Optimal storage of invariant sets of patterns in neural network memories

1991 J. Phys. A: Math. Gen. 24 4197

(http://iopscience.iop.org/0305-4470/24/17/031)

J. Phys. A: Math. Gen. 24 (1991) 4197–4217. Printed in the UK

Optimal storage of invariant sets of patterns in neural network memories

Wojciech Tarkowski†, Marek Komarnicki‡ and Maciej Lewenstein†

† Institute for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

‡ Institute of Fundamental Technical Research, Polish Academy of Sciences, Świętokrzyska 21, 00-049 Warsaw, Poland

Received 6 February 1991, in final form 2 May 1991

Abstract. We investigate optimal conditions for the storage of invariant sets of patterns in neural network memories. The sets of patterns that we consider are highly correlated, non-random and invariant with respect to some symmetry transformations. We generalize Gardner's method to study the fractional volume in the space of neural interactions that allow for storage of invariant patterns. We demonstrate that optimal storage conditions correspond to non-trivial relations between the number of stored patterns, their stability, and other, symmetry-specific characteristics of patterns.

1. Introduction

Recent progress in the theory of attractor neural networks [1, 2] has became possible due to an extensive use of statistical physics methods (see [3] for an extensive presentation of the theory over the last five years). In the recent years a lot of effort has been devoted to applications of such methods to the theory of learning (for a review see [4]), i.e. the determination of interactions for which network dynamics leads to desired attractors. Such works have been initiated by Gardner [5] who derived the optimal conditions for which storage of a set of random patterns is possible. We apply Gardner's program in the present paper to the problem learning sets of patterns that are invariant with respect to a group of symmetry transformations.

The standard models of attractor neural networks consist of a set of N binary neurons. We denote the neurons states by $\sigma_i = \pm 1$. Neural dynamics in the absence of noise takes the form of the deterministic updating rule,

$$\sigma_i(t + \Delta t) = \operatorname{sign}\left(\sum_{j \neq i} J_{ij} \sigma_j(t)\right) \tag{1}$$

where J_{ij} denotes the synaptic connection matrix. A set of p specified patterns, ξ_j^μ, where $\mu = 1, \ldots, p$, and $j = 1, \ldots, N$ is said to be stored in the memory described by the dynamics (1) if ξ are the stationary states of this dynamics. The matrix J_{ij} for which ξ are fixed points of the dynamics should satisfy for each μ and i the inequalities

$$\xi_i^\mu \sum_{j \neq i} J_{ij} \xi_j^\mu > \kappa \sqrt{N} \tag{2}$$

0305-4470/91/174197+21$03.50

provided the connection matrix is, for each i, normalized according to

$$\sum_{j\neq i} J_{ij}^2 = N \tag{3}$$

The states ξ are stable even for $\kappa = 0$. Larger values of the stability parameter κ ensure larger stability domains of the patterns.

The probability that solution of the inequalities (2) exists is, according to Gardner [5], given by the fractional volume in the interaction space of the matrices J_{ij},

$$V_T = \frac{\prod_i\{\int\prod_{j\neq i} \mathrm{d}J_{ij} \prod_\mu \Theta[\xi_i^\mu \sum_{j\neq i}(J_{ij}/\sqrt{N})\xi_j^\mu - \kappa]\delta(\sum_{j\neq i} J_{ij}^2 - N)\}}{\prod_i[\int\prod_{j\neq i} \mathrm{d}J_{ij}\,\delta(\sum_{j\neq i} J_{ij}^2 - N)]}. \tag{4}$$

Gardner's approach allows to determine in the statistical sense whether or not the solutions of the inequalities (2) exist. The existence of such solutions may then serve as a necessary condition for convergence of various learning algorithms [5–8].

The original idea of Gardner has been developed in various directions recently (for a review see [9]). The existence of optimal solutions of inequalities (2), that is, those that contain the smallest possible amount of errors, can be formulated as a problem of canonical ensemble calculation [10]. Different restrictions of the matrices J_{ij} have been considered. Gardner's program has in particular been applied to matrices with $J_{ij} = \pm 1$ [10–12], and to matrices characterized by a definite ratio of symmetric and asymmetric parts [13]. All these generalization of Gardner's approach may also be studied in the context of the problem considered in the present paper.

One of the basic questions in the theory of learning concerns learning of structured sets of data. Correlated biased patterns have already been investigated by Gardner [5] and have been shown to allow better information capacity of the network. Specific learning rules have been derived for hierarchically organized data [14, 15]. Learning algorithms for strongly correlated sets of patterns have been proposed by several authors [4]. Some of these algorithms assure optimal stability (maximal values of the κ parameter at the given capacity, cf works by Krauth and Mezard [6], Schmitz *et al* [8]). The problem of learning of structured patterns within a given context has been addressed recently [16, 17].

Among the problems of storing and retrieval of structured sets of data one of the most important is that of *perceptual invariance*, i.e. the ability of cognitive systems to recognize transformed versions of remembered patterns. Among possible transformations there are translations, rotations, scale transformations, etc. The problem of perceptual invariance is central for cognitive psychology and has been widely discussed for multi-layered feed-forward networks (see for instance [18]). In the context of attractor neural networks this problem has been addressed by von der Malsburg and Bienenstock [19, 20], and Kree and Zippelius [21]. These authors considered networks that recognize topological features of images. All homeomorphically invariant images are recognized in such networks as prototype graphs.

Our aim in the present paper is to study *optimal storage of invariant sets of data* in attractor neural networks in the other sense. We shall demand that *elements of a given set of patterns invariant with respect to some group of transformations are fixed points of the network dynamics (1)*. All invariant images are recognized as such in our models. Such a property may turn out to be useful in several problems of pattern recognition.

The plan of the paper is the following. In section 2 we formulate the problem and define invariant sets of patterns. We consider sets of highly correlated, non-random states. Section 3 contains a detailed description of an alternative approach to the evaluation of fractional volume (4) for the invariant sets of patterns. Analytic evaluation of the storage capacity for such sets is discussed in section 4. We derive variationally exact boundaries of the region in the control parameter space in which solutions of equation (2) exist. The results are presented and discussed in section 5. We demonstrate that optimal storage conditions correspond to non-trivial relations between the number of stored patterns, their stability, and other, symmetry-specific characteristics of patterns. Three appendices at the end of the paper contain technical details of the calculations.

2. Invariant sets of patterns

In most of the neural networks models one considers random patterns, i.e. patterns that at least to some extend are constructed according to stochastic rules [3]. This fact implies the necessity of use of sophisticated methods of the theory of disordered systems, such as replica trick. In this paper we consider *non-random patterns*, however, since we want to extract possible effects of group invariance. For this reason the use of replica trick is not needed.

One possible implementation of our idea is to represent the pattens in d-dimensional lattices and to consider patterns invariant with respect to translations. The simplest example of such sets of patterns corresponds to $d = 1$.

We consider a set of N neurons located for simplicity on a one-dimensional torus rather than a one-dimensional line. Torus topology allow easier treatment of boundary effects but is by no means necessary for our approach. As one of the patterns ξ^{μ}, the one with $\mu = 1$, we take

$$\xi_i^1 = +1 \tag{5}$$

for $1 \leqslant i \leqslant d$, and $\xi_i^1 = -1$ otherwise. Such a state represents a one-dimensional 'pixel' of d active ($\xi = 1$), neurons surrounded by a 'sea' of inactive neurons ($\xi = -1$). We consider now all the translated versions of the pattern ξ^1, where the smallest possible translation is by a neurons. Evidently $1 \leqslant a \leqslant N$. The pattern ξ^2, for instance, is given by

$$\xi_i^2 = +1 \tag{6}$$

for $1 + a \leqslant i \leqslant d + a$, and $\xi^2 = -1$ otherwise, etc.

For the torus topology translations are the same as rotations, so that we choose a to divide N. The number of different patterns in the set of all 'translated' versions of ξ^1 is therefore $p = \alpha N = N/a$, and is a natural number†. We denote this set of patterns by S. We ask in the present paper whether it is possible to store this set of patterns in the neural network memory of the Hopfield-type. We also investigate optimal storage conditions like, for instance, the conditions for maximal stability of stored patterns at the given capacity of the network etc.

