



**HAL**  
open science

## Dynamics of a multi-layered perceptron model : a rigorous result

A.E. Patrick, V.A. Zagrebnov

► **To cite this version:**

A.E. Patrick, V.A. Zagrebnov. Dynamics of a multi-layered perceptron model : a rigorous result. Journal de Physique, 1990, 51 (11), pp.1129-1137. 10.1051/jphys:0199000510110112900 . jpa-00212435

**HAL Id: jpa-00212435**

**<https://hal.science/jpa-00212435>**

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 — 05.90 — 87.30

## Dynamics of a multi-layered perceptron model : a rigorous result

A. E. Patrick (\*) and V. A. Zagrebnov (\*)

Centre de Physique Théorique, Luminy, 13288 Marseille Cedex 9, France

(Reçu le 24 novembre 1989, accepté le 26 janvier 1990)

**Résumé.** — Nous dérivons exactement et rigoureusement les systèmes d'équations dynamiques pour un perceptron multi-couches proposé par Domany, Meir et Kinzel (DMK-model). Ces équations décrivent simultanément l'évolution des fonctions de recouvrement essentielles et résiduelles.

**Abstract.** — We derive exactly and rigorously the system of dynamical equations for a multi-layered perceptron proposed by Domany, Meir and Kinzel (DMK-model). They describes both the main and the residual overlaps evolution.

1. The DMK-model [1] is a layered feed-forward neural network without couplings within layers. In this sense it can be considered as a generalization of the idea of the perceptron (*multi-layered perceptron*), see e.g. [2]. The dynamics of the original DMK-model, and of its various extensions, have been solved in [3-5] by a (hard controlled) saddle point integration method, proposed in [6] for the studying of the parallel dynamics in fully connected neural networks.

The aim of the present paper is to derive (*exactly and rigorously*) the *system* of dynamics equations for DMK-model exploiting some main ideas of the probability theory and the theory of dynamical systems with noise. As a result we get not only rigorous but even shorter (then in [3-5]) way to exact system of dynamics equations.

Below we follow the line of reasoning formulated in [7] for the Hopfield model. This program is focussed on the accurate calculations of stochastic properties of the internal noise in the limit of large neural networks. This approach have been improved in the subsequent paper [8]. There we explain the importance to consider together with the *main overlap* the corresponding stochastic equations for evolution of the internal noise (*residual overlaps*). This allows to rederive (exactly, rigorously and in shorter way) the *Gardner-Derrida-Mottishaw equation* [6] for the main overlap (in the parallel dynamics) for the second step

---

(\*) *Permanent address* : Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna 141980, U.S.S.R.

$t = 2$  and to derive the corresponding equation for  $t = 3$ . This approach gives a more clear view on a possible behaviour of the main overlap for arbitrary  $t$  and allows to present a pure dynamical few-lines derivation of the *Amit-Gutfreund-Sompolinsky formula* [9] for the main overlap at  $t = \infty$ .

Here we exploit this approach to DMK-model where it gives, besides the well-known exact result for dynamics of the main overlap [3-5], the exact equation for the evolution of the noisy terms (residual overlaps).

2. We recall that DMK-model can be formulated as follows. Each of the layer  $t$  ( $= 1, 2, \dots, T$ ) contains  $N$  binary variables (spin configurations)  $\mathcal{S}(t) = \{S_i(t) = \pm 1\}_{i=1}^N$  and the set of  $M$  fixed realizations  $\{\xi_i^p(t)\}_{p=1}^M$  of the sequence of independent identically distributed random variables (i.i.d.r.v.)  $\{\xi_i = \pm 1\}_{i=1}^N$ ,  $\Pr \{\xi_i = \pm 1\} = \frac{1}{2}$ , of the length  $N$ . Therefore, each  $\xi_i^p(t) \in \Omega_N(t) = \{-1, 1\}^N$  — the probability space with the natural  $\sigma$ -algebra and probability  $\mathbb{P}_N\{\xi\}$  defined as the product-measure by  $\Pr \{\xi_i = \pm 1\} = \frac{1}{2}$ .

