

# Phase transitions in dilute, locally connected neural networks

Katherine J. Strandburg

*Argonne National Laboratory, Argonne IL 60439*

*Northwestern University, Evanston IL 60208*

Michael A. Peshkin

*Northwestern University, Evanston IL 60208*

Daniel F. Boyd

Christopher Chambers

Brennan O'Keefe

*Argonne National Laboratory, Argonne IL 60439*

**Abstract:** We report numerical studies of the "memory-loss" phase transition in Hopfield-like symmetric neural networks in which the neurons are connected to all other neurons within a local neighborhood (dense, short-range connectivity). The number of connections per neuron,  $K$ , scales as the number of neurons,  $N$ , raised to a power less than one (i. e.,  $K \sim N^\eta$ ,  $\eta < 1$ ). We use the recently developed Lee-Kosterlitz finite size scaling technique to determine the critical value of  $\eta$  below which the first-order phase transition disappears.

Phase transitions in dilute, locally connected neural networks

Katherine J. Strandburg, Michael A. Peshkin, Daniel F. Boyd, Christopher Chambers, Brennan O'Keefe

Physical Review A, 45(8) 6135-6138, 1992

Abstract: We report numerical studies of the "memory-loss" phase transition in Hopfield-like symmetric neural networks in which the neurons are connected to all other neurons within a local neighborhood (dense, short-range connectivity). The number of connections per neuron,  $K$ , scales as the number of neurons,  $N$ , raised to a power less than one (i. e.,  $K \sim N^\eta$ ,  $\eta < 1$ ) We use the recently developed Lee-Kosterlitz finite size scaling technique to determine the critical value of  $\eta$  below which the first-order phase transition disappears.

Using a recently developed finite-size scaling approach<sup>1</sup>, we have determined the phase diagram of limited-connectivity Hopfield-type<sup>2</sup> networks, in which each of  $N$  neurons is symmetrically connected to the  $K$  neurons closest to it. We observe empirically a new relationship of these Hopfield-type neural networks to standard statistical mechanical models. This connection is distinct from that between Hopfield networks and spin glasses which has been of great importance in the mean field theory of Hopfield networks<sup>3</sup>. The Hopfield network and its limited connectivity variants are observed to obey an approximate Boltzmann-like energy distribution where a function of the number of patterns plays the role of temperature. We exploit this observation in an investigation of the memory-loss phase transitions of networks where the number of connections increases with  $N$

with a power less than 1 (i.e.  $K \sim N^\eta$ ,  $\eta < 1$ ). We find a critical value of  $\eta \sim .65$  below which the first order transition disappears. Surprisingly, this critical value is independent (or at least nearly so) of the dimensionality of the connection arrangements.

The connection structure we explore is thus dense ( $K \sim r^d$ , where  $r$  is the range of the connections), but local. Our model shares with biological networks this locally dense but globally dilute interconnectivity<sup>3</sup>. For computational applications the case of dense but local connections is also of interest. (Optimal connection structure is determined in part by a competition between the increase in memory capability for larger  $K$  and the increased wiring requirements for larger  $r$ .)

In previous publications others have studied densely connected symmetric networks with  $K \propto N$  using primarily a mean field theory approach<sup>3, 4</sup>. The case of highly dilute asymmetric connections ( $K \ll r^d$ ) has also been recently addressed by exact arguments for the long range<sup>5</sup> ( $r \sim N^{1/d}$ ) and local<sup>6</sup> ( $r \sim N^{\eta/d}$ ,  $\eta < 1$ ) cases. The behavior of networks with dense, local connections cannot be investigated using these analytical techniques.

In what follows we begin by reviewing the Hopfield model and introducing the modifications used here. We then describe a finite-size scaling technique recently developed by Lee and Kosterlitz and apply it to our limited-connectivity networks.

