Storage of sets of correlated data in neural network memories

1993 J. Phys. A: Math. Gen. 26 2453

(http://iopscience.iop.org/0305-4470/26/11/001)

J. Phys. A: Math. Gen. 26 (1993) 2453–2469. Printed in the UK

Storage of sets of correlated data in neural network memories

Wojciech Tarkowski† and Maciej Lewenstein‡§

† Centrum Fizyki Teoretycznej Polskiej Akademii Nauk, Al. Lotników 32/46, 02–668 Warsaw, Poland

‡ Centre d'Etudes de Saclay, Service des Photons, Atomes et Molécules, Bat. 22, 91191 Gif sur Yvette Cedex, France

Received 6 August 1992, in final form 11 January 1993

Abstract. In most applications of neural network models one has to consider sets of correlated data. We study the problem of the storage of associated patterns in neural network memories. We consider two basic types of correlation: 'semantic' ones—among the different patterns—and 'spatial' ones—among the different sites of the network. We apply Gardner's program to evaluate optimal storage conditions for both kinds of correlation. The 'spatial' ones worsen, generally speaking, the storage properties of a simple perceptron, while they improve them for the Hopfield network. In the case of 'semantic' correlations we obtain bounds of the critical capacity for both kinds of networks; the storage ratio of the perceptron may be significantly increased in this case.

1. Introduction

One of the most important problems in the theory of attractor neural networks concerns learning. In particular, this problem consists of answering the question concerning the critical storage capacity. The best approach to its study has been proposed by Gardner [1]. Originally, Gardner's program was applied to statistically independent, unbiased patterns. In this case the critical capacity for errorless storage is $\alpha_c = 1/G(\kappa)$, where

$$G(\kappa) = \int_{-\kappa}^{+\infty} \mathrm{d}t\, \frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}} (t+\kappa)^2 \tag{1}$$

whereas κ determines the stability of the stored patterns. Gardner also studied the case of biased patterns, and showed that in the limit of maximal bias or, in another words, sparse coding [2], α_c tends to infinity (see also Willshaw *et al* [3]).

Gardner's program has since been extended to various models, cf [6–13]. There have been, however, practically no studies of statistically correlated patterns (see, e.g., [14, 15]). Therefore it is very important to study the problem of the learning of correlated patterns both in the context of single-layer perceptrons [16], as well as multilayered networks and auto-associative ones of the Hopfield–Little type [17, 18].

In recent work [19] we have extended Gardner's program to 'semantically' correlated patterns. Using this approach we have calculated α_c, and showed that in the limit of large correlation length $L_c \to \infty$, α_c scales as L_c.

§ Permanent address: Centrum Fizyki Teoretycznej Polskiej Akademii Nauk, Al. Lotników 32/46, 02–668 Warsaw, Poland.

0305-4470/93/112453+17$07.50

In this paper we treat the storage of sets of correlated data in a more general perspective: we divide the discussion of this problem into the two cases—'semantically' and 'spatially' correlated patterns. Let us denote them by ξ_j^μ, where μ enumerates patterns and j sites in the network, and $\xi_j^\mu = +1$ or -1. Throughout this paper we consider only the unbiased patterns for which

$$\left\langle\!\left\langle \xi_j^\mu \right\rangle\!\right\rangle = 0. \tag{2}$$

On the other hand, we introduce 'spatial' correlations in the form

$$\left\langle\!\left\langle \xi_j^\mu \xi_{j'}^{\mu'} \right\rangle\!\right\rangle = \delta_{\mu\mu'} C_{jj'} \tag{3}$$

where $C_{jj'}$ is a correlation matrix, and 'semantic' ones

$$\left\langle\!\left\langle \xi_j^\mu \xi_{j'}^{\mu'} \right\rangle\!\right\rangle = C_{\mu\mu'} \delta_{jj'} \tag{4}$$

where $C_{\mu\mu'}$ is a matrix of 'semantic' correlations. Both these matrices are symmetric and positively defined.

Many examples of 'spatial' correlations are known in the theory of visual information processing (cf [20]), whereas 'semantic' ones arise when we consider the learning of categories, subcategories of patterns, etc. It has recently been demonstrated that attractors observed in neurophysiological experiments are 'semantically' correlated [21, 22]. Here the amount of association depends on the temporal interval between the learning times of the corresponding patterns.

The main aim of this paper is to present and discuss the storage of correlated patterns within the framework of Gardner's approach. The plan of the paper is as follows. In sections 2 and 3 we present and discuss the problem of 'spatially' correlated patterns. The results concerning a simple perceptron are described in section 2, while in section 3 we investigate the Hopfield neural network. Similar considerations, but for 'semantically' correlated data in the case of a simple perceptron are presented in sections 4 and 5. This theory is extended to the case of Hopfield networks in section 6. Section 7 contains the final results and discussions.

2. Storage of 'spatially' correlated patterns in a perceptron

In this section we investigate the patterns that are correlated in a 'spatial' way (3). We can find the exact curve which determines the optimal storage in the control parameters (α, κ) space. We show that the critical capacity does not exceed Cover's limit [23], $\alpha_c = 2$ for $\kappa = 0$. The results of this section have been reported by Monasson [24]. We therefore skip most of the calculations and present only final expressions for the storage capacity in the limit of small correlations. These general expressions have been discussed by us in [25].

We consider here a perceptron with N binary neurons in the input layer, $\sigma_j = \pm 1$, and a single output unit, η. The propagation rule takes the form $\eta = \mathrm{sign}\left(\sum_{j=1}^N J_j \sigma_j\right)$. We investigate the existence conditions for a set of connections J_j that lead from a given set of αN input patterns ξ_j^μ, $\mu = 1, \ldots, \alpha N$, to a given output η_μ, with a given stability κ.

In order to do this we calculate the fractional volume in the interaction space of this set of connections,

$$V = \frac{\int \prod_j \mathrm{d}J_j \prod_\mu \Theta\left(\eta_\mu \sum_j (J_j/\sqrt{N})\xi_j^\mu - \kappa\right)\delta\left(\sum_j J_j^2 - N\right)}{\int \prod_j \mathrm{d}J_j\, \delta\left(\sum_j J_j^2 - N\right)}. \tag{5}$$

Note that we normalize J_j according to $\sum_j J_j^2 = N$. The critical storage $\alpha_c(\kappa)$ is attained when the averaged logarithm of V tends to $-\infty$ [1]. The expectation $\langle\langle \ln V \rangle\rangle$ is obtained with the help of the replica method

$$\langle\langle \ln V \rangle\rangle = \lim_{n\to 0} \frac{\langle\langle V^n \rangle\rangle - 1}{n}. \tag{6}$$

In the evaluation of expression (5) one proceeds in a way described in detail in [1, 9]. After performing the average (3), the mean of the nth power of the fractional volume V reads

$$\langle\langle V^n \rangle\rangle = C \int \mathcal{D}E^\alpha \mathcal{D}J_j^\alpha \mathcal{D}x_\mu^\alpha \int_\kappa^\infty \mathcal{D}\left(\frac{\lambda_\mu^\alpha}{2\pi}\right) \times \exp\left\{ \mathrm{i} \sum_{\alpha,\mu} \lambda_\mu^\alpha x_\mu^\alpha - \frac{1}{2N} \sum_{\alpha,\beta,\mu,j,j'} x_\mu^\alpha J_j^\alpha C_{jj'} J_{j'}^\beta x_\mu^\beta - \sum_\alpha \frac{E^\alpha}{2}\left(\sum_j (J_j^\alpha)^2 - N\right)\right\} \tag{7}$$

where C is a constant, which plays no role in the further considerations. Note that result (7), as in [1], does not depend on outputs η_μ. This property of $\langle\langle \ln V \rangle\rangle$ is characteristic for unbiased 'spatially' correlated data. It does not hold, as we shall see, for 'semantically' correlated ones.