Then for the zero temperature ( $\theta = 0$ ) evolution of the DMK-model corresponds to updating of the spin configuration at the  $(t + 1)$ -th layer according to the effective fields

$$h_i(t + 1) = \sum_{j=1}^N J_{ij}(t + 1) S_j(t), \quad i = 1, 2, \dots, N \quad (1)$$

by the law

$$S_i(t + 1) = \text{sign}(h_i(t + 1)). \quad (2)$$

Here fields (1) are defined by the spin configuration on the previous  $t$ -th layer and by the (directed) couplings between this two layers, which have Hebbian form [1]

$$J_{ij}(t + 1) = \frac{1}{N} \sum_{p=1}^M \xi_i^p(t + 1) \xi_j^p(t). \quad (3)$$

For non-zero temperature ( $\theta \neq 0$ ) evolution of the DMK-model is defined by stochastic equation

$$S_i(t + 1) = \text{sign}[h_i(t + 1) + \phi_i(t + 1)]. \quad (4)$$

Here i.i.d.r.v.  $\{\phi_i(t)\}_{i=1}^N$  reproduce a heat-bath temperature noise if

$$\Pr \{\phi_i(t) \leq x\} = \left[ 1 + \text{th} \left( \frac{x}{\theta} \right) \right] / 2. \quad (5)$$

So, in contrast to the quenched random vectors  $\{\xi_i^p(t)\}_{p=1}^M, T_{t=1}$  (key patterns at each layer) the random variables  $\{\phi(t)\}_{t=1}^T$  are unquenched. We have to average over this thermal noise variables to consider dynamics in the presence of the heat-bath.

3. The problem of the neural networks dynamics is usually specified as follows. Let initial configuration  $\mathcal{S}(t = 1)$  has a macroscopic overlap with one of the key patterns, e.g., with  $p = q$ . Then we are interesting in the evolution of this overlap ( $t \rightarrow \infty$ ) after thermodynamic

limit  $N \rightarrow \infty$ . For the DMK-model this means that for a fixed (quenched) random system of uncorrelated key patterns  $\{\xi_i^p(t)\}_{p=1, T, i=1}^M$  one chooses initial configuration in such a way that

$$m_N^q(t = 1) = \frac{1}{N} \sum_{i=1}^N S_i(t = 1) \xi_i^q(t = 1) \tag{6}$$

has non-zero limit  $m^q(t = 1)$  for  $N \rightarrow \infty$  (*main overlap*) and  $m_N^p(t = 1) = \left(\frac{1}{\sqrt{N}}\right)$  for  $p(\neq q) = 1, 2, \dots, M$  even if  $M = \alpha N(\alpha - \text{lim})$ .

From equations (1), (3) and (6) one gets

$$m_N^q(t + 1) = \frac{1}{N} \sum_{i=1}^N \text{sign} \left[ m_N^q(t) + \xi_i^q(t + 1) \frac{1}{\sqrt{N}} \sum_{\substack{p=1 \\ (p \neq q)}}^M \xi_i^p(t + 1) r_N^p(t) \right], \tag{7}$$

where we introduce *residual overlaps*

$$r_N^p(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \xi_i^p(t) S_i(t), \quad p(\neq q) = 1, 2, \dots, M. \tag{8}$$

In contrast to  $m_N^p(t)$ , the residual overlaps (8) do not converge to any limits for the quenched system of key patterns. But they have limits as sequences of random variables if we consider patterns as random vectors in the space  $(\Omega, \sigma, \mathbb{P})$ .

Suppose that at the moment  $t \quad \alpha - \text{lim} \ m_N^q(t) = m^q(t)$  almost surely (a.s.) with respect to the probability  $\mathbb{P}$ . Then by the individual ergodic (Birkhoff-Khinchin) theorem ([10], V, Sect. 3) we get that: (1)  $\alpha - \text{lim}$  for the arithmetic mean in the right-hand side of (7) exists for almost all (by the measure  $\mathbb{P}$ ) quenched systems of key patterns; (2) this limit is a.s. independent of the choice of the fixed system of patterns (self-averaging) and (3) it is equal to expectation

$$m^q(t + 1) = E_\eta \text{sign} [m^q(t) + \eta(t + 1)]. \tag{9}$$

Here noisy term  $\eta(t + 1) = \eta_i(t + 1)$  and

$$\eta_i(t + 1) \stackrel{d}{=} \xi_i^q(t + 1) \alpha - \text{lim} \frac{1}{\sqrt{N}} \sum_{\substack{p=1 \\ (p \neq q)}}^M \xi_i^p(t + 1) r_N^p(t) \tag{10}$$

is the stationary ( $i = 1, 2, \dots$ ) ergodic sequence of random variables. The convergence in (10) occurs in the sense of distribution.