† The discrete symmetry group in question is isomorphic to Z_p, i.e. group of natural numbers with addition modulo p.

Obviously the presented construction of non-random patterns may easily be generalized to multiple dimensions and to multidimensional symmetry groups, such as two-dimensional translations on a plane. Torus topology, as we mentioned does not have to be used, when the effects of plane boundaries are taken into account. Invariance with respect to a whole group of isometries (translations, rotations, reflections) may be considered.

Another important generalization is to consider the basic states of arbitrary form, instead of 'pixel' shape described by equation (5). One can also include an element of randomness, and consider patterns that have 'fuzzy' shape.

For any case the sets of data that we consider are highly correlated. When storing them one usually encounters the appearance of *stable spurious memories*. Instead of having elementary 'pixel' states as the only stationary states, one encounters two-'pixel' states, three-'pixel' states, and so forth. We do not study this problem in the present paper for two reasons. First of all, the spurious states that we described may in fact be useful for some applications and their existence may even be desired. Secondly, the question that we ask is how to store optimally the set of patterns $\mathcal{S}$ in the memory, independently of the appearance of spurious states. The questions of typical performance of the memory in optimal conditions in general and of the stability of spurious states in particular belong, on the other hand, to a different category. They may be answered by studying applications of particular learning algorithms (such as various versions of the minimal overlap algorithm [6,8]) in order to store the set $\mathcal{S}$. We leave the discussion of those problems to a separate paper. Here we concentrate on the interplay between the group invariance and optimal storage properties in the frame of Gardner's program.

3. Fractional volume in the space of interactions

The calculation of the fractional volume in the space of interactions (4) for the set of non-random patterns invariant with respect to one-dimensional translations and defined as is section 2, equations (5), (6), does not require the use of the replica trick. It does require, however, some modifications to the original approach of Gardner that we descibe below.

The fractional volume V_T may be written as

$$V_T = \prod_{i=1}^{N} V_i \tag{7}$$

where each of the partial fractional volumes V_i is defined as

$$V_i = \frac{\int \prod_{j\neq i} \mathrm{d}J_{ij} \prod_{\mu} \Theta\left[\xi_i^\mu \sum_{j\neq i}(J_{ij}/\sqrt{N})\xi_j^\mu - \kappa\right] \delta\left(\sum_{j\neq i} J_{ij}^2 - N\right)}{\mathcal{N}_i}. \tag{8}$$

The normalization constant in the above expression is given by

$$\mathcal{N}_i = \int \prod_{j\neq i} \mathrm{d}J_{ij}\, \delta\Big(\sum_{j\neq i} J_{ij}^2 - N\Big). \tag{9}$$

After representing the Dirac δ function in the form of a Fourier integral, the integral over J_{ij} becomes Gaussian and can be performed analytically. The remaining integration over the Fourier variable can be evaluated in the $N \to \infty$ limit using the saddle point technique,

$$\mathcal{N}_i = \frac{1}{2\pi \mathrm{i}} \int_C \mathrm{d}s \exp\left[N\left(s - \frac{1}{2}\ln s + \frac{1}{2}\ln\pi\right)\right]$$
$$= \mathcal{C} \exp\left[\frac{1}{2}N\left(1 - \ln\frac{1}{2}\right)\right]. \tag{10}$$

C denotes in the above expression integration contour for s going from $-\mathrm{i}\infty$ to $+\mathrm{i}\infty$. $\mathcal{C}$ is a constant that for large N behaves as $\ln \mathcal{C}/N \to 0$. The asymptotic behaviour of the integral (10) is governed by the value of the integrand at the saddle point $s = \frac{1}{2}$.

Similar techniques can be used to evaluate the numerator of the expression (8), which we denote as Φ_i. We represent the Dirac δ and Θ functions as Fourier integrals [5] and use the following representation

$$\Phi_i = \frac{1}{2\pi \mathrm{i}} \int_C \mathrm{d}s \int \prod_{j\neq i} \mathrm{d}J_{ij} \exp\left[-s\left(\sum_{j\neq i} J_{ij}^2 - N\right)\right]$$
$$\times \int_\kappa^\infty \prod_\mu \mathrm{d}\lambda_\mu \int \prod_\mu (\mathrm{d}x_\mu/2\pi)$$
$$\times \exp\left[\mathrm{i}\sum_{\mu,j} x_\mu\left(\xi_i^\mu \sum_{j\neq i}\frac{J_{ij}}{\sqrt{N}}\xi_j^\mu - \lambda_\mu\right)\right]. \tag{11}$$

Gaussian integrations over J_{ij}, and then over x_μ can be performed explicitly. We obtain then the following expression

$$\Phi_i = \frac{1}{2\pi \mathrm{i}}\frac{1}{(2\pi)^{\alpha N}}\frac{(\sqrt{4\pi})^{\alpha N}}{(\det \mathbf{M}^i)^{1/2}} \int_C \mathrm{d}s \exp\left[N\left(s - \frac{1-\alpha}{2}\ln s + \frac{1}{2}\ln\pi\right)\right]$$
$$\times \int_\kappa^\infty \prod_\mu \mathrm{d}\lambda_\mu \exp(-s\sum_{\mu,\mu'} \lambda_\mu (M^i)^{-1}_{\mu\mu'}\lambda_{\mu'}) \tag{12}$$

where $(M^i)^{-1}_{\mu\mu'}$ are the matrix elements of the inverse matrix of the overlap matrix $\mathbf{M}^i$. The elements of the overlap matrix are, in turn, given by

$$M^i_{\mu\mu'} = \frac{1}{N}\sum_{j\neq i} \xi_i^\mu \xi_j^\mu \xi_i^{\mu'} \xi_j^{\mu'}. \tag{13}$$

Note the overlap matrix $\mathbf{M}^i$ is i-dependent. It has the dimension $p = \alpha N$, and, as we shall discuss later, is positively defined for the considered invariant sets of data (see appendix A)†. For this reason we may consider the contour C to lie to the right of the imaginary axis ($\mathrm{Re}\, s \geqslant 0$). In particular we may consider positive real values of s as candidates for saddle points.

† It is worth stressing that even for random, statistically independent patterns the matrix $\mathbf{M}$ can be shown to be non-negative [22].

Let us introduce the function

$$f(s,\alpha,\kappa)=\int_{\kappa}^{\infty}\mathcal{D}\lambda\exp\Big(-s\sum_{\mu,\mu'}\lambda_\mu(M^i)^{-1}_{\mu\mu'}\lambda_{\mu'}\Big) \tag{14}$$

where we used the shortened notation for the multiple integral over λ_μ. We also define

$$g(s,\alpha,\kappa)=s-\frac{1-\alpha}{2}\ln s+\frac{1}{N}\ln f(s,\alpha,\kappa). \tag{15}$$

Using the above introduced notation we may write down

$$\Phi_i=\frac{1}{2\pi\mathrm{i}}\frac{1}{(2\pi)^{\alpha N}}\frac{(\sqrt{4\pi})^{\alpha N}}{(\det\mathbf{M}^i)^{1/2}}\int_C \mathrm{d}s\,\mathrm{e}^{Ng(s,\alpha,\kappa)}. \tag{16}$$

The integral over s may now be evaluated using the saddle point method. We show below that for some values of the parameters α, κ the saddle point corresponds to real s.