## **Model**

The Hopfield model<sup>3</sup> is an approach to producing a neural network with an associative memory property. In this model the neurons are simple on/off elements which can be thought of as Ising spins. A memory is a specification of the

state of each of the neurons. The network is run from some initial state by proceeding randomly through the neurons and turning each neuron on or off depending on whether the input it receives from the neurons to which it is connected is positive or negative. In spin language, we flip each spin if it decreases the energy, where the energy function is of the usual Ising form

$$E = - \sum J_{ij} s_i s_j \quad (1)$$

where each spin  $s_i = \pm 1$ . This procedure has the effect of moving the system to the nearest (first-encountered) local minimum of the energy. The  $J_{ij}$  are chosen in an attempt to make the desired memories stable states (or local energy minima) of the system. The particular prescription used in this work is standard and is given by

$$J_{ij} = \sum \xi_i^\mu \xi_j^\mu, \quad J_{ii} = 0 \quad (2)$$

where  $\xi_i^\mu$  is the value of the spin at site  $i$  in memory  $\mu$  and  $\mu = 1, 2, \dots, p$  labels the desired memories. In the original Hopfield model and in the work described here these memories are taken to be uncorrelated random patterns of the spins. Besides local minima near the programmed patterns, the prescription (2) leads also to the generation of large numbers of spurious minima (not close to any of the programmed memories) even for low values of  $p$ . As  $p$  is increased the minima corresponding to the programmed memories are eventually destabilized and the memory property of the network is lost.

In our study of the phase transitions in limited connectivity networks we need to define a connection scheme. Here we take our neurons as arranged on a  $d$ -dimensional hypercubic lattice and connect them to the nearest  $K$  neurons. The

coupling constants are determined as in (2) and then the  $J_{ij}$  for  $ij$  pairs not connected in our system are set to zero. We have studied the behavior for  $d=1, 2$ , and  $\infty$  (where  $d=\infty$  is implemented by using a random choice of connected neurons<sup>4</sup>).

A standard test to check the associative memory property of the network may be performed by initializing the network in one of the programmed memories  $\mu$ , and then running the network. The resulting state is compared to the memory  $\mu$  and the number of bit errors counted. If the number of bit errors is small, then we say that the network remembers the pattern  $\mu$ . A large number of errors means that the pattern has been forgotten. If we perform this stability test for a large number of choices of the set of programmed patterns we may determine the average number of bit errors  $\varepsilon$  characterizing a particular number of patterns,  $p$ . In practice<sup>7</sup>, for a fully connected Hopfield model, a first-order memory loss phase transition characterized by a discontinuous jump in  $\varepsilon$  occurs when the number of programmed patterns exceeds a critical value given by  $p = .14N$ . (In a finite system the discontinuity is, of course, rounded.) Thus the quantity  $\varepsilon$  functions like an order parameter for the memory-loss phase transition and the quantity  $\alpha$ , defined as  $p/N$ , determines the transition point. When the number of connections,  $K$ , is not equal to  $N$ , the appropriate parameter<sup>4</sup> is  $\alpha = p/K$ .

### **Boltzmann-like behavior**

We have observed empirically that these Hopfield-type networks are characterized by a Boltzmann-like distribution of the energy of the final states, where number of programmed patterns plays a role analogous to that of temperature. *A priori* one would not expect such behavior. Indeed, the differences between our neural network system and a standard statistical mechanical system are great. In a statistical mechanical system a fixed energy landscape is sampled ergodically at a given

temperature. In our neural system we have instead an  $\alpha$ -dependent energy surface, which is sampled in a highly non-ergodic manner: an irreversible network dynamics causes  $\varepsilon$  to be sampled only over those minima which are closest to our programmed patterns. In spite of these significant differences we find empirically that the  $\alpha$ -dependence of our energy histograms,  $P_\alpha(E)$ , is approximated surprisingly well by a Boltzman-like factor  $e^{-f(\alpha)E}$ . As a consequence we can use a reweighting scheme in the style of Ferrenberg-Swendsen to predict the energy distribution at a given  $\alpha$  from the distribution at a nearby  $\alpha$ . The reweighting scheme relies strongly on the exponential dependence of  $P_\alpha(E)$  on the quantity  $f(\alpha)E$ . Reweighting works well in these networks as illustrated in Fig. 1.

