ANNALES DE L'I. H. P., SECTION A

PIERRE PICCO

Artificial neural networks. A review from physical and mathematical points of view

Annales de l'I. H. P., section A, tome 64, nº 3 (1996), p. 289-307
<http://www.numdam.org/item?id=AIHPA_1996__64_3_289_0>

Ann. Inst. Henri Poincaré,

Vol. 64, n° 3, 1996, p. 289-307. *Physique théorique*

Artificial neural networks A review from Physical and Mathematical points of view

by

Pierre PICCO*

CPT CNRS, Luminy Marseille
Case 907, 13288 Marseille Cedex 09
E.mail: picco@cptsu2.univ-mes.fr

ABSTRACT. – This survey summarises recent mathematical results on the statistical mechanics of the Hopfield and Kac-Hopfield models.

Key words: Neural Networks, Hopfield model.

RÉSUMÉ. – Nous donnons ici une revue, sans entrer dans les détails de preuves, des résultats mathématiques obtenus récemment pour les modèles de Hopfield et de Kac-Hopfield.

Mots clés : Réseaux de neurones, modèle de Hopfield.

0. INTRODUCTION

Neural Networks are models made for understanding some features of the brain. We do not think that biological neurons work as neural networks. However, an interesting consequence of the study of Neural Networks is that now, Neural Networks are used for many applications such as

* Work partially supported by EU grant CHRX-CT93-0411.

pattern classification and pattern recognition. They are used in robotics and computer vision. The range of possible applications seems to become wider and wider. New journals dedicated to Neural Networks appeared as *Neural Network* in 1988, *Neural Computation* in 1989, *Network* in 1990, *Artificial Neural Networks* in 1991. People coming from different scientific communities are working in this subject and we will restrict ourself in this review to a rather small part of these communities, namely the Mathematical Physics one. From the Mathematical Statistics point of view a review by Cheng and Titterington [CT] appeared in February 1994. From the discrete dynamical point of view the book by Goles and Martinez [GM] is a good reference. From the Theoretical Physics point of view the book by Amit [A] is now a classic. The one by Minsky and Papert [MP] is historically fundamental.

Neural Networks are collections of simple units called neurons and connections that link neurons. In the models we consider, the neuron can be in two states, fired or unfired and is described by a variable taking the values ± 1 that we call σ. For some applications, it could be useful to consider the case where the number of possible states of a given neuron is an integer between 1 and q. This is the case of colored image recognition. Models where σ takes continuous values between -1 and $+1$ are also considered. The connections are usually intricate and in some models any given neuron is connected with all the other ones. However, in the brain there is not as much connection and a way to model this fact is to consider the so-called diluted Neural Networks, where the connections are cut in a random way and stability with respect to dilution of the connections is studied.

The first model of Neural Networks is the McCulloch-Pitts neuron [McCP], an input-output model with a single unit. It is described as follows: The inputs are $(\sigma_j)_{j=1}^N$, the output is x where

$$x = \text{sgn}(w_0 + \sum_{j=1}^{N} w_j \sigma_j) \tag{0.1}$$

with $\sigma_j = \pm 1$ and $x = \pm 1$. The w_j are called connection weights and w_0 is a bias, sometimes called a threshold. Modified versions of this model exist where the inputs and output are not restricted to be ± 1 but can take continuous values between 0 and 1 or between -1 and $+1$, in this case the *sgn* function is replaced by a convenient *sigmoid.*

A multidimensional extension of this model is defined in the following way: a state of the system is charaterized by a vector: σ_i with $i \in \{1, \ldots, N\}$ for some integer N which is usually very large. N represents the number of neurons, which in the brain is of the order of 10^{11}. In this model, the outputs are σ'_i with

$$\sigma'_i = \text{sgn}\left(\theta_i + \sum_{j=1}^{N} W_{ij}\sigma_j\right) \tag{0.2}$$

The matrix W is called the connection matrix and θ the treshold vector.

Multilayered Neural Networks are models where units are arranged in a series of layers and connections are present only between consecutive layers. Multilayered versions are used for practical applications and finding the architecture that have some optimality for a given problem is usually extremely difficult. As an example, the Le Cun *et al.* network for hand written Zip-code recognition involves 1256 units, 63,660 connections and 9760 independents parameters [LeC1] [LeC2]. All these models and possible generalisations belong to the family of input-output models, that is given an input the Network gives an answer, the output. They are very well adapted to recognition problems.

1. THE HOPFIELD MODEL

An important conceptual step in Neural Network theory was done by Hopfield [Ho] in 1982. The objective was to associate with a binary ± 1 vector ξ one of a set of m examplars that have been memorized in the system. The association is not done on the form of an output as before but as the result of an iterative procedure, as in a discrete time dynamical system. Starting with ξ as an initial condition, the iteration will bring the system to the stable points (or limit cycles) of the dynamics that are the associated vectors. The advantage of this formulation is the possibility to model the fact that the human memory does not need a perfect image or sound to associate it with previously memorized image or sound but is able to correct errors. In dynamical system theory, this is just the fact that starting in the basin of attraction of a stable point the trajectory converges to this stable point. Let us describe how Hopfield introduced his model: Take N-dimensional ± 1 vectors $(\xi_i)_{i=1}^{N} \equiv \xi$, and then take M of such objects called *patterns* ξ^μ for $\mu = 1, \ldots, M$, and consider the following iterative procedure:

Given a configuration $\sigma(t)$ at a time t, on the space $\{-1, +1\}^N$, the configuration at the time $t + 1$ is

$$\sigma_i(t+1) = \text{sgn}\left(\sum_{j=1}^{N} W_{ij}\sigma_j(t)\right) \tag{1.1}$$

where the connections are chosen according to the Hebb [He] rule:

$$W_{ij} \equiv \frac{1}{N} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu \tag{1.2}$$

The dynamics could be synchronous (parallel), *i.e.* , the updating defined by (1.1) is done simultaneously for all the sites i, or asynchonous (sequential), *i.e.*, the updating is done on one site and then on another one and so on by using an exhaustive procedure.

There are a lot of choices for the connection matrix W. An important property of Hopfield's one is the fact that W is a symmetric matrix. Hopfield noticed that the function

$$H(\sigma) \equiv -\frac{1}{2} \sum_{(i,j)} W_{ij} \sigma_i \sigma_j \tag{1.3}$$

is a Lyapunov function of the asynchronous dynamics, that is a decreasing function along the orbits, so the asynchronous dynamics will converge to local minima of this function and there are not limit cycles in this case. This is true for general matrix connections as long as the diagonal elements are strictly positive and the matrix is symmetric, as was proved by Goles, Fogelman-Soulie and Pellegrin [GFP]. For the synchronous dynamics, when the connection matrix is symmetric , we can have fixed points or limit cycles of period at most two. This follows from the fact that the functional

$$G(t) = -\sum_{ij} W_{ij} \sigma_i(t+1) \sigma_j(t) \tag{1.4}$$

is a Lyapunov fonction of the synchronous dynamics [GFP][GM]. Notice that the Hebb rule with a zero diagonal can have limit cycles of period at most two. If the connection matrix is not symmetric, limit cycles of period more than two may exist.

