors [12,13] from approximate analytic treatments and numerical simulations. These results suggest that the presence of asymmetry in the synaptic connections may improve the retrieval performance of Hopfield-like models of associative memory by eliminating some of the spurious attractors.

Parisi [15] has proposed another positive role of asymmetry in the context of neural network modeling. In the Hopfield model, it is impossible for the network to discriminate between the two cases (a) retrieval of one of the stored patterns and (b) "confusion" as represented by convergence to a spurious fixed-point attractor, because in both cases the network settles down in a time-independent state. Parisi argues that asymmetry may be functional in converting spin-glass-like fixed points into chaotic trajectories. It would then be possible to discriminate between the two cases through the temporal behavior of the network. Such discrimination would be necessary for the prevention of Hebbian learning of spurious patterns.

In this paper, we consider a Hopfield model with a random antisymmetric part added to the otherwise symmetric synaptic connection matrix. This model may have some biological relevance in the *tabula non rasa* scenario proposed by Toulouse, Dehaene, and Changeux [16] according to which memories are iteratively added by a Hebbian learning process in a network which starts out with random asymmetric synaptic connections. The resulting synaptic matrix would have a symmetric Hebbian component and a random asymmetric one which may be divided into a symmetric and an antisymmetric part. It is well known [17] that the addition of a random symmetric part to the usual Hebbian synaptic matrix does not produce any qualitative change in the behavior of the network. One may then ignore the symmetric part of the random component and consider the effects of the random antisymmetric part on the performance of the network. The model we consider is essentially the same as one of the models studied in Ref. [8] by approximate analytic methods.

Analytic studies of the behavior of networks with asymmetric connections are difficult because methods of equilibrium statistical mechanics cannot be used in such studies. As a first step towards the development of an understanding of the behavior of such networks, we have carried out an analytic calculation of the average number of locally stable fixed points (i.e., states which are stable to all single spin flips) of the model described above. This calculation uses a method which was originally developed [18–21] for counting the number of metastable states in the Sherrington-Kirkpatrick model of Ising spin glass. This method was later extended to calculations of the average number of metastable states of the Hopfield model [6], the Sherrington-Kirkpatrick model with random asymmetry [12,13], and an asymmetrically diluted version of the Hopfield model [22]. The main results obtained from our study are summarized below. We find that the addition of the antisymmetric part to the synaptic matrix leads to a computational advantage over the original Hopfield model in that the asymmetric network has exponentially fewer spurious fixed points. In particular, we show that the expected number of fixed points in the asymmetric neural network behaves as $e^{NF(\alpha,k)}$ where N is the number of neurons, α is the memory loading level (defined as $\alpha \equiv p/N$, where p is the number of random uncorrelated memories), and the asymmetry parameter k measures the strength of the antisymmetric component of the synaptic matrix relative to the symmetric one. For a given value of α, $F(\alpha,k)$ decreases monotonically with k, going to zero at a "critical" value $k_1(\alpha)$ which increases with α. A calculation of the dependence of the average number of fixed points on the fractional Hamming distance measured from a particular memory state shows, in accordance with previous results [6,22], that the fixed points form two distinct "bands" for small values of k: a narrow "retrieval" band of fixed points which are strongly correlated with a memory state, and a wide "spurious" band which is centered around states having no macroscopic overlap with the chosen memory state. As the value of k is increased from zero, the height of the retrieval band increases initially, indicating that the introduction of a small amount of asymmetry actually leads to an exponential *increase* in the number of fixed points correlated with a stored memory. Further increase in the asymmetry parameter leads to a disappearance of the retrieval band. This happens in two different ways, depending on the value of α. For values of α smaller than $\alpha_0 \approx 0.0776$, the value of F at the peak of the retrieval band decreases after an initial increase as k is increased from zero, and crosses zero at a second "critical" value, $k_2(\alpha) < k_1(\alpha)$. Thus, for $\alpha < \alpha_0$ and $k > k_2(\alpha)$, retrieval fixed points do not exist in the thermodynamic limit. The value of $k_2(\alpha)$ increases with α. For $\alpha > \alpha_0$, the value of F at the peak of the retrieval band remains positive and the disappearance of the retrieval band occurs through its merger with the spurious one at $k = k_2'(\alpha)$, which decreases with increasing α. In all cases, the center of the retrieval band shifts towards higher values of the fractional Hamming distance as k is increased from zero, indicating that the quality of retrieval deteriorates under the introduction of asymmetry. For small values of α, the deterioration in the quality of retrieval is minimal for moderate values of k. The degradation of the quality of retrieval increases with α for a fixed value of k, and becomes substantial as α approaches the saturation value for the symmetric network, $\alpha_c \approx 0.14$. We also find that most of the fixed points which are destabilized by the introduction of asymmetry have relatively high values of an "energy" function which may be defined in the usual way [4] for the symmetric part of the interaction matrix.

We have also carried out numerical simulations of the number and properties of the fixed points of this model. Since the number of fixed points increases exponentially with the number of neurons (N) present in the network, a complete enumeration of all the fixed points can be carried out only for small values of N. For this reason, the simulations described here are limited to networks with $N \leq 80$. Despite the small size of the simulated networks, we find good qualitative agreement between the numerical results and the predictions of the analytic calculation.

For some of the quantities of interest [e.g., the quantity $F(\alpha,k)$ defined above], it is possible to extrapolate the results obtained for several small values of N to the $N\to\infty$ limit. The results obtained from such extrapolations are found to be in quantitative agreement with those obtained from the analytic calculation.

The remaining part of this paper is organized as follows. Section II begins with a description of the model we consider and contains an account of our analytic calculations on the number of fixed points of this model. The results obtained from these calculations are described in detail in this section. In Sec. III, the results of our numerical simulations are described in detail and compared with the predictions of the analytic calculations. Section IV contains a summary of the main results obtained from this study, a comparison of these results with existing ones obtained for similar models, and a discussion of the implications of these results for the retrieval properties of asymmetric neutral network models of associative memory. Technical details of the analytic calculation are presented in the Appendix.

II. ANALYTIC CALCULATION OF THE NUMBER OF FIXED POINTS

We consider a network of N two-state neurons. The state of the ith neuron at discrete time t is described by $\sigma_i(t)$, where σ_i is an Ising variable ($\sigma_i=\pm 1$). We choose a synaptic connection matrix J_{ij} of the following form:

$$J_{ij}=J_{ij}^{\mathrm{s}}+kJ_{ij}^{\mathrm{as}} , \tag{1}$$

where J_{ij}^{as} is an antisymmetric matrix with independent Gaussian elements, i.e.,

$$P(J_{ij}^{\mathrm{as}})=\frac{\sqrt{N}}{\sqrt{2\pi}}\exp[-N(J_{ij}^{\mathrm{as}})^2/2], \quad i<j , \tag{2}$$

and $J_{ji}^{\mathrm{as}}=-J_{ij}^{\mathrm{as}}$. The symmetric part is given by the usual Hebb rule [1]:

$$J_{ij}^{\mathrm{s}}=\frac{1}{N}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu, \quad i\neq j, \quad J_{ii}^{\mathrm{s}}=0 , \tag{3}$$

where the $\{\xi_i^\mu\}, i=1,\dots,N;\mu=1,\dots,p$ are the patterns (memories) stored in the network. Each ξ_i^μ is an independent random variable which takes the values ± 1 with equal probability. The memory loading level $\alpha\equiv p/N$ is the ratio of the number of stored patterns to the number of neurons.