In the limit of $s\to 0$ we introduce new integration variables $y_\mu=\sqrt{s}\lambda_\mu$, so that

$$f(s,\alpha,\kappa)=\frac{1}{(\sqrt{s})^{\alpha N}}\int_{\sqrt{s}\kappa}^{\infty}\mathcal{D}y\exp\Big(-\sum_{\mu,\mu'}y_\mu(M^i)^{-1}_{\mu\mu'}y_{\mu'}\Big) \tag{17}$$

From (17) we easily obtain the asymptotic behaviour of f for $s\to 0$,

$$\ln f(s,\alpha,\kappa)=-\frac{\alpha N}{2}\ln s+\text{constant}. \tag{18}$$

The above result implies that for $s\to 0$

$$g(s,\alpha,\kappa)\to\infty. \tag{19}$$

Direct inspection of the definition of f suggests on the other hand that for $s\to\infty$

$$f(s,\alpha,\kappa)\to\mathrm{e}^{-NA(\alpha,\kappa)s} \tag{20}$$

where $A(\alpha,\kappa)$ is an intensive function of the control parameters α, κ and others if there are any. This function is given by

$$A(\alpha,\kappa)=\frac{1}{N}\min_{\lambda_\mu\geqslant\kappa}\Big(\sum_{\mu,\mu'}\lambda_\mu(M^i)^{-1}_{\mu\mu'}\lambda_{\mu'}\Big). \tag{21}$$

From the expression (20) we obtain asymptotic behaviour of the function g for $s\to\infty$,

$$g(s,\alpha,\kappa)\to s(1-A(\alpha,\kappa))-B(\alpha,\kappa)\ln s+\cdots \tag{22}$$

where $B(.,.)$ is another intensive function of control parameters. Evidently, g tends to ∞ provided $A(\alpha,\kappa)<1$. In such a case the function $g(s,\alpha,\kappa)$ has a minimum for some s in the interval $(0,\infty)$. This value of s is a saddle point value for the contour integral (16). Obviously, the value of s at the saddle point tends to infinity as $A(\alpha, \kappa)$ approaches 1. At the same time the value of g tends logarithmically to $-\infty$. Fractional volume shrinks to zero when

$$A(\alpha, \kappa) = 1 \tag{23}$$

or, in another form,

$$\min_{\lambda_\mu \geqslant \kappa} \left(\sum_{\mu, \mu'} \lambda_\mu (M^i)^{-1}_{\mu\mu'} \lambda_{\mu'} \right) = N \tag{24}$$

From (24) one obtains the analogues of Gardner's critical curve $\alpha_c(\kappa)$ that determines the boundary of the phase in the control parameters' space in which solutions of inequalities (2) exist.

There are specific aspects of our method of calculation of the volume in the interaction space that require more discussion. First of all, note that until now we have not attempted to discuss the possible randomness of the patterns ξ_i^μ. In another words, we have not calculated the average over such randomness. There are basically two ways of treating random patterns within our method.

The first way consists in calculating the average over ξ^μ in (15). This requires averaging of the logarithm of $f(s, \alpha, \kappa)$, and may be performed using the standard replica method. As a result one easily recovers the original Gardner result. Our method works well in such a case, but is nothing more but a modification of the standard approach, as is for instance the cavity approach of Mezard [23].

The second way of averaging the logarithm of the fractional volume consist in the direct use of expression (24). This is, however, a very difficult task, since for typical random realizations of the matrix $\mathbf{M}^i$ the configuration λ_μ that minimizes the quadratic form on the LHS of (24) itself depends on ξ^μ and is very hard to find. Nevertheless, one may try to find it for each realization of the quenched variables using variational methods. For a given α and a given realization of ξ^μ, one may then find a value of κ that fulfils (24), but is still ξ^μ dependent. Finally, repeating such a procedure for different realizations of ξ^μ, one may calculate the average and recover the Gardner result. In this case our method is much more complicated than the replica or the cavity method, but gives some new insight into the problem discussed. It relates, namely, the problem of calculation of the fractional volume to properties of the random matrix $\mathbf{M}^i$. This relation, in turn, allows us to apply random matrix theory [24] to evaluate the critical curve.

In particular, by looking at the averaged density of the eigenvalues of the matrix $\mathbf{M}^i$, one can show that typical realizations of $\mathbf{M}^i$ have spectra that are bounded from above and from below. This fact us allows immediately to derive upper and lower bounds on Gardner's critical curve [22]. Such bounds are quite precise for not too small values of κ. Moreover, such bounds may be quite easily derived for various distributions of ξ^μ and various network architectures. In [22] we have derived such bounds for biased patterns, as well as for randomly distorted versions of the 'pixel' patterns that are discussed in the present paper. Such applications of our method allow to determine critical curves only approximately (within the obtained bound). The advantage of our approach, however, is that we are able to calculate precise estimates of critical curves for a much more general class of random patterns that is accessible to the replica or the cavity method.

Finally, our method may be directly used for non-random patterns, since in such a case no averaging over the quenched disorder is needed. In the present paper, in fact, we apply this method to non-random 'pixel' patterns (see the next section). Similarly, we may apply the method to any other highly 'structured' and organized set of data. The investigations of optimal storage conditions and critical curves in such cases allow us to isolate and study the interplay between the structure of the set of patterns and properties of the memory. Such problems cannot be easily resolved with the help of the standard methods.

4. Calculation of critical storage conditions

In order to derive explicit form of the critical condition (24) we have to specify the overlap matrix **M**. We consider here the particular example of the set of invariant patterns — the set of 'pixel' patterns introduced in section 2. For this case the matrix **M** defined as

$$M^i_{\mu\mu'} = \frac{1}{N}\sum_{j\neq i}\xi^\mu_i\xi^\mu_j\xi^{\mu'}_i\xi^{\mu'}_j = \frac{1}{N}\sum_{j=1}^{N}\xi^\mu_i\xi^\mu_j\xi^{\mu'}_i\xi^{\mu'}_j - \frac{1}{N} \tag{25}$$

has the following form

$$M^i_{\mu\mu'} = \frac{4d}{N}\delta_{\mu,\mu'} + \xi^\mu_i\xi^{\mu'}_i\left(\frac{4(d-a)}{N}(\delta_{\mu,\mu'+1}+\delta_{\mu,\mu'-1}) + \frac{4(d-2a)}{N}(\delta_{\mu,\mu'+2}+\delta_{\mu,\mu'-2})+\cdots+\frac{4(d-ra)}{N}(\delta_{\mu,\mu'+r}+\delta_{\mu,\mu'-r})\right) + \left(1-\frac{4d}{N}\right)\xi^\mu_i\xi^{\mu'}_i - \frac{1}{N}. \tag{26}$$

The matrix **M** has the form generic for sets of data invariant with respect to some group of geometrical transformations. First two rows of (26) describe a 'short range' part that is non-zero for neighbouring μ and μ' only. Note that for the case of sets of patterns that have a compact shape and are invariant with respect to translations or rotations, the concept of neighbouring μ and μ' can always be defined precisely. The short-range part of **M** consists in the presently considered case of r terms, where $r = [d/a]_*$, while the function $[.]_*$ is defined as follows: for any real, non-integer x, $[x]_*$ is the integer part of x. For integer x, $[x]_* = x - 1$. If we take any of the 'pixel' patterns and start to translate it in one direction (say to the left), exactly r of the translated 'pixel' configurations will partially cover the starting configuration. Each of the neurons is active ($\xi^\mu_i = 1$) for r or $r + 1$ patterns.

The remaining part of the matrix **M** is a 'long-range' part. This part (last row of (26)) has the form of the sum of several (in the present case two) projection operators and does not have any trace of the metric structure. The generic form of the matrix **M** allows us to find its eigenvectors, eigenvalues and its inverse analytically for a broad class of invariant sets of data using Fourier transform techniques.

In order to get an insight into the critical condition (24) we have to achieve the two tasks.

(i) We have to evaluate the action of the inverse of the matrix $\mathbf{M}$ on the vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{\alpha N})$.
(ii) We have to find the minimum of the function $\boldsymbol{\lambda}\mathbf{M}^{-1}\boldsymbol{\lambda}$.

The first task is done by solving the linear equation

$$\lambda_\mu = \sum_{\mu'} M^i_{\mu\mu'} x_{\mu'} \tag{27}$$

with respect to x_μ. This is done easily after introduction of two unitary transformations.