In Neural Network theory it is said that the *memory works* if, roughly speaking, starting "near" a pattern the dynamics converge as the time tends to ∞, to a pattern. Here "near" means that the number of discrepancies between the initial point and the pattern is smaller than ϵN, for some $\epsilon > 0$, that is w.r.t. the Hamming distance. We can define the radius of attraction of a stable (for the dynamics) state x as the largest value of ρ such that every vector at a distance not more than ρN eventually reaches the state x. However the usual notion of stability, that is a fixed point of the dynamics, could be too strong and we can define a weaker one: given two non-negative numbers ϵ and δ we call a vector x (δ, ϵ)-stable if starting from an initial point within a distance δN the system ends up within a distance ϵN of x. A stable state is just a $(\delta, 0)$-stable state. Notice also that $(0, \epsilon)$-stable states are just configurations such that starting from it the system ends up in a neighbourhood of radius ϵN. As a dynamical system this model is involved, but its discrete nature has however the advantage to make it suitable for numerical simulations.

The numerical results of Hopfield are the following: He chose for testing : $\xi_i^\mu = \pm 1$ randomly with probability $p = \frac{1}{2}$. If $\frac{M}{N} \equiv \alpha$ and $\alpha < \alpha_c \approx 0.14$, he found that the memory works, in a sense that can be interpretated as the fact the original patterns are $(0, \epsilon)$-stable for some $\epsilon(\alpha) > 0$ small enough. If $\alpha > \alpha_c$ then the memory does not work, in a sense that can be interpreted as $\epsilon \geq 1/2$ for any $\delta > 0$. This is because the typical Hamming distance between two patterns is $1/2$.

What is now known as the Hopfield model is merely this particular randomized version. Let us remark at this point that as a dynamical system this is a model in a random environment and the knowledge of properties of such kinds of dynamical systems is rather poor. An important feature to notice is that N and M are supposed to be very large and we will be interested in the case where they tend to infinity. The main difficulty is now that the state space of our dynamical system is growing. The advantage is that in this limit we can expect to have results that are either true with probablity one, or else with probability approching one with N and M or we could have just convergence in Law.

Knowing that the function H is a Lyapunov function, it is natural to study the corresponding energy landscape, and to check that in some regime of the parameters, the stored patterns are minima of this function. However, since H is a function on a discrete space, a point σ^* is a local minimum if for any single flip at the site k, $T_k\sigma \equiv (T_k\sigma(j))_{j=1}^N$ where $T_k\sigma(j) = \sigma(j)$ if $j \neq k$ and $T_k\sigma(k) = -\sigma(k)$, we have $\Delta_k H(s^*) \equiv H(T_k s^*) - H(\sigma^*) \geq 0$. In this form, the memory capacity, that is the largest α such that every one of the pattern is a (local) minima, was given by McEliece *et al.* [McE] who show that if $\alpha = \frac{1}{4 \ln N}$ then, with a probability tending to one when $N \uparrow \infty$, the patterns are local minima. This was extended to almost sure convergence if $\alpha = \frac{1}{6 \ln N}$ by Martinez [M].

A weaker notion was implicitly introduced by Amit, Gutfreund and Sompolinsky [AGS] and mathematicaly analyzed by Newman [N] in a fundamental work. Let $\mathcal{S}(\sigma, \delta) \equiv \{\sigma' | d(\sigma, \sigma') = [\delta N]\}$, where $d(\sigma, \sigma') \equiv \frac{1}{2}[N - (\sigma, \sigma')]$ is the Hamming distance, be the *sphere* of radius $[\delta N]$ centered at σ. Let us call $h_N(\sigma, \delta)$ the minimum of the energy on the sphere $\mathcal{S}(\sigma, \delta)$. We will say that there exists an energy barrier of height ϵN if for some $\delta \in (0, 1/2)$ we have

$$h_N(\xi^\mu, \delta) \geq H_N(\xi^\mu) + \epsilon N \tag{1.5}$$

The main result of Newman is

THEOREM 1. – *There exists* $\alpha_0 > 0$ *such that if* $M \leq \alpha_0 N$ *then there exists* $\epsilon > 0$ *and* $0 < \delta < 1/2$ *such that*

$$\liminf_{N\uparrow\infty} \inf_{0\leq\mu\leq M} (h_n(\xi^\mu, \delta) - H_N(\xi^\mu) - \epsilon N) \geq 0 \tag{1.6}$$

almost surely.

Newman's estimate for $\alpha_0 \approx 0.056$ is far, also not too much so from the expected 0.14. This result was recently improved by Loukianova [Lo], who get $\alpha_0 \approx 0.071$.

Let us notice that this implies that there exist at least one minimum in a *ball* of radius δ centered at ξ^μ. The study of the energy landscape is far from being complete, such as the expected local minima that are mixture of the patterns. However Newman proved that vectors of the form

$$\eta_i = \operatorname{sgn}\left(\sum_{\mu=1}^{M} \nu_\mu \xi_i^\mu\right) \tag{1.7}$$

with all but a finite number of the ν_μ equal to zero and such that $\sum_{\mu=1}^{M} \epsilon_\mu \nu_\mu \neq 0$ for all possible choices of $\epsilon_\mu = \pm 1$, the ν_μ are some particular dyadic rationals, are surrounded by energy barriers.

It is expected, in fact that the number of local minima is growing exponentialy with N, *see* [McE]. The convergence of the previous iterative procedure was studied by Komlós and Paturi [KPa]. They proved that in the regime $\alpha = 1/(4 \ln N)$ the synchronous dynamics takes $O(\ln \ln N)$ steps for convergence. In the regime where $\alpha \leq \alpha_0$, they exhibit a real valued function $\lambda(.)$ such that, starting within a Hamming distance $\lambda(\alpha)N$ from a pattern ξ^η, then uniformly with respect to the chosen pattern, in a time $O(\ln \frac{1}{\alpha})$, the system will end up within a distance $\epsilon(\alpha)N$ of ξ^η, and this holds with a probability converging to one when $N \uparrow \infty$.

