



HAL
open science

Storage capacity of a Potts-perceptron

Jean-Pierre Nadal, Albrecht Rau

► **To cite this version:**

Jean-Pierre Nadal, Albrecht Rau. Storage capacity of a Potts-perceptron. Journal de Physique I, 1991, 1 (8), pp.1109-1121. 10.1051/jp1:1991104 . jpa-00246396

HAL Id: jpa-00246396

<https://hal.science/jpa-00246396>

Submitted on 4 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Classification
 Physics Abstracts
 05.20 — 87.30

Storage capacity of a Potts-perceptron

Jean-Pierre Nadal ⁽¹⁾ and Albrecht Rau ⁽²⁾

⁽¹⁾ Lab. de Physique Statistique (*), Ecole Normale Supérieure, 24 rue Lhomond, F-75231 Paris Cedex 05, France

⁽²⁾ Dept. of Theoretical Physics, University of Oxford, 1 Keble Rd., GB-Oxford OX1 3NP, G.B.

(Received 22 January 1991, accepted in final form 30 April 1991)

Abstract. — We consider the properties of « Potts » neural networks where each neuron can be in Q different states. For a « Potts-perceptron » with N Q -states input neurons and one Q' states output neuron, we compute the maximal storage capacity for unbiased patterns. In the large N limit the maximal number of patterns that can be stored is found to be proportional to $N(Q-1)f(Q')$, where $f(Q')$ is of order 1.

1. Potts-perceptrons, linear machines and winner take all systems.

The Hopfield model [1] of a formal neural network used neurons having two possible states, and was based on an analogy with spin glass systems. A natural extension of this model of associative memory is to consider neurons taking more than two states: instead of having neuronal state described by an Ising like variable, one has a Potts [2] like variable. One then obtains a « Potts neural network » [3], which is to the Potts Glass [4] what the Hopfield model is to the Spin Glass model. The statistical physics of Potts neural networks has been studied for attractor neural networks with Hebbian learning rules ([3, 5-7]). In these studies, one is considering a Potts-Attractor Neural Network (PANN): each neuron, whose activity can take Q different values, is connected to every other neuron, and the response to a stimulus (a given initial configuration of activities) is the attractor to which the natural dynamics of the net leads. Clearly one can also consider non Hebbian learning rules — and in particular non symmetric couplings — as for the case of binary neurons. One can also consider Potts-Perceptrons (PP), that is (possibly multilayer) feedforward networks. In that case, the number of possible states may differ from layer to layer. In the following we will consider only the simplest case, that is one input layer with Q -states neurons and one output layer with Q' states neurons (and no hidden layer). Such systems are well known in data analysis literature under the name of « linear machines » [8]. The binary perceptron [9] is one particular case. In fact the perceptron algorithm, as well as its variants, is easily adapted to multistates neurons ([8, 10]).

(*) Laboratoire associé au CNRS (URA 1306) et aux Universités Paris VI et Paris VII.

The motivations for considering such systems are manifolds. i) Consider the case of a feedforward network. As just mentioned, such system allows to deal with multiclass classification tasks, and with strings of data made of N items, each one taking Q possible different values. Here are few examples. In image processing, $Q = Q'$ would be the number of grey levels. In the analysis of DNA sequences, each input node would be in one of the $Q = 4$ letters, A, T, G or C and the output might be binary for codon *versus* exon (see e.g. [11]). In the analysis of proteins, if one is interested in predicting the secondary structure from the sequence of amino acids, Q would be 20, the number of different amino acids, and Q' would be 3, the number of different possible structures (α helix, β sheet or random) ii) Consider a feedforward network with N inputs and one Q' -state output neuron. One can consider the output neurons as Q' binary neurons, with a « winner take all » updating rule, and only the neuron receiving the highest input field will fire. Such systems are commonly found in the literature, both for biological modeling and engineering applications. There are in particular models of self-organisation based on the competition between neurons (see e.g. [12]). iii) We have already mentioned the fact that such system are a natural extension of binary neural networks from the physicist point of view. In addition, for a multiclass classification task (including the two class, binary neuron case), the use of a Potts perceptron (as precisely defined in the next section) can be justified in Bayes decision theory (see e.g [8] p. 16 and [11]). iv) Lastly we note that, on the biological side, there are indications that cortical columns behave coherently, having a small number of possible coherent states [13]. Hence a Potts-attractor neural network may give some insight on the behavior of a network of cortical columns.

Here we will be concerned with Potts neural networks for associative memory (PANN) or classification tasks (PP), without specifying the learning algorithm. In particular we will follow the approach initiated by Gardner [14] for computing the maximal storage capacity of a Potts neural network. Namely, we will compute the fractional volume of the weights which realize the learning of a set of randomly chosen patterns. We will limit our study to unbiased patterns, although the computation can be extended to biased patterns. In section 2 we present the model and discuss gauge invariance properties. We will show its relevance to the dynamics of the network. In section 3 we present the storage capacity results. The details of the computation will be given in the Appendix.