The network evolves according to the iterated map

$$\sigma_i(t+1)=\mathrm{sgn}\left[\sum_{j=1}^{N}J_{ij}\sigma_j(t)\right] , \tag{4}$$

with either parallel or sequential updating. Fixed points of this dynamics correspond to the configurations $\{\sigma_i\}$ which satisfy

$$\sigma_i=\mathrm{sgn}\left[\sum_{j=1}^{N}J_{ij}\sigma_j\right], \quad i=1,\dots,N . \tag{5}$$

Equation (5) can be written as

$$\lambda_i=\sigma_i\sum_j J_{ij}\sigma_j, \quad \lambda_i>0, \quad i=1,\dots,N , \tag{6}$$

so that the average number of fixed points is given by

$$\langle N_{fp}(N,\alpha,k)\rangle=\int_0^\infty\prod_{i=1}^{N}d\lambda_i\mathrm{Tr}_{\{\sigma_i\}} \times\left\langle\prod_{i=1}^{N}\delta\left[\lambda_i-\sigma_i\sum_j J_{ij}\sigma_j\right]\right\rangle . \tag{7}$$

In Eq. (7), the brackets $\langle\ \rangle$ denote an average over the Np random variables $\{\xi_i^\mu\}$ and $N(N-1)/2$ random variables $\{J_{ij}^{\mathrm{as}}, i<j\}$. We note that $\ln(N_{fp})$ (rather than N_{fp} itself) is an extensive quantity. Therefore, we should, in principle [23], calculate the disorder average of $\ln(N_{fp})$, rather than that of N_{fp}. A calculation of $\langle\ln N_{fp}(N,\alpha,k)\rangle$ would involve the use of the replica method [20,21], which would introduce many complications. For this reason, we instead calculate $\langle N_{fp}(N,\alpha,k)\rangle$ which gives an upper bound for the expected number of fixed points [6]. Previous studies on similar models [13,20,21] suggest that the quantity N_{fp} is self-averaging in the thermodynamic limit, so that the two averages are expected to be essentially the same. This point is discussed in more detail in Sec. IV below.

Using a straightforward generalization of the method of Ref. [6], we find (see the Appendix for details) that in the $N\to\infty$ limit, the average number of fixed points is given by

$$\langle N_{fp}(N,\alpha,k)\rangle\approx e^{NF(\alpha,k)} , \tag{8}$$

where $F(\alpha,k)$ is given by the saddle point (over the three parameters a, b, and x) of the following function:

$$F(\alpha,k,a,b,x)=-\tfrac{1}{2}x^2+\alpha\left[b-\frac{1}{2}+\frac{(1-b)^2}{2a}+\frac{1}{2}\ln a\right] +\ln 2\phi(t) , \tag{9}$$

where

$$t=\frac{\alpha b-ikx}{\sqrt{a\alpha+k^2}} \tag{10}$$

and

$$\phi(t)\equiv\int_t^\infty\frac{d\lambda}{\sqrt{2\pi}}e^{-\lambda^2/2} . \tag{11}$$

The saddle-point conditions yield the following equations:

$$\frac{\partial F}{\partial x}=-x-\frac{ik}{\sqrt{a\alpha+k^2}}\frac{\phi'(t)}{\phi(t)}=0 , \tag{12}$$

$$\frac{\partial F}{\partial b}=1-\frac{1-b}{a}+\frac{1}{\sqrt{a\alpha+k^2}}\frac{\phi'(t)}{\phi(t)}=0 , \tag{13}$$

$$\frac{\partial F}{\partial a}=\frac{1}{a}\left[1-\frac{(1-b)^2}{a}\right]-\frac{t}{(a\alpha+k^2)}\frac{\phi'(t)}{\phi(t)}=0 . \tag{14}$$

In these equations, ϕ' represents a derivative of ϕ with respect to its argument. Using Eqs. (12) and (13), we obtain

$$x = ik\left[1-\frac{(1-b)}{a}\right] . \tag{15}$$

With this value for the parameter x, Eqs. (13) and (14) can be solved numerically for the parameters a and b to obtain $F(\alpha,k)$. We recover existing results on related models in several special limits of the variables k and α. For $k=0$ (no asymmetry), our results for $F(\alpha,k)$ reduce to those obtained by Gardner [6] for the original Hopfield model. In the limit $\alpha\to\infty$, $k\to\infty$, $k'\equiv k/\sqrt{\alpha}$ finite, we recover the results obtained previously [12,13] for an asymmetric version of the infinite-range Ising spin glass, with k' playing the role of the asymmetry parameter k of Ref. [12,13]. This is as would be expected because in the $\alpha\to\infty$ limit, the correlations among the elements of the matrix $\mathbf{J}^s$ may be neglected, and the Hopfield model becomes essentially identical to the Sherrington-Kirkpatrick model of infinite-range Ising spin glass. Finally, in the limit $\alpha\to\infty$, $k=0$, our result for F reduces to that obtained by Tanaka and Edwards [19] for the infinite-range Ising spin glass.

Figure 1 shows $F(\alpha,k)$ as a function of the asymmetry parameter k for three values (0.01, 0.05, and 0.1) of the memory loading level α. As k is increased $F(\alpha,k)$ decreases monotonically. Beyond a critical value $k_1(\alpha)$ of the asymmetry parameter, F becomes negative. Since e^{NF} goes to zero in the large-N limit if F is negative, there are no fixed points for values of k higher than $k_1(\alpha)$. The value of the critical asymmetry parameter $k_1(\alpha)$ increases with α (see Fig. 5). The calculated values of k_1 are higher than $\sqrt{\alpha}$, the value expected [12,13] in the $\alpha\to\infty$ limit. This result shows that the presence of correlations among the elements of the matrix $\mathbf{J}^s$ increases the value of k at which all the fixed points are suppressed.

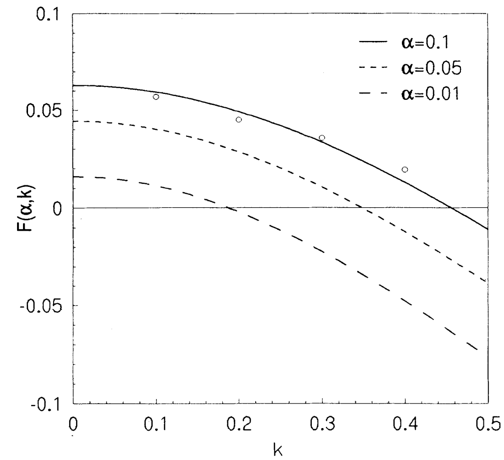

FIG. 1. Analytic results for the variation of $F(\alpha,k)$ with the asymmetry parameter k for three different values of the memory loading parameter α: $\alpha=0.1$ (top curve), $\alpha=0.05$ (middle curve), and $\alpha=0.01$ (bottom curve). Our numerical results for $F(\alpha,k)$, obtained for $\alpha=0.1$, are also shown (open circles).