Therefore, in the  $\alpha - \text{lim}$  by ergodic theorem one reduces the initial problem for quenched key patterns to calculation the expectation with respect to distribution of the limit random variable (10). Hence, to construct dynamics for the main overlap we have to consider together with equation (7) the evolution of the residual overlaps (8). Using equations (1), (3) and (8) one gets

$$r_N^p(t + 1) = \frac{1}{\sqrt{N}} \sum_{j=1}^N \text{sign} \left\{ \frac{1}{\sqrt{N}} r_N^p(t) + \xi_j^p(t + 1) [\xi_j^q(t + 1) m_N^q(t) + \frac{1}{\sqrt{N}} \sum_{\substack{\ell=1 \\ (\ell \neq p, q)}}^M \xi_j^\ell(t + 1) r_N^\ell(t)] \right\} \quad p(\neq q) = 1, 2, \dots, M. \tag{11}$$

This is a stochastic equation (recurrence relation) for random variables  $\{r_N^p(t+1)\}_{p=1}^M$  generated by the different realizations of the key patterns  $\{\xi^p(t+1)\}_{p=1}^M$  and the noise  $\{r_N^p(t)\}_{p=1}^M$ . In the  $\alpha$  - lim relations (7) and (11) form a system of coupled equations which describes dynamics of the DMK-model.

4. According to above observation at *first* we have to derive in  $\alpha$  - lim the stochastic recursion relations for evolution of the residual overlaps  $\{r^p(t)\}_{p \neq q}$  (8) and *then* use the distribution of the noisy term (10) for calculation the recursion relation for the main overlap  $m^q(t)$  (9). We do this by induction in  $t$ .

For  $t = 1$  the initial configuration  $\xi(t = 1)$  is chosen in such a way that

$$\alpha - \lim m_N^p(t) = \begin{cases} m^q(t = 1) > 0, & p = q \\ O\left(\frac{1}{\sqrt{N}}\right), & p \neq q. \end{cases} \quad (12)$$

For unquenched key patterns this means that  $\xi(t = 1)$  and  $\xi^q(t = 1)$  are correlated:  $\Pr(S_i(t = 1) = \xi_i^q) = \frac{1}{2}(1 + m^q(t = 1))$ , but for  $p \neq q$   $\{\xi_i^p(t = 1) S_i(t = 1)\}_{i=1}^\infty$  is a sequence of i.i.d.r.v. with  $\Pr\{\xi_i^p(t = 1) S_i(t = 1) = \pm 1\} = \frac{1}{2}$ . Consequently, by the central limit theorem (CLT) residual overlaps (as random variables) converge in the sense of distribution to the Gaussian random variable:

$$\alpha - \lim r_N^p(t = 1) \stackrel{d}{=} N(0, 1), \quad (13)$$

with mean 0 and variance 1. Moreover, for the different  $p (\neq q)$  they are *independent*, i.e.,  $\{r^p(t = 1)\}_{p=1}^\infty$  is the sequence of i.i.d. Gaussian r.v. Note that for  $t = 2$  random variables  $\{\xi_i^p(t)\}$  are independent of  $\{r_N^p(t = 1)\}$  (their arguments belong to different layers!). Then by the symmetry of  $\xi_i^p(t = 2) = \pm 1$  and of the distribution for  $N(0, 1)$  we get that  $\{\xi_i^p(t = 2) r^p(t = 1)\}_{p=1}^\infty$  is again a sequence of i.i.d. Gaussian r.v. This means that by the