Newman's result was extended to the case of variables ξ taking discrete values in $\{1, 2, \ldots, q\}$ with a uniform distribution, by Ferrari, Martinez and Picco [FMP]. The Hamiltonian is now

$$H(\sigma) = \frac{1}{N} \sum_{\mu=1}^{M} \left(\sum_{i=1}^{N} \delta(\sigma_i, \xi_i^\mu) - \frac{1}{q} \right)^2 \tag{1.8}$$

where $\delta(.,.)$ is the Kronecker symbol. Their results are exactly the same as Theorem 1 except that now all the constants are functions of q.

The Newman result was extended by Bovier and Gayrard [BG1] in the case of the diluted Hopfield model. They consider the Hamiltonian

$$H_N(\sigma) = -\frac{1}{Np} \sum_{\substack{i,j \\ i \neq j}} \epsilon_{ij} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu \sigma_i \sigma_j \tag{1.9}$$

where the ϵ_{ij}'s are a family of N^2 i.i.d. random variable with $\mathbb{P}(\epsilon_{ij} = 1) = 1 - \mathbb{P}(\epsilon_{ij} = 0) = p(N)$. Their result is

THEOREM 1.2. – *Suppose $p \geq c\sqrt{\frac{\ln N}{N}}$. Then there exists $\alpha_c \geq 0$, such that if $M \leq \alpha_c pN$, then there exists $\epsilon > 0$ and $0 < \delta < 1/2$ such that*

$$\liminf_{N\uparrow\infty} inf_{0\leq\mu\leq M}(h_N(\xi^\mu, \delta) - H_N(\xi^\mu) - \epsilon N) \geq 0 \tag{1.10}$$

almost surely.

Notice that here pN is the mean number of connections per site, and is in fact the correct scale of the parameter α.

An very important open problem that remains to prove is the converse of Newman's theorem.

Another interesting problem is to consider correlated patterns and to try to extend at least Newman result's to these cases. The problem of correlated patterns was studied in slightly different contexts. One is related to optimal storage properties. The problem is to find the connections W_{ij} that give the maximal possible α_c. This was studied first by Gardner [Gar] and Gardner and Derrida [GD] from the Theoretical Physics point of view. Monasson [Mo1], [Mo2] and Tarkowski and Lewenstein [TL] studied the problem of optimal storage properties in the case of correlated patterns. The patterns can have spatial correlations *i.e.* $\mathbb{E}(\xi_i^\mu \xi_j^\nu) = \delta^{\mu\nu} C(i,j)$ or semantic correlations *i.e.* $\mathbb{E}(\xi_i^\mu \xi_j^\nu) = C'(\mu,\nu)\delta_{i,j}$. Here $\delta(.,.)$ is the Kronecker symbol and C, C' are correlation matrices. It could be interesting to study how the α_0 of Newman depends on the correlations. The case with exponentially decreasing spatial correlations does not seems too difficult but the case of semantic correlations could be a little more involved. We can expect to get at least an α_0' which is larger than the one in Newman's theorem. Let us notice that there are no rigorous results in the problem of optimal storage properties. Another context where correlated pattern are studied is a work of Griniasty, Tsodyks and Amit [GTA] who model the results of Miyashita's experiment [Mi] in which patterns presented in a certain order in time leads to attractors with space correlations. Let us note that various modifications of the Hopfield hamiltonian exist, including the following:

Higher order Networks:

$$H(\sigma) = -\sum_{\mu=1}^{M}\left(\sum_{i=1}^{N}\xi_i^\mu\sigma_i\right)^d \tag{1.11}$$

For $d \geq 2$ this was studied from Theoretical Phyics point of view by Lee *et al.* [L] and by Newman [N] from the Mathematical Physics point of view. Here the scaling relation $\alpha = M/N$ valid for $d = 2$ has to be replaced by $\alpha = M/N^{d-1}$ and $\alpha_0(d)$ is a decreasing function of *d*. Notice that the extension to $d > 2$ of the result of Kömlos and Paturi is an interesting problem to study. Here also there is no limit cycle for the dynamics if *d* is even, as was shown by Lee *et al.* [L].

Model for short time memory:

This was introduced by Kohonen [Koh] and studied by Mézard, Nadal and Toulouse [MNT]. The Hamiltonian is

$$H(\sigma) = -1/N\sum_{\mu=1}^{M}\Lambda(\mu/N)\left(\sum_{i=1}^{N}\xi_i^\mu\sigma_i\right)^2 \tag{1.12}$$

where $\Lambda(x)$ is $\epsilon e^{-x\epsilon^2/2}$. It will be interesting to have the analogues of the Newman Theorem and the Kómlos and Paturi Theorem in this case and also for the corresponding higher order models.

The pseudo inverse model: If the ξ^μ were orthogonal vectors, the connection matrix would satisfy

$$W\xi^\mu = \frac{N}{M}\xi^\mu \tag{1.13}$$

and the ξ^μ would be fixed points of the dynamics. In the random case the patterns are not orthogonal, and the idea is just to project on the space generated by the $(\xi^\mu)_{\mu=1}^M$. The Hamiltonian is

$$H(\sigma) = 1/N\sum_{i,j}\sigma_i\left(\sum_{\mu=1}^{M}\sum_{\nu=1}^{M}\xi_i^\mu\left(\frac{1}{N}\sum_k\xi_k\xi_k\right)^{-1}_{\mu,\nu}\xi_j^\nu\right)\sigma_j \tag{1.14}$$

here $\left(\frac{1}{N}\sum_k\xi_k\xi_k\right)^{-1}_{\mu,\nu}$ is the μ,ν element of the $M \times M$ matrix which is the inverse of the positive definite $\frac{1}{N}\sum_k\xi_k\xi_k$. This model was introduced by Personnaz, Guyon and Dreyfus [PGD]. and studied by Kanter and Sompolinsky [KS] from the Statistical Physics point of view. Here also the analogues of Newman's Theorem and the Komlós and Paturi Theorem are interesting problems.