The concept of an "energy function" [4] plays an important role in the theory of neural networks with symmetric interactions. Although such a function does not exist for asymmetric networks, one may still define an energylike function in terms of the symmetric part of the synaptic matrix alone. As discussed below, this function is useful in the development of physical insight into the behavior of the network in the presence of asymmetry. We define the energy function as

$$E=-\tfrac{1}{2}\sum_{i,j(i\neq j)} J_{ij}^s\sigma_i\sigma_j \equiv N\sqrt{\alpha}\epsilon . \tag{16}$$

The average number of fixed points with energy parameter ϵ, $\langle N_{fp}(N,\alpha,k,\epsilon)\rangle$, is given by

$$\langle N_{fp}(N,\alpha,k,\epsilon)\rangle = \mathrm{Tr}_{\{\sigma_i\}}\left\langle \int_0^\infty \prod_i d\lambda_i \prod_i \delta\left[\lambda_i-\sigma_i\sum_j J_{ij}\sigma_j\right]^s \left[N\sqrt{\alpha}\epsilon+\tfrac{1}{2}\sum_{i,j(i\neq j)} J_{ij}^s\sigma_i\sigma_j\right]\right\rangle . \tag{17}$$

Proceeding in the same way as above, we obtain, in the thermodynamic limit,

$$\langle N_{fp}(N,\alpha,k,\epsilon)\rangle \approx e^{NF(\alpha,k,\epsilon)} , \tag{18}$$

where

$$F(\alpha,k,\epsilon)=\frac{t^2}{2}-\frac{2t\epsilon\sqrt{\alpha}}{\sqrt{k^2+\alpha-2\epsilon\sqrt{\alpha}}} - \frac{k^2\epsilon^2\sqrt{\alpha}}{(\sqrt{\alpha}-2\epsilon)(k^2+\alpha-2\epsilon\sqrt{\alpha})} + \frac{\alpha}{2}\left[\frac{2\epsilon}{\sqrt{\alpha}-2\epsilon}+\ln\left[1-\frac{2\epsilon}{\sqrt{\alpha}}\right]\right]+\ln 2\phi(t) \tag{19}$$

and t is given by the saddle-point equation

$$t-\frac{2\epsilon\sqrt{\alpha}}{\sqrt{k^2+\alpha-2\epsilon\sqrt{\alpha}}}+\frac{\phi'(t)}{\phi(t)}=0 . \tag{20}$$

Figure 2 shows the variation $F(\alpha,k,\epsilon)$ with ϵ at different values of the asymmetry parameter k for $\alpha=0.1$. It is clear that the introduction of asymmetry causes a pronounced reduction in the number of fixed points with high values of ϵ. Consequently, the peak of the energy distribution shifts towards lower energy as the value of k is increased. This trend can be understood qualitatively if we consider the role of the energy parameter in the stability of a fixed point. For $k=0$, a configuration $\{\sigma_i\}$ is stable against all single spin flips if

$$\lambda_i' \equiv \sigma_i\sum_j J_{ij}^s\sigma_j > 0, \quad i=1,\ldots,N . \tag{21}$$

It follows from Eq. (16) that

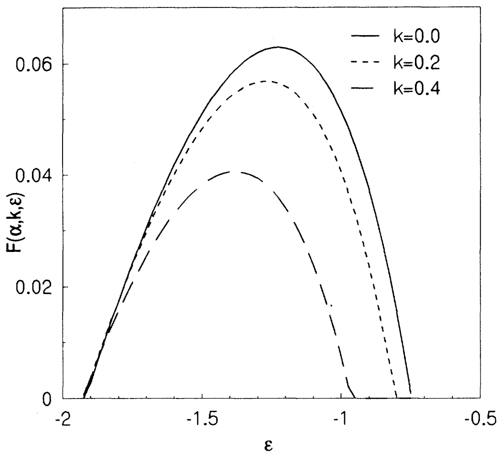

FIG. 2. Analytic results for the variation of $F(\alpha,k,\epsilon)$ with the energy parameter ϵ for $\alpha=0.1$ and three different values of the asymmetry parameter k: $k=0$ (top curve), $k=0.2$ (middle curve), and $k=0.4$ (bottom curve).

$$E=-\tfrac{1}{2}\sum_i \lambda_i' \; . \tag{22}$$

Thus, a fixed point of the symmetric model with low (high) energy has, on the average, large (small) values of the "stability parameters" $\{\lambda_i'\}$. Therefore, fixed points with relatively high energy are more likely to be destabilized by the introduction of the random asymmetric component of the interaction. This is precisely the trend we find (see Fig. 2). In the special case, $k=0$ and $\alpha\to\infty$, we recover, as expected, the results obtained previously [20,21] for the Sherrington-Kirkpatrick model of Ising spin glass.

To study the behavior of fixed points correlated with a stored pattern in the presence of asymmetry, we generalize the calculation of Gardner [6] on the number of fixed points at Hamming distance Ng from a stored pattern. We consider a configuration $\{\sigma_i\}$ which is at a Hamming distance Ng from the νth stored pattern $\{\xi_i^\nu\}$. This configuration will be a fixed point of the network dynamics if the following conditions are satisfied:

$$R_i^\nu\equiv\sigma_i\sum_j J_{ij}\sigma_j=\lambda_i,\quad \lambda_i>0,\quad i=1,\ldots,N \; . \tag{23}$$

Thus, the average number of fixed points at a Hamming distance Ng from the νth stored pattern is given by

$$\langle N_{fp}(N,\alpha,k,g)\rangle=\int_0^\infty \prod_i d\lambda_i \mathrm{Tr}_{\{\sigma_i\}}\left\langle \prod_i \delta(\lambda_i-R_i^\nu)\right\rangle \; . \tag{24}$$

Separation of the term coming from the νth pattern and the interference term coming from the other patterns gives

$$R_i^\nu=1-2g+\frac{1}{N}\sum_{j\neq i}\sum_{\mu\neq\nu}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j+k\sum_j J_{ij}^{\mathrm{as}}\sigma_i\sigma_j \tag{25a}$$

for $N(1-g)$ values of i for which $\xi_i^\nu=\sigma_i$, and

$$R_i^\nu=2g-1+\frac{1}{N}\sum_{j\neq i}\sum_{\mu\neq\nu}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j+k\sum_j J_{ij}^{\mathrm{as}}\sigma_i\sigma_j \tag{25b}$$

otherwise. Manipulating Eq. (24) in a manner similar to that used in the calculation of $\langle N_{fp}(N,\alpha,k)\rangle$ we get

$$\langle N_{fp}(N,\alpha,k,g)\rangle\approx e^{NF(\alpha,k,g)} \; . \tag{26}$$

$F(\alpha,k,g)$ is obtained as the saddle point (over the three parameters x, a, and b) of the function

$$\begin{aligned}F(\alpha,k,g,x,a,b)=&-\frac{x^2}{2}+\alpha\left[b-\frac{1}{2}+\frac{(1-b)^2}{2a}+\frac{1}{2}\ln a\right]\\&+(1-g)\ln\phi(t)+g\ln\phi(u)\\&-g\ln g-(1-g)\ln(1-g)\;,\end{aligned} \tag{27}$$

where

$$t=\frac{2g-1+b\alpha-ikx}{\sqrt{k^2+\alpha a}}\;, \tag{28a}$$

$$u=\frac{1-2g+\alpha b-ikx}{\sqrt{k^2+\alpha a}}\;. \tag{28b}$$

It can be checked explicitly that $F(\alpha,k)=F(\alpha,k,g=0.5)$. This implies that in the thermodynamic limit, the total number of fixed points of the network is dominated by the fixed points uncorrelated with the chosen memory.