2. Gauge invariances, encoding and dynamics.

Let us first define the model for the simplest feedforward case, the Potts perceptron with N inputs (Q -state neurons) and one Q' -state output neuron. The local field at state s' is given by

$$h_{s'} = \sum_{j,s} J_j(s', s) \delta_{s, n_j} . \quad (1)$$

The synaptic matrix $J_j(s', s)$ indicates the weight of a signal coming from node j , which is in state s , on the state s' of the processing unit. Unless otherwise indicated $j \in \{1.. N\}$, $s \in \{1.. Q\}$ and $s' \in \{1.. Q'\}$. A pattern of input activities will be denoted $(n_j, j = 1.. N)$, with $n_j \in \{1.. Q\}$. The probabilist decision rule at a « temperature » $T = 1/\beta$, is defined by

$$s'_{\text{out}} = s' \text{ with probability } \frac{\exp \beta (h_{s'} - \theta_{s'})}{\sum_s \exp \beta (h_{s'} - \theta_s)} \quad (2)$$

where the $\theta_{s'}$ are the thresholds. At zero temperature, the decision rule is

$$s'_{\text{out}} = \{s'_0 | h_{s'_0} - \theta_{s'_0} \rangle h_{s'} - \theta_{s'} \forall s' \neq s'_0\} . \quad (3)$$

This is a « winner-takes-all » rule : one can consider the Q' state output neuron as Q' binary neurons ; neuron s' receives the input field $h_{s'}$ defined by (1), and only the neuron with the highest local field will become active. The generalization to N' output neurons or to a fully connected net (PANN) — for which $N' = N$ and $Q = Q'$ — is straightforward. Indeed, as for the case of binary neurons, for the basic properties considered in this section as well for the storage analysis in the next section, one can always focus on one particular (output) neuron, which is equivalent to consider a perceptron with only one output neuron.

Updating rules (2) and (3) are invariant under the following translations

$$\begin{aligned} J_j(s', s) &\rightarrow J_j(s', s) + u_j(s) \\ \theta_{s'} &\rightarrow \theta_{s'} + u_0 \end{aligned} \quad (4)$$

and under

$$J_j(s', s) \rightarrow J_j(s', s) + v_j(s') \quad (5)$$

$$\theta_{s'} \rightarrow \theta_{s'} - \sum_j v_j(s') \quad (6)$$

where u_0 , the $u_j(s)$ and the $v_j(s')$ are arbitrary real numbers. The first transformation (4) add to each local field a term which, although a function of the input pattern, is independent of the output states. The second set of translations on the couplings (5) modifies the local fields by a term independent of the input pattern, which then can be absorbed in the redefinition (6) of the thresholds. In addition the updating rule is invariant under a global rescaling :

$$J_j(s', s) \rightarrow \lambda J_j(s', s) , \quad \theta_{s'} \rightarrow \lambda \theta_{s'} , \quad \beta \rightarrow \beta/\lambda \quad (7)$$

for any strictly positive real number λ .

These gauge invariances allows to reduce the number of parameters by « fixing the gauge ». One possible choice is to pick one particular input state s_0 and one particular output state s'_0 , and to set to zero all the couplings $J_j(s'_0, s)$ and $J_j(s', s_0)$. This is a choice made by several authors ([8, 11]). It is clear however that any gauge choice will do. In the next section we will make another choice namely :

$$\sum_s J_j(s', s) = 0 ; \quad \forall s' \quad (8)$$

$$\sum_{s'} J_j(s', s) = 0 ; \quad \forall s \quad (9)$$

which, as we argue now by considering the dynamics, is more natural.

The dynamics of the network, whether it is an attractor neural net or a feedforward net, should only depend on gauge invariant quantities. This can be easily shown for the one step dynamics. For simplicity, consider a perceptron with NQ -states inputs and one binary output ($Q' = 2$). In that case one can write the updating rule (3) as

$$\sigma = \text{sgn} (h) \quad h = \sum_j w_j(n_j) - \theta_0 \quad (10)$$

with $w_j(s) = J_j(2, s) - J_j(1, s)$ and $\theta_0 = \theta_2 - \theta_1$, and $\sigma = \pm 1$ is the binary output. Suppose

now that the presented input pattern is a noisy version of a learned pattern, say pattern 1 :

$$\begin{aligned} n_j &= n_j^1 \quad \text{with probability } f \\ n_j &= s \neq n_j^1 \quad \text{with probability } \frac{1-f}{Q-1} \end{aligned} \quad (11)$$