CLT random variables  $v_{i,N}^q(t = 2) = \frac{1}{\sqrt{N}} \sum_{\substack{p=1 \\ (p \neq q)}}^M \xi_i^p(t = 2) r_N^p(t = 1)$  (deviation of  $r_N^p(t = 1)$ )

from  $r^p(t = 1)$  can be controlled by the Berry-Esseen theorem ([10], III, Sect. 6), i.e., we use CLT in the scheme of series) converge in distribution:

$$\alpha - \lim v_{i,N}^q(t = 2) \stackrel{d}{=} \sqrt{\alpha} N(0, 1), \quad (14)$$

to the Gaussian random variable with the variance  $\alpha$ . Finally, taking into account equations (13), (14), symmetry of  $N(0, 1)$ , independence of random variables  $\{v_{i,N}^q(t = 2)\}$ ,  $\{\xi_i^q(t = 2)\}$  and again using the Berry-Esseen theorem, we get

$$\begin{aligned} \Pr\{m_N^q(t = 1) + \xi_i^q(t = 2) v_{i,N}^q(t = 2) \leq x\} &= \\ &= \frac{1}{\sqrt{2\pi\alpha}} \int_{-\infty}^{x - m^q(t = 1)} dz \exp(-z^2/2\alpha) + O\left(\frac{1}{\sqrt{N}}\right). \end{aligned} \quad (15)$$

Using distribution (15) and ergodic theorem one gets (see Eqs. (7) and (9))

$$m^q(t = 2) = \operatorname{erf} \left( \frac{m^q(t = 1)}{\sqrt{\alpha}} \right),$$

$$\operatorname{erf} z = \sqrt{\frac{2}{\pi}} \int_0^z dt \exp(-t^2/2). \tag{16}$$

5. To go to the next step  $t = 3$  for the  $m^q(t)$  we have first to derive the probability distribution for  $r^p(t = 2)$ , see (9), (10). To this end we calculate the  $\alpha$  – lim for stochastic recurrence relation (11), i.e., stochastic characteristics of random variables  $\{r^p(t = 2)\}_{p=1}^\infty$  conditioned by arbitrarily fixed realization  $\{r^p(t = 1)\}_{p=1}^\infty$ .

Now application of the CLT to the sequence of random variables  $v_{i,N}(t = 2)$  (see (11) and (14)) is not very straightforward. For fixed Gaussian realizations  $\{r_N^p(t = 1)\}_p$  (see (13)) one gets that  $\{\xi_j^\ell(t = 2) r_N^\ell(t = 1)\}_{\ell \neq q}$  is the sequence of independent *but not identically* distributed random variables : in general  $r_N^p(t = 1) \neq r_N^{p'}(t = 1)$  for  $p \neq p'$ . Therefore, one has either to check for this sequence the Lindeberg conditions or to use Lyapunov’s method of characteristic functions for the direct proof ([10], III, Sect. 3).

From  $\Pr \{\xi_j^\ell(t = 2) r_N^\ell(t = 1) = \pm r_N^\ell(t = 1)\} = \frac{1}{2}$  it follows that for the random variable  $S_M = \sum_{\substack{\ell=1 \\ (\ell \neq q)}}^M \xi_j^\ell(t = 2) r_N^\ell(t = 1)$  one gets  $ES_M = 0$ ,  $\operatorname{Var} S_M = \sum_{\substack{\ell=1 \\ (\ell \neq q)}}^M (r_N^\ell(t = 1))^2$ . So, the characteristic function  $f_M(\tau)$  for the random variable  $S_M/\sqrt{\operatorname{Var} S_M}$  has the form

$$f_M(\tau) = E \exp \left( i \tau \frac{S_M}{\sqrt{\operatorname{Var} S_M}} \right) = \prod_{\substack{\ell=1 \\ (\ell \neq q)}}^M \cos \left( \tau \frac{r_N^\ell(t = 1)}{\sqrt{\sum_{\substack{\ell=1 \\ (\ell \neq q)}}^M (r_N^\ell(t = 1))^2}} \right). \tag{17}$$

Now, by the ergodic theorem (or the law of the large numbers) for the almost all realizations  $\{r^\ell(t = 1)\}_{\ell \neq q}$  we get that  $\alpha$  – lim  $M^{-1} \sum_{\substack{\ell=1 \\ (\ell \neq q)}}^M (r_N^\ell(t = 1))^2 = E(r^\ell(t = 1))^2 = 1$ , see equation

(13) and remark after it. This means that by definition (8) and by the iterated logarithm law ([10], IV, Sect. 4) one gets

$$\alpha - \lim \left( \frac{r_N^\ell(t = 1)/\sqrt{M}}{\sqrt{\frac{1}{M} \sum_{\substack{\ell=1 \\ (\ell \neq q)}}^M (r_N^\ell(t = 1))^2}} \right) \stackrel{\operatorname{Pr}=1}{=} 0. \tag{18}$$