In order to disentangle the second term from the exponent we introduce the order parameters $q^{\alpha\beta}$, Q^α defined as follows

$$q^{\alpha\beta} = \frac{1}{N}\sum_{j,j'} C_{jj'} J_j^\alpha J_{j'}^\beta \qquad Q^\alpha = \frac{1}{N}\sum_{j,j'} C_{jj'} J_j^\alpha J_{j'}^\alpha \tag{8}$$

and their conjugated counterparts $f^{\alpha\beta}$ and F^α, respectively. Note that $\alpha, \beta = 1, \ldots, n$ denote here the replica indices—see [1].

In order to perform the integration over J_j^α in (7) we change the integration variables introducing the set of eigenvectors $\{\epsilon_j^k\}_k$ of the correlation matrix $\widehat{C}$. They then fulfil

$$\sum_{j'} C_{jj'} \epsilon_{j'}^k = C_k \epsilon_j^k \tag{9}$$

for each $j = 1, \ldots, N$ and $k = 1, \ldots, N$, where C_k denote eigenvalues of the correlation matrix. The C_k are real and non-negative. When neurons are arranged in a torus, and the correlation matrix is translationally (rotationally) invariant,

$$C_{jj'} = C\left(|j - j'|\right) \tag{10}$$

the eigenvectors ϵ_j^k are Fourier components:

$$\epsilon_j^k = \sqrt{(2/N)}\cos(\omega_k j) \tag{11}$$

or

$$\epsilon_j^k = \sqrt{(2/N)}\sin(\omega_k j) \tag{12}$$

with $\omega_k = (2\pi/N)k$, $k = 1, \ldots, N$ and

$$\epsilon_j^0 = 1/\sqrt{N}. \tag{13}$$

We introduce new variables $\widetilde{J}_k$, expanding

$$J_j = \sum_k \widetilde{J}_k \epsilon_j^k. \tag{14}$$

The order parameters $q^{\alpha\beta}$, Q^α thus become:

$$q^{\alpha\beta} = \frac{1}{N}\sum_k C_k \widetilde{J}_k^\alpha \widetilde{J}_k^\beta \tag{15}$$

$$Q^\alpha = \frac{1}{N}\sum_k C_k \left(\widetilde{J}_k^\alpha\right)^2. \tag{16}$$

We proceed then along the lines of [1, 24] and perform the Gaussian integrals over the variables x and $\widetilde{J}$. Note that the latter integration may now be done separately for each $\widetilde{J}_k$. We also anticipate a replica-symmetric saddle point and substitute $q^{\alpha\beta} = q$, $Q^\alpha = Q$, $f^{\alpha\beta} = f$ and $F^\alpha = F$. The free energy density function $\mathcal{F}$, which has to be extremized at the saddle point, takes in the limit $n \to 0$ the form

$$\mathcal{F} = -\alpha G_1 - \frac{1}{N}G_2 - \tfrac{1}{2}fq + \tfrac{1}{2}FQ - \tfrac{1}{2}E \tag{17}$$

where

$$G_1 = \int \mathrm{D}t \ \ln H\left[\frac{\kappa + t\sqrt{q}}{\sqrt{Q-q}}\right] \tag{18}$$

$$G_2 = -\frac{1}{2}\sum_k \ln(E + C_k f - C_k F) + \frac{1}{2}\sum_k \frac{C_k f}{E + C_k f - C_k F}$$
$$\equiv -\frac{N}{2}\langle\!\langle \ln(E + Cf - CF)\rangle\!\rangle_C + \frac{N}{2}\left\langle\!\!\left\langle \frac{Cf}{E + Cf - CF}\right\rangle\!\!\right\rangle_C. \tag{19}$$

The symbol $\langle\!\langle\cdot\rangle\!\rangle_C$ denotes here the averaging over the spectrum of the correlation matrix $C_{jj'}$, the Gaussian measure

$$\mathrm{D}t = \frac{\mathrm{d}t}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2} \tag{20}$$

and the function H is defined as follows

$$H(x) = \int_x^\infty \mathrm{D}t. \tag{21}$$

After this calculation we may apply the saddle-oint technique to the function $\mathcal{F}$ in order to evaluate (7). One may notice that the critical point of vanishing of V is attained when the order parameters q and Q become equal. In the limit $q \to Q$ the free energy density, and also the saddle-oint equations, may be simplified enormously. We simply anticipate that at the criticallity the order parameters behave as

$$E = \widetilde{E}/(Q-q) \tag{22}$$

$$f = \widetilde{f}/(Q-q)^2 \tag{23}$$

$$f - F = \widetilde{g}/(Q-q). \tag{24}$$

We also evaluate the asymptotic form of the function G_1 for $q \to Q$

$$G_1 = -\frac{1}{2}\left\langle\!\!\left\langle \min_{\lambda \geqslant \kappa} \frac{(\lambda + t\sqrt{q})^2}{Q-q} \right\rangle\!\!\right\rangle_t \tag{25}$$

where the symbol $\langle\!\langle \cdot \rangle\!\rangle_t$ now denotes averaging over the Gaussian variable t with the measure Dt. The saddle-oint equations then become

$$1 = \left\langle\!\!\left\langle \frac{C}{\widetilde{E} + \widetilde{g}C} \right\rangle\!\!\right\rangle_C \tag{26}$$

$$q = \left\langle\!\!\left\langle \frac{\widetilde{f}C^2}{(\widetilde{E} + \widetilde{g}C)^2} \right\rangle\!\!\right\rangle_C \tag{27}$$

$$1 = \left\langle\!\!\left\langle \frac{\widetilde{f}C}{(\widetilde{E} + \widetilde{g}C)^2} \right\rangle\!\!\right\rangle_C \tag{28}$$

$$\widetilde{f} = \alpha P \tag{29}$$

$$\widetilde{g} = \alpha R \tag{30}$$

where the functions P, R are given by

$$P = \int \mathrm{D}t\, \left(\kappa + t\sqrt{q}\right)^2 \Theta\left(\kappa + t\sqrt{q}\right) \tag{31}$$

$$R = \int \mathrm{D}t\, \left(\kappa + t\sqrt{q}\right) \frac{t}{\sqrt{q}} \Theta\left(\kappa + t\sqrt{q}\right). \tag{32}$$

Note that the variable $(Q-q)$ does not enter these equations in the asymptotic limit. Since there are five equations for four variables this means that (26)–(30) imply the additional constraint for a control parameter, that is the condition for the critical capacity.