- multiplication by ξ_i^μ

$$\bar{\lambda}_\mu = \xi_i^\mu \lambda_\mu \tag{28}$$

$$\bar{x}_\mu = \xi_i^\mu x_\mu \tag{29}$$

- discrete Fourier transform

$$\lambda(k) = \frac{1}{\sqrt{\alpha N}} \sum_{\mu=0}^{\alpha N-1} \bar{\lambda}_\mu e^{i\omega_k \mu} \tag{30}$$

$$x(k) = \frac{1}{\sqrt{\alpha N}} \sum_{\mu=0}^{\alpha N-1} \bar{x}_\mu e^{i\omega_k \mu} \tag{31}$$

with the Fourier frequencies given by

$$\omega_k = \frac{2\pi}{\alpha N} k$$

for $k = 0, \ldots, \alpha N - 1$.

Elementary calculation then yields

$$\lambda(k) = f(k)x(k) - \frac{1}{N}\xi_i(k) \sum_{k'} \xi_i^*(k')x(k') \tag{32}$$

where

$$f(k) = \frac{4d}{N} + \left(1 - \frac{4d}{N}\right)\alpha N \delta_{k,0} + \sum_{r'=1}^{r} \frac{8(d - r'a)}{N} \cos(r'\omega_k) \tag{33}$$

while $\xi_i(k)$ is the Fourier transform of ξ_i^μ. The matrix $\mathbf{M}$ after performing the transformations (28) – (31) becomes a sum of a diagonal matrix and a projection operator. The solution of (32) can be easily found and reads

$$x(k) = \frac{\lambda(k)}{f(k)} + \frac{1}{N}\frac{\xi_i(k)}{f(k)} \frac{\sum_{k'} \xi_i^*(k')\lambda(k')/f(k')}{1 - \frac{1}{N}\sum_{k'} \xi_i^*(k')\xi_i(k')/f(k')}. \tag{34}$$

In the above formula $f(k)$ must be different from 0. As we discuss in appendix A, $f(k)$ is typically greater than zero, except for even N, $\omega_k = \pi$, even r and for

$d = ra$. Fortunately, for this particular case $|\xi_i(k)|^2$ vanishes as well. The detailed discussion of this apparent singularity is contained in appendix B.

Using the solution (34) we may now turn to the second task — evaluation of the minimum over $\lambda_\mu \geqslant \kappa$ of the function

$$\boldsymbol{\lambda}\mathbf{M}^{-1}\boldsymbol{\lambda} = \sum_k \lambda^*(k)x(k).$$

We shall construct this minimum using the variational approach. It is obvious that the minimum is obtained on the boundary of the set S_κ of $\lambda_\mu \geqslant \kappa$. A natural candidate for the minimum is therefore the 'edge' point $\boldsymbol{\lambda}^0_{\rm pr} = (\kappa, \ldots, \kappa)$. Note that $\lambda^0_{\rm pr}(k) = \kappa\xi_i(k)$. Inserting the probe vector $\boldsymbol{\lambda}^0_{\rm pr}$ into (24) we obtain after some algebra the approximate formula for the critical storage condition. Firstly, denoting

$$\mathcal{A} = \frac{1}{N}\sum_{k'=1}^{\alpha N-1}\frac{\xi_i^*(k')\xi_i(k')}{f(k')} = \frac{1}{\kappa N}\sum_{k'=1}^{\alpha N-1}\frac{\lambda^*_{\rm pr}(k')\xi_i(k')}{f(k')} \tag{35}$$

we immediately obtain from (24) the expression

$$\frac{1}{\kappa^2} = \mathcal{A} + \frac{\mathcal{A}^2}{1-\mathcal{A}} \tag{36}$$

so that finally

$$\frac{1}{\kappa^2} \geqslant \frac{4\sum_{k=1}^{\alpha N-1}\frac{1}{\alpha N^2 f(k)}\frac{\sin^2\omega_k(r+\delta_i)/2}{\sin^2\omega_k/2}}{1-4\sum_{k=1}^{\alpha N-1}\frac{1}{\alpha N^2 f(k)}\frac{\sin^2\omega_k(r+\delta_i)/2}{\sin^2\omega_k/2}} \tag{37}$$

In the above formula we have introduced the parameter δ_i which is equal 0 or 1 depending on whether the ith neuron belongs to, i.e. is active ($\xi_i^\mu = 1$) in exactly r or $r+1$ 'pixel' patterns, respectively. We have also neglected in (37) the term corresponding to $k = 0$ that enters the definition of $\mathcal{A}$. For $k = 0$, $|\xi_i^\mu(0)|^2$ as well as $f(0)$ are both intensive, so that the ratio $|\xi_i^\mu(0)|^2/Nf(0)$ that enters (35) is negligible for $N \to \infty$. Direct inspection of (33) indicates, on the other hand, that $Nf(k)$ is an intensive quantity for $k > 0$ and has a well defined limit for $N \to \infty$. That fact indicates also that the series on the RHS of (37) has also a well defined limit and becomes the integral

$$\frac{1}{\kappa^2} \geqslant \frac{\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\omega\,\frac{1}{\tilde{f}(\omega)}\frac{4\sin^2\omega(r+\delta_i)/2}{\sin^2\omega/2}}{1-\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\omega\,\frac{1}{\tilde{f}(\omega)}\frac{4\sin^2\omega(r+\delta_i)/2}{\sin^2\omega/2}} \tag{38}$$

where

$$\tilde{f}(\omega) = 4d + \sum_{r'=1}^{r} 8(d - r'a)\cos(r'\omega). \tag{39}$$

Unfortunately, it turns out that the 'edge' point does not always correspond to the minimum of the expression (24). The inequality (38), nevertheless, describes the exact critical curve quite well.

In order to see that $\lambda^0_{\rm pr}$ does not correspond to the minimum of the quadratic form $\boldsymbol{\lambda}\mathbf{M}^{-1}\boldsymbol{\lambda}$ over $\lambda \in S_\kappa$, we easily observe that the necessary and sufficient condition for this is:

$$(\mathbf{M}^{-1}\boldsymbol{\lambda})_\mu = x_\mu \geqslant 0 \tag{40}$$

for all μ. Direct calculation of x_μ (see appendix C) shows that it is true only for $d \leqslant a$ (i.e. for $r = 0$) and it is not the case in general. With increasing d some components of the vector x_μ become more negative. The negative components, however, decrease exponentially when μ varies, and one can expect that their 'effective' number remains, in a sense, finite. For instance, for the case $r = 1$ and $\delta_i = 0$, the negative components appear for $\mu = \mu(i) \pm 2$, ± 4, etc, where $\mu(i)$ denotes the single value of μ for which $\xi^\mu_i = -1$. A better approximation for the minimum of the expression (24) is obtained for

$$\boldsymbol{\lambda}^2_{\rm pr} = (\kappa, \ldots, \lambda_{\mu(i)-2}, \kappa, \ldots, \lambda_{\mu(i)+2}, \kappa, \ldots) \tag{41}$$

then for

$$\boldsymbol{\lambda}^4_{\rm pr} = (\kappa, \ldots, \lambda_{\mu(i)-4}, \kappa, \lambda_{\mu(i)-2}, \kappa, \ldots, \lambda_{\mu(i)+2}, \kappa, \lambda_{\mu(i)+4}, \kappa, \ldots) \tag{42}$$

etc. The values of $\lambda_{\mu(i)\pm 2}$ in the case (41), and $\lambda_{\mu(i)\pm 2,\pm 4}$ in the case (42) have to be determined using standard differential calculus (for the detailed discussion see appendix C). In this way for every value of d we are able to construct a systematic approximation to the minimum of the quadratic form (24) using elementary methods.

The detailed numerical analysis of the above result is presented in the next section.

5. Discussion of the results

We start this section with the detailed discussion of the approximate result (37) and (38) that corresponds to the simplest variational probe function that describes the minimum of the form (24).

The critical storage conditions (37), (38) define the boundary of the region of parameters $\alpha = 1/a$, d and κ for which there exists at least one solution J_{ij} in the space of interactions that fulfil the condition (2) for a given i. Since in general there is always an extensive number of neurons that are covered by $r + 1$ 'pixels' and correspond to $\delta_i = 1$, as well as those that are covered by r 'pixels' and correspond to $\delta_i = 0$, the inequalities (37) and (38) must be fulfilled for every i, i.e. for both values $\delta_i = 0$, 1.