As for the binary perceptron [15] one can show that, for large N , the input field is a Gaussian variable, so that here the probability f' for getting the correct output is :

$$f' = H \left(- \frac{\langle h \rangle}{\sqrt{\langle h^2 \rangle_c}} \right) \quad (12)$$

where

$$H(x) = \int_x^\infty Dy, \quad (13)$$

the Gaussian measure being indicated by

$$Dy = \exp(-y^2/2) / \sqrt{2\pi}. \quad (14)$$

The two first cumulants of h , $\langle h \rangle$ and $\langle h^2 \rangle_c$ (the averages being over all possible initial conditions with the same bias f), are given by

$$\begin{aligned} \langle h \rangle &= m \sum_j J_j(n_j^1) - \theta \\ \langle h^2 \rangle_c &= m(1-m) \sum_i [J_j(n_j^1)]^2 + (1-m) \frac{1}{Q} \sum_{i,s} [J_j(s)]^2 \end{aligned} \quad (15)$$

and the renormalized couplings and threshold are given by

$$\begin{aligned} J_j(s) &\equiv w_j(s) - \frac{1}{Q} \sum_i w_j(i) \\ \theta &\equiv \theta_0 - \frac{1}{Q} \sum_i w_j(i). \end{aligned} \quad (16)$$

The parameter m which appears above characterize the correlation between the pattern and the noisy version :

$$\langle Q\delta_{n_i, s} - 1 \rangle = m(Q\delta_{n_i^1, s} - 1), \quad (17)$$

and takes the value

$$m = \frac{fQ-1}{Q-1} \quad (18)$$

It is the natural generalization of the usual « overlap » between the input string and the pattern in the case of binary neurons ($Q = 2$). It takes the value 1 if there is no noise in the input, and 0 if the input is uncorrelated with the pattern. It is important to notice that all the dependance in the initial conditions is concentrated in this parameter. The above formula (12) can easily be generalized to finite temperature. Note also that, in the case of an attractor

neural network, (12) fully characterizes the dynamics in the case of a highly diluted network [15].

To conclude, it is thus more natural to define the local fields as

$$h_{s'} = \sum_{j,s} J_j(s',s) (Q\delta_{s,n_j} - 1) \tag{19}$$

and to explicitly fixe the gauges by (8) and (9). This choice, in the case of the binary perceptron, that is for $Q = Q' = 2$, corresponds to the standard choice of representing the neuron activity by a spin like variable, $S_j = \pm 1$. With these choices, the case « without threshold » is the only situation which preserves the symmetry between the Q states : in particular, when learning random, unbiased patterns, the optimal thresholds will be zero.

At this point it is useful to make some comments on the encoding problem, that is on the way data are represented in the input layer. The definition (1) of the local field can be understood as the field for a neural net with NQ entries ; for each j , the activity is encoded as a string of Q nodes. The direct interpretation of (1) is that the s -th state is represented by a Q -bit string with only the n_j -th bit on. In that case $J_j(s',s)$ is the coupling from node s at site j . This is the most common choice of encoding found in the literature. But one can also consider other representations. In the general case, state n is defined by a given configuration of the Q nodes

$$\mathbf{x}^n = (x_s^n, s = 1 \dots Q),$$

with the Q vectors $(\mathbf{x}^n, n = 1 \dots Q)$ forming a base of R^Q . A typical example is the « thermometric » representation, where $x_s^n = 1$ for $s = 1$ to n_j , and 0 otherwise. If we note $\mathbf{J}_j(s')$ the vector whose s component is the coupling coming from node s at site j , then $J_j(s',n) = \mathbf{J}_j(s') \cdot \mathbf{x}^n$. Otherwise stated, all what matters is that there are Q different values for a given s' and a given j . What we have just shown is that all possible choices are equivalent, *via* a reinterpretation of the couplings and of the thresholds. Note that these considerations apply as well for the first layer of any feedforward neural network for which the input data are strings of « letters ». The choice, that we adopt in the following, thus corresponds to take the updating rule (3) with the definition (19) for the local fields, and the gauge choices (8) and (9).

3. Storage capacity of a Potts-perceptron.

We now consider the storage capacity of a Potts perceptron with N inputs, each one having Q possible states, and one Q' state output neuron. We will compute the maximal number of random input-output pairs $\{(n_j^\mu, j = 1 \dots N), n'^\mu\}$, $\mu = 1 \dots p$, for which there exists a set of couplings such that, for every μ ($\mu = 1, \dots p$), (3) is true with $s'_{out} = n'^\mu$ if the input is pattern number μ . All Potts states occur in the set of patterns with the same probability, which is equivalent to

$$\langle Q\delta_{s,n_j^\mu} - 1 \rangle_\xi = 0.$$

Here the pattern average is indicated by $\langle \dots \rangle_\xi$. Since the patterns are unbiased, we can set all the thresholds to zero.