Equations (29) and (30) may be solved with respect to $\widetilde{f}$ and $\widetilde{g}$. Inserting the solution into the three remaining equations for $\widetilde{E}$, q and α one obtains

$$1 = \left\langle\!\!\left\langle \frac{C}{\widetilde{E} + \alpha RC} \right\rangle\!\!\right\rangle_C \tag{33}$$

$$q = \left\langle\!\!\left\langle \frac{\alpha PC^2}{(\widetilde{E} + \alpha RC)^2} \right\rangle\!\!\right\rangle_C \tag{34}$$

$$1 = \left\langle\!\!\left\langle \frac{\alpha PC}{(\widetilde{E} + \alpha RC)^2} \right\rangle\!\!\right\rangle_C . \tag{35}$$

From these equations we may obtain the functional (α, κ) dependence. This task must be achieved using numerical methods. One may, however, easily check that in Cover's limit [23] $\kappa \to 0$, $\alpha = 2$, just as in the uncorrelated case [1, 19]. In this limit $P = q/2$, $R = 1/2$. The corresponding solution of (33)–(35) has the form $\widetilde{E} = 0$, $q = 1/\langle\!\langle 1/C \rangle\!\rangle_C$. The fact that $\alpha_c(\kappa = 0) = 2$ agrees with the result of Cover's original work, in which it was proven that whenever the outputs remain unbiased and independent random variables (uncorrelated to the inputs), then one always obtains $\alpha_c = 2$ for the zeroth value of the stability parameter κ.

The critical curves for exponentially correlated sites were plotted by Monasson in his recent work [24]. They indicate that for the case of exponential correlations $\alpha_c(\kappa)$ is a decreasing function of κ and that $\alpha_c(\kappa) \leqslant \alpha_c^G(\kappa)$, where $\alpha_c^G(\kappa)$ denotes Gardner's original result, which is valid for uncorrelated patterns. The analysis presented here is for an arbitrary form of the correlation matrix. The critical curves depend significantly on the width of the distribution of C, but we expect (cf [26]) that they do not depend strongly on its shape.

Here, instead of solving (33)–(35) numerically, we present an analytical expression for the solutions, valid in the limit of weak correlations. In order to do this we observe that the uncorrelated case corresponds to the density of the eigenvalues of the correlation matrix simply equal to 1. In the limit of weak correlations we may then expand the averages over the distribution of C on the right-hand side of (33)–(35) in $\delta C = C - 1$ up to the lowest non-vanishing terms of the order of $\langle\!\langle \delta C^2 \rangle\!\rangle$. We may then solve equations (33)–(35) perturbatively in $\langle\!\langle \delta C^2 \rangle\!\rangle$. As a result we obtain

$$\alpha_c(\kappa) = [1/G(\kappa)] \left\{1 - S(\kappa)\,[1 - S(\kappa)]\, \langle\!\langle \delta C^2 \rangle\!\rangle_C \right\} \tag{36}$$

where the function $S(\kappa)$ is defined as follows

$$S(\kappa) = (1/G(\kappa))\, \langle\!\langle (\kappa + t) t \Theta(\kappa + t) \rangle\!\rangle_t . \tag{37}$$

The form of the dependences (36) and (37) implies that $\alpha_c(\kappa) \leqslant \alpha_c^G(\kappa)$, because the inequality $0 \leqslant S(\kappa) \leqslant 1$ is fulfilled. Expression (36) provides an elegant estimate of $\alpha_c(\kappa)$ which is valid for arbitrary correlation matrices.

3. 'Spatial' correlations of data in the Hopfield network

As in the previous section, calculations may be done for the Hopfield neural network. This problem has recently been solved by Monasson [27] and independently by us [25]. We again skip most of the technical details of the analysis since they can be found in [27]. We present, however, an elegant final expression for the storage capacity that holds for arbitrary correlation matrices (see also [25]).

Let us consider the network with N binary nodes, $\sigma_j = \pm 1$, which follows a propagation rule of the form

$$\sigma_i(t+\Delta t) = \mathrm{sign}\left[\sum_{j\neq i} J_{ij}\sigma_j(t)\right] \tag{38}$$

for $i = 1, \ldots, N$, while the fractional volume in the interaction space reads

$$V = \frac{\prod_i \left[\int \prod_{j\neq i} \mathrm{d}J_{ij} \prod_\mu \Theta\left(\xi_i^\mu \sum_{j\neq i}(J_{ij}/\sqrt{N})\xi_j^\mu - \kappa\right) \delta\left(\sum_{j\neq i} J_{ij}^2 - N\right)\right]}{\prod_i \left[\int \prod_{j\neq i} \mathrm{d}J_{ij} \delta\left(\sum_{j\neq i} J_{ij}^2 - N\right)\right]} \tag{39}$$

and the spherical normalization constraint for each $i = 1, \ldots, N$ is assumed as well

$$\sum_{j\neq i} J_{ij}^2 = N. \tag{40}$$

It has been stressed [27] that the proper quenched averages over the patterns $\{\xi_j^\mu\}$ cannot be defined by (2) and (3) only. The reason is that for 'spatially' correlated patterns—in contrast to the usual case, when J_{ij} are of the order of one—the optimal connection matrices behave as $J_{ij} \simeq \mathrm{O}(\sqrt{N})$ for i and j close enough. The averaging procedure, as performed in the previous section, remains, strictly speaking, valid only for continuous Gaussian inputs [28], or in the limit of small correlation length. Keeping these restrictions in mind, we proceed as usual and apply the replica method. In order to do this we define the three order parameters

$$q^{\alpha\beta} = \frac{1}{N} \sum_{j\neq i, j'\neq i} C_{jj'} J_{ij}^\alpha J_{ij'}^\beta \tag{41}$$

$$Q^\alpha = \frac{1}{N} \sum_{j\neq i, j'\neq i} C_{jj'} J_{ij}^\alpha J_{ij'}^\alpha \tag{42}$$

$$m^\alpha = \frac{1}{\sqrt{N}} \sum_{j\neq i} C_{ij} J_{ij}^\alpha \tag{43}$$

and their conjugated counterparts $f^{\alpha\beta}$, F^α and M^α.

We assume as before the torus topology of the network and the rotational invariance of the correlation matrix $\widehat{C}$ (see (10)). Taking the replica-symmetric ansatz and performing all the Gaussian integrals, one arrives at the free energy density function $\mathcal{F}$, which must be extremized with respect to all its variables (order parameters and their conjugated counterparts)

$$\mathcal{F} = -\alpha G_1 - \frac{1}{N} G_2 - \tfrac{1}{2} f q + \tfrac{1}{2} F Q + M m - \tfrac{1}{2} E. \tag{44}$$