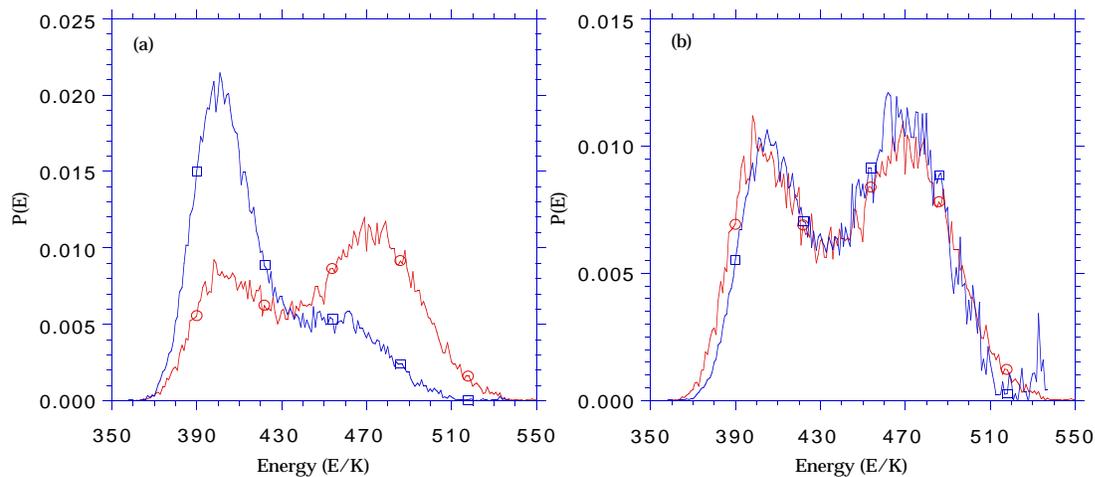


Figure 1. Histograms of  $P(E)$  vs.  $E$  for one-dimensional networks with  $N=400$ ,  $K=320$ ,  $p=60$  (squares) and  $p=70$  (circles): (a) as measured, (b) after reweighting by a factor  $e^{.024 E/K}$  for  $p=60$  and  $e^{-.004 E/K}$  for  $p=70$ .

### Lee-Kosterlitz finite size scaling

Lee and Kosterlitz<sup>1</sup> have recently presented a finite-size-scaling procedure which is useful for determining the order of a phase transition when the system size is too small to use the standard finite-size-scaling analysis. The method has been used in only a few cases so far. The work described here thus tests the practicality of the procedure as well as showing its power in analyzing our current problem. We describe the method briefly here.

First order phase transitions are characterized by two-phase coexistence. In our neural network system<sup>7</sup> the coexisting phases correspond to (i) the collection of minima close to the programmed memories (the low  $\alpha$  phase) and (ii) the collection of "spin glass" minima far from the programmed memories (the high  $\alpha$  phase). It is important to remember that the network dynamics samples only the *local* minima obtained by starting in a programmed memory. The memory-loss transition occurs when these local minima become *unstable*. The spin glass minima have already become the *global* minima at a value of  $\alpha$  well below our transition value  $\alpha_c$ .

We may also characterize the local minima which we reach in performing the stability test by their energy distribution  $P_\alpha(E)$ . A distribution  $P_\alpha(E)$  for a finite system near a first order phase transition will be double-peaked. This signature of two-phase coexistence would be observed only exactly at the transition point in an infinite system. For a finite system the coexistence is smeared over a region around the transition. Given the distributions  $P_\alpha(E)$  (or a distribution of some other relevant quantity such as the order parameter) a finite size approximation to  $\alpha_c$  may be obtained by finding the point at which the two peaks are of equal height. This point  $\alpha^*$  will converge to  $\alpha_c$  as the system size is increased.

If we observe the distribution  $P^*(E)$  or  $P^*(\epsilon)$  as the system size is increased, we will see that if there is a first order phase transition the peaks will sharpen and become more distinct as the system size increases. Figure 2 shows a series of histograms obtained from a 1-d neural network with a fixed connectivity fraction  $K/N$  (i.e.  $\eta = 1$ ).