We remember once more that d and a are natural numbers, and that a must divide N. d is the width of the 'pixel' (the number of is such that $\xi^\mu_i = 1$) and must be smaller than N. Since the theory is invariant with respect to the simultaneous flip of all neurons, the results must be invariant with respect to the exchange $d \to N - d$. We shall consider in the following only the case of $d \ll N$ which is interesting from the point of view of pattern recognition theory and for which the limit $N \to \infty$ was

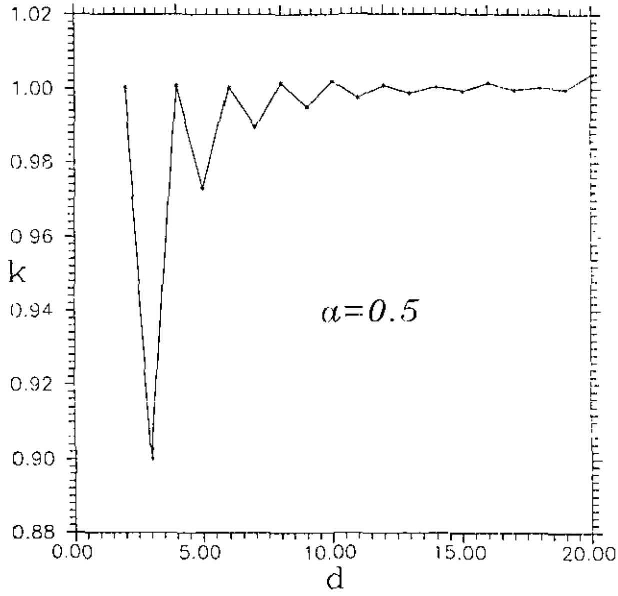

Figure 1. The maximal stability parameter κ as a function of 'pixel' size d for the network capacity $\alpha = 0.5$; $N = 5000$.

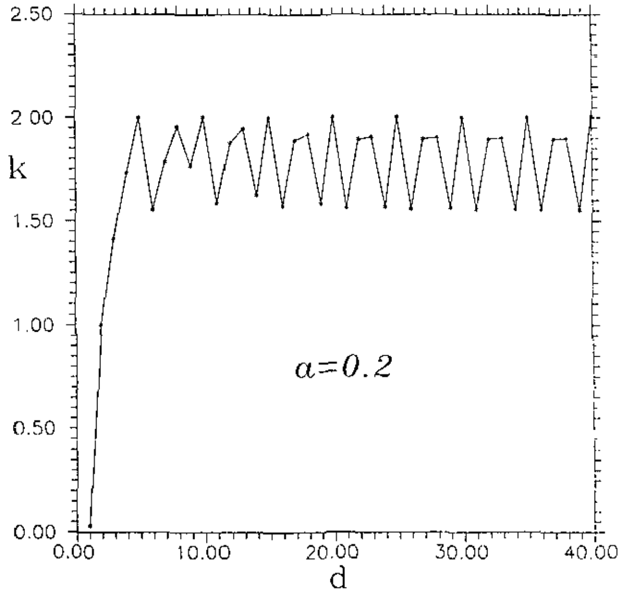

Figure 2. Same as figure 1, but for $\alpha = 0.2$

considered. We present results of (37) for $N = 5000$ which corresponds already to the asymptotic limit of (38). For larger values of N results are the same.

In figures 1, 2 and 3 we present the curves that correspond to the maximal stability parameter κ as a function of d for a given value of α. For large values of α (see for instance figure 1 for $\alpha = 0.5$), close to the upper capacity limit $\alpha_{\max} = 1/a_{\min} = 1$,

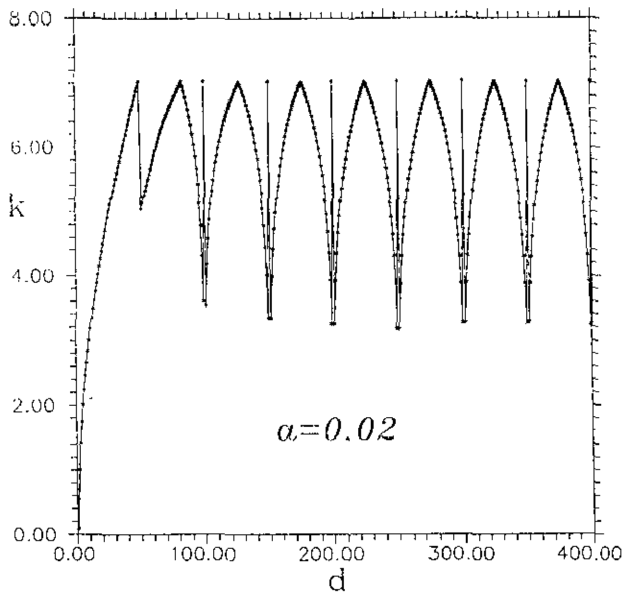

Figure 3. Same as figure 2, but for $\alpha = 0.02$

the maximal values of κ behave more or less regularly. κ_{max} grows from practically zero, and saturates at some value for larger d. Note, however, that the approach to the asymptotic limit is oscillatory. The oscillations are relatively weak and have period 2. One may expect that even weak randomness will tend to smooth out these effect [22]. The oscillatory structure is visible much better for smaller $\alpha = 0.2$ in figure 2; oscillations are here relatively larger and have the period $a = 5$. After somewhat irregular behaviour for small d the curve aproaches regular oscillatory phase that does not seem to be damped. Still one may argue that even weak disorder (such as for instance fluctuations of d) will destroy the observed structure.

This structure is fully developed for small values of $\alpha = 0.02$, as presented in figure 3. Again, after initial growth, κ_{max} depends periodically on d. Oscillations are large (50% of amplitude) and have the period equal to $a = 1/\alpha = 50$. They seem to be stable with respect to pattern randomness [22]. The two parts of the 'gothic arcs' come from the two different conditions (37) for $\delta_i = 0$ (left 'arc') and $\delta_i = 1$ (right 'arc'), respectively. The maxima at the values of d which are multiples of a are the result of the fact that at this point there are no neurons with $\delta_i = 0$. This points are derived from the condition (37) for $\delta_i = 1$ which has a jump when going from $d = ka$ to $d = ka + 1$.

Neglecting the maxima in $d = ka$, the interpretation of the curve in figure 3 is the following. There are optimal values of the size parameter d that allow for higher stability of the stored patterns at the same level of capacity α. High stability is obtained when the distance d_s between the two 'pixels' that are separate, i.e. do not have any common active neurons, is maximal. This condition is obtained for $d_s \simeq a/2$, i.e. for $d = ka + a/2$. On the other hand, maximal stability is much lower for when d_s is small, i.e. for $d = ka$. One should stress that the above statement, which is the central result of our paper, is very general and turns out to be valid for higher dimensions for different symmetry groups, as well as in the presence of disorder [22].

The periodicity of the optimal stability curves in figures 2 and 3 is an interesting effect itself. It means that when we construct basic 'pixel' states using various numbers of the blocks of the size a plus one block of the size d^*, so that $d = ma + d^*$, the resulting optimal stability is m independent.

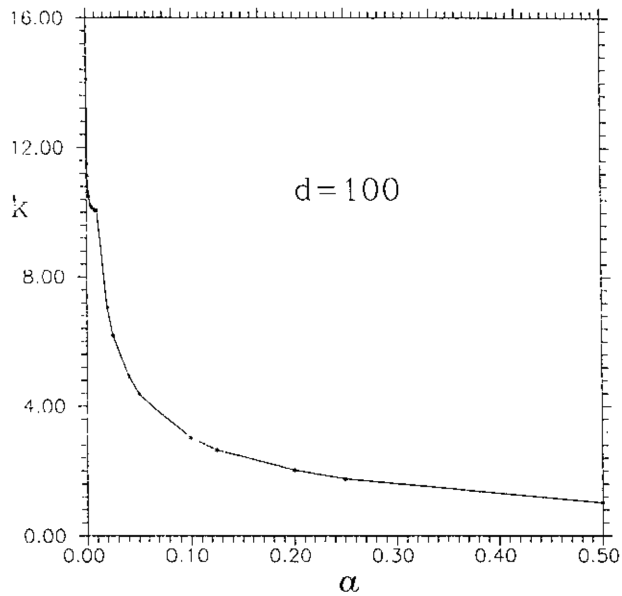

Figure 4. The maximal stability parameter κ as a function of network capacity α for $d = 100$; $N = 5000$.