The functions G_1, G_2 are given by

$$G_1 = \int \mathrm{D}t \ln H\left[\frac{\kappa - m + t\sqrt{q-m^2}}{\sqrt{Q-q}}\right] \tag{45}$$

$$\begin{aligned} G_2 &= -\frac{1}{2}\sum_k \ln(E + C_k f - C_k F) + \frac{1}{2}\sum_k \frac{C_k f}{E + C_k f - C_k F} + \frac{1}{2}M^2 \sum_k \frac{N b_k}{E + C_k f - C_k F} \\ &\equiv -\frac{N}{2}\langle\langle \ln(E + Cf - CF)\rangle\rangle_C \\ &\quad + \frac{N}{2}\left\langle\left\langle \frac{Cf}{E + Cf - CF}\right\rangle\right\rangle_C + \frac{N}{2}M^2 \left\langle\left\langle \frac{N b_C}{E + Cf - CF}\right\rangle\right\rangle_C \end{aligned} \tag{46}$$

where all the symbols (i.e. $\langle\langle\cdot\rangle\rangle_C$) and functions (i.e. $H[\cdot]$) are defined as in the previous section. The quantity b_C reads

$$b_C = \frac{1}{N}\left[\sum_{j\neq i} C_{ij}\epsilon_j^C\right]^2 \tag{47}$$

with $\{\epsilon_j^C\}_C$ being the Fourier components (see definitions (11)–(13)). Introducing the new variables $g \equiv f - F$, $r \equiv Q - q$ and anticipating, as usual, that at criticallity the order parameters behave as $E = \widetilde{E}/r$, $f = \widetilde{f}/r^2$, $g = \widetilde{g}/r$ and $M = \widetilde{M}/r$, one obtains the seven saddle-oint equations

$$1 = \left\langle\left\langle \frac{C}{\widetilde{E} + \widetilde{g}C}\right\rangle\right\rangle_C \tag{48}$$

$$m = \widetilde{M}\left\langle\left\langle \frac{N b_C}{\widetilde{E} + \widetilde{g}C}\right\rangle\right\rangle_C \tag{49}$$

$$q = \widetilde{f}\left\langle\left\langle \frac{C^2}{(\widetilde{E} + \widetilde{g}C)^2}\right\rangle\right\rangle_C + \widetilde{M}^2\left\langle\left\langle \frac{N C b_C}{(\widetilde{E} + \widetilde{g}C)^2}\right\rangle\right\rangle_C \tag{50}$$

$$1 = \widetilde{f}\left\langle\left\langle \frac{C}{(\widetilde{E} + \widetilde{g}C)^2}\right\rangle\right\rangle_C + \widetilde{M}^2\left\langle\left\langle \frac{N b_C}{(\widetilde{E} + \widetilde{g}C)^2}\right\rangle\right\rangle_C \tag{51}$$

$$\widetilde{f} = \alpha\int Dt\left(\kappa - m + t\sqrt{q-m^2}\right)^2 \Theta\left(\kappa - m + t\sqrt{q-m^2}\right) \tag{52}$$

$$\widetilde{g} = \alpha\int Dt\left(\kappa - m + t\sqrt{q-m^2}\right)\frac{t}{\sqrt{q-m^2}}\Theta\left(\kappa - m + t\sqrt{q-m^2}\right) \tag{53}$$

$$\widetilde{M} = \alpha\int Dt\left(\kappa - m + t\sqrt{q-m^2}\right)\left(1 + \frac{mt}{\sqrt{q-m^2}}\right)\Theta\left(\kappa - m + t\sqrt{q-m^2}\right). \tag{54}$$

These equations may be solved numerically—see [28], in which we obtain the critical capacity curves for the case of the continuous Gaussian inputs. For binary ones, however, one has to use the perturbative approach (in the small correlation limit, i.e. when $C_{jj'} \simeq \delta_{jj'}$).

This is a consequence of the fact that (48)–(54) are systematically valid in that limit only. The approximate value of the storage ratio α_c for the minimal stability parameter $\kappa = 0$ (see also [25]) reads then

$$\alpha_c \cong 2 + 2 \left\langle\!\!\left\langle \frac{Nb_C}{C} \right\rangle\!\!\right\rangle_C \bigg/ \left\langle\!\!\left\langle \frac{1}{C} \right\rangle\!\!\right\rangle_C . \tag{55}$$

This formula indicates that in the case of the storage of the 'spatially' correlated patterns in the Hopfield network the critical capacity ratio α_c somewhat exceeds Cover's limit $\alpha_c\,(\kappa = 0) = 2$.

4. 'Semantically' correlated patterns in perceptron

In this section we consider the problem of the critical storage capacity for 'semantically' correlated patterns which can be, in general, described by (4). We will proceed along the lines of [19], and investigate unbiased patterns.

We consider a perceptron with N binary neurons in the input layer, $\sigma_j = \pm 1$ and a single output unit, η. The propagation rule then takes the form $\eta = \mathrm{sign}\left(\sum_{j=1}^{N} J_j \sigma_j\right)$. We investigate, as before, the existence conditions for a set of connections J_j that lead from a given set of αN input patterns ξ_j^μ, $\mu = 1, \ldots, \alpha N$, to a given output η_μ, with a given stability κ. Once more we calculate the fractional volume in the interaction space of this set of connections,

$$V = \frac{\int \prod_j \mathrm{d}J_j \prod_\mu \Theta\left(\eta_\mu \sum_j (J_j/\sqrt{N}) \xi_j^\mu - \kappa\right) \delta\left(\sum_j J_j^2 - N\right)}{\int \prod_j \mathrm{d}J_j\, \delta\left(\sum_j J_j^2 - N\right)} . \tag{56}$$

The J_js are again normalized according to $\sum_j J_j^2 = N$. The critical storage $\alpha_c(\kappa)$ is attained when the averaged logarithm of V tends to $-\infty$ [1].

We consider a very general class of 'semantically' correlated patterns that allow for input–output correlations. We write the correlation function (4) in the form

$$\left\langle\!\!\left\langle \xi_j^\mu \xi_{j'}^{\mu'} \right\rangle\!\!\right\rangle = \delta_{jj'} \eta_\mu \zeta_\mu C_{\mu\mu'} \zeta_{\mu'} \eta_{\mu'} \tag{57}$$

where $\zeta_\mu = \pm 1$ for $\mu = 1, \ldots; \propto N$ are not necessarily all equal. The motivations for writing the previous expression are as follows.

(i) $C_{\mu\mu'}$ is a matrix with positive elements and it takes care of the correlations decay with the increase of $|\mu - \mu'|$.

(ii) We factor out $\eta_\mu \eta_{\mu'}$, since only the quantity $\eta_\mu \left\langle\!\!\left\langle \xi_j^\mu \xi_{j'}^{\mu'} \right\rangle\!\!\right\rangle \eta_{\mu'}$ enters the calculations.

(iii) Finally we have introduced the additional term $\zeta_\mu \zeta_{\mu'}$ which attends to the possible correlations of the signs of inputs and outputs.

Expression (57) may be interpreted in various ways depending on the specific relations between η_μ and ζ_μ. In particular,

(i) when $\zeta_\mu = \zeta_{\mu'}$ we have the case of optimal input–output correlations in a perceptron;

(ii) when $\eta_\mu = \zeta_\mu$ and $\eta_\mu = \xi_i^\mu$ for some i, we deal with the attractor network of the Hopfield type;

(iii) finally, when ζ_μ are statistically independent and unrelated to η_μ we have the case of minimal input–output correlation in a perceptron.

In this section we calculate the volume (56) for exponentially correlated patterns. We assume that the patterns ξ_j^μ are statistically independent for all js and have the mean zero. On the other hand, we shall take

$$\left\langle\!\left\langle \xi_j^\mu \xi_j^{\mu'} \right\rangle\!\right\rangle = \eta_\mu \zeta_\mu \mathrm{e}^{-b|\mu-\mu'|} \zeta_{\mu'} \eta_{\mu'} \equiv \eta_\mu M_{\mu\mu'} \eta_{\mu'} \tag{58}$$

where b denotes the bandwidth of the correlation matrix $M_{\mu\mu'}$ and is equal to the inverse of the correlation length L_c.