For $k=0$, we recover, as expected, the results of Ref. [6]. In particular, we find that there is a narrow band of fixed points corresponding to the retrieval states. The peak of this band is at a finite Hamming fraction g_0 which implies that the recall of memory is always accompanied by some error, as a consequence of the extensive loading of the memory. The retrieval band is followed by a gap and thereafter by a broad band of fixed points corresponding to the spurious states. This band is centered around $g=0.5$ (i.e., states uncorrelated with the memory) and extends up to states having macroscopic correlations with the memory. As k is increased from zero, we find the following behavior in addition to the aforementioned reduction of the total number of fixed points.

(1) The height and the width of the retrieval band increases, signaling an exponential increase in the number of retrieval states.

(2) The peak of the retrieval band shifts towards higher values of the Hamming fraction, i.e., the value of g_0 increases with k.

While (1) is beneficial from the point of view of associative recall, (2) adversely affects the retrieval quality of the network. The result that increasing k leads to increased error in retrieval may be understood qualitatively if we consider the source of retrieval error in the symmetric network itself [3,4]. From Eq. (6), the condition that the ith bit of the νth pattern is correctly retrieved in the symmetric network when all the other neurons are fixed in the states corresponding to the νth pattern ($\sigma_j=\xi_j^\nu$ for $i\neq j$) is

$$\frac{N-1}{N}+\frac{1}{N}\sum_{j\neq i}\sum_{\mu\neq\nu}\xi_i^\nu\xi_i^\mu\xi_j^\nu\xi_j^\mu>0 \ . \tag{29}$$

In the limit of large N, the first term on the left-hand side of Eq. (29) tends to unity. The second term involves a sum of $(N-1)(p-1)\approx Np$ random terms, each of which takes the values $+1$ and -1 with equal probability. Neglecting correlations among these terms, the sum may be approximated by a Gaussian random variable with zero mean and variance $\sigma^2=\alpha$. Hence, the probability that the ith bit is retrieved correctly is given by

$$P\approx\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-1}^{\infty}dx\,\exp\left[-\frac{x^2}{2\sigma^2}\right]$$

$$=\frac{1}{2}\left[1+\mathrm{erf}\left[\frac{1}{2\sigma^2}\right]^{1/2}\right] . \tag{30}$$

From this equation, it is clear that a finite amount of retrieval error is present for any nonzero value of α. For the asymmetric case, Eq. (29) has an additional term, $k\sum_{j,j\neq i}J_{ij}^{\mathrm{as}}\xi_i^\nu\xi_j^\nu$, on its left-hand side. Consequently, σ in Eq. (30) is given by

$$\sigma^2=\alpha+k^2 \ . \tag{31}$$

Thus, the introduction of asymmetry increases the strength of the "noise" term, leading to increased error in the retrieval of a memory. The increase in the peak height of the retrieval band is due to the increase in the phase space factor associated with the increased error in the retrieval.

Our results for $F(\alpha,k,g)$ for two different values of the memory loading level, $\alpha=0.05$ and 0.1, and several values of the asymmetry parameter k are shown in Figs. 3 and 4. Since F is invariant under $g\to 1-g$, its value for $g\leq 0.5$ only are shown in these figures. For relatively small values of α (see Fig. 3), the retrieval band is very well separated from the spurious band. The value of F at the peak of the retrieval band increases initially as k is increased from zero, but subsequently begins to decrease as k is increased further. Eventually, it crosses zero at a "critical" value, $k_2(\alpha)$, of k, which increases with α. Thus, retrieval fixed points do not exist in the thermodynamic limit if $k>k_2(\alpha)$. The peak of the retrieval band for $k<k_2(\alpha)$ occurs at a small value of g, indicating that the quality of retrieval is not seriously affected by the introduction of asymmetry if the value of α is small. The qualitative behavior of $F(\alpha,k,g)$ as a function of k changes as the value of α is increased from $\alpha_0\approx 0.0776$. For such values of α, the disappearance of the retrieval band occurs through its merger with the spurious band (i.e., a vanishing of the gap between the two bands). This

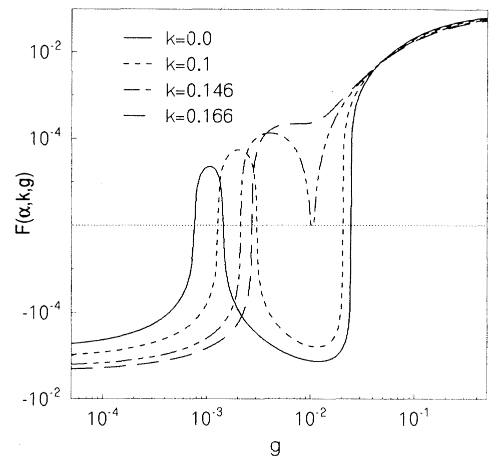

FIG. 4. Analytic results for the dependence of $F(\alpha,k,g)$ on the Hamming fraction g, for $\alpha=0.1$ and four different values (0, 0.1, 0.146, and 0.166) of the asymmetry parameter k. The retrieval band merges with the spurious band near $k=0.146$ and the low-g peak of F disappears near $k=0.166$.

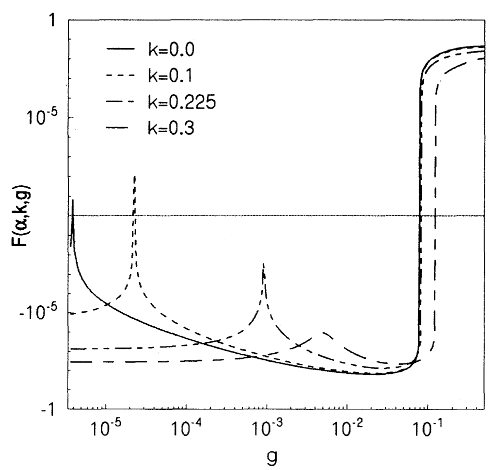

FIG. 3. Analytic results for the dependence of $F(\alpha,k,g)$ on the Hamming fraction g, for $\alpha=0.05$ and four different values (0, 0.1, 0.225, and 0.3) of the asymmetry parameter k. The peak of the retrieval band crosses zero near $k=0.225$.

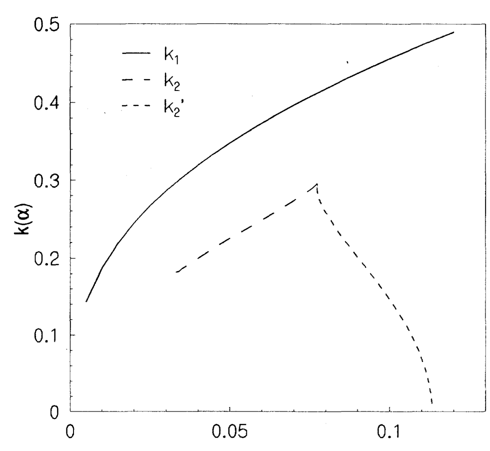

FIG. 5. Analytic results for the three "critical" values of the asymmetry parameter k. The upper curve shows the variation of k_1 with α. The lower curve shows the α dependence of k_2 for $\alpha<\alpha_0\approx 0.0776$ and that of k_2' for $\alpha\geq\alpha_0\approx 0.0776$.

is illustrated in Fig. 4. The gap between the two bands goes to zero at $k=k_2'(\alpha)$, which decreases as α is increased and goes to zero at $\alpha \approx 0.11$, in agreement with the result derived in Ref. [6]. The variations of the "critical" values, k_1 (the value of k at which all fixed points disappear), k_2 (for $\alpha<\alpha_0$), and k_2' (for $\alpha>\alpha_0$) with α are shown in Fig. 5. A comparison of the results shown in Figs. 3 and 4 shows that the degradation in the quality of retrieval for a fixed value of k increases as α is increased. This result and the observation that the suppression of the retrieval band occurs at relatively high values of k for small values of α suggest that the introduction of asymmetry will have a beneficial effect on the performance of the network only if the value of α is small ($\alpha \le \alpha_0 \approx 0.0776$).