Consequently, from equations (17) and (18) it follows that

$$\alpha - \lim f_M(\tau) = \exp \left( -\frac{1}{2} \tau^2 \right), \tag{19}$$

i.e.,  $S_M/\sqrt{\text{Var } S_M} \rightarrow N(0, 1)$  in the  $\alpha$  - lim in distribution. Then as a corollary (cf. (11)), we get

$$\alpha - \lim \frac{1}{\sqrt{N}} \sum_{\substack{\ell=1 \\ (\ell \neq p, q)}}^M \xi_j^\ell(t=2) r_N^\ell(t=1) \stackrel{d}{=} \sqrt{\alpha} N(0, 1). \tag{20}$$

Using the independence  $\{\xi_j^q(t=2)\}$  of random variables (20) and equations (12), (20) we can calculate limit distribution for the random variables (cf. (11))

$$w_{j,N}^{q,p}(t=2) = \left[ \xi_j^q(t=2) m_N^q(t=1) + \frac{1}{\sqrt{N}} \sum_{\substack{\ell=1 \\ (\ell \neq p, q)}}^M \xi_j^\ell(t=2) r_N^\ell(t=1) \right]. \tag{21}$$

One gets

$$F_{w_N}(x) = \Pr \{w_{j,N}^{q,p}(t=2) \leq x\} = \frac{1}{2\sqrt{2\pi\alpha}} \int_{-\infty}^x dz \times \\ \times \left\{ \exp\left[-\frac{(z - m^q(t=1))^2}{2\alpha}\right] + \exp\left[-\frac{(z + m^q(t=1))^2}{2\alpha}\right] \right\} + O(1/\sqrt{N}). \tag{22}$$

From the independence  $\{\xi_j^p(t=2)\}_{p \neq q}$  of  $\{w_{j,N}^{p,q}(t=2)\}$  and the symmetry of distribution (22) we obtain that distribution of its products (cf. (11))  $\{\xi_j^p(t=2) w_{j,N}^{p,q}(t=2)\}_{j=1}^N$  coincides with (22) :

$$F_{\xi w}(x) = F_w(x). \tag{23}$$

Therefore, random variables  $\left\{ \eta_{j,N}^p = \frac{1}{\sqrt{N}} r_N^p(t=1) + \xi_j^p(t=2) w_{j,N}^{p,q}(t=2) \right\}_{j=1}^N$  (for a fixed realization  $\{r_N^\ell(t=1)\}_{\ell \neq q}$  !) form i.i.d.r.v. Now we can apply CLT (again in the scheme of series ([10], III, Sect. 4)) to the sequence  $\{\text{sign } \eta_{j,N}^p\}_{j=1}^N$ , controlling the rate of convergence in (22) by the Berry-Esseen theorem ([10], III, Sect. 6) :

$$\alpha - \lim \frac{\sum_{j=1}^N \text{sign } \eta_{j,N}^p - E \sum_{j=1}^N \text{sign } \eta_{j,N}^p}{\sqrt{\text{Var } \sum_{j=1}^N \text{sign } \eta_{j,N}^p}} \stackrel{d}{=} N(0, 1). \tag{24}$$

Let us stress that it is key patterns from the next layer  $t=2$  that guarantee the independence  $\{\xi_j^p(t=2) w_{j,N}^{p,q}(t=2)\}_{j=1}^N$  of realization  $\{r_N^\ell(t=1)\}_p$  and finally the independence  $N(0, 1)$ , in (24), of  $r^p(t=1)$ . For the Hopfield model they are *dependent* : this creates main difficulties in derivation of dynamical equations for this model [8].

By definition of the variance one gets

$$\text{Var } \sum_{j=1}^N \text{sign } \eta_{j,N}^p = \sum_{j=1}^N (1 - (E \text{sign } \eta_{j,N}^p)^2). \tag{25}$$

Using equations (22) and (23) we can estimate expectation in (24) and (25) :

$$\begin{aligned}
 E \operatorname{sign} \eta_{j,N}^p &= \frac{1}{2 \sqrt{2 \pi \alpha}} \int_{-r_N^p(t=1)/\sqrt{N}}^{r_N^p(t=1)/\sqrt{N}} dF_w(x) + O(1/N) \\
 &= \frac{1}{\sqrt{N}} r^p(t=1) \sqrt{\frac{2}{\pi \alpha}} \exp\left[-\frac{(m^q(t=1))^2}{2 \alpha}\right] + O(1/N).
 \end{aligned}
 \tag{26}$$