2. STATISTICAL MECHANICS OF THE HOPFIELD MODEL

The main problem of the deterministic dynamical system defined by Hopfield is the fact that the system is trapped in local minima of the Lyapounov functional. Since there is an exponential (in N) number of such minima this situation will happen quite frequently. The usual way to be able to avoid this situation is to use stochastic dymanics, that is to introduce a random noise that allows the system to go away from these minima and reach the absolute minima that could have or not, depending on the value of α, some relationship with the stored patterns. There are many possibilities to choose such a dynamics. A possible choice is the heat bath algorithm, that is to introduce a temperature in such a way that the previous dynamics corresponds to the zero temperature dynamics. Let us call

$$H_i(\sigma(t)) \equiv \sum_j W_{ij}\sigma_j(t) \tag{2.1}$$

and

$$\pi_i \equiv \frac{e^{\beta H_i(\sigma(t)}}{e^{(\beta H_i(\sigma(t))} + e^{(-\beta H_i(\sigma(t))}} \tag{2.2}$$

The stochastic dynamics is defined by

$$\begin{aligned} \sigma_i(t+1) &= +1 \quad \text{with} \quad \text{Prob } \pi_i \\ \sigma_i(t+1) &= -1 \quad \text{with} \quad \text{Prob } 1-\pi_i \end{aligned} \tag{2.3}$$

By construction, for all finite N, the Gibbs measure, a measure valued random variable, defined by

$$\mathcal{G} \equiv \frac{\exp -\beta\mathcal{H}(\sigma)}{\mathcal{Z}} \tag{2.4}$$

is reversible with respect to this dynamics. Therefore a preliminary question is to study the Gibbs measures in the limit when N tends to ∞. At this point the Hopfield model can be either this Statistical Mechanics model or the previous deterministic Dynamical System. In 1985-87 Amit, Gutfreund and Sompolimsky [AGS1] and [AGS2] studied the thermodynamics of the Hopfield model from the Theoretical Physics point of view. They found a very rich structure in the regime where $\alpha > 0$, with a spin glass phase [MPV] if β and α are large enough. The Hopfield model is, as α varies, intermediate between an ordered model of Curie-Weiss type (when $M = 1$) and the Sherrington and Kirkpatrick model (when $\alpha \uparrow \infty$). All the techniques introduced by Mezard, Parisi and Virasoro [MPV] were applied to the Hopfield model. Unfortunately for the Mathematical point of view the only rigorous results for the Sherrington and Kirkpatrick model are in the high temperature regime, *cf* Aizenman, Lebowitz and Ruelle [ALR]. The way used by Mathematical Physicists is in fact the opposite of the Theoretical Physicists, namely the study of the Hopfield model could help us to understand the Sherrington and Kirkpatrick model and the spin-glass phase. An important historical point to notice is that the Hopfield model as a model of Statistical Mechanics is not a new one, it was introduced in 1977 by Pastur and Figotin [FP1], [FP2] and [FP3], with the main difference that M was a finite number. It was considered as a model of spin glasses, since a basic property of spin-glasses is the presence of an infinite number of Gibbs States, this model which has M Gibbs states at low temperature did not receive too much interest at that moment. Let us notice that this model has the two other properties that seems to characterise a spin glass, namely frustration and disorder.

The first problem to solve in Statistical Mechanic is the existence of thermodymanics. Namely the existence of the infinite volume limit of the free energy. In mean field models, this is not a trivial problem. The first reason is that you cannot use a sub-additive argument to prove it, as in the *usual* Statistical Mechanics models where the strengh of the interaction does not depend on the volume. In random mean field models the situation is worse as the models are usually not even stable in the sense that the energy is not bounded from below by some constant times the volume, for all realisations of the randomness. However such a lower bound does happen to be true for almost all the realisations. The main difficulty is to prove the existence of the infinite volume limit *i.e.* in our case $N \uparrow \infty$. Since the finite volume free energy is a random quantity, the infinite volume limit could occur in probability (that is with a probability that goes to one with the volume) or almost surely (that is when the probability of the realisations for which the free energy converges is one). In usual models this is not a problem since the existence of the thermodynamics follows from the sub-additive ergodic theorem and the convergence occurs always almost surely. The second problem, given that the existence of the thermodynamics limit is proven, is to show that the limiting free energy is a non random function. At first sight, the two problems seem unrelated. This is not the case. We proved recently [BGP3] a concentration result for the Hopfield model and the Sherrington and Kirkpatrick model that implies

THEOREM 2.1. – *If the mean of the free energy converges then the free energy converges almost surely.*

We stated this result for these two models but any model which depends in a Lipschitzian way on the random parameters should also have this property. This is the case for all the Hamiltonians introduced in the previous section. However, in some cases it could be difficult to prove the analogue of the previous theorem.

Historically the first rigorous result on the Hopfield model was given by Pastur and Figotin in 1977, [FP1] [FP2] [FP3]. They proved the existence of the infinite volume limit for the free energy in the regime of finite M. The value of the limit was given as the solution of a variational problem.

The second result is contained in a very nice paper which did not receive the success it deserves. In 1989 Koch and Piasko [KP] studied the case of M such that $2^M = o(N)$ using a special representation of the model found by Grensing and Kuhn [GK]. They got convergence in probability for the free energy, but almost everywhere results can be obtained by easy modification of their proof. An important fact is that the value of the limiting free energy is the same as the usual Curie-Weiss model [E]. This result can be found in [FP1]. This was not noticed at that moment.

This was extended in 1991 to the case of the Potts-Hopfield model by Gayrard [Gay] who gave almost sure results for the free energy and studied the structure of the Gibbs states. In 1991 Scacciatelli and Tirozzi [ST] extended the result of Aizenman, Lebowitz, and Ruelle [ALR] for the SK model to the Hopfield model and proved the existence of thermodynamics for all the paramagnetic phase.

At the end of 1991 a very nice article by Shcherbina and Tirozzi [ShT] appeared where the limit $\alpha = 0$ was considered. They proved that the limiting (in probability) free energy is the Curie-Weiss free energy. They proved also that the free energy and the overlap parameters are self averaging in the following sense: a thermodynamics quantity $g(\beta, N)$ is *self-averaging* if

$$\lim_{N\uparrow\infty} \mathcal{E}(g(\beta, N) - \mathcal{E}(g(\beta, N))^2 = 0 \tag{2.5}$$

a notion which implies only convergence in Probability. This estimate was extended to some weak exponential moments in [BGP3].

At the beginning of 1992 Koch [Ko] gave another proof of the same result. All these results were extended by Bovier and Gayrard to the case of the diluted Hopfield model [BG2]. In 1992 Pastur, Shcherbina, and Tirozzi studied the cavity method for the Hopfield model. This method was introduced by Mézard, Parisi and Virasoro [MPV] for the study of the Sherrington and Kirkpatrick model as an alternative to the replica symetry breaking scheme. Pastur, Shcherbina and Tirozzi [PST] proved that if the Parisi replica symetry breaking scheme holds then the Edwards and Anderson parameter is not self averaging. An important fact that does not seems to be known by people working in this subject.

In 1993 Bovier, Gayrard and Picco [BGP1] studied the limiting Gibbs states of the Hopfield model. Since we are in a mean field model the notion of Gibbs states is not very well defined. This has been a source of mistake and misunderstanding that seems to persist in the literature related to disordered mean field models. There is however a very complete analysis of the basic mean field model, the Curie-Weiss model, done by Ellis and Newman [E] that give us a framework to start with. The extension to the Hopfield model is not trivial, owing to the fact that the number of *order parameters* depends on the volume in a very sensitive way. This creates new difficulties. To describe the results, we need to introduce some notation.