The translational freedom of the local field is fixed by

$$\sum_s J_j(s',s) = 0; \quad \forall s' = 1 \dots Q' - 1 \tag{20}$$

$$\sum_{s'} J_j(s', s) = 0; \quad \forall s = 1 \dots Q'. \quad (21)$$

It should be noted that the constraint of (20) for $s' = Q'$ is included in (21). We will consider only the case when the scalar degree of freedom is fixed independently for each s' by

$$\sum_{j,s} (J_j(s', s))^2 = N\gamma. \quad (22)$$

The constant γ is of course arbitrary, but for later convenience we take :

$$\gamma = (Q - 1) (Q' - 1) / Q'. \quad (23)$$

From the gauge invariance we see that there are only $N(Q - 1) (Q' - 1)$ free parameters. For each pattern there are $(Q' - 1)$ inequalities which need to be satisfied. Thus one can expect that the maximal number p_{\max} of patterns that can be stored will be of order $N(Q - 1) (Q' - 1) / (Q' - 1)$, that is

$$p_{\max} = \alpha (Q - 1) N. \quad (24)$$

We will see that this is the case, with α a slowly varying, bounded, function of Q' only.

We follow Gardner's [14] maximum entropy method, and we define the partition function of the system as

$$Z = \int d\mu(\mathbf{J}) \prod_{\mu} \prod_{s (\neq n^{\mu})} \Theta[h^{\mu}(n^{\mu}) - h^{\mu}(s') - \kappa \sqrt{NQ(Q-1)}]. \quad (25)$$

Here $h^{\mu}(s')$ is the local field when the input is pattern μ ,

$$h^{\mu}(s') = \sum_{j,s} J_j(s', s) (Q\delta_{s,n^{\mu}} - 1),$$

and $d\mu(\mathbf{J})$ is the natural measure in the space of interactions which is compatible with (20)-(22). As in [14], we are asking for a minimal stability κ . (In the case of $Q = Q' = 2$, the stability κ as defined in (25) differs by a factor $\sqrt{2}$ from the standard stability parameter introduced by Gardner [14]). In order to evaluate the quenched average over the distribution of patterns of the entropy $S \sim \langle \ln Z \rangle_{\xi}$ we use the *replica trick*. Assuming the validity of a replica symmetric solution and shrinking the volume of interactions to zero we find for the maximum storage capacity α (as defined by (24))

$$\alpha^{-1} = \int Dy \left[\frac{H_2}{2} + (2(H_1)^2 - H_0 H_2) \frac{(1 - H_0)^{Q'} - 1 + Q' H_0 - (H_0)^2 Q' (Q' - 1)^{1/2}}{Q' (Q' - 1) (H_0)^3} \right]. \quad (26)$$

Here Dy is the Gaussian measure (14) and we have introduced the functions

$$H_i(y) \equiv \int_{-\kappa}^{+\infty} Dt (t + y + \kappa)^i; \quad i = 0, 1, 2. \quad (27)$$

Details of the derivation of (26) is given in the Appendix. The justification for restricting to the replica symmetric calculation is the same as for the usual, $Q = Q' = 2$, case. Indeed, any solution can be continuously transformed into any other solution, as can be seen from the geometrical picture (one particular solution is a set of Q' vectors with well defined angles between them, and one go through any solution by global rotations).

For $Q' = 2$ one recovers the critical capacity found by Gardner, with in particular $\alpha(\kappa = 0) = 2$. For $\kappa = 0$ and $Q' = 3$ one can evaluate the storage capacity analytically, and one finds :

$$\alpha(Q' = 3) = 2 \left(1 - \frac{\sqrt{3}}{4\pi} \right)^{-1} \approx 2.320 . \tag{28}$$

In the large Q' limit one gets

$$\alpha(Q' \gg 1) = \int Dy [H_2 - (H_1)^2/H_0] \approx 3.850 . \tag{29}$$

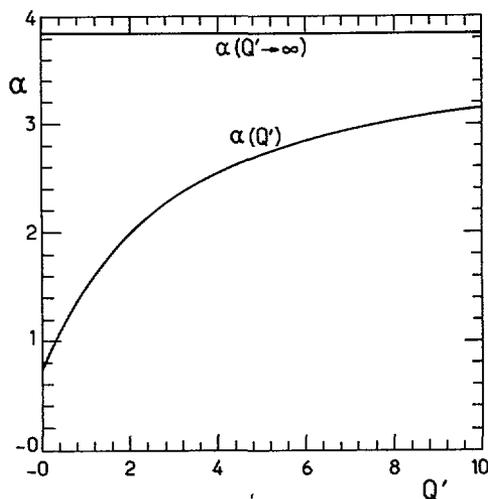


Fig. 1. — Maximal capacity for $\kappa = 0$ as a function of Q' .

The results for solving the integral in (26) numerically for $\kappa = 0$ are shown in figure 1 and table I. In figure 2 is shown the maximal capacity as a function of κ for $Q' = 2, 3$ and 4.

Table I. — Maximal storage capacity at $\kappa = 0$.