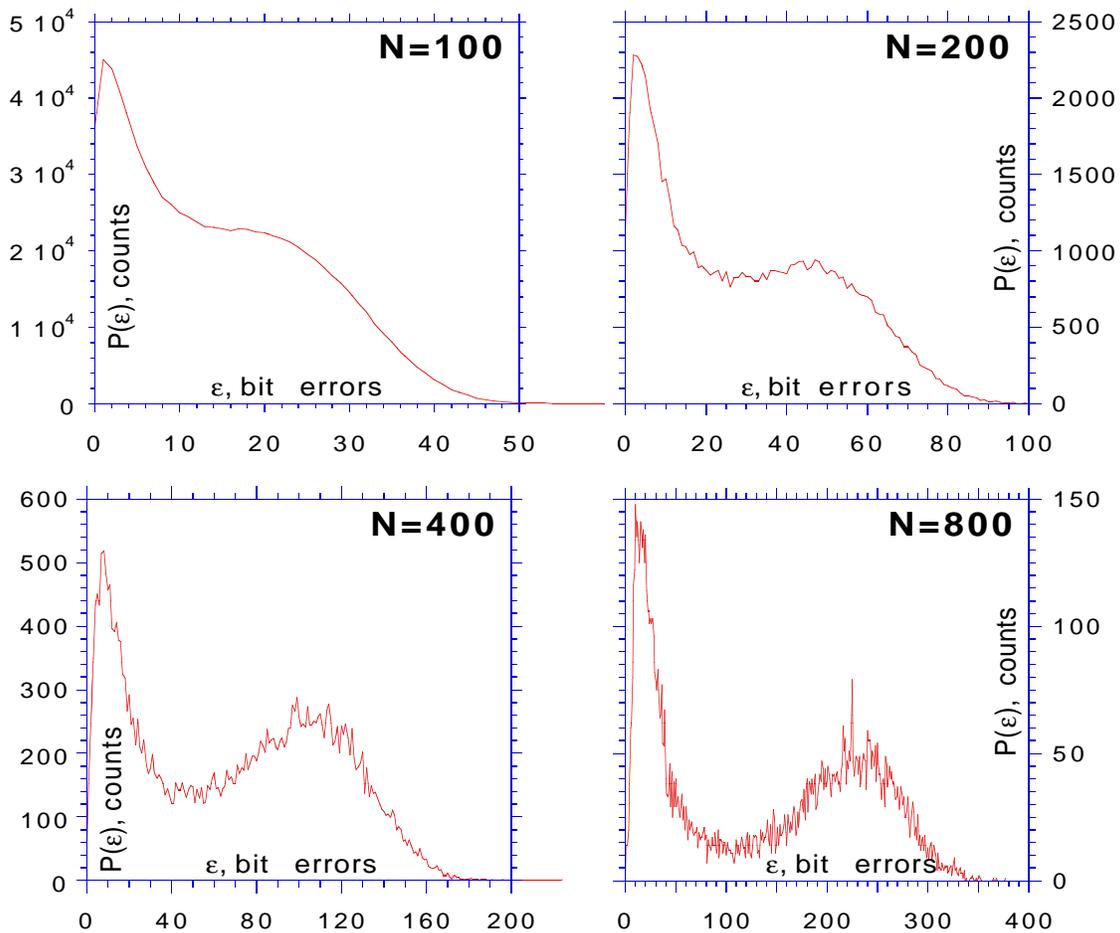


Figure 2: Histograms of  $P(\epsilon)$  vs.  $\epsilon$  for one-dimensional networks with  $N=100$ , 200, 400, 800. Connectivity  $K$  is proportional to  $N$ , i.e.  $\eta = 1$ , displaying the increasing peak to valley ratio which occurs for  $\eta > \eta_c$ . For these curves  $K = 70, 140, 280, 560$  respectively, and the number of stored patterns  $p = 22, 36, 64, 112$ . The vertical axis is the probability of observing a number of bit-errors  $\epsilon$ . The vertical axis is not normalized, so the numbers are the total number of observations.

Here we know from the mean field results of Canning and Gardner<sup>4</sup> that a first-order transition is expected. The sharpening of the peaks as the system size is increased is easily observed. In order to quantify this sharpening we follow Lee and Kosterlitz and measure the logarithm of the ratio of the probability at the peaks of the distribution to that at the valley between them. At a first order transition this quantity will diverge as system size is increased. At the critical value of  $\eta$  where the first order transition disappears it will tend to some nonzero constant, and below this critical value it will tend to zero, as the two-peaked structure disappears. Thus the behavior of the peak-to-valley ratio can be used to probe the phase diagram of a system.

As a practical matter it is difficult to pinpoint the value of  $\alpha$  at which the peaks are of equal height. To obtain a good approximation of the equal-peak-height distribution

we use the reweighting scheme described above and illustrated in Figure 1.

### **Peak to valley ratios**

We determined peak-to-valley ratios for both energy,  $P(E)$ , and bit error,  $P(\epsilon)$ , histograms. The results we present were obtained using the  $\epsilon$  histograms because, for the system sizes accessible to us, the separation between peaks was more pronounced in the  $\epsilon$  histograms. Results obtained from the energy histograms were consistent with those we will present.