For $\eta \in \mathbb{N}$, we denote by $\mathcal{G}^{\eta}_{N,\beta,h}$ the random probability measure that assigns to each configuration σ the mass

$$\mathcal{G}^{\eta}_{N,\beta,h}(\sigma) \equiv \frac{1}{Z^{\eta}_{N,\beta,h}} \exp\left(-\beta H_N(\sigma) - \beta h \sum_{i\in\Lambda} \xi^{\eta}_i \sigma_i\right) \tag{2.6}$$

where $Z^{\eta}_{N,\beta,h}$ is the *partition function*. $\mathcal{G}^{\eta}_{N,\beta,h}$ is called a *finite volume Gibbs state with magnetic field*. An important observation is that the value of the measure $\mathcal{G}^{\eta}_{N,\beta,h}(\sigma)$ depends on σ only through the quantities

$$m^{\mu}_N(\sigma) \equiv \frac{1}{N}\sum_{i=1}^{N} \xi^{\mu}_i \sigma_i, \qquad \mu = 1,\ldots,M \tag{2.7}$$

called *overlap parameters*, since the Hamiltonian may be written in the form

$$H_N(\sigma) = -\beta N \sum_{\mu=1}^{M} \left(m^{\mu}_N(\sigma)\right)^2 \tag{2.8}$$

We define the random map

$$\begin{aligned} \mathcal{M}_N : \mathcal{S}_\Lambda &\to \mathbb{R}^M \\ \sigma &\to \mathcal{M}_N(\sigma) \equiv \left(m^1_N(\sigma),\ldots,m^M_N(\sigma)\right) \end{aligned} \tag{2.9}$$

and the measures $\mathcal{Q}^{\eta}_{N,\beta,h}$ on $(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M))$ that are induced by $\mathcal{G}^{\eta}_{N,\beta,h}$ through the map $\mathcal{M}_N$, *i.e.*

$$\mathcal{Q}^{\eta}_{N,\beta,h} \equiv \mathcal{G}^{\eta}_{N,\beta,h} \circ \mathcal{M}_N^{-1} \tag{2.10}$$

These measures are the natural definition of the Gibbs measures of the Hopfield model. In mean field models, the measures on the space of spin configurations are usually trivial: they are just products of Bernoulli measures. As we will see later this is also the case for the Hopfield model.

THEOREM 2.4. – *Assume that M is non-decreasing and satisfies $\lim_{N\uparrow\infty} \frac{M(N)}{N} = 0$. Let $a^{\pm}(\beta)$ denote the largest (resp. smallest) solution of $a = \tanh(\beta a)$. Then, for all $\beta \geq 0$,*

(i) $$w - \lim_{h\to 0\pm} \lim_{N\uparrow\infty} \mathcal{Q}^{\eta}_{N,\beta,h} = \delta_{a^{\pm}(\beta)e^{\eta}}, \quad \mathbb{P} - \textit{almost surely} \tag{2.11}$$

where the limits are understood in the sense of weak convergence of probability distributions; $\delta_{a^{\pm}(\beta)e^{\eta}}$ denotes the Dirac-measure concentrated on $a^{\pm}(\beta)e^{\eta}$ and e^{η} is the η-th unit vector in $\mathbb{R}^{\mathbb{N}}$.

(ii) *Moreover, any limiting induced measure is a convex combination of the measures in* (2.11)

Note that for $\beta \leq 1$, $a^{+}(\beta) = a^{-}(\beta) = 0$ so that in this case there is a unique limiting measure.

In the case where $\lim_{N\uparrow\infty} \frac{M}{N} > 0$, some modification have to be done: For $\delta > 0$, we will write $a(\delta, \beta)$ for the largest solution of the equation

$$\delta a = \tanh(\beta a) \tag{2.12}$$

We denote by $\|\cdot\|$ the ℓ^2-norm on $\mathbb{R}^{\mathbb{N}}$. Given that $\lim_{N\uparrow\infty} \frac{M(N)}{N} = \alpha$, we set, for fixed β,

$$B_{\rho}^{(\nu,s)} \equiv \left\{x \in \mathbb{R}^{\mathbb{N}} \,\middle|\, \|x - sa(1-2\sqrt{\alpha}, \beta)e^{\nu}\| < \rho\right\} \tag{2.13}$$

Using results of [BGP1] and [BGP3] we have

THEOREM 2.5. – *There exists $\alpha_0 > 0$ such that if $\lim \frac{M(N)}{N} = \alpha$, with $\alpha \leq \alpha_0$, then, for all $\beta > 1 + 3\sqrt{\alpha}$, if $\rho^2 > C(a(1-2\sqrt{\alpha}, \beta))^{3/2}\alpha^{1/8}|\ln \alpha|^{1/4}$, for almost all ω,*

$$\lim_{h\downarrow 0} \lim_{N\uparrow\infty} \mathcal{Q}^{\eta}_{N,\beta,h}\left(B_{\rho}^{(\eta,1)}\right) = 1 \tag{2.14}$$

Theorem 2.5 *excludes* in particular that any of the so-called *mixed states* (which are associated to *local* minima of the Hamiltonian; *see* [AGS], [N], [KPa]) give rise to Gibbs states in this regime of parameters.

As we mentioned before the measures induced on the spin configurations are rather trivial. Here we have:

THEOREM 2.6. – *Under the assumptions and with the notation of Theorem* 2.4,

$$w - \lim_{h \to \pm 0} \lim_{N \uparrow \infty} \mathcal{G}^{\eta}_{\beta,N,h} = B^{\eta}_{\pm a(\beta)}, \qquad \mathbb{P} - \textit{almost surely} \tag{2.15}$$

where B^{η}_{a} *denotes the product measure on* $\{-1, 1\}^{\mathbb{N}}$ *with the marginal measure on* σ_i *given by the Bernoulli measure on* $\{-1, 1\}$ *with mean* $\xi^{\eta}_{i} a$.

3. THE KAC-HOPFIELD MODEL

The main problem of mean field models like the Curie-Weiss is the unphysical property that the interaction energy between two sites depends on the volume. This fact has various consequences that are rather pathological. The first one is that the Canonical Free energy (that is the free energy with fixed magnetisation in the case of the Magnetic systems) is not a convex function of the magnetisation. This nonconvexity can be seen from the fact that starting from a spin configuration with all spins $\sigma_i = 1$, the cost in energy to create a droplet of -1 is proportional to the volume of the droplet, instead of being proportional to the length of the boundary of the droplet. This fact is responsible also fo the impossibility to use a sub-addiditive argument to prove the existence of the infinite volume free energy. Moreover the definition of Gibbs States is not given by DLR equations [D], [LR] and the notion of extremal measures is not well defined. The advantage is however the possibility to make explicit computation and to have non trivial critical Thermodynamical properties, *see* [E].