We stress that the specific exponential form of the correlation function (58) or (57) is quite generic for 'semantically' associated patterns. Exponentially correlated patterns may also be easily generated in numerical simulations. They correspond to thermal equilibrium states of the one-dimensional Ising model with the Hamiltonian

$$\mathcal{H} = -\sum_\mu \eta_\mu \xi^\mu \zeta_\mu \zeta_{\mu+1} \xi^{\mu+1} \eta_{\mu+1}. \tag{59}$$

The temperature T is then related to b through

$$b = 1/L_c = -\ln\tanh(1/T). \tag{60}$$

For simplicity, we assume a torus geometry, so that $M_{\mu\mu'}$ is translationally invariant. The matrix $\hat{M}$ is particularly convenient [29], since its inverse is tridiagonal. Neglecting exponentially small terms

$$\left(M^{-1}\right)_{\mu\mu'} = \zeta_\mu \left[(A+2B)\delta_{\mu\mu'} - B\left(\delta_{\mu,\mu'+1} + \delta_{\mu,\mu'-1}\right)\right]\zeta_{\mu'} \tag{61}$$

where $A = \left(1-\mathrm{e}^{-b}\right)/\left(1+\mathrm{e}^{-b}\right)$ and $B = \mathrm{e}^{-b}/\left(1-\mathrm{e}^{-2b}\right)$.

The calculation of $\langle\!\langle \ln V \rangle\!\rangle$ is performed with the help of replica method

$$\ln V = \lim_{n\to 0}\left(V^n - 1\right)/n. \tag{62}$$

We calculate $\langle\!\langle V^n \rangle\!\rangle$, which after performing average over ξs takes the form

$$\begin{aligned}\left\langle\!\left\langle V^n \right\rangle\!\right\rangle = \int \mathcal{D}E^\alpha \mathcal{D}J_j^\alpha \mathcal{D}r_\mu^\alpha \int_\kappa^\infty \mathcal{D}\left(\frac{\lambda_\mu^\alpha}{2\pi}\right) \\ \times \exp\left\{ \mathrm{i}\sum_{\alpha,\mu} \lambda_\mu^\alpha r_\mu^\alpha - \frac{1}{2N}\sum_{\alpha,\beta,\mu,\mu',j} J_j^\alpha r_\mu^\alpha M_{\mu\mu'} r_{\mu'}^\beta J_j^\beta - \sum_\alpha \frac{E^\alpha}{2}\left(\sum_j (J_j^\alpha)^2 - N\right)\right\} \\ \times \left(\int \mathcal{D}E^\alpha \mathcal{D}J_j^\alpha \exp\left\{-\sum_\alpha \frac{E^\alpha}{2}\left(\sum_j (J_j^\alpha)^2 - N\right)\right\}\right)^{-1}.\end{aligned} \tag{63}$$

Now, as in [1] and [19] we introduce the overlap

$$q^{\alpha\beta} = \frac{1}{N}\sum_{j=1}^N J_j^\alpha J_j^\beta. \tag{64}$$

Critical capacity is attained when $q^{\alpha\beta} \to 1$, which means that there exists only one set of J_js that fulfils the desired input–output relations. We also introduce the conjugated variables $F^{\alpha\beta}$ to ensure the constraint (64). After performing integrations over r_μ^α and J_j^α we may execute the remaining integrations of E^α, $q^{\alpha\beta}$, $F^{\alpha\beta}$ with the help of the saddle-oint technique. Assuming replica-symmetric solution $E^\alpha = E$, $q^{\alpha\beta} = q$, $F^{\alpha\beta} = F$ in the limit $n \to 0$, the free energy density function $\mathcal{F}$, which remains to be extremized becomes

$$\mathcal{F} = -\frac{\alpha H(q)}{2} + \frac{\alpha}{2}\ln(1-q) + \frac{\alpha q}{2(1-q)} - \frac{1}{2}Fq - \frac{E}{2} + \frac{1}{2}\ln(E+F) - \frac{1}{2}\frac{F}{E+F} \tag{65}$$

where $H(q)$ is defined through the relation

$$H(q) = \lim_{n\to 0}\frac{2}{\alpha n N}\ln\int_\kappa^\infty \mathcal{D}\lambda_\mu^\alpha \exp\left\{-\frac{1}{2}\mathrm{Tr}\left[\lambda\left(\widehat{M}^{-1}\otimes\widehat{d}^{-1}\right)\lambda\right]\right\} \tag{66}$$

where the elements of the matrix $\widehat{d}$ are defined as follows

$$d_{\alpha\beta} = \delta_{\alpha\beta} + \left(1-\delta_{\alpha\beta}\right)q \tag{67}$$

with $\delta_{\alpha\beta}$ being the Kronecker delta.

Now we determine the saddle point from (65) and after simple calculations we obtain in the limit $q \to 1$ the general expression for α_c

$$\alpha_c = \left(1 - \lim_{q\to 1}\frac{(1-q)^2}{q}H'(q)\right)^{-1}. \tag{68}$$

Let us make a direct inspection of the function H. Denoting

$$I = \exp\left\{\tfrac{1}{2}Nn\alpha H\right\} \tag{69}$$

and using the expression for $\widehat{M}^{-1}$ we obtain

$$\begin{aligned} I = \int_\kappa^\infty \mathcal{D}\lambda_\mu^\alpha \exp\Bigg\{ &-\frac{1}{2(1-q)}\left[A\sum_{\alpha,\mu}\left(\lambda_\mu^\alpha\right)^2 + B\sum_{\alpha,\mu}\left(\zeta_\mu\lambda_\mu^\alpha - \zeta_{\mu+1}\lambda_{\mu+1}^\alpha\right)^2\right] \\ &+ \frac{q}{2(1-q)(1-q+nq)}\Bigg[A\sum_\mu\left(\sum_\alpha\lambda_\mu^\alpha\right)^2 \\ &+ B\sum_\mu\left(\sum_\alpha\zeta_\mu\lambda_\mu^\alpha - \sum_\alpha\zeta_{\mu+1}\lambda_{\mu+1}^\alpha\right)^2\Bigg]\Bigg\}. \end{aligned} \tag{70}$$

We introduce two white noise variables x_μ, y_μ in order to disentangle terms in (70) that contain squares of sums over α. Both x_μ, y_μ are normally distributed random variables with mean zero and variance one. The limit $n \to 0$ may then be performed explicitly and we get

$$\begin{aligned} H(q) = \frac{4}{\alpha N}\Bigg\langle\!\!\Bigg\langle \ln\int_\kappa^\infty \mathcal{D}\lambda_\mu \exp\Bigg\{ &-\frac{1}{2(1-q)}\Bigg[A\sum_\mu\lambda_\mu^2 + B\sum_\mu\left(\zeta_\mu\lambda_\mu - \zeta_{\mu+1}\lambda_{\mu+1}\right)^2 \\ &+ 2\sqrt{q}\sum_\mu\lambda_\mu\left(\sqrt{A}x_\mu + \sqrt{B}\left(\zeta_\mu y_\mu - \zeta_{\mu-1}y_{\mu-1}\right)\right)\Bigg]\Bigg\}\Bigg\rangle\!\!\Bigg\rangle_{x,y}. \end{aligned} \tag{71}$$