Our goal in this study was to determine the effect on the phase transition of having  $K \sim N^\eta$  for various values of  $\eta < 1$ . Because we had limited values of  $K$  and  $N$  available to us it was not efficient to perform our calculations by sweeping  $K$  and  $N$

for every value of  $\eta$ . Instead we took our available values of  $K$  and  $N$  and interpolated the resulting data to obtain contours of constant peak-to-valley ratio. The resulting contour plots for  $d=1$  and  $\infty$  are shown on a log-log scale in Figures 3 and 4.

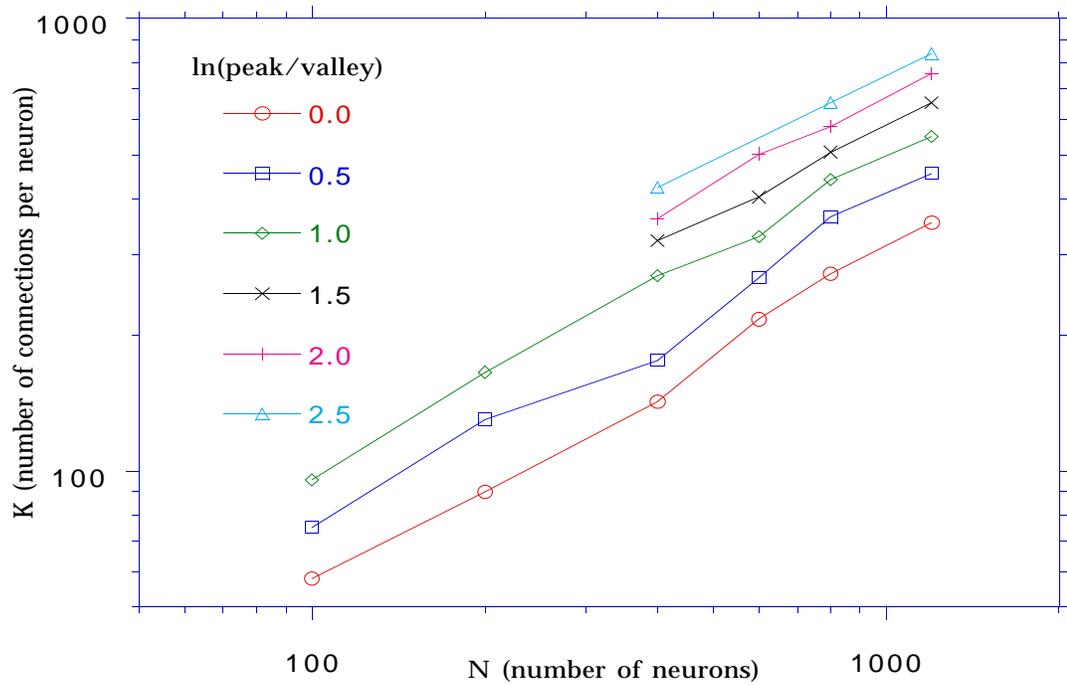


Figure 3. Contours of constant peak-to-valley ratio for one-dimensional networks.