However there is a model invented by Kac [K] in the fifties, that has no such pathologies, which is explicitly solvable, and is related to the Curie Weiss model. The one dimensional ferromagnetic version of this model, is the following: The Hamiltonian is

$$H^{\mathrm{Kac}}(\sigma) \equiv -\frac{\gamma}{2} \sum_{i,j} e^{-\gamma|i-j|} \sigma_i \sigma_j \tag{3.1}$$

Therefore, for all fixed γ, there is no phase transition. However if we perform the limit $\lim \gamma \downarrow 0$ after the thermodyamic limit, it was proved that the free energy is the same as the Curie-Weiss free energy, the canonical free energy is the convex hull of the canonical free energy of the Curie-Weiss model. This is know as the Lebowitz and Penrose theorem [LP] and there is spontaneous magnetisation at low temperature (which is $\beta = 1$ here). The ferromagnetic model is now relatively well understood, although the complete study in one dimension is rather recent and was done by Cassandro, Orlandi and Presutti [COP].

To avoid the previouly mentioned pathologies of the mean field models, it is important to consider the Kac version of the Hopfield model. Let us define it precisly.

We denote by Λ the set of integers $\Lambda \equiv \{-N, -N+1, \ldots, N\}$. We define a random Hamiltonian as follows. Let $\xi \equiv \{\xi_i^\mu\}_{i\in\mathbb{Z},\mu\in\mathbb{N}}$ be a two-parameter family of independent, identically distributed random variables such that $\mathbb{P}(\xi_i^\mu = 1) = \mathbb{P}(\xi_i^\mu = -1) = \frac{1}{2}$. The Hamiltonian with free boundary conditions on Λ is then given by

$$H_\Lambda(\sigma) = -\frac{1}{2} \sum_{(i,j)\in\Lambda\times\Lambda} \sum_{\mu=1}^{M(\gamma)} \xi_i^\mu \xi_j^\mu J_\gamma(i-j)\sigma_i\sigma_j \tag{3.2}$$

where $J_\gamma(i-j) \equiv \frac{\gamma}{2} J(\gamma|i-j|)$, and

$$J(x) = \begin{cases} 1, & \text{if } |x| \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.3}$$

It is slightly different from the model with exponentially decaying interactions, however the two versions are equivalent. We are interested in the case where $M(\gamma) \uparrow \infty$, as $\gamma \downarrow 0$. We will set $\alpha(\gamma) \equiv \gamma M(\gamma)$. The first result concern the free energy, $F_{\Lambda,\beta,\gamma} \equiv -\beta\Lambda^{-1} \ln Z_{\Lambda,\beta,\gamma}$:

PROPOSITION 3.1. – *Assume that* $\lim_{\gamma\downarrow 0} \gamma M(\gamma) = 0$. *Then, for allmost all* ω,

$$\lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma} = \lim_{N\uparrow\infty} F_{N,\beta}^{Hopf} = F_\beta^{CW} \tag{3.4}$$

with $F_\beta^{CW} \equiv \inf_{x\in\mathbb{R}} \left(\frac{x^2}{2} - \beta^{-1} \ln\cosh\beta x\right)$, *the free energy of the Curie-Weiss model.*

The finite volume Gibbs measure for our model is defined by assigning to each $\sigma \in \mathcal{S}_\Lambda$ the mass

$$\mathcal{G}_{\Lambda,\beta,\gamma}(\sigma) \equiv \frac{1}{Z_{\Lambda,\beta,\gamma}} e^{-\beta H_{\Lambda,\gamma}(\sigma)} \tag{3.5}$$

where $Z_{\Lambda,\beta,\gamma}$ is the *partition function.* For any subset $\Delta \subset \mathbb{Z}$, we define the M-dimensional vector of 'overlaps' $m_\Delta(\sigma)$ whose components are

$$m_\Delta^\mu(\sigma) = \frac{1}{|\Delta|} \sum_{i\in\Delta} \xi_i^\mu \sigma_i, \qquad \mu = 1, \ldots, M \tag{3.6}$$

The main object to study are the distributions of $m_\Lambda(\sigma)$ under the Gibbs measure, *i.e.*

$$\mathcal{Q}_{\Lambda,\beta,\gamma}(m) = \mathcal{G}_{\Lambda,\beta,\gamma}(\{m_\Lambda(\sigma) = m\}) \tag{3.7}$$

$\mathcal{Q}_{\Lambda,\beta,\gamma}$ defines a random probability measure on $(\mathbb{R}^{M(\gamma)}, \mathcal{B}(\mathbb{R}^{M(\gamma)}))$. For fixed $\gamma > 0$, this sequence of probability measures satisfies a large deviation principle in the sense that for instance the limit

$$\lim_{\epsilon\downarrow 0}\lim_{N\uparrow\infty} \frac{1}{2N+1} \ln\left(\mathcal{Q}_{\Lambda,\beta,\gamma}\left[\|m_\Lambda - \tilde{m}\|_2^2 \leq \epsilon\right]\right) \equiv -\beta F_{\beta,\gamma}(\tilde{m}) \tag{3.8}$$

exists almost surely by the subadditive ergodic theorem. Moreover, $F_{\beta,\gamma}(\tilde{m})$ is a convex function of its argument. We are interested in the limiting behaviour of $F_{\beta,\gamma}$ as $\gamma \downarrow 0$. Since the domain of this function depends on γ via $M(\gamma)$, it is natural to consider restrictions to finite dimensional cylinders. Thus, let $I \subset \mathbb{N}$ be a finite set and denote by $\Pi_I : \mathbb{R}^M \to \mathbb{R}^I$, for any M such that $I \subset \{1, \ldots, M\}$, the orthogonal projector on the components m^μ, with $\mu \in I$, of a vector $m \in \mathbb{R}^M$. We set, for $\tilde{m} \in \mathbb{R}^I$,

$$F^I_{\beta,\gamma}(\tilde{m}) \equiv -\beta^{-1} \lim_{\epsilon\downarrow 0}\lim_{N\uparrow\infty} \frac{1}{2N+1} \ln\left(\mathcal{Q}_{\Lambda,\beta,\gamma}\left[\|\Pi_I m_\Lambda - \tilde{m}\|_2^2 \leq \epsilon\right]\right) \tag{3.9}$$

The Lebowitz-Penrose theory relates the limit of these quantities to the analogous ones in the corresponding mean-field model, *i.e.* the Hopfield model.