III. SIMULATIONS

We have carried out a number of numerical simulations to test predictions of the analytic calculations described in the preceding section. Since the total number of fixed points increases exponentially with N, an exhaustive enumeration of all the fixed points is numerically feasible only if the value of N is small. For this reason, the simulations described in this paper were restricted to networks of rather small sizes, $N=40$, 60, and 80. These values of N are in the same range as the sample sizes considered in all existing numerical studies [13,25,26] of the number and properties of fixed points of similar models. For asymmetric networks, four different values 0.1, 0.2, 0.3, 0.4, of the asymmetry parameter k were considered and simulations for each realization of the symmetric component of the synaptic matrix were carried out for two different sets of antisymmetric interconnections. For $N=80$, only the asymmetric network was studied. The memory loading level α of the networks was fixed at 0.1. Random sequential updating was used in all these simulations. The number of realizations used for $N=40$, 60, and 80 was 100, 20, and 10, respectively.

We first describe the procedure used in the enumeration of all the fixed points of a network. For each realization of the synaptic interaction matrix, we started with a large number of randomly chosen initial configurations of the variables $\{\sigma_i\}$ and allowed the network to evolve until it reached a fixed point or a preset time limit (see below) was exceeded. A fixed point is referred to as "paired" if both the fixed point and its complement (obtained by reversing the signs of all the σ_i's) have been found in the search. Otherwise, it is called "unpaired." In the simulation we found that $N_{\rm up}$, the number of "unpaired" fixed points, increases with the increase of the total number of found fixed points at the beginning of the search procedure. Then, as the search proceeds further, $N_{\rm up}$ fluctuates around a characteristic value as new fixed points as well as the complements of already found fixed points are encountered in the search. Eventually, $N_{\rm up}$ decreases towards zero. After it reaches zero, no new fixed point is found even if the search is continued further for a very large number of random inputs (e.g., 10 times the number of inputs used before $N_{\rm up}$ reaches zero). Therefore, the vanishing of $N_{\rm up}$ may reasonably be taken to be an indication that almost all the fixed points have already been found in the search. Another indicator that may be used is the frequency of finding "new" fixed points. At the beginning of the search, almost every fixed point found is a new one. As the search proceeds, the frequency of appearance of new fixed points gradually decreases and finally approaches zero. In this situation also, we may assume that almost all the fixed points have been found. In our simulation, both these indicators were used. If all of the existing fixed points are paired, and no new fixed point is encountered for a sufficiently large number of subsequent inputs (e.g., 6000 for $N=60$ and $k=0$), the search is terminated and the total number of fixed points is taken to be two times the number of paired ones. If the number of unpaired fixed points becomes very small but does not reach zero, and no new fixed point appears for a much larger number of inputs (e.g., 60 000 inputs for $N=60$ and $k=0$), then the search is terminated and the total number of fixed points is taken to be two times the sum of the numbers of both "paired" and "unpaired" ones. In our simulations, the total number of inputs used to find all the fixed points of a network ranges from 2000 to more than 5 000 000. With the increase of N and decrease of k, the number of inputs needed for searching out all the fixed points increases rapidly. For $N=40$ and $k=0.4$, the average number of inputs is 2631. It reaches the value 2 480 000 for $N=60$ and $k=0$, and 2 570 000 for $N=80$ and $k=0.1$. The largest number of input for one set of couplings is 5 356 625, obtained in the case $N=60$ and $k=0$. The average number of fixed points found in our simulations ranges from 7 to 302, the smallest number being obtained for $N=40$ and $k=0.4$, and the largest one for $N=60$ and $k=0$. The largest number of fixed points found in our simulations for a particular network is 774, obtained in a network with $N=60$ and $k=0$. For every fixed point, its energy [as defined in Eq. (16)], its Hamming distance from each memory state of the network, and the number of random inputs which converged to the fixed point were recorded.

Limit cycles and chaotic trajectories are known [12,27] to occur in the dynamics of networks with asymmetric connections. To identify such trajectories, the average time needed to reach a fixed point from a randomly chosen initial configuration was first estimated from simulations on several samples. Then, a time limit was set at a value greater than ten times this average. If a given input did not reach a fixed point within this time limit, it was considered to have been trapped into a limit cycle or a chaotic trajectory.

Two methods of averaging were used in the calculation of $F(\alpha,k)$. In the first method, $\ln N_{fp}$ was calculated for each realization of the network, and its mean and variance over all the realizations were evaluated. Then, $F(\alpha,k)$ was determined by fitting the results obtained for different values of N to the linear equation

$$\langle \ln N_{fp} \rangle = NF(\alpha,k)+C \ , \tag{32}$$

where C is a constant. In the second method, the average of N_{fp} over all sets of couplings was calculated at first, and $F(\alpha,k)$ was determined from a weighted least-squares fit to the linear equation

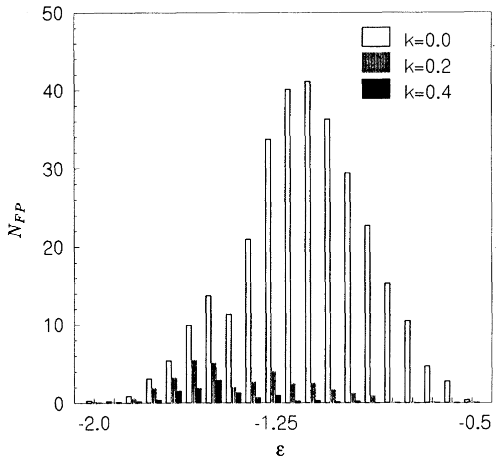

FIG. 6. Numerical results for the dependence of N_{fp}, the average number of fixed points, on the energy parameter ϵ, for networks with $N=60$, $\alpha=0.1$, and three different values (0, 0.2, and 0.4) of the asymmetry parameter k.

$$\ln\langle N_{fp}\rangle = NF(\alpha,k)+C \ . \tag{33}$$

We found that the results obtained by the two methods are essentially identical within error bars. This observation strongly suggests that the quantity N_{fp} is self-averaging. The numerical results for $F(\alpha,k)$ for $\alpha=0.1$, and a comparison with the analytic ones are shown in Fig. 1. The results of the simulation are found to be in good agreement with those obtained from the analytic study.

Simulation results for the dependence of the number of fixed points on the energy parameter ϵ and the Hamming fraction g are shown in Figs. 6 and 7, respectively. These data were all obtained from simulations of networks with $N=60$. To obtain the data shown in Fig. 6, the entire range of possible energy values was divided into a number of intervals ("bins") of equal width. For each coupling matrix, the number of stable states whose energies fall into the different bins were counted. The numbers obtained in this way for all the bins were then averaged over all realizations of the coupling matrix. A comparison of Fig. 6 with Fig. 2 shows that all the qualitative features predicted by the analytic calculation (e.g., selective elimination of fixed points with high energy by the introduction of asymmetry and a shift of the peak of the energy distribution to lower values of the energy as the asymmetry parameter k is increased) are reproduced in the numerical results.