Q'	$\alpha(Q')$
2	2.000
3	2.320
4	2.546
5	2.714
6	2.844
\vdots	\vdots
∞	3.850

The critical capacity as measured by α is thus an increasing function of Q' , saturating at a value of about 3.85. However, the information content $i(Q')$, in nat per (free) parameters, that is

$$i(Q') = \alpha(Q') \ln Q' / (Q' - 1) , \tag{30}$$

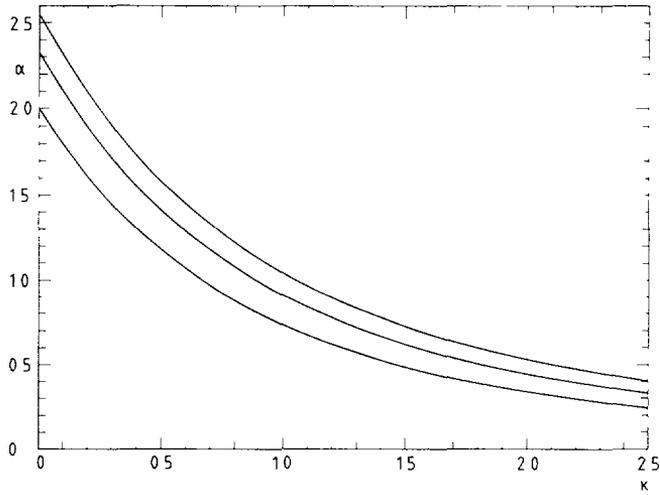


Fig. 2. — Maximal capacity as a function of the stability parameter κ for $Q' = 2, 3$ and 4

is a decreasing function of Q' , going to zero when Q' goes to infinity. Hence we find that, qualitatively, the optimal behavior with Q' is similar to the one found by Kanter in the case of a Hebb rule. It is important to note that, as mentioned in the introduction, there are perceptron type algorithms which allow to find a set of couplings whenever there exists at least one solution ([8, 10]). Moreover, such algorithms allow to respect any particular gauge choice. To see this, let us give the perceptron algorithm for the choice of the gauges (20) and (21). One starts with zero couplings. Then, the following is repeated until convergence :

Pick a pattern μ at random. For any $s' \neq n'^{\mu}$ such that $h_{s'}^{\mu} > h_{n'^{\mu}}^{\mu}$, make a learning step for every j and every s by :

$$\begin{aligned} J_j(n'^{\mu}, s) &\rightarrow J_j(n'^{\mu}, s) + (Q\delta_{s, n_j^{\mu}} - 1) \\ J_j(s', s) &\rightarrow J_j(s', s) - (Q\delta_{s, n_j^{\mu}} - 1). \end{aligned} \quad (31)$$

It is clear that at each time step the current couplings will satisfy (20) and (21).

Conclusion.

In this article we analysed the storage capacity of a Potts perceptron. We have obtained that the maximal number of pattern that can be stored is proportional to $N(Q - 1)$, the prefactor α being a slowly increasing and bounded function of Q' . Our calculation can easily be generalized to other cases, e.g. the case of biased patterns. More generally, it is clear that most of the analysis done for the binary perceptron ($Q = Q' = 2$) could be generalized to the Potts perceptron. In particular it would be interesting to consider other cost function than the number of errors, as considered here, and to consider the properties above saturation. (We note that very recently ([19]) a method, simpler than the replica techniques, has been derived in order to study binary perceptrons at and above criticality, for any choice of cost function. We have checked that it can be easily applied to Potts-perceptrons.) In practical applications of neural networks, winner take all systems are currently used. This is presently the main reason for pursuing the analytical study of such systems.

Acknowledgements.

One of us (A.R.) wants to thank the group at the Ecole Normale Supérieure for their kind hospitality during his stay there. He further acknowledges financial support by the *Studienstiftung des deutschen Volkes*, the *SERC* and *Corpus Christi College, Oxford*. This work has been supported by EEC BRAIN twinning contracts (ST2J.0312.C(EDB) and ST2J.0422.C(EDB)). We would like to thank Marc Mézard for several discussions.

Appendix.

In order to evaluate $\langle \ln Z \rangle_\xi$ we use the identity

$$\langle \ln Z \rangle_\xi = \lim_{n \rightarrow 0} \frac{\langle Z^n \rangle_\xi - 1}{n} \tag{i}$$

We thus consider

$$Z^n = \int d\mu(\mathbf{J}^a) \prod_{\mu, a} \prod_{s' (\neq n^\mu)} \left[\int dx_{\mu a}^{s'} \int_{\kappa \sqrt{Q(Q-1)}}^{\infty} d\lambda_{\mu a}^{s'} \times \exp \left\{ i x_{\mu a}^{s'} \left(\frac{\Delta_{\mu a}^{s'}}{\sqrt{N}} - \lambda_{\mu a}^{s'} \right) \right\} \right] \tag{ii}$$

where $a = 1 \dots n$ and we introduced

$$\Delta_{\mu a}^{s'} \equiv \sum_{j, s} (J_j^a(n^\mu, s) - J_j^a(s', s)) (Q\delta_{s, n^\mu} - 1). \tag{iii}$$