We denote by $\mathcal{G}^{Hopf}_{N,\beta}$ the corresponding Gibbs measure, and by $\mathcal{Q}^{Hopf}_{N,\beta}$ the induced distribution of the overlap parameters. We also write

$$F^{Hopf,I}_\beta(\tilde{m}) \equiv -\beta^{-1} \lim_{\epsilon\downarrow 0}\lim_{N\uparrow\infty} \frac{1}{N} \ln(\mathcal{Q}^{Hopf}_{N,\beta}\left[\|\Pi_I m_N - \tilde{m}\|_2^2 \leq \epsilon\right]) \tag{3.10}$$

provided this limit exists. Notice that the case of a finite number of patterns was studied by Comets [Co].

Let us define the convex functions $C^{Hopf,I}_\beta$, which, if $F^{Hopf,I}_\beta$ exists, are the convex hulls of these functions. We set

$$C^{Hopf,I}_\beta(\tilde{m}) \equiv \beta^{-1} \lim_{\epsilon\downarrow 0}\lim_{N\uparrow\infty} \frac{1}{N}\text{Conv}(-\ln(\mathcal{Q}^{Hopf}_{N,\beta}\left[\|\Pi_I m_\Lambda - \tilde{m}\|_2^2 \leq \epsilon\right])) \tag{3.11}$$

Our first result concerns the existence of the functions $C^{Hopf,I}_\beta$.

THEOREM 3.1. – *Suppose that $p(N)$ is such that* $\lim_{N\uparrow\infty} p(N) = +\infty$ *and* $\lim_{N\uparrow\infty} \frac{p(N)\ln N}{N} = 0$. *Then,*

(i) *For almost all realisation of the patterns,* $C_\beta^{Hopf,I}(\tilde{m})$ *defined through* (3.11) *exists for any finite set* $I \subset \mathbb{N}$ *and is independent of* ϵ, *and the function* $p(N)$.

(ii) *If, moreover,* $\lim_{N\uparrow\infty} \frac{2^{p(N)}}{N} = 0$, *then, almost surely,* $F_\beta^{Hopf,I}(\tilde{m})$ *defined through* (3.11) *exists and is independent of* $p(N)$.

For the Kac-Hopfield model, we obtain the analogue of the Lebowitz-Penrose Theorem:

THEOREM 3.2. – *Assume that* $M(\gamma)$ *satisfies* $\lim_{\gamma\downarrow 0} M(\gamma) = +\infty$ *and* $\lim_{\gamma\downarrow 0} \gamma|M(\gamma) = 0$. *Then, for any* β, *and any finite subset* I, *for almost all realisations of the patterns,*

$$- \beta^{-1} \lim_{\epsilon\downarrow 0} \lim_{\gamma\downarrow 0} \lim_{N\uparrow\infty} \frac{1}{2N+1} \ln \left(\mathcal{Q}_{\Lambda,\beta,\gamma} \left[\|\Pi_I m_\Lambda - \tilde{m}\|_2^2 \leq \epsilon \right] \right)$$
$$= C_\beta^{Hopf,I}(\tilde{m}) \qquad (3.12)$$

ACKNOWLEDGEMENTS

I would like to thank Flora Koukiou for giving me the opportunity to write this review and the referee for his suggestions.

REFERENCES

[ALR] M. AIZENMAN, J. LEBOWITZ and D. RUELLE, Some rigorous results on the Sherrington- Kirkpatrick model, *Com. Math. Phys.* Vol. **112**, 1987, pp. 3-20.

[A] D. J. AMIT, Modeling Brain Function, Cambridge Univ. Press. Cambridge, 1988.

[AGS1] D. J. AMIT, H. GUTFREUND and H. SOMPOLINSKY, Spin-glass models of neural networks, *Phys. Rev.*, Vol. **A32**, 1985, pp. 1007-1018.

[AGS] D. J. AMIT, H. GUTFREUND and H. SOMPOLINSKY, Statistical Mechanics of Neural Networks near Saturation, *Ann. of Phys.*, Vol. **173**, 1987, pp. 30-67.

[BG1] A. BOVIER and V. GAYRARD, Rigorous bounds on the storage capacity of the dilute Hopfield model, *J. Stat. Phys.*, Vol. **69**, 1992, pp. 597-627.

[BG2] A. BOVIER and V. GAYRARD, Rigorous results on the thermodynamics of the dilute Hopfield model, *J. Stat. Phys.*, Vol. **69**, 1993, pp. 597-627.

[BGP1] A. BOVIER, V. GAYRARD and P. PICCO, Gibbs states of the Hopfield model in the regime of perfect memory, to appear in *Prob. Theor. Rel. Fields*, 1995.

[BGP2] A. BOVIER, V. GAYRARD and P. PICCO, Large deviation principles for the Hopfield and the Kac-Hopfield model, to aappear in *Probability Theory and related fields*, 1995.

[BGP3] A. BOVIER, V. GAYRARD and P. PICCO, Gibbs states for the Hopfield model with extensively many patterns, to appear in *J.S.P.*, 1995.

[CT] B. CHENG and D. M. TITTERINGTON, Neural Networks: A Review from a Statistical Perspective, *Statistical Science*, Vol. **9**, 1994, pp. 2-54.

[Co] F. COMETS, Large deviation estimates for a conditional probability distribution. Applications to random Gibbs measures, *Prob. Theor. Rel. Fields*, Vol. **80**, 1989, pp. 407-432.

[COP] M. CASSANDRO, E. ORLANDI and E. PRESUTTI, Interfaces and typical Gibbs configurations for one-dimensional Kac potentials, *Prob. Theor. Rel. Fields*, Vol. **96**, 1993, pp. 57-96.

[E] R. S. ELLIS, Entropy, large deviations, and statistical mechanics, Springer-Verlag, Berlin, 1985.

[FMP] P. FERRARI, S. MARTINEZ and P. PICCO, A lower bound for the memory capacity in the Hopfield-Potts model, *J. Stat. Phys.*, Vol. **66**, 1991, pp. 1643-1651.

[FP1] L. A. PASTUR and A. L. FIGOTIN, Exactly soluble model of a spin glass, *Sov. J. Low Temp. Phys.*, Vol. **3(6)**, 1977, pp. 378-383.

[FP2] L. A. PASTUR and A. L. FIGOTIN, On the theory of disordered spin systems, *Theor. Math. Phys.*, Vol. **35**, 1978, pp. 403-414.

[FP3] L. A. PASTUR and A. L. FIGOTIN, Infinite range limit for a class of disordered spin systems, *Theor. Math. Phys.*, Vol. **51**, 1982, pp. 564-569.

[Gay] V. GAYRARD, The thermodynamics limit of the q-state Potts-Hopfield model with infinitely many patterns, *J. Stat. Phys.*, Vol. **68**, 1992, pp. 977-1011.