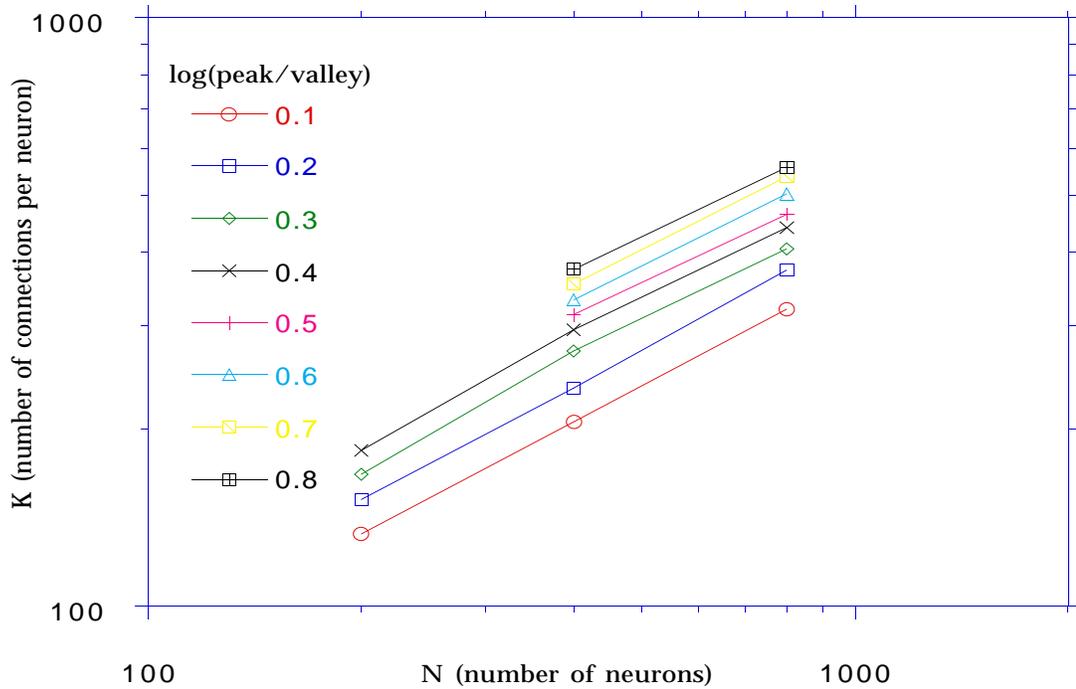


Figure 4. Contours of constant peak-to-valley ratio for  $\infty$ -dimensional (randomly connected) networks.

This graphical method provides good insight into the behavior of the system. We see that the contours of constant peak-to-valley ratio are parallel lines of slope  $\sim .65$ . This slope corresponds to the critical value of  $\eta_c$  below which the transition disappears. We can see this by inspecting Figures 3 and 4. If, for example, we imagine increasing  $N$  at constant  $K$  (i.e.  $\eta=0$ ) we see that contours of lower and lower peak-to-valley ratio will be crossed, indicating the disappearance of the two-peaked structure. This decrease will occur for any power  $\eta < \eta_c$ . On the other hand, if the line  $K = N$  (i.e.  $\eta=1$ ) shown on the plot is followed, contours of increasing peak-to-valley ratio are crossed, indicating a first order transition in that case. Such an

increase occurs for any power  $\eta > \eta_c$ . For  $K \sim N^{\eta_c}$  we find a constant value of peak-to-valley ratio, thus identifying the critical value of  $\eta$ . The different contours of slope  $\eta_c$  correspond to different values of the prefactor in  $K = CN^{\eta_c}$ .

Measuring the slopes of the contours to within our numerical accuracy we find  $\eta_c = .68 \pm .05$  for  $d = 1$ ,  $.71 \pm .03$  for  $d = 2$ , and  $\eta_c = .61 \pm .03$  for  $d = \infty$ . These values indicate that the value of  $\eta_c$  is surprisingly insensitive to (and perhaps even independent of) the dimensionality of the connection arrangement.

In conclusion, we have shown that the zero-temperature Hopfield network and its limited connectivity variants obey an approximate Boltzmann-like distribution where a function of the number of patterns plays the role of temperature. We have applied the Lee-Kosterlitz finite-size scaling scheme to the limited connectivity networks and have shown that the memory-loss phase transition disappears for  $\eta < \eta_c \sim .65$  where  $K \sim N^\eta$ . The value of  $\eta_c$  is nearly independent of dimension.

Finally we comment that the Lee-Kosterlitz technique is a very powerful tool for studying the memory-loss phase transition. Unlike the mean-field technique and the exact arguments used for dilute asymmetric connectivity, its applicability is not severely restricted by the choice of connection scheme. It can therefore be used to investigate the importance of variations on the connection scheme use here.

This work was supported by the Division of Educational Programs, ANL (D.B., C.C.) U. S. Department of Energy, BES-Materials Sciences, under contract W-31-109-ENG-38 and NSF contract RII-9003018 (K.J.S), NSF contract DMC-8857854 and FAA contract SC-91-209 (M.P.). Conversations with S. A. Solla and J. Lee are gratefully acknowledged.

## References

1. Lee, J. and J. M. Kosterlitz, Phys. Rev. Lett. **65**, 137 (1990).
  2. Hopfield, J. J., Proc. Nat. Acad. **79**, 2554 (1982).
  3. Amit, D. J., *Modelling Brain Function: The World of Attractor Neural Networks*, Cambridge University Press, Cambridge (1989).
  4. Canning, A. and E. Gardner, J Phys A **21**, 3275 (1988).
  5. Derrida, B., E. Gardner and A. Zippelius, Europhys Lett. **4**(2), 167 (1987).
  6. Noest, A. J., Phys Rev L **63**(16), (1989).
  7. Amit, D. J., H. Gutfreund and H. Sompolinski, Phys. Rev. A **32**(2), 1007 (1985).
- p