So, using equations (24), (26) we get the recurrence relation for residual overlaps (cf. (11)) :

$$r^p(t=2) = \alpha - \lim \frac{1}{\sqrt{N}} \sum_{j=1}^N \operatorname{sign} \eta_{j,N}^p \stackrel{d}{=} N(0, 1) + r^p(t=1) \sqrt{\frac{2}{\pi \alpha}} \exp\left[-\frac{(m^q(t=1))^2}{2 \alpha}\right].
 \tag{27}$$

Here, in accordance to that we have stressed above, the random Gaussian variable  $N(0, 1)$  is *independent* of  $r^p(t=1)$ . So, one gets for  $\operatorname{Var} r^p(t) = D(t)$  the following relations (see (13) and (27)) :

$$\begin{aligned}
 D(t=1) &= 1, \\
 D(t=2) &= 1 + \frac{2}{\pi \alpha} \exp\left[-\frac{(m^q(t=1))^2}{\alpha}\right].
 \end{aligned}$$

**6.** From the explicit formula for  $\eta_{j,N}^p$  and the limit distribution (22) it follows that random variables  $\{\eta_j^p\}_p$  are uncorrelated for  $p \neq p'$  :  $E(\xi_j^p(t=2) \xi_j^{p'}(t=2)) = 0$ , and variables  $\{r^p(t=1)\}_p$  are even independent for  $p \neq p'$ . Therefore, random variables  $\{r^p(t=2)\}_p$  are uncorrelated. From (11) it follows that this property is reproducible. If  $\{r^p(t)\}_p$  are *uncorrelated* at the moment  $t$ , then by above arguments the same property has the sequence  $\{r^p(t+1)\}_p$ .

Therefore, we can use induction in  $t$  to derive the recurrence relations for the main and residual overlaps for arbitrary  $t$ . The first step we have proved in 4 and 5.

Now let  $\{r^p(t)\}_p$  be uncorrelated identically distributed Gaussian random variables with zero mean and  $\operatorname{Var} r^p(t) = D(t)$ . From the symmetry and the independence  $\xi_j^p(t+1)$  of  $r^p(t)$  we get that  $\frac{1}{\sqrt{N}} \sum_{p \neq q}^M \xi_j^p(t+1) r^p(t)$  is again Gaussian variable (as a sum of Gaussians)

which in the  $\alpha - \lim$  is equal to  $\sqrt{\alpha} N(0, D(t))$ . Then, taking into account  $r^p(t) \stackrel{d}{=} \alpha - \lim r_N^p(t)$  and equations (9), (10), one gets

$$m^q(t+1) = \operatorname{erf}\left(\frac{m^q(t)}{\sqrt{\alpha D(t)}}\right).
 \tag{28}$$

Using for uncorrelated sequence  $\{r_N^p(t)\}_p$  the individual ergodic theorem we can state that

$$\alpha - \lim \frac{1}{M} \sum_{p \neq q}^M (r_N^p(t))^2 = D(t)
 \tag{29}$$



for the almost all realizations. Let realization  $\{r^p(t)\}_p$  be fixed. Then, using (29) and the line of reasoning of the section 5, we get (see (11) and (20)-(22)) :

$$\alpha - \lim \frac{1}{\sqrt{N}} \sum_{\substack{\ell=1 \\ (\ell \neq p, q)}^M} \xi_j^\ell(t+1) r_N^\ell(t) \stackrel{d}{=} \sqrt{\alpha} N(0, D(t)). \quad (30)$$

This means that we can copy the proof in the section 5 to get (cf. (27)) :

$$r^p(t+1) = N(0, 1) + r^p(t) \sqrt{\frac{2}{\pi \alpha D(t)}} \exp\left[-\frac{(m^q(t))^2}{2 \alpha D(t)}\right] \quad (31)$$

where again the random variables  $N(0, 1)$  and  $r^p(t)$  are independent, i.e.,

$$D(t+1) = 1 + \frac{2}{\pi \alpha} \exp\left[-\frac{(m^q(t))^2}{\alpha D(t)}\right]. \quad (32)$$

This finishes the proof for the zero temperature  $\theta = 0$ .