The measure on the couplings is defined by

$$d\mu(\mathbf{J}^a) \equiv \left(\prod_{j, a, s'} \int dJ_j^a(s', s) \right) \prod_{a, s'}^{Q-1} \delta \left(\sum_s J_j^a(s', s) \right) \prod_{a, s} \delta \left(\sum_{s'} J_j^a(s', s) \right) \times \\ \times \prod_{a, s'} \delta \left(\frac{\sum_{j, s} J_j^a(s', s)^2}{N\gamma} - 1 \right) / V \tag{iv}$$

where V is a normalization constant and γ is the constant fixing the normalization (22). We will compute $\langle Z^n \rangle_\xi$, the average of Z^n over the pattern distribution, in the small n limit :

$$\exp nN (Q - 1) f = \langle Z^n \rangle_\xi. \tag{v}$$

Since the patterns are unbiased one has :

$$\langle Q\delta_{s, n^\mu} - 1 \rangle_\xi = 0 \tag{vi}$$

$$\langle (Q\delta_{s, n^\mu} - 1) (Q\delta_{t, n^\mu} - 1) \rangle_\xi = Q\delta_{s, t} - 1. \tag{vii}$$

Then performing the average over the n^μ one gets

$$\exp nN (Q - 1) f = \prod_{a, b} \prod_{s', t'} \int dq_{ab}^{s' t'} \exp (G_p + G_j) \tag{viii}$$

with :

$$\exp G_p = \prod_{\mu} \prod_{s \neq n^{\mu}, a} \int dx_{\mu a}^{s'} \int_{\kappa}^{\infty} \frac{d\lambda_{\mu a}^{s'}}{\sqrt{Q(Q-1)}} \exp \{-i x_{\mu a}^{s'} \lambda_{\mu a}^{s'}\} \times \\ \times \exp -\frac{Q\gamma}{2} \sum_{a,b} \sum_{(s,t \neq n^{\mu})} x_{\mu a}^{s'} x_{\mu b}^{t'} [q_{ab}^{n^{\mu} n^{\mu}} - q_{ab}^{n^{\mu} t'} - q_{ab}^{s' n^{\mu}} + q_{ab}^{s' t'}] \quad (\text{ix})$$

and :

$$\exp G_J = \left(\prod_a \int d\mu(J^a) \right) \times \\ \times \prod_{a \ s < t'} \delta \left(q_a^{s' t'} - \frac{\sum_{j^s} J_j^a(s', s) J_j^a(t', s)}{N\gamma} \right) \\ \times \prod_{a < b, s} \delta \left(q_{ab}^{s' s'} - \frac{\sum_{j^s} J_j^a(s', s) J_j^b(s', s)}{N\gamma} \right) \\ \times \prod_{a \neq b, s' < t} \delta \left(q_{ab}^{s' t'} - \frac{\sum_{j^s} J_j^a(s', s) J_j^b(t', s)}{N\gamma} \right) \quad (\text{x})$$

Here we have introduced the generalization of the Edwards-Anderson order parameter

$$q_{ab}^{s' t'} \equiv \frac{1}{N\gamma} \sum_{j^s} J_j^a(s', s) J_j^b(t', s). \quad (\text{xi})$$

Assuming a replica symmetric saddle point we write :

$$q_{ab}^{s' t'} = \delta_{a,b} (\delta_{s',t'} + (1 - \delta_{s',t'}) q_2) + (1 - \delta_{a,b}) (\delta_{s',t} q_1 + (1 - \delta_{s',t'}) q_3). \quad (\text{xii})$$

The first, diagonal, term corresponds to the normalization condition (22). The physical interpretation of the order parameters q_1 , q_2 and q_3 is the following :

$$q_1 = \frac{1}{N\gamma} \sum_{sj} \langle [J_j(s', s)]_T^2 \rangle_{\xi} \quad (\text{xiii})$$

$$q_2 = \frac{1}{N\gamma} \sum_{sj} \langle [J_j(s', s) J_j(t', s)]_T \rangle_{\xi} \quad (\text{xiv})$$

$$q_3 = \frac{1}{N\gamma} \sum_{sj} \langle [J_j(s', s)]_T [J_j(t', s)]_T \rangle_{\xi} \quad (\text{xv})$$

where $[\cdot]_T$ denotes the « thermal » average as defined from the partition function Z .

To compute the J_j^a -independent part, G_p , one first note that $\exp G_p$ is a product of p independent random variables, and using the fact that the n^{μ} are equally distributed, one arrives at :

$$\exp G_p = \exp p \ln \prod_{s=1}^{Q'} \prod_{a=1}^n \int dx_a^{s'} \int_{\kappa}^{\infty} \frac{d\lambda_a^{s'}}{\sqrt{Q(Q-1)}} \exp -i x_a^{s'} \lambda_a^{s'} - \frac{1}{2} Q\gamma A$$

where

$$A = (1 - q_1 + q_3 - q_2) \left[\sum_{s'} \sum_a (x_a^{s'})^2 + \sum_a \left(\sum_{s'} x_a^{s'} \right)^2 \right] + (q_1 - q_3) \left[\sum_{s'} \left(\sum_a x_a^{s'} \right)^2 + \left(\sum_a \sum_{s'} x_a^{s'} \right)^2 \right]. \quad (\text{xvi})$$