The different look at our results is presented in figures 4 and 5 that present $\kappa_{\max}$ as a function of α at the given value of d. These figures are direct analogues of Gardner's curve $\alpha_c(\kappa)$ [5]. Here, however, the discreteness of our model plays an essential role. The values of α correspond to divisors of $N = 5000$. For $d = 100$ which is a multiple of $a = 1/\alpha$ for most of the larger values of α presented in figure 4, the curve has a regular shape — it decreases with increasing α. For $d \leqslant a$ the curve enters another regime, fully described by the expression (37) with $r = 0$. The situation is much more complicated for $d = 99$. Although the behaviour for $a = 1/\alpha \leqslant d$ is analogous, the curve exhibits irregular oscillations for smaller values of α. The points of the curve lie either close to minima or to maxima of the curves represented in figures 1–3, depending on the value d modulo a.

The major question now is how reliable are the above results. We remind the reader that they are exact only for $d \leqslant a$ and otherwise they are based on an *approximate variational method* of finding the minimum of the quadratic form (24). To answer this question we construct the *exact* minimum using the method sketched at the end of the previous section and presented in detail in the appendix C. Figure 6 represents the essential features of the exact solution. We limit ourselves here to $d \leqslant 2a$, and $\alpha = 0.02$. The analysis described in the appendix C shows that the point

$$\lambda^0_{\mathrm{pr}} = (\kappa, \ldots, \kappa) \tag{43}$$

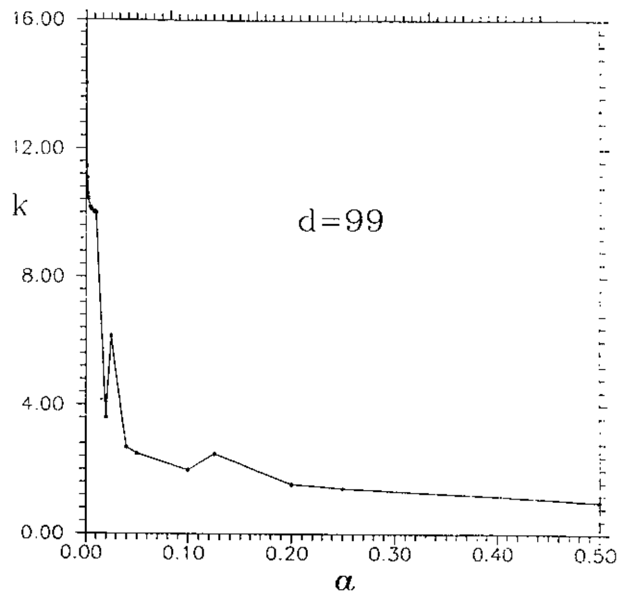

Figure 5. Same as figure 4, but for $d = 99$.

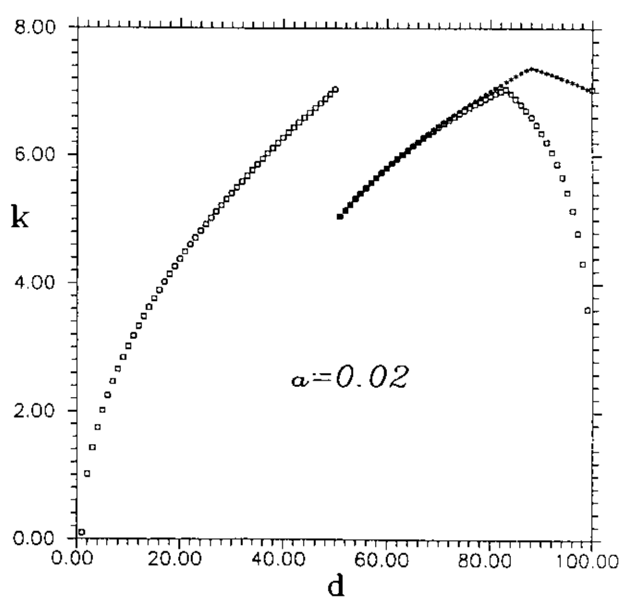

Figure 6. The maximal stability parameter κ as a function of 'pixel' size d for the network capacity $\alpha = 0.02$; $N = 5000$. Only the region $d \leqslant 2a$ is shown. Squares correspond to the results of figure 3 and describe exact results for $d \leqslant a$ and approximate result for $a < d \leqslant 2a$. Stars denote the 'exact' result.

does not give the exact minimum of (24) (the corresponding critical condition is represented by squares in figure 6). A better approximation for the exact minimum is obtained for

$$\boldsymbol{\lambda}^2_{\rm pr} = (\kappa, \ldots, \lambda_{\mu(i)-2}, \kappa, \ldots, \lambda_{\mu(i)+3}, \kappa, \ldots). \tag{44}$$

Note that this is an analogue of (41) for the case $\delta_i = 1$. The indices $\mu(i)$ and $\mu(i)+1$ correspond to $\xi_{\mu(i)} = \xi_{\mu(i)+1} = -1$. Although the results for (44) (stars in *figure 6*) differ quantitatively from those obtained for (43), their qualitative character remains the same. As we expect the difference grows with d (see appendix C). We have also tried another probe vector

$$\boldsymbol{\lambda}^4_{\rm pr} = (\kappa, \ldots, \lambda_{\mu(i)-4}, \kappa, \lambda_{\mu(i)-2}, \kappa, \ldots, \lambda_{\mu(i)+3}, \kappa, \lambda_{\mu(i)+5}, \kappa, \ldots). \tag{45}$$

In principle, more and more components of $\boldsymbol{\lambda}$ should become rigorously larger than κ and should be determined using differential calculus when d approaches $2a$. Amazingly, however, the results obtained for (45) do not differ at all from those obtained for (44). Although we could not prove it, this suggests strongly that the results obtained with (44) are exact.

In the next 'gothic arc' ($2a < d \leqslant 3a$) the situation is analogous and most probably a finite number of components of $\lambda_r > \kappa$ suffices to estimate the minimum very precisely.

We conclude from this analysis that:

(i) simple probe vectors, such as $\boldsymbol{\lambda}^0_{\rm pr}$ or $\boldsymbol{\lambda}^2_{\rm pr}$, give for smaller values of d the exact description of the critical curve; the same vectors give a very good approximate description of the critical curve for all values of d.

(ii) The main difference between the exact results and those obtained for $\boldsymbol{\lambda}^0_{\rm pr} = (\kappa, \ldots, \kappa)$ consists in smoothing and damping of the 'gothic arc' oscillations in figures 2 and 3.

Nevertheless, the oscillations observed for approximate solution (figures 2 and 3) survive for the exact solution at least to some extent. Therefore our main conclusion concerning non-trivial optimal conditions for higher stability of patterns remains valid in general.

In the conclusion we would like to stress that we have presented here an example of the application of Gardner's program to the problem of storage of invariant data. The theory presented may be easily generalized to various other cases and the results have generic character. They indicate that at the given capacity level of the memory there are usually optimal stability conditions for storage of invariant sets of patterns. Such conditions are typically related to the size of the stored patterns and to the difference between close, but separate patterns.

Acknowledgments

This paper has been supported by the Polish Academy of Sciences Grant PR 1. We thank Professor H Gutfreund, who introduced us to the current problems of Gardner's theory. We acknowledge illuminating dicussions with Professor H Gutfreund, Professor D Amit and Professor B Derrida.