In the limit $q \to 1$, the maximum of the integrand gives the leading contribution, so that we obtain

$$\alpha_c = \Bigg\{ 1 + \Bigg\langle\!\!\Bigg\langle \frac{1}{\alpha N} \min_{\lambda_\mu \geqslant \kappa} \Bigg[A \sum_\mu \lambda_\mu^2 + B \sum_\mu \left(\zeta_\mu \lambda_\mu - \zeta_{\mu+1} \lambda_{\mu+1} \right)^2 + 2 \sum_\mu \lambda_\mu \left(\sqrt{A} x_\mu + \sqrt{B} \left(\zeta_\mu y_\mu - \zeta_{\mu-1} y_{\mu-1} \right) \right) \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{x,y} \Bigg\}^{-1}. \tag{72}$$

Denoting $h_\mu = \sqrt{A} x_\mu + \sqrt{B} \left(\zeta_\mu y_\mu - \zeta_{\mu-1} y_{\mu-1} \right)$ and $\lambda'_\mu = \lambda_\mu + \sum_{\mu'} M_{\mu\mu'} h_{\mu'} = \lambda_\mu + \sum_{\mu'} \mathrm{e}^{-b|\mu-\mu'|} h_{\mu'}$ and using the fact that $\left\langle\!\left\langle \sum_{\mu,\mu'} h_\mu M_{\mu\mu'} h_{\mu'} \right\rangle\!\right\rangle_{x,y} = \alpha N$ we may rewrite the expression (72) in the form

$$\alpha_c = \Bigg\{ \Bigg\langle\!\!\Bigg\langle \frac{1}{\alpha N} \min_{\lambda'_\mu \geqslant \kappa + (\widehat{M} h)_\mu} \Bigg[A \sum_\mu \left(\lambda'_\mu \right)^2 + B \sum_\mu \left(\zeta_\mu \lambda'_\mu - \zeta_{\mu+1} \lambda'_{\mu+1} \right)^2 \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{x,y} \Bigg\}^{-1}. \tag{73}$$

As we noted in [19] the calculation of exact minimum on the right-hand side of (73) is by no means trivial. It is easy, however, to derive variational bounds on α_c. A similar kind of bounds on α_c have recently been discussed by us [13, 19]. We get

$$\Bigg\{ \Bigg\langle\!\!\Bigg\langle \frac{1}{\alpha N} \Bigg[A \sum_\mu \left(\lambda_\mu^{\mathrm{pr}} \right)^2 + B \sum_\mu \left(\zeta_\mu \lambda_\mu^{\mathrm{pr}} - \zeta_{\mu+1} \lambda_{\mu+1}^{\mathrm{pr}} \right)^2 \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{x,y} \Bigg\}^{-1} \leqslant \alpha_c \leqslant \Bigg\{ \Bigg\langle\!\!\Bigg\langle \frac{1}{\alpha N} \min_{\lambda'_\mu \geqslant \kappa + (\widehat{M} h)_\mu} \Bigg[A \sum_\mu \left(\lambda'_\mu \right)^2 \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{x,y} \Bigg\}^{-1}. \tag{74}$$

The minimum on the right-hand side of (74) may be found exactly and is attained for $\lambda'_\mu = 0$, provided $\kappa + \sum_{\mu'} M_{\mu\mu'} h_{\mu'} \leqslant 0$ and $\lambda'_\mu = \kappa + \sum_{\mu'} M_{\mu\mu'} h_{\mu'}$ otherwise. The configuration $\lambda_\mu^{\mathrm{pr}}$ on the left-hand side denotes, on the other hand, any probe configuration which fulfils $\lambda'_\mu \geqslant \kappa + \sum_{\mu'} M_{\mu\mu'} h_{\mu'}$ for all μ.

Expression (74) is the main result of this section since it gives precise bounds of critical capacity α_c.

5. Storage capacity for optimal input–output correlations

This case is obtained when the input correlations are in the mean equal to the output ones,

$$\left\langle\!\left\langle \xi_j^\mu \xi_{j'}^{\mu'} \right\rangle\!\right\rangle = \delta_{jj'} \eta_\mu \mathrm{e}^{-b|\mu-\mu'|} \eta_{\mu'}. \tag{75}$$

This means that $\zeta_\mu = \zeta_{\mu'}$, and the bounds (74) take the simpler form

$$\Bigg\{ \Bigg\langle\!\!\Bigg\langle \frac{1}{\alpha N} \Bigg[A \sum_\mu \left(\lambda_\mu^{\mathrm{pr}} \right)^2 + B \sum_\mu \left(\lambda_\mu^{\mathrm{pr}} - \lambda_{\mu+1}^{\mathrm{pr}} \right)^2 \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{x,y} \Bigg\}^{-1} \leqslant \alpha_c \leqslant \Bigg\{ \Bigg\langle\!\!\Bigg\langle \frac{1}{\alpha N} \min_{\lambda'_\mu \geqslant \kappa + (\widehat{M} h)_\mu} \Bigg[A \sum_\mu \left(\lambda'_\mu \right)^2 \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{x,y} \Bigg\}^{-1}. \tag{76}$$

Let us now consider the two limiting cases of small and large correlation lengths (or $b \to \infty$ and $b \to 0$, respectively).

For $b \to \infty$ we simply obtain the upper bound for α_c, using the fact that for each μ, $g_\mu = \sum_{\mu'} e^{-b|\mu-\mu'|} h_{\mu'}$ is a Gaussian random variable of the mean zero and variance one,

$$\alpha_c \leqslant \frac{1 + e^{-b}}{\left(1 - e^{-b}\right) G(\kappa)}. \tag{77}$$

The lower bound can be obtained by using as λ^{pr} the same vector, which minimizes the function on the right-hand side of (74).

The resulting formula is slightly complicated but can be simplified when $\kappa = 0$. In this limit and for $b \to \infty$ we neglect the correlation between g_μ and $g_{\mu'}$ for $\mu \neq \mu'$, which is exponentially small. The quantity g_μ is then positive with probability 1/2 and the contribution of 'kinetic' (proportional to B) term on the left-hand side of (74) can easily be evaluated. As a result we obtain:

$$2\left(1 + \frac{2}{\pi} e^{-b}\right) \leqslant \alpha_c \leqslant 2\left(1 + 2e^{-b}\right). \tag{78}$$

We see that α_c takes a value slightly larger than 2 which is Gardner's result for uncorrelated patterns (see also [23]).

The case $b \to 0$ is much more interesting. The upper bound (77) is still valid. The effective lower bound can be obtained with the help of the probe configuration

$$\lambda_\mu^{pr} = \kappa + g_{\mu_0} \tag{79}$$

where μ_0 is the value of μ for which g_μ attains maximum. Note that for such a probe configuration the 'kinetic' term (proportional to B) on the left-hand side of (74) does not contribute at all!