[Gar] E. GARDNER, The space of interactions in neural network models, *J. Phys. A: Math. Gen.*, Vol. **21**, 1988, pp. 257-270.

[GD] E. GARDNER and B. DERRIDA, Optimal storage properties of neural network models, *J. Phys. A: Math. Gen.*, Vol. **21**, 1988, pp. 271-284.

[GFP] E. GOLES, F. FOGELMAN-SOULIE and D. PELLEGRIN, Decreasing energy functons as a tool for studying threshold networks, *Discrete App. Math.*, Vol. **12**, 1985, pp. 261-277.

[GM] E. GOLES and S. MARTINEZ, Neural and automata networks. Dynamical behavior and applications, Kluwer Ac. Pub. Dordrecht, Boston, London, 1990.

[GK] D. GRENSING and K. KÜHN, On classical spin-glass models, *J. Phys. Paris*, Vol. **48**, 1987, pp. 713-721.

[GTA] M. GRINIASTY, M. TSODYKS and D. J. AMIT, *Neural Comp.*, Vol. **5**, 1993, pp. 1-17.

[He] D. O. HEBB, The organisation of behavior, J. Willey & Son, New York, 1949.

[Ho] J. J. HOPFIELD, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA*, Vol. **79**, 1982, pp. 2554-2558.

[KS] I. KANTER and H. SOMPOLINSKY, Associative recall of memory without errors, *Phys. Rev. A*, Vol. **35**, 1987, pp. 380-392.

[Ko] H. KOCH, A free energy bound for the Hopfield model, *J. Phys. A: Math Gen.*, Vol. **26**, 1993, pp. L353-L355.

[Koh] T. KOHONEN, Self-organisation and associative memory, Springer Verlag, Berlin, 1984.

[KP] H. KOCH and J. PIASKO, Some rigorous results on the Hopfield neural network model, *J. Stat. Phys.*, Vol. **55**, 1989, pp. 903-928.

[KUH] M. KAC, G. UHLENBEC and P. C. HEMMER, On the van der Waals theory of vapour-liquid equilibrium. I. Discussion of a one-dimensional model, *J. Math. Phys.*, Vol. **4**, 1963, pp. 216-228; II. Discussion of the distribution functions, *J. Math. Phys.*, Vol. **4**, 1963, pp. 229-247; III. Discussion of the critical region, *J. Math. Phys.*, Vol. **5**, 1964, pp. 60-74.

[KPa] J. KOMLOS and R. PATURI, Convergence results in an associative memory model, *Neural Network.*, Vol. **1**, 1988, pp. 239-250.

[LP] J. LEBOWITZ and O. PENROSE, Rigorous treatment of the Van der Waals Maxwell theory of the liquid-vapour transition, *J. Math. Phys.*, Vol. **7**, 1966, pp. 98-113.

[LeC1] Y. LE CUN, B. BOSER, J. S. DENKER, D. HENDERSON, R. E. HOWARD, W. HUBBARD and L. D. JACKEL, Backpropagation applied to handwritten Zip code recognition, *Neural Computation*, Vol. **1**, 1989, pp. 541-551.

[LeC2] Y. LE CUN, B. BOSER, J. S. DENKER, D. HENDERSON, R. E. HOWARD, W. HUBBARD and L. D. JACKEL, Handwritten digit recognition with a backpropagation network, Advances in *Neural Information Processing Systems II* (D. S. Touretzky, eds.), Morgan Kaufmann, San Mateo, CA, 1990, pp. 396-404.

[L] Y. C. LEE, G. DOOLEN, H. H. CHEN, G. Z. SUN, T MAXWELL, H. Y. LEE and C. L. GILES, Machine learning using a higher order correlaton network, *Physica D*, Vol. **22**, 1987, pp. 276-867.

[Lo] D. LOUKIANOVA, Capacité de mémoire dans le modèle de Hopfield, *C. R. Acad. Sc. Paris*, Serie 1, Vol. **31**, 1994, pp. 157-160.

[McE] R. J. MCELIECE, E. C. POSNER, E. R. RODEMICH and S. S. VENKATESH, The capacity of the Hopfield associative memory, *IEEE Trans. Inform. Theory*, Vol. **33**, 1987, pp. 461-482.

[McCP] W. S. MCCULLOCH and W. A. PITTS, A logical calculus of the ideas immanent in nervous activity, *Bul. Math. Biophysics*, Vol. **5**, 1943, pp. 115-133.

[M] S. MARTINEZ, Temuco summer school. To appear, 1992.

[MP] M. MINSKY and S. PAPERT, Perceptrons, MIT Press, Cambridge MA, 1969.

[MNT] M. MÉZARD, J. P. NADAL and G. TOULOUSE, Solvable models of working memories, *J. Phys. Paris*, Vol. **47**, 1986, pp. 1457-1462.

[MPV] M. MÉZARD, G. PARISI and M. VIRASORO, Spin Glass Theory and Beyond, World Scientific, Singapore, 1987.

[Mi] Y. MIYASHITA, Neuronal correlate of visual associative long-term memory in primate temporal cortex, *Nature*, Vol. **335**, 1988, p. 817.

[Mo1] R. MONASSON, Properties of neural networks storing spatially correlated patterns, *J. Phys. A: Math. Gen.*, Vol. **25**, 1992, pp. 3701-3720.

[Mo2] R. MONASSON, Storage of spatially correlated patterns in auto-associative memories, *J. Phys. Paris*, Vol. **3**, 1993, pp. 1141-1152.

[N] C. M. NEWMAN, Memory capacity in neural network models: rigorous lower bounds, *Neural Networks*, Vol. **1**, 1988, pp. 223-238.

[PST] L. PASTUR, M. SHCHERBINA and B. TIROZZI, The replica symmetric solution without the replica trick for the Hopfield model, *J. Stat. Phys.*, Vol. **74**, 1994, pp. 1161-1183.

[PGD] L. PERSONNAZ, I. GUYON and G. DREYFUS, Information storage and retrivial in Spin-glass like neural networks, *J. Phys. Lett.*, Vol. **49**, 1985, pp. 359-365.

[ST] E. SCACCIATELLI and B. TIROZZI, Fluctuation of the free energy in the Hopfield model, *J. Stat. Phys.*, Vol. **67**, 1992, pp. 981-1008.

[ShT] M. SHCHERBINA and B. TIROZZI, The free energy for a class of Hopfield models, *J. Stat. Phys.*, Vol. **72**, 1992, pp. 113-125.

[TL] W. TARKOWSKI and M. LEWENSTEIN, Storage of sets of correlated data in neural network memories, *J. Phys A: Math. Gen.*, Vol. **26**, 1993, pp. 2453-2469.

(Manuscript received February 2, 1995.)