Appendix A

In this appendix we discuss in detail properties of the matrix $M^i_{\mu\mu'}$. The diagonal part of this matrix in the basis obtained after applying two unitary transformations (equations (29) and (31)) is related to the function $f(k)$. For odd N all the values of $f(k)$, apart from $k = 0$, are doubly degenerated for $\omega_k = 2\pi - \omega_k$. When N is even, additional non-degenarate value of $f(k)$ appears, since k may be equal to $\alpha N/2$, i.e. $\omega_k = \pi$. It is easy to check then that for $(r+1)a \geqslant d \geqslant ra$

$$f(\alpha N/2) = \frac{4}{N}(d - ra), \tag{46}$$

for even r, and

$$f(\alpha N/2) = \frac{4}{N}[(r+1)a - d] \tag{47}$$

for odd r. Evidently for $d = ra$ and even r, $f(\alpha N/2) = 0$. For the particular case when d is a multiple of a, all neurons correspond to $\delta_i = 1$. Therefore, it is easy to check that for each i, $\xi_i(\alpha N/2) = 0$. The vector $x(k) = \delta_{\alpha N/2,k}$ is the eigenvector of the matrix $\mathbf{M}$ with the non-degenerate eigenvalue zero. In the original basis it is real and has the form $x_\mu = \xi^\mu_i(-1)^\mu$. That means that the integral over x_μ in (11) cannot be performed. The discussion how to deal with this complication is contained in appendix B. Here we stress only that apart from this singular case $f(k)$ is strictly positive for all k and all values of the parameters d and a. It is elementary to prove this fact for low values of r, and we carefully checked it using numerical methods for all other cases considered .

The eigenvales of the matrix $\mathbf{M}$ may be found easily from the equation

$$\lambda x(k) = f(k)x(k) - \frac{1}{N}\xi_i(k)\sum_{k'}\xi^*_i(k')x(k') \tag{48}$$

The above eigenequation has following solutions.

- For all k such that $\xi_i(k) = 0$, the eigenvalues are

  $$\lambda = f(k) > 0 \tag{49}$$

  while the eigenvector

  $$x(k') = \delta_{k',k}. \tag{50}$$

  Note that this eigenvalue is doubly degenerate. There are two real eigenvectors that correspond to the vector (50), namely

  $$x_\mu \propto \frac{\xi^\mu_i \cos(\omega_k \mu)}{\sqrt{\alpha N}}$$

  and

  $$x_\mu \propto \frac{\xi^\mu_i \sin(\omega_k \mu)}{\sqrt{\alpha N}}.$$

- For all k such that $\xi_i(k) \neq 0$ some eigenvalues can be determined from the equation

$$\frac{1}{N}\sum_k^{*} \frac{|\xi_i(k)|^2}{\lambda - f(k)} = 1 \tag{51}$$

The sum $\sum^*$ is restricted to the set S^* of k, i.e. $k = 0$ and N^* values of $k \neq 0$ such that $\xi_i(k) \neq 0$. There are exactly N^* non-degenerate solutions of the (51) located between the different values of $f(k)$. The largest solution lies above $f(0) = O(N)$ and is non-degenarate. It is important that all eigenvalues fulfil

$$\lambda > \min_{k\in S^*} f(k) > 0. \tag{52}$$

The corresponding eigenvectors are given by

$$x(k') = \frac{\xi_i(k')}{\lambda - f(k')} \tag{53}$$

and correspond to real x_μ, since $f(k) = f(\alpha N - k)$ and $\xi_i(k) = \xi_i^*(\alpha N - k)$.

- Finally, for each of the k such that $\xi_i(k) \neq 0$ nad $k \neq 0$ there exists one eigenvector of the form

$$x(k') = \xi_i(k)\delta_{k',k} - \xi_i(\alpha N - k)\delta_{k',\alpha N-k} \tag{54}$$

so that

$$\sum_{k'} \xi_i^*(k')x(k') = 0. \tag{55}$$

There are exactly N^* such eigenvectors and they correspond to the eigenvalues

$$\lambda = f(k). \tag{56}$$

It is easy to check that the above constructed eigenvectors constitute the whole basis set. It is also clear that the matrix $\mathbf{M}^i$ is indeed positively defined and has all eigevalues of the order of $f(k)$, i.e. of the order of $1/N$, except for one non-degenerate eigenvalue that is of the order of N.

Appendix B

In this appendix we discuss the method of handling the zero eigenvalue of the matrix $\mathbf{M}^i$ that appears for even N, even r, and $d = ra$. The existence of zero eigenvalue that corresponds to the eigenvector $x_\mu^0 \propto (-1)^\mu \xi_i^\mu$, means that the integrand of the integral (11), after performing integration over the J_{ij} does not effectively depend on the coordinate that measures the projection of x_μ onto x_μ^0. The integration over this particular coordinate introduces new Dirac's δ. That in turn means that the integral over λ contains an additional constrain,

$$\sum_\mu (-1)^\mu \xi_i^\mu \lambda_\mu = 0. \tag{57}$$

The minimum that is calculated in (24) must be modified in two ways.

(i) The range of λ must be limited to $\lambda_\mu \geqslant \kappa$ and $\sum_\mu (-1)^\mu \xi_i^\mu \lambda_\mu = 0$.
(ii) The matrix $\mathbf{M}$ must be substituted by the matrix $\mathbf{M}^*$ which is identical to $\mathbf{M}$ when acting on any vector perpendicular to $x_\mu^0 = (-1)^\mu \xi_i^\mu$.

The proper formula reads then

$$\min_{\lambda_\mu \geqslant \kappa, (\boldsymbol{\lambda} \cdot \boldsymbol{x}^0) \neq 0} \left(\sum_{\mu,\mu'} \lambda_\mu (M^i)^{-1}_{\mu\mu'} \lambda_{\mu'} \right) = N \tag{58}$$

The expression (37) should in principle contain sums that do not include the term $k = \alpha N/2$ and the division by the factor $f(\alpha N/2) = 0$. Fortunately the terms in question are proportional to the factors $|\xi_i(k)|^2$ which vanish as well for $k = \alpha N/2$. The ratio of the two factors has a well defined limit as $\omega_k \to \infty$. The dangerous term, when calculated in this limiting sense, is therefore of the order of $1/N$, and can be left intact in the limit $N \to \infty$. In effect the expressions (37) and (38) may be regarded as generally valid, provided we treat the singular term in the above described manner.

Appendix C

In this appendix we construct exact minimum of the form

$$\min_{\lambda_\mu \geqslant \kappa} (\boldsymbol{\lambda}(\mathbf{M}^i)^{-1}\boldsymbol{\lambda}). \tag{59}$$

Elementary calculations show that for $d \leqslant a$ the exact minimum is obtained $\boldsymbol{\lambda} = \boldsymbol{\lambda}^0_{\mathrm{pr}}$. Here we consider $a < d \leqslant 2a$ (i.e. for $r = 1$). For larger values of d the calculations are technically more complicated but otherwise can be done along the same lines.

Let us first turn back to the vector $\boldsymbol{\lambda}^0_{\mathrm{pr}} = (\kappa, \ldots, \kappa)$. This vector would minimize the form (59) if for any vector $\Delta\boldsymbol{\lambda}$, such that $\Delta\boldsymbol{\lambda}_\mu \geqslant 0$ the following were true

$$(\boldsymbol{\lambda}^0_{\mathrm{pr}} + \Delta\boldsymbol{\lambda})\mathbf{M}^{-1}(\boldsymbol{\lambda}^0_{\mathrm{pr}} + \Delta\boldsymbol{\lambda}) \geqslant \boldsymbol{\lambda}^0_{\mathrm{pr}}\mathbf{M}^{-1}\boldsymbol{\lambda}^0_{\mathrm{pr}}. \tag{60}$$

This condition is fulfiled, provided

$$x_\mu = (\mathbf{M}^{-1}\boldsymbol{\lambda}^0_{\mathrm{pr}})_\mu \geqslant 0 \tag{61}$$

for all μ. From the (34) we obtain

$$x(k) = \frac{\kappa \xi_i(k)}{(1 - \mathcal{A}) f(k)} \tag{62}$$

so that

$$x_\mu = \frac{1}{\sqrt{N\alpha}} \xi_i^\mu \sum_k \exp(-\mathrm{i}\omega_k \mu x(k)). \tag{63}$$

In the limit of $N \to \infty$ the sums may be replaced by integrals. For instance

$$x_\mu = \frac{\kappa N}{2\pi} \sum_{\mu'} \int_0^{2\pi} \frac{\exp[-i\omega(\mu - \mu')\xi_i^{\mu'}]}{(1-\mathcal{A})\tilde{f}(\omega)} - \frac{\kappa N \xi_i^\mu}{(1-\mathcal{A})f(0)} \tag{64}$$

The integral over ω can be easily evaluated using the complex variable $z = e^{\pm i\omega}$ and the Cauchy theorem. The function $\tilde{f}(\omega)$ becomes then the polynom of the $2r$th order in z. The zeros of this polynom become simple poles of the integrand.