After simple algebra we obtain for $b \to 0$

$$\frac{2}{b\left(\kappa^2 + 1\right)} \leqslant \alpha_c \leqslant \frac{2}{bG(\kappa)}. \tag{80}$$

Alternatively we may write

$$\frac{2L_c}{\kappa^2 + 1} \leqslant \alpha_c \leqslant \frac{2L_c}{G(\kappa)}. \tag{81}$$

This is the main result of this section. As we can see, when $L_c \to \infty$, so does α_c. In particular, when L_c scales with N as $L_c \sim N^x$, while $0 \leqslant x \leqslant 1$, so does α_c. This provides an elegant generalization of Willshaw *et al* [3] result that predicts $\alpha_c \to \infty$ in the sparse coding case for maximally biased patterns. Our result, however, holds for unbiased patterns that are characterized by exponential correlations (i.e. also by a non-vanishing overlap).

We should stress that our result is very general. Our approach may be generalized to the case when the function on the right-hand side of (73) contains other terms proportional to $\left(\lambda_\mu - \lambda_{\mu+2}\right)^2$, $\left(\lambda_\mu - \lambda_{\mu+3}\right)^2$, etc. Such terms arise in a natural way when we consider a very general class of correlation matrices of the form $M(\mu-\mu') = W(|\mu-\mu'|)\exp(-b|\mu-\mu'|)$, where $W(\cdot)$ denotes a polynomial function of its argument. All terms containing differences of λs vanish for constant probe functions, such as (79). This means that our result holds for practically arbitrary (not only exponential) shapes of the correlation function (58). In such a general case, the role of the coefficient A is undertaken by

$$A = \sum_{\mu'} M_{\mu\mu'}. \tag{82}$$

If $M_{\mu\mu'}$ decays with the increase of $|\mu-\mu'|$ on a characteristic length scale L_c, then $A \propto L_c$ for $L_c \to \infty$. Thus, the bounds (80) and (81) would indeed hold.

Unfortunately, the lower bound in (81) cannot be obtained when the input–output correlations are not maximal. The $\xi_\mu\xi_{\mu+1}$ cannot then be equal to 1 for all μ and the kinetic term on the left-hand side of (74) contributes to this bound in such a case.

In the next section we consider a simple example of such a situation.

6. 'Semantically' correlated sets of data in the Hopfield network

In the case of Hopfield neural network [17] the fractional volume V may be written as

$$V = \prod_{i=1}^{N} V_i \tag{83}$$

where each of the partial fractional volumes V_i is given by

$$V_i = \frac{\int \prod_{j\neq i} \mathrm{d}J_{ij} \prod_\mu \Theta\left(\xi_i^\mu \sum_{j\neq i}(J_{ij}/\sqrt{N})\xi_j^\mu - \kappa\right)\delta\left(\sum_{j\neq i} J_{ij}^2 - N\right)}{\int \prod_{j\neq i} \mathrm{d}J_{ij}\delta\left(\sum_{j\neq i} J_{ij}^2 - N\right)}. \tag{84}$$

As we see, expression (84) corresponds to a perceptron with inputs ξ_j^μ and outputs ξ_i^μ for $j \neq i$. We consider the correlation of the sort

$$\left\langle\!\left\langle \xi_j^\mu \xi_{j'}^{\mu'} \right\rangle\!\right\rangle = \delta_{jj'} \mathrm{e}^{-b|\mu-\mu'|}. \tag{85}$$

Such correlations are described by our general expression (58) provided $\zeta_\mu = \xi_i^\mu$, $\eta_\mu = \xi_i^\mu$. It can be seen that the lower bound in (80) or (81) changes dramatically because the input–output correlations are not explicit, as in expression (57). The 'kinetic' term in (73) in the present case undergoes the change

$$\sum_\mu B\left(\lambda_\mu - \lambda_{\mu+1}\right)^2 \;\to\; \sum_\mu B\left(\xi_i^\mu \lambda_\mu - \xi_i^{\mu+1}\lambda_{\mu+1}\right)^2 \tag{86}$$

and cannot, in general, be put equal to zero, since ξ_i^μ and $\xi_i^{\mu+1}$ may have different signs.

All the further calculations are very similar to those from the previous section with only the change of (86). For the small correlation length ($b \to \infty$) we obtain the same upper bound on α_c by setting $\lambda'_\mu = 0$ for $\kappa + \sum_{\mu'} M_{\mu\mu'}h_{\mu'} < 0$ and $\lambda'_\mu = \kappa + \sum_{\mu'} M_{\mu\mu'}h_{\mu'}$ otherwise. We take as $\lambda_\mu^{\mathrm{pr}}$ on the left-hand side of (74) the same configuration as used for upper bound. Note that this configuration is independent from ξ_i^μ. The lower bound in (74) thus self-averages over ξ_i^μ and becomes

$$\left\{\left\langle\!\left\langle \frac{1}{\alpha N}\left[A\sum_\mu \left(\lambda_\mu^{\mathrm{pr}}\right)^2 + \sum_\mu \left(2B\left(\lambda_\mu^{\mathrm{pr}}\right)^2 - 2B\mathrm{e}^{-b}\lambda_\mu^{\mathrm{pr}}\lambda_{\mu+1}^{\mathrm{pr}}\right)\right]\right\rangle\!\right\rangle\right\}^{-1} \leqslant \alpha_c. \tag{87}$$

For $b \to \infty$ the last term on the left-hand side of (87) may be neglected and we obtain

$$\frac{1}{G(\kappa)} \leqslant \alpha_c \leqslant \frac{1+2e^{-b}}{G(\kappa)}. \tag{88}$$

As we see, the lower bound becomes simply equal to Gardner's result.

For the large correlation length $b \to 0$ the upper bound does not change. The lower bound can be obtained again with $\lambda_\mu^{pr} = \kappa + g_{\mu_0}$, where μ_0 is the value of μ for which g_μ attains a maximum. Unfortunately, the 'kinetic' term (86) does contribute to such a lower bound. One may see, however, that its contribution is not very large, since, as $b \to 0$, ξ_i^μ and $\xi_i^{\mu+1}$ become equal for most neighbouring pairs. We then obtain

$$\left\{\left\langle\!\left\langle A\left(\kappa+g_{\mu_0}\right)^2 + 2B\left(1-e^{-b}\right)\left(\kappa+g_{\mu_0}\right)^2\right\rangle\!\right\rangle\right\}^{-1} \leqslant \alpha_c \tag{89}$$

and, finally,

$$\frac{1}{\kappa^2+1} \leqslant \alpha_c \leqslant \frac{1}{bG(\kappa)}. \tag{90}$$

We stress that the lower bound is finite, whereas the upper one diverges with $b \to 0$.