To perform the integration over the x 's one has to decouple the replicas by introducing Gaussian additional variables. Then, in the small n limit, one finally gets

$$\exp G_p = \exp \left\{ pn \left(\prod_{s'=1}^{Q'-1} \int Dz_{s'} \right) \frac{1}{\sqrt{Q'}} \exp \left[\frac{1}{2Q'} \left(\sum_{s'} z_{s'} \right)^2 \right] \times \right. \\ \left. \times \ln \left[\prod_{s'=1}^{Q'-1} \int_{X_{s'}}^\infty Dt_{s'} \exp \frac{1}{2Q'} \left(\sum_{s'} t_{s'} \right)^2 \right] \right\} \quad (\text{xvii})$$

where

$$X_{s'} = \frac{\sqrt{Q'/(Q'-1)} \kappa + z_{s'} \sqrt{q_1 - q_3}}{\sqrt{1 - q_2 - q_1 + q_3}} \quad (\text{xviii})$$

Now we perform the integration over the J_s^a . Introducing integral representations of the δ distributions and using the fact that n is small, one arrives at the following expression for (x) :

$$\exp G_J = \exp \left\{ \frac{N}{2} n(Q'-1)(Q-1) [-\ln(E + \hat{q}_1 + \hat{q}_2 - \hat{q}_3) + E + q_1 \hat{q}_1] + \right. \\ \left. + \frac{\hat{q}_1 - \hat{q}_3}{E + \hat{q}_1 + \hat{q}_2 - \hat{q}_3} - (Q'-1)q_2 \hat{q}_2 + (Q'-1)q_3 \hat{q}_3 \right\}. \quad (\text{xix})$$

Here $\hat{q}_i, i = 1, 2, 3$ are the conjugate momenta of q_i and E is the momentum introduced to enforce the normalization (22). Note that G_J is of order $Nn(Q-1)(Q'-1)$ and G_p of order $pn(Q'-1)$, which is consistent with the scaling (24) for p .

The order parameters q_i, \hat{q}_i and E are determined by the saddle point equations. One can first express all the parameters as functions of q_1 using the following saddle point equations :

$$q_2 = -\frac{1}{Q'-1}; \quad q_3 = -\frac{q_1}{Q'-1} \quad (\text{xx})$$

$$\hat{q}_1 - \hat{q}_3 = \frac{q_1}{(1 - q_1)^2}; \quad E + \hat{q}_1 + \hat{q}_2 - \hat{q}_3 = \frac{1}{1 - q_1} \quad (\text{xxi})$$

The relations (xx) were to be expected from geometrical considerations: the couplings associated with each output state s' form a vector in a $N(Q-1)$ dimensional space (20) means that the sum of all these Q' vectors is zero, and all these vectors have the same norm (22). Thus their two by two scalar products are fixed to $-\frac{1}{Q'-1}$.

Inserting (xx) and (xxi) in (xvii) and (xix), one gets

$$X_{s'} = \frac{\sqrt{q_1} z_{s'} + \kappa}{\sqrt{1 - q_1}} \quad (\text{xxii})$$

and (omitting non essential constants)

$$f = \frac{1}{2} \left[\frac{1}{1 - q_1} + \ln (1 - q_1) \right] + \alpha \Psi (\kappa, q_1) \tag{xxiii}$$

$$\Psi (\kappa, q_1) = \prod_{s'=1}^{Q'-1} \int Dz_s \frac{\exp \frac{1}{2Q'} \left(\sum_s z_s \right)^2}{\sqrt{Q'}} \ln \left[\prod_{s=1}^{Q'-1} \int_{X_s}^{\infty} Dt_s \exp \frac{1}{2Q'} \left(\sum_s t_s \right)^2 \right] \tag{xxiv}$$

Here Q_1 is the solution of the saddle point equation .

$$\frac{q_1}{(1 - q_1)^2} = -2 \alpha \frac{\delta}{\delta q_1} \Psi (\kappa, q_1) . \tag{xxv}$$

The maximal capacity α will be obtained by taking the limit $q_1 \rightarrow 1$. In the asymptotic behaviour of (xxiv) one can see that there will be a non trivial contribution from the integration on the z_s , on any sector such that at least one $X_{s'}$ is positive, and thus going to plus infinity. If k of them — say $s' = 1$ to k — go to infinity, we find the following asymptotic expression as $q_1 \rightarrow 1$ for the argument of the logarithm in (xxiv) :

$$\rightarrow \frac{\exp \left\{ -\frac{1}{2} \Sigma' X_{s'}^2 + \frac{1}{2(k+1)} (\Sigma' X_s)^2 \right\}}{\Pi' \left[X_s - \frac{1}{k+1} (\Sigma' X_{s'}) \right]} \tag{xxvi}$$