Figure 7 shows the average number of fixed points as a function of g, the fractional Hamming distance measured from a specific memory state. The data shown in this figure were obtained in the following way. For a given coupling matrix, the numbers of stable states with specific Hamming distances from a particular memory state were counted. The numbers for different Hamming distances were then averaged over all the memories of the network. Finally, by averaging over all realizations of the coupling matrix, the results shown in Fig. 7 were obtained. From a comparison with Fig. 4, it is clear that all the main results obtained from the analytic study (e.g., the existence of two bands of fixed points for small values of k, a decrease in the height of the spurious band as k is increased, and a merger of the retrieval band with the spurious one for large values of k) are qualitatively reproduced in the simulation. The spurious band is found to exhibit a secondary peak in the region $g\approx 0.2-0.3$. As discussed below, this is possibly an artifact of the smallness of sample size.

Figure 8 shows the dependence of the average number of fixed points (for $N=60$) on the "smallest" Hamming

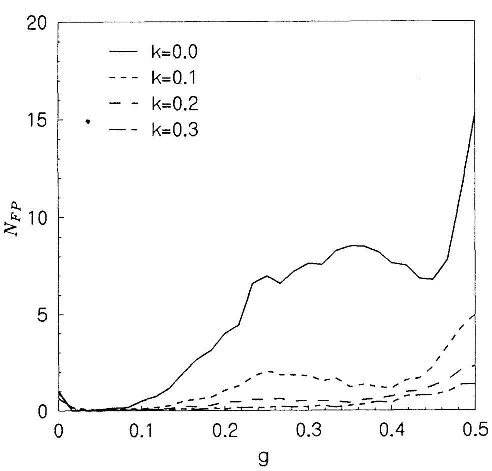

FIG. 7. Numerical results for the dependence of N_{fp}, the average number of fixed points, on the fractional Hamming distance g measured from a particular memory. The results shown are for networks with $N=60$, $\alpha=0.1$, and four different values (0, 0.1, 0.2, and 0.3) of the asymmetry parameter k.

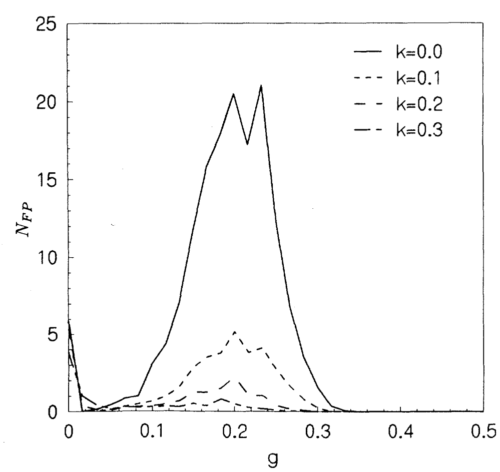

FIG. 8. Numerical results for the dependence of N_{fp}, the average number of fixed points, on g_m, the minimum fractional Hamming distance from the memory states. The results shown are for networks with $N=60$, $\alpha=0.1$, and four different values (0, 0.1, 0.2, and 0.3) of the asymmetry parameter k.

distance (i.e., the least one of the Hamming distances measured from all the memories stored in the network). In obtaining the data shown in this figure, the smallest Hamming distance for each fixed point was selected from the recorded Hamming distances from all the memory states. The number of fixed points with a particular value of the smallest Hamming distance was then counted and averaged over different realizations of the coupling matrix. The presence of two distinct bands of fixed points and a selective elimination of spurious fixed points by the introduction of asymmetry are clearly seen in this figure. An interesting feature of the data shown here is that the band of spurious states has its peak at $g \approx 0.2$, which is very different from the expected value, $g = 0.5$. It is not clear at this stage whether this feature is an artifact of the smallness of the size of the network. Preliminary results obtained for larger networks ($N \leq 500$) indicate that the peak of the spurious band does move to larger values of g as N is increased. Our results, however, bring out the important fact that nearly all of the spurious fixed points in networks of small size are substantially correlated with at least one of the memory states.

IV. SUMMARY AND DISCUSSIONS

In summary, we have studied, using analytic methods and numerical simulations, the number and properties of the fixed points of a Hopfield model in which a random antisymmetric part is added to the otherwise symmetric synaptic matrix. We find that the introduction of the antisymmetric component leads to an exponential decrease in the number of spurious fixed points, whereas the retrieval fixed points remain mostly unaffected by the introduction of asymmetry if the relative strength of the antisymmetric component of the interaction matrix and the memory loading level are small. The fixed points which are eliminated due to the introduction of asymmetry have relatively high values of an energy parameter defined in terms of the symmetric part of the interaction matrix. All the qualitative predictions and a few quantitative ones of the analytic calculations are confirmed by the results obtained from simulations.

As mentioned in the Introduction, calculations similar to the ones described here have been carried out by Treves and Amit [22] for an asymmetrically diluted Hopfield network. The main differences between the model studied by them and the one considered here are as follows.

(1) The symmetric component of the interaction matrix in the model with asymmetric dilution is not identical to the synaptic matrix of the original Hopfield model. Therefore, the random component introduced in the synaptic matrix by asymmetric dilution has both symmetric and antisymmetric parts. For this reason, it is difficult to determine which of the changes caused by asymmetric dilution are consequences of the presence of an antisymmetric component in the interaction matrix. In contrast, the symmetric part of the synaptic matrix of the model we consider is the same as that of the original Hopfield model. Therefore, any observed difference between the behavior of our model and that of the Hopfield model may unambiguously be attributed to the presence of the antisymmetric component.

(2) In the asymmetrically diluted model, one (and only one) of J_{ij} and J_{ji} is zero if they are not equal. This implies that the magnitudes of the symmetric and antisymmetric components of J_{ij} are equal if $J_{ij} \neq J_{ji}$. Thus, the elements of the antisymmetric part of the interaction matrix of this model are not completely independent of those of the symmetric part. In our model, in contrast, the elements of J^{as} are completely independent of the corresponding elements of J^{s}.

A comparison of our results with those obtained in Ref. [22] shows that the behavior of spurious fixed points is essentially the same in the two models—these fixed points are exponentially suppressed by the introduction of asymmetry. The main difference between the results obtained for the two models lies in the behavior of the retrieval band. In the asymmetrically diluted model, the width and the height of this band decrease with increasing asymmetry, while both these quantities increase with asymmetry in our model when the asymmetry parameter k is below a critical value. Furthermore, in the model of Ref. [22], the dilution does not cause any significant change in the retrieval quality, whereas this is true in our model only for low loading of the memory.

It was mentioned earlier that although we have calculated $\langle N_{fp} \rangle$ in the analytic part of this study, the quantity that should actually be calculated is $\langle \ln N_{fp} \rangle$. A calculation of the latter quantity would involve the use of replicas. Such a calculation was formulated by Bray and Moore [20,21] for the infinite-range Ising spin glass. They found that the calculation of $\langle \ln N_{fp} \rangle$ reduces to that of $\ln \langle N_{fp} \rangle$ if all saddle-point parameters which are off-diagonal in replica space are set to zero. This "diagonal" saddle point was found to be locally stable, which suggests (but does not prove because local stability of this saddle point is a necessary but not sufficient condition for N_{fp} to be self-averaging) that the results for $\langle \ln N_{fp} \rangle$ and $\ln \langle N_{fp} \rangle$ should be the same if N_{fp} is the total number of fixed points (metastable states). Our numerical results [approximate equality of $\langle \ln N_{fp} \rangle$ and $\ln \langle N_{fp} \rangle$, and good agreement between the analytic and numerical results for $F(\alpha,k)$] suggest that a similar conclusion applies to our calculation of the total number of fixed points. We did not attempt a stability analysis similar to that of Ref. [20,21] because such an analysis is much more complicated for the model we consider here. We have formulated the replica analysis for a calculation of $\langle \ln N_{fp} \rangle$ for the purely symmetric Hopfield model. The resulting saddle point equations involve ten parameters and several multidimensional integrals. An analytic or numerical solution of these saddle point equations would be a formidable task. A similar calculation for the model we consider here would be even more complicated. For this reason, a calculation along these lines was not attempted.