One should notice that we can apply a better approach in order to evaluate the upper bound of α_c in (74) for $b \to 0$. It can be obtained by neglecting the positive 'kinetic' term from (73). A more careful analysis suggests that we neglect only part of the 'kinetic' term. We may, for instance, leave

$$\alpha_c \leqslant \left\{\left\langle\!\left\langle \frac{1}{\alpha N} \min_{\lambda_\mu \geqslant \kappa + g_\mu} \left[A \sum_{\mu=0}^{[\alpha N/2]} \left(\lambda_{2\mu}^2 + \lambda_{2\mu+1}^2 + B\left(\xi^{2\mu}\lambda_{2\mu} - \xi^{2\mu+1}\lambda_{2\mu+1}\right)^2 \right) \right] \right\rangle\!\right\rangle\right\}^{-1}. \tag{91}$$

In this way we divide the set of λs into a set of independent pairs. Still, the minimum on the right-hand side of (91) can, in principle, be determined exactly. Moreover, in the limit of strong correlations $b \to 0$ one can neglect terms proportional to A on the right-hand side of (91). One then obtains

$$\alpha_c \leqslant \left\{\left\langle\!\left\langle \frac{1}{2} \min_{\lambda_\mu \geqslant \kappa + g_\mu} B\left(\lambda_1\xi^1 - \lambda_2\xi^2\right)^2 \right\rangle\!\right\rangle\right\}^{-1}. \tag{92}$$

Since in the limit $b \to 0$, g_μ becomes in 100% correlated, i.e. equal to $g_{\mu+1}$, the upper bound is given by

$$\alpha_c \leqslant \left\{\left\langle\!\left\langle B\Theta\left(\kappa+g_1\right)\left(\kappa+g_1\right)^2\left(1-\xi^1\xi^2\right)\right\rangle\!\right\rangle\right\}^{-1}. \tag{93}$$

Finally, we obtain

$$\frac{1}{\kappa^2+1} \leqslant \alpha_c \leqslant \frac{2}{G(\kappa)}. \tag{94}$$

One should stress that the both bounds undergo the dramatic change for the Hopfield network in comparison with the case of the perceptron with optimal input–output correlations. Both bounds remain finite, as $b \to 0$.

At the end let us recall the result of our paper [26], in which we have shown that the bounds of the critical curve do not, practically, depend on the type of distribution of C, but only on its width. We expect that this observation will also hold for the situations considered in the present paper.

7. Conclusions

In this paper we have considered the storage of correlated patterns in neural network memories. We applied Gardner's program [1] combined with the variational approach. We stressed that the sets of the correlated data correspond much closer to realistic physical and biological situations than simple uncorrelated patterns. Two kinds of correlations were considered: 'spatial' and 'semantic' ones.

The problem of 'spatial' correlations may be solved exactly within the framework of Gardner's program [1] for a simple perceptron. The critical curves do not differ too much from Gardner's original one [1]. Especially, for $\kappa = 0$ the maximal capacity takes the value $\alpha_c = 2$. On the other hand, for the Hopfield neural network, the storage ratio may exceed Gardner's result. The situation is much more complicated in the case of the 'semantically' correlated data. The exact critical curve is hard to obtain. After applying the variational method it is quite easy, however, to investigate the bounds of this curve. We generalized the result of Willshaw *et al* [3] for a simple perceptron and showed that the storage capacity α_c tends to infinity as the correlation length L_c increases, provided the optimal input–output correlations are present. For the Hopfield network such a situation does not occur and α_c remains finite as L_c grows.

There are several further questions concerning the storage of correlated data. It is, for instance, interesting to consider the so-called generalization error for correlated training sets [15, 30] and to look for the optimal diluted network architectures in such a case [28]. As we expect, the structure of the set of patterns will determine the optimal structure of the network connections.

Acknowledgments

We thank R Monasson, who studied 'spatially' correlated unbiased patterns stored in a simple perceptron [24] and the Hopfield network [27] for providing us with preprints of his work prior to publication.

This paper has been supported by the KBN Grant No. 2–0207–91–01. ML thanks Centre d'Etudes de Saclay for their hospitality and the French Ministry of Research and Technology for financial support.

References

[1] Gardner E 1988 The space of interactions in neural network models *J. Phys. A: Math. Gen.* **21** 257–70
[2] Nadal J-P and Toulouse G 1990 Information storage in sparsely coded memory nets *Network* **1** 41–60
[3] Willshaw D J, Buneman O P and Longuet-Higgins H C 1969 Non-holographic associative memory *Nature* **222** 960
[4] Krauth W and Mezard M 1989 Storage capacity of memory networks with binary couplings *J. Physique* **50** 3057–66
[5] Gutfreund H and Stein Y 1990 Capacity of neural networks with discrete couplings *J. Phys. A: Math. Gen.* **23** 2613–30
[6] Barkai E, Hansel D and Kanter I 1990 Statistical mechanics of a multilayered neural network *Phys. Rev. Lett.* **65** 2312–5
[7] Kanter I 1992 Information theory of a multilayer neural network with discrete weights *Europhys. Lett.* **17** 181–6
[8] Barkai E, Hansel D and Sompolinsky H 1992 Broken symmetries in multilayered perceptrons *Phys. Rev.* A **45** 4146–61
[9] Griniasty M and Grossman T 1992 Two-layer perceptrons at saturation *Phys. Rev.* A **45** 8924–37

[10] Gardner E and Derrida B 1988 Optimal storage properties of neural network models *J. Phys. A: Math. Gen.* **21** 271–84

[11] Amit D J, Evans M R, Horner H and Wong K Y M 1990 Retrieval phase diagrams for attractor neural networks with optimal interactions *J. Phys. A: Math. Gen.* **23** 3361–81

[12] Wong K Y M and Sherrington D 1990 Optimally adapted attractor neural networks in the presence of noise *J. Phys. A: Math. Gen.* **23** 4659

[13] Tarkowski W, Komarnicki M and Lewenstein M 1991 Optimal storage of invariant sets of patterns in neural network memories *J. Phys. A: Math. Gen.* **24** 4197–217

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

[15] Watkin T L H, Rau A and Biehl M 1993 The statistical mechanics of learning a rule *Rev. Mod. Phys.* in press

[16] Minsky M and Papert S 1969 *Perceptrons* (Cambridge, MA: MIT)

[17] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational capabilities *Proc. Natl Acad. Sci. USA* **79** 2554–8

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

[19] Lewenstein M and Tarkowski W 1992 Optimal storage of correlated patterns in neural network memories *Phys. Rev.* A **46** 2139–42

[20] Fassnacht C and Zippelius A Recognition and categorization in a structured neural network with attractor dynamics *Network* **2** 63–84

[21] Miyashita Y 1988 Neuronal correlate of visual associative long-term memory in the primate temporal cortex *Nature* **335** 817–9

[22] Griniasty M, Tsodyks M V and Amit D J 1992 Conversion of temporal correlations between stimuli to spatial correlations between attractors *Preprint* Universitá di Roma, 856

[23] Cover T M 1965 Geometrical and statistical properties of systems of linear inequalities with application in pattern recognition *IEEE Trans. Electron. Comput.* **EC-14** 326–34

[24] Monasson R 1992 Properties of neural networks storing spatially correlated patterns *J. Phys. A: Math. Gen.* **25** 3701–20

[25] Lewenstein M and Tarkowski W 1992 Optimal storage of correlated patterns in neural networks *Proc. 18th IUPAP Int. Conf. Statistical Physics (Berlin, August 1992)*

[26] Tarkowski W and Lewenstein M 1992 Estimates of optimal storage conditions in neural network memories based on random matrix theory *J. Phys. A: Math. Gen.* **25** 6251–64

[27] Monasson R 1993 Storage of spatially correlated patterns in autoassociative memories *J. Physique I* submitted

[28] Tarkowski W, Lewenstein M and Nowak A 1993 Optimal network architecture for storage of 'spatially' correlated patterns in the Hopfield neural network *in preparation*

[29] Fyodorov Y V and Mirlin A D 1991 Scaling properties of localization in random band matrices: A σ-model approach *Phys. Rev. Lett.* **67** 2405–9

[30] Tarkowski W and Lewenstein M 1993 Learning from correlated examples in a perceptron *J. Phys. A: Math. Gen.* in press