For $r = 1$ the calculations are particularly simple. Of the two poles, only one lies inside the integration contour (i.e. the unit circle)

$$z = -s = -\gamma + \sqrt{\gamma^2 - 1} \tag{65}$$

where $\gamma = d/2(d-a)$. We then obtain

$$x_\mu = \frac{2\kappa N \xi_\mu^i}{8(d-a)(\gamma - s)(1-\mathcal{A})} \left[(-s)^{|\mu - \mu(i)|} + \delta_i (-s)^{|\mu - \mu(i) - 1|}\right]. \tag{66}$$

An analogous method leads to

$$\mathcal{A} = \frac{1 - \delta_i}{2(d-a)(\gamma - s)} + \frac{\delta_i(1-s)}{(d-a)(\gamma - s)} \tag{67}$$

and

$$M^{-1}_{\mu\mu'} = \frac{N \xi_i^\mu \xi_i^{\mu'}}{8(d-a)(\gamma - s)} \left((-s)^{|\mu - \mu'|} - \frac{(\gamma - s)}{N\alpha(\gamma + 1)}\right) + \frac{4N\xi_i^\mu \xi_i^{\mu'}}{(1-\mathcal{A})[8(d-a)(\gamma - s)]^2} \left[(-s)^{|\mu - \mu(i)|} + \delta_i(-s)^{|\mu - \mu(i) - 1|}\right]. \tag{68}$$

Direct inspection into the expression (66) indicates that in the case $\delta_i = 1$, $x_\mu \leqslant 0$ for $\mu = \mu(i) + 3, +5, \ldots$ and $\mu = \mu(i) - 2, -4, \ldots$. For $\delta_i = 0$ the components of x_μ are negative for $\mu = \mu(i) \pm 2, \pm 4, \ldots$. This proves that λ^0_{pr} is not a minimum of (59) for $a < d \leqslant 2a$.

It is obvious, however, that the true minimum can be found by allowing more and more components of the vector λ to be strictly greater than κ. For $\delta_i = 0$ the natural choice is the probe vector

$$\boldsymbol{\lambda}^2_{\mathrm{pr}} = (\kappa, \ldots, \lambda_{\mu(i)-2}, \kappa, \ldots, \lambda_{\mu(i)+2}, \kappa, \ldots). \tag{69}$$

Using the standard methods of the differential calculus we find that

$$\boldsymbol{\lambda}^2_{\mathrm{pr}} \mathbf{M}^{-1} \boldsymbol{\lambda}^2_{\mathrm{pr}} = \boldsymbol{\lambda}^0_{\mathrm{pr}} \mathbf{M}^{-1} \boldsymbol{\lambda}^0_{\mathrm{pr}} - \mathbf{M}^{-1} \boldsymbol{\lambda}^0_{\mathrm{pr}} (\mathcal{P}\mathbf{M}^{-1}\mathcal{P})^{-1} \mathbf{M}^{-1} \boldsymbol{\lambda}^0_{\mathrm{pr}} \tag{70}$$

where $\mathcal{P}$ denotes the projection onto the 2-dimensional subspace spanned by the coordinates $\mu(i) \pm 2$.The inversion of the projected matrix $\mathcal{P}\mathbf{M}^{-1}\mathcal{P}$ refers to the subspace of projection only. In the next step we construct the next approximation for the extremum by considering

$$\boldsymbol{\lambda} = (\kappa, \ldots, \lambda_{\mu(i)-4}, \kappa, \lambda_{\mu(i)-2}, \kappa, \ldots, \lambda_{\mu(i)+2}, \kappa, \lambda_{\mu(i)+4}, \kappa, \ldots) \tag{71}$$

and so on. Amazingly, it turns out that the result obtained for the probe vector (69) do not differ from those obtained for (71). This suggests strongly that (69) describes exact form of the minimum. An analogous situation occurs for $\delta_i = 1$.

References

[1] Little W A 1974 The existence of persistent states in the brain *Math. Biosci.* **19** 101

[2] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational capabilities *Proc. Natl Acad. Sci. USA* **79** 2554–2558

[3] Amit D J 1989 *Modelling Brain Function: The World of Attractor Neural Networks* (Cambridge: Cambridge University Press)

[4] Abbott L F 1990 Learning in neural network memories *Network* **1** 105–122

[5] Gardner E 1988 The space of interactions in neural network models *J. Phys. A: Math. Gen.* **21** 257–270

[6] Krauth W and Mézard M 1987 Learning algorithms with optimal stability for neural networks *J. Phys. A: Math. Gen.* **20** L745–52

[7] Pöppel G and Krey U 1987 Dynamical learning process for recognition of correlated patterns in symmetric spin glass models *Europhys. Lett.* **4** 979–85

[8] Schmitz H J, Pöppel G, Wünsch F and Krey U 1990 Fast recognition of real objects by an optimized hetero-associative neural network *J. Physique* **51** 167–183

[9] Gutfreund H 1990 From statistical mechanics to neural networks and back *Proc. 17th Int. Conf. on Statistical Mechanics*

[10] Gardner E and Derrida B 1988 Optimal storage properties of neural network models *J. Phys. A: Math. Gen.* **21** 271–284

[11] Krauth W and Opper M 1989 Critical storage capacity of the $j = \pm 1$ neural networks *J. Phys. A: Math. Gen.* **22** L519

[12] Gutfreund H and Stein Y 1990 Preprint

[13] Gardner E, Gutfreund H and Yekutieli I 1989 The phase space of interactions in neural networks with definite symmetry *J. Phys. A: Math. Gen.* **22** 1995–2008

[14] Gutfreund H 1987 Neural networks with hierarchically correlated patterns *Phys. Rev.* A **37** 570

[15] Feigelman M V and Ioffe L B 1988 The augmented models of associative memory: asymmetric interaction and hierarchy of patterns *Int. J. Mod. Phys.* B **1** 51

[16] Pöppel G and Krey U 1989 Learning of structured patterns and the generation of context in spin glass models *Neural Networks: From Models to Applications* ed L Personnaz and G Dreyfus pp 217–26

[17] Krey U and Pöppel G 1989 On the termodynamics of associative recall of structured patterns within a given context *Z. Phys.* B **76** 513–20

[18] Rumelhart D E and McClelland J L (eds) 1986 *Parallel Distributed Processing: Explorations in Microstructure of Cognition* vols I and II (Cambridge, MA: MIT Press)

[19] von der Malsburg C and Bienenstock E 1987 A neural network for the retrieval of superimposed connection patterns *Europhys. Lett.* **3** 1243

[20] Bienenstock E and von der Malsburg C 1987 A neural network for invariant pattern recognition *Europhys. Lett.* **4** 121

[21] Kree R and Zippelius A 1989 Recognition of topological features of graphs and images in neural networks *J. Phys. A: Math. Gen.* **22** L813

[22] Tarkowski W and Lewenstein M 1990 Estimates of optimal storage conditions in neural network memories based on random matrix theory in preparation.

[23] Mezard M 1989 The space of interactions in neural networks: Gardner's computation with the cavity method *J. Phys. A: Math. Gen.* **22** 2181

[24] Mehta M L 1967 *Random Matrices and Statistical Theory of Energy Levels* (New York: Academic)