Existing studies [6,20–22] of the number and properties of fixed points of spin-glass and neural network models, which are special cases of the model we consider here, suggest that the quantities $N_{fp}(N,\alpha,k,\epsilon)$ and

$N_{fp}(N,\alpha,k,g)$ are probably not self-averaging for all values of ϵ and g. For this reason, a comparison between our analytic and numerical results for these quantities would be meaningful only at a qualitative level. The observed similarity between our analytic and numerical results for these quantities suggests that most of the qualitative features of the actual behavior of these quantities are rendered correctly in our (possibly approximate) analytic calculation. Since a quantitative agreement between analytic and numerical results is not expected, we did not attempt to carry out a detailed finite-size scaling analysis of the numerical data for these quantities.

We close with a discussion on the implications of the results of this study on the retrieval properties of Hopfield-type models with random asymmetry. The observation that the introduction of moderate amounts of random asymmetry causes an exponential reduction in the number of spurious fixed points, but does not affect the retrieval states seriously suggests that the presence of such asymmetry may improve the performance of the network as an associative memory by enlarging the basins of attraction of the memory states. Of course, the results obtained in Sec. II tell us that random asymmetry alone cannot completely eliminate the spurious fixed points without eliminating the retrieval states too. In fact, for any value of the asymmetry parameter k, the height of the spurious band of fixed points is larger than that of the retrieval band. Since the number of spurious fixed points is always exponentially large in comparison with that of retrieval states, the fraction of randomly chosen initial configurations which converge to retrieval states is expected to go to zero in the large-N limit for any value of k. However, the reduction in the number of spurious fixed points by the introduction of asymmetry may still have some beneficial effect in networks of relatively small size. For example, our numerical calculations show that for a network with $N=60$, $\alpha=0.1$, the average number of fixed-point attractors decreases from ≈ 300 to ≈ 20 as k is increased from 0 to 0.3. It is reasonable to expect that such a reduction in the number of spurious attractors would lead to a substantial enlargement of the basins of attraction of the memory states in such a network. One should, however, be cautious in making such a prediction. Our results indicate that the attractors which are eliminated by the introduction of asymmetry have relatively high values of the energy parameter ϵ. Although any general relation between the energy and the size of the basin of attraction of a fixed point has not yet been established, it is generally believed [28] that fixed points with high energies have small basins of attraction. If this is true in general, then the fixed points which are eliminated by the presence of asymmetry would have small basins of attraction and their elimination would not have any significant effect on the retrieval performance of the network. It would be interesting to check whether this is the case.

Information about the number and properties of the fixed points of the network is, of course, not sufficient for the development of a complete understanding of its performance as an associative memory: it is also necessary to analyze its dynamics. At the present time, many aspects of the dynamics of neural networks with asymmetric interactions remain incompletely understood [29]. A question of particular interest in the present context is how the time taken by the network to converge to a fixed point is modified when asymmetry is introduced in the synaptic interactions. The results reported in Ref. [15] suggest that the presence of asymmetry greatly increases the time of convergence to spurious attractors while leaving the convergence time for retrieval states mostly unaffected. However, no systematic investigation of the dependence of the time of convergence on the nature of the attractor, the memory loading level, the relative strength of the antisymmetric part of the interactions etc. has been reported so far. Another interesting question is whether a network with $N\rightarrow\infty$ and asymmetric interactions ever settles down to a fixed point attractor. Several authors [1,7,8] have suggested that the presence of an antisymmetric part in the connection matrix effectively makes the deterministic update rule of Eq. (4) a stochastic one, similar to the update rule corresponding to a nonzero temperature. If this analogy is correct, then a network with asymmetric interactions would never settle down into a time-independent state. In this scenario, the retrieval of a memory would correspond to the system embarking on a trajectory which remains strongly correlated with a memory state at all times. It is not clear whether trajectories which remain strongly correlated with one of the spurious fixed point attractors of the symmetric model would also occur in the dynamics when asymmetry is introduced. Detailed studies of some of these outstanding issues would be most interesting.

ACKNOWLEDGMENTS

Z.C. would like to thank the Department of Physics, Indian Institute of Science and Professor T. V. Ramakrishnan for hospitality for the period during which this work was carried out. M.P.P. would like to thank the Jawaharlal Nehru Centre for Advanced Scientific Research for support and hospitality for a visit during which a part of this work was carried out. Facilities for performing the numerical work were provided by the Supercomputer Education and Research Center of Indian Institute of Science.

APPENDIX: DERIVATION OF THE SADDLE-POINT EQUATIONS

Using an integral representation of the δ function, Eq. (7) can be written as

$$\langle N_{fp}(N,\alpha,k)\rangle=\int_0^\infty\prod_i d\lambda_i\int_{-\infty}^{\infty}\prod_i\frac{d\phi_i}{2\pi}\mathrm{Tr}_{\{\sigma_i\}}\exp\left[\sum_i i\phi_i\lambda_i\right]\left\langle\exp\left[-i\sum_{i,j}\phi_iJ_{ij}\sigma_i\sigma_j\right]\right\rangle , \tag{A1}$$

where

$$\left\langle \exp\left[-i\sum_{i,j}\phi_i J_{ij}\sigma_i\sigma_j\right]\right\rangle \equiv \left\langle \exp\left[-i\sum_{i,j(i\neq j)}\frac{1}{N}\sum_\mu \xi_i^\mu\xi_j^\mu\sigma_i\sigma_j\phi_i\right]\right\rangle_{\{\xi_i^\mu\}}\left\langle \exp\left[-ik\sum_{i,j(i<j)}J_{ij}^{\rm as}\sigma_i\sigma_j(\phi_i-\phi_j)\right]\right\rangle_{\{J_{ij}^{\rm as}\}} . \quad \text{(A2)}$$

Let us denote the first term on the right-hand side of Eq. (A2) by X and the second term by Y. Then we have

$$X=\exp\left[i\alpha\sum_i\phi_i\right]\left\langle \exp\left\{\frac{-i}{N}\sum_\mu\left[\sum_i\sigma_i\phi_i\xi_i^\mu\right]\left[\sum_i\sigma_i\xi_i^\mu\right]\right\}\right\rangle_{\{\xi_i^\mu\}} . \quad \text{(A3)}$$

Using the transformation [24]

$$\exp(-iA_\mu B_\mu/N)=\int_{-\infty}^{\infty}\frac{da_\mu}{(2\pi/N)^{1/2}}\int_{-\infty}^{\infty}\frac{db_\mu}{(2\pi/N)^{1/2}}\exp\left[i\frac{N}{2}(a_\mu^2-b_\mu^2)-\frac{i}{\sqrt{2}}A_\mu(a_\mu+b_\mu)-\frac{i}{\sqrt{2}}B_\mu(a_\mu-b_\mu)\right] , \quad \text{(A4)}$$