For  $\theta \neq 0$  one has to use evolution defined by stochastic equation (4). As far as the heat-bath temperature noise (5) is independent of internal evolution we have simply to average equations (28) and (30) over this noise. As a result one gets

$$m^q(t+1) = \frac{1}{\sqrt{2 \pi \alpha D(t)}} \int_{-\infty}^{+\infty} dx \operatorname{th} [\beta (x + m^q(t))] \exp\left[-\frac{x^2}{2 \alpha D(t)}\right], \quad (33)$$

$$r^p(t+1) = N(0, 1) + r^p(t) \frac{\beta}{\sqrt{2 \pi \alpha D(t)}} \int_{-\infty}^{+\infty} dy \frac{\exp[-y^2/2 \alpha D(t)]}{\operatorname{ch}^2 \beta (y + m^q(t))}, \quad (34)$$

where random variables  $N(0, 1)$  and  $r^p(t)$ ,  $p \neq q$ , are uncorrelated,  $r^p(t=1) = N(0, 1)$  and  $D(t) = \operatorname{Var} r^p(t)$ , i.e.,

$$D(t+1) = 1 + \frac{\beta^2}{2 \pi \alpha} \left[ \int_{-\infty}^{+\infty} dy \frac{\exp[-y^2/2 \alpha D(t)]}{\operatorname{ch}^2 \beta (y + m^q(t))} \right]^2. \quad (35)$$

7. The equations (28), (32) and (33), (35) have been derived at the first time in [3, 4] for the case of linear Hebbian couplings (3) and for the various extensions of the basic DMK-model in reference [5]. There one can find also a detailed analysis of the fixed point equation and the critical behavior of the model.

Function  $D(t)$  in the above papers in an auxiliary saddle point order parameter. The advantage of our method is, in addition, in decoding of the meaning of the parameter  $D(t)$  as the variance of the residual overlaps responsible for the intrinsic noise in dynamics of the main overlap.

*Note added.* — After this paper has been finished (Dubna, July 1989) we become aware of the recent work by the DMK-model inventors [11] where very similar (except of the equation for the residual overlaps and the rigorous proofs) ideas were proposed, see also [12].

#### Acknowledgments.

This (English) version of the paper has been written when one of the authors (V. A. Z.) visited Centre de Physique Théorique, CNRS-Luminy, Marseille. He would like to thank R. Lima, A. Messager, P. Picco, J. Ruiz and other members of the Centre for their interest

and useful discussions. He also thanks R. Kühn and J. Slawny for helpful remarks and comments. V. A. Z. thanks CPT-CNRS and the Faculté des Sciences de Luminy, Université d'Aix-Marseille II for kind hospitality and financial support. The authors are grateful to Antoinette Sueur for efficient typing of the manuscript.

#### References

- [1] DOMANY E., MEIR R. and KINZEL W., *Europhys. Lett.* **2** (1986) 175.
- [2] MINSKY M. and PAPERT S., *Perceptrons* (Cambridge, Mass. : MIT Press) 1969.
- [3] MEIR R. and DOMANY E., *Phys. Rev. Lett.* **59** (1987) 359 ; *Europhys. Lett.* **4** (1987) 645.
- [4] MEIR R. and DOMANY E., *Phys. Rev. A* **37** (1988) 608.
- [5] MEIR R., *J. Phys. France* **49** (1988) 201.
- [6] GARDNER E., DERRIDA B. and MOTTISHAW P., *J. Phys. France* **48** (1987) 741.
- [7] ZAGREBNOV V. A. and CHVYROV A. S., *Sov. Phys. JETP* **68** (1989) 153.
- [8] ZAGREBNOV V. A. and PATRICK A. E., Probabilistic Approach to Parallel Dynamics for Hopfield Model, JINR-Dubna, Preprint 1989 (in Russian).
- [9] AMIT D. J., GUTFREUND H. and SOMPOLINSKY H., *Ann. Phys. NY* **173** (1987) 30.
- [10] SHIRYAYEV A. N., *Probability* (New York : Springer-Verlag) 1984.
- [11] DOMANY E., MEIR R. and KINZEL W., *J. Phys. A : Math. Gen.* **22** (1989) 2081.
- [12] DERRIDA B. and MEIR R., *Phys. Rev. A* **38** (1988) 3116.