The primed sums and products are taken from $s' = 1$. k . One has to sum over all the possible choices of such sectors, which gives :

$$\Psi (\kappa, q_1) \stackrel{q_1 \rightarrow 1}{\sim} \sum_{k=1}^{Q'-1} \binom{Q'-1}{k} \left(\prod_{s'=1}^k \int_{-\kappa}^{\infty} Dz_{s'} \right) \left(\prod_{s=k+1}^{Q'-1} \int_{-\infty}^{-\kappa} Dz_{s'} \right) \times \frac{\exp \left\{ \frac{1}{2Q'} \left(\sum_{s'=1}^{Q'-1} z_{s'} \right)^2 \right\}}{\sqrt{Q'}} \ln \left[\frac{\sqrt{\frac{Q'}{k+1}} \exp \left\{ -\frac{1}{2} \Sigma' X_{s'}^2 + \frac{1}{2(k+1)} (\Sigma' X_s)^2 \right\}}{\Pi' \left[X_{s'} - \frac{1}{k+1} (\Sigma' X_s) \right]} \right] \tag{xxvii}$$

with $\binom{Q'-1}{k} = \frac{(Q'-1)!}{k!(Q'-1-k)!}$ From this we get the following expression for the critical storage capacity :

$$\frac{N(Q-1)}{p_{\max}} = \frac{1}{Q'-1} \sum_{k=1}^{Q'-1} \binom{Q'-1}{k} \left(\prod_{s=1}^k \int_{-\kappa}^{\infty} Dz_{s'} \right) \left(\prod_{s=k+1}^{Q'-1} \int_{-\infty}^{-\kappa} Dz_{s'} \right) \times \frac{\exp \left\{ \frac{1}{2Q'} \left(\sum_{s'=1}^{Q'-1} z_{s'} \right)^2 \right\}}{\sqrt{Q'}} \left[\sum_{s=1}^k z_{s'}^2 - \frac{1}{k+1} \left(\sum_{s=1}^k z_{s'} \right)^2 \right] . \tag{xxviii}$$

Introducing a Gaussian variable in order to decouple the z_s 's, and using the identities

$$\sum_{k=1}^K \binom{K}{k} \frac{\Lambda^k}{k+1} = \frac{(1+\Lambda)^{K+1} - 1}{\Lambda(K+1)} - 1;$$

$$\sum_{k=1}^K \binom{K}{k} k \Lambda^k = K \Lambda (1+\Lambda)^{K-1} \quad (\text{xxix})$$

one can sum over k and arrives finally at the expression (26) for the critical storage capacity.

References

- [1] HOPFIELD J. J., *Proc. Natl. Acad. Sci. (USA)* **79** (1982) 2554.
- [2] POTTS R. B., *Proc. Camb. Phil. Soc.* **48** (1952) 106 ;
WU F. Y., *Rev. Mod. Phys.* **54** (1982) 235.
- [3] KANTER I., *Phys. Rev. A* **37** (1988) 2739.
- [4] GROSS D. J., KANTER I. and SOMPOLINSKY H., *Phys. Rev. Lett.* **55** (1985) 304.
- [5] COOK J., *J. Phys. A* **22** (1989) 2057.
- [6] BOLLÉ D. and DUPONT P., *Proceedings of the Sitges Conference on Neural networks (Sitges, Spain, June 1990)* ;
BOLLÉ D., DUPONT P. and VAN MOURIK J., *J. Phys. A* **24** (1991) 1065.
- [7] SHIM G. M., KIM D. and CHOI M. Y., 1990 Seoul National Univ. preprint (SNUTP 90/33).
- [8] DUDA R. O. and HART P. E., *Pattern Classification and Scene Analysis* (New York : Wiley, 1973) Chap. 5.
- [9] MINSKY M. and PAPERT S., *Perceptrons* (M.I.T. Press, Cambridge Mass., 1988).
- [10] GALLANT S. I., *IEEE Trans. Neural Net.* **1** (1990) 179.
- [11] LAPEDES 1990 *preprint* ;
BRUNAK S., ENGELBRECHT J. and KNUDSEN S., *Nucl. Acids Res.* **18** (1990) 4787.
- [12] RUMELHART D. E. and ZIPSER D., in *Parallel Distributed Processing* (MIT Press, Cambridge, 1986) pp. 151-193 ;
DAYHOFF J., *Neural Network Architectures* (van Nostrand Reinhold, New York, 1990) Chap. 6.
- [13] PERETTO P., *The Modeling of Neural networks* (Cambridge Univ. Press, 1991) ; BURNOD Y.
- [14] GARDNER E., *J. Phys. A* **21** (1988) 257.
- [15] KRAUTH W., MÉZARD M. and NADAL J.-P., *Complex Systems* **2** (1988) 387 ;
KEPLER T. B. and ABBOTT L. F., *J. Phys. France* **49** (1988) 1657.
- [16] GUTFREUND H., 1990, private communication and in preparation.