Eq. (A3) becomes

$$X=\exp\left[i\alpha\sum_i\phi_i\right]\int_{-\infty}^{\infty}\prod_\mu\frac{da_\mu}{(2\pi/N)^{1/2}}\int_{-\infty}^{\infty}\prod_\mu\frac{db_\mu}{(2\pi/N)^{1/2}}\exp\left[i\frac{N}{2}\sum_\mu(a_\mu^2-b_\mu^2)+\sum_{\mu,i}\ln\cos\frac{1}{\sqrt{2}}\{\phi_i(a_\mu+b_\mu)+(a_\mu-b_\mu)\}\right] . \quad \text{(A5)}$$

Since $\sum_i\sigma_i\xi_i^\mu$ is $O(\sqrt{N})$ for an overwhelming majority of fixed points, a_μ and b_μ would be $O(1/\sqrt{N})$. Therefore, expanding the last term in the exponent in Eq. (A5) and neglecting terms of the order $1/N$ and higher orders, we have

$$X=\exp\left[i\alpha\sum_i\phi_i\right]\int\prod_\mu\frac{da_\mu}{(2\pi/N)^{1/2}}\int\prod_\mu\frac{db_\mu}{(2\pi/N)^{1/2}}\exp\left[i\frac{N}{2}\sum_\mu(a_\mu^2-b_\mu^2)-\tfrac{1}{4}\sum_{\mu,i}\{\phi_i^2(a_\mu+b_\mu)^2+2\phi_i(a_\mu^2-b_\mu^2)+(a_\mu-b_\mu)^2\}\right] . \quad \text{(A6)}$$

Introducing the variables

$$a=\frac{1}{2\alpha}\sum_\mu(a_\mu+b_\mu)^2 , \quad \text{(A7a)}$$

$$b=\frac{i}{2\alpha}\sum_\mu(a_\mu^2-b_\mu^2)+1 , \quad \text{(A7b)}$$

and the Lagrange multipliers A and B, respectively conjugate to the constraints in Eqs. (A7a) and (A7b), so that,

$$1=\int\frac{da\,dA}{2\pi/N\alpha}\exp\left[iNA\left[\alpha a-\sum_\mu\frac{(a_\mu+b_\mu)^2}{2}\right]\right] , \quad \text{(A8a)}$$

$$1=\int\frac{db\,dB}{2\pi/N\alpha}\exp\left[iNB\left[\alpha b-i\sum_\mu\frac{(a_\mu^2-b_\mu^2)}{2}-\alpha\right]\right] , \quad \text{(A8b)}$$

the integrations over a_μ and b_μ are decoupled and can be done easily. For large N, saddle-point integration over the variables A and B can be performed explicitly as the saddle-point equations for A and B are algebraic. One then finds that

$$X=\int\frac{da\,db}{(4\pi/N\alpha)}\exp\left[N\alpha\left\{b-\frac{1}{2}+\frac{1}{2}\frac{(1-b)^2}{a}+\tfrac{1}{2}\ln a\right\}-\tfrac{1}{2}\alpha a\sum_i\phi_i^2+i\alpha b\sum_i\phi_i\right] . \quad \text{(A9)}$$

Performing averages over the Gaussian variables $\{J_{ij}^{\rm as}\}$ and ignoring $O(1/N)$ terms in Eq. (A2), we get

$$Y=\exp\left[-\frac{k^2}{2}\sum_i\phi_i^2+\frac{k^2}{2N}\left[\sum_i\phi_i\right]^2\right] . \quad \text{(A10)}$$

Using the Hubbard-Stratonovich identity Eq. (A10) can be written as

$$Y=\exp\left[-\frac{k^2}{2}\sum_i\phi_i^2\right]\frac{1}{\sqrt{2\pi/N}}\times\int dx\,\exp\left[-\frac{Nx^2}{2}+xk\sum_i\phi_i\right] . \quad \text{(A11)}$$

Thus the integrand in Eq. (A1) becomes independent of $\{\sigma_i\}$ and therefore $\mathrm{Tr}_{\{\sigma_i\}}$ gives 2^N. Performing Gaussian integration over the variables $\{\phi_i\}$ in Eq. (A1), we finally get

$$\langle N_{fp}(N,\alpha,k)\rangle\approx e^{NF(\alpha,k)} , \quad \text{(A12)}$$

where the quantity $F(\alpha,k)$ is the saddle-point value of the function defined in Eq. (9).

[1] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
[2] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **81**, 3088 (1984).
[3] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985); Phys. Rev. Lett. **55**, 1530 (1985); Ann. Phys. (N.Y.) **173**, 30 (1987).
[4] D. J. Amit, *Modeling Brain Functions* (Cambridge University Press, Cambridge, 1989).
[5] J. C. Eccles, *The Physiology of Synapses* (Springer-Verlag, Berlin, 1964).
[6] E. G. Gardner, J. Phys. A **19**, L1047 (1986).
[7] J. A. Hertz, G. Grinstein, and S. A. Solla, in *Heidelberg Colloquium on Glassy Dynamics,* edited by L. N. van Hemmen and I. Morgenstern (Springer-Verlag, Heidelberg, 1987).
[8] A. Crisanti and H. Sompolinsky, Phys. Rev. A **36**, 4922 (1987).
[9] M. V. Feigelman and L. B. Ioffe, Europhys. Lett. **1**, 197 (1986).
[10] M. V. Feigelman and L. B. Ioffe, Int. J. Mod. Phys. **1**, 51 (1987).
[11] W. Kinzel, in *Heidelberg Colloquium on Glassy Dynamics,* edited by L. N. van Hemmen and I. Morgenstern (Springer-Verlag, Heidelberg, 1987).
[12] A. Crisanti and H. Sompolinsky, Phys. Rev. A **37**, 4865 (1988).
[13] H. Gutfreund, J. D. Reger, and A. P. Young, J. Phys. A **21**, 2775 (1988).
[14] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[15] G. Parisi, J. Phys. A **19**, L675 (1986).
[16] G. Toulouse, S. Dehaene, and J. P. Changeux, Proc. Natl. Acad. Sci. USA **83**, 1695 (1986).
[17] H. Sompolinsky, in *Heidelberg Colloquium on Glassy Dynamics* (Ref. [7]).
[18] S. F. Edwards and F. Tanaka, J. Phys. F **10**, 2471 (1980).
[19] F. Tanaka and S. F. Edwards, J. Phys. F **10**, 2769 (1980).
[20] A. J. Bray and M. A. Moore, J. Phys. C **13**, L469 (1980).
[21] A. J. Bray and M. A. Moore, J. Phys. C **14**, 1313 (1981).
[22] A. Treves and D. J. Amit, J. Phys. A **21**, 3155 (1988).
[23] R. Brout, Phys. Rev. **115**, 824 (1959).
[24] A. D. Bruce, E. J. Gardner, and D. J. Wallace, J. Phys. A **20**, 2909 (1987).
[25] F. R. Waugh, C. M. Marcus, and R. M. Westervelt, Phys. Rev. Lett. **64**, 1964 (1990).
[26] P. S. Spitzer and W. Kinzel, Z. Phys. B **74**, 539 (1989).
[27] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).
[28] J. J. Hopfield, D. L. Feinstein, and R. G. Palmer, Nature **304**, 158 (1983).
[29] See W. Krauth, J.-P. Nadal, and M. Mezard, J. Phys. C **21**, 2995 (1988) for an attempt in this direction.