Dynamics of Ising random-bond models: neural network and random-anisotropy-axis model

1995 J. Phys. A: Math. Gen. 28 63

(http://iopscience.iop.org/0305-4470/28/1/013)

Dynamics of Ising random-bond models: neural network and random-anisotropy-axis model

D R C Dominguez and W K Theumann

Instituto de Física, Universidade Federal do Rio Grande do Sul Caixa Postal 15051, 91501-970, Porto Alegre, RS, Brazil

Received 24 June 1994

Abstract. A discrete-time retrieval dynamics for a class of random-bond models with infinite-range interactions is studied in a unified picture. The Hopfield model of neural networks with the Hebb learning rule is considered together with the Ising random-anisotropy-axis model in the strong-anisotropy limit. The main overlap (magnetization), the residual overlap and its dispersion, as well as the correlation between the Gaussian component of the residual overlap and the initial value of the latter, obtained in the first step of a recursion relation, are used to infer the structure of the recursion for large times. A crucial assumption is the strong stationarity of the Gaussian component. The dynamics are discussed for finite $\alpha = p/N$ (the storage ratio in the neural network problem or the ratio of random-axis components per site) in the limit where both p and N go to infinity. The long-time behaviour of the theory is shown to yield the equilibrium solution of an earlier work in mean-field theory, for a tri-modal distribution of random-axis components. Explicit results for the basins of attraction of either a ferromagnetic or a spin-glass phase are obtained, as well as the relaxation time with a square-root power-law decay near saturation.

1. Introduction

The equilibrium properties of the Hopfield model of neural networks (NN) with the Hebb learning rule [1–3] and those of the isomorphic Ising random-anisotropy-axis model (IRAM), in the strong anisotropy limit [4–8] have been studied extensively in recent years and are now well understood. These are systems which can be described by a quenched Ising random-bond Hamiltonian

$$H_N(\{\sigma_i\}) = \sum_{i \neq j} J_{ij}\sigma_i\sigma_j \tag{1}$$

with spins $\sigma_i = \pm 1$ (active or inactive neuron, respectively) on the sites (neurons) $i = 1, \ldots, N$. The interaction J_{ij} (synaptic matrix) has the generalized Hebbian form

$$\begin{aligned} J_{ij} &= J\boldsymbol{n}_i \cdot \boldsymbol{n}_j \\ &= \frac{1}{N}\sum_{\mu=1}^{p} n_i^\mu n_j^\mu \end{aligned} \tag{2}$$

in which the components $\{n_i^\mu, \mu = 1, \ldots, p\}$ are random numbers representing the embedded patterns ($n_i^\mu = \pm 1$) in NN or the random-axis components in the IRAM.

The IRAM describes the unusual properties of amorphous intermetallic compounds through the strong-anisotropy limit of the Hamiltonian [9]

$$H_N(\{s_i\}) = -J\sum_{i\neq j}^{N} s_i \cdot s_j - D\sum_{i=1}^{N}(n_i \cdot s_i)^2 \tag{3}$$

when $D/J \to \infty$, in which case the spins tend to align with the anisotropy axis, such that $s_i = n_i\sigma_i$ and (3) reduces, up to a constant, to (1) with J_{ij} given by (2).

Different probability distributions for the random vectors are of interest in the two models. Whereas a diagonal distribution is appropriate for NN, isotropic and cubic-anisotropic distributions have been used for the IRAM. In a recent work in mean-field theory (MFT) on the IRAM we discussed the effects of the tri-modal distribution [10]

$$p(n_i^\mu) = b\left[\delta(n_i^\mu - a) + \delta(n_i^\mu + a)\right]/2 + (1-b)\delta(n_i^\mu) \tag{4}$$

of independent identically distributed random variables (IIDRV), where $ba^2 = 1$ and in which $b = 1$ corresponds to the diagonal case. It was shown that magnetic ordered states with a finite magnetization appear for both finite p and in the large-component limit below a critical value of $\alpha = p/N$ (the storage ratio of the NN problem) and that the size of the magnetic region shrinks with decreasing b in a (α, T) phase diagram, T being the temperature (synaptic noise in NN). The magnetic states correspond to finite overlaps with the stored patterns in the NN analogue and they can be characterized by locally or globally stable states.

Whereas MFT is appropriate for the NN problem (below the critical storage capacity α_c) because of the large number of neurons in synaptic contact with a given neuron ($\sim 10^4$) it is certainly only an approximation (albeit a valuable one) for the IRAM with finite-range interaction. Careful numerical simulations for the NN problem above α_c yield results which are not explained by MFT at $T = 0$ [3].

Nevertheless, it is of interest to have a dynamics to check the MFT results in the long-time limit, and numerous works have been devoted to the NN problem [11–21] (although none, to our knowledge, to the IRAM).

Parallel (synchronous) dynamics, studied in the absence of external noise, is one in which the spins are updated according to the rule

$$\sigma_i(t+1) = \mathrm{sgn}[h_i(t)] \tag{5}$$

where the local field is defined as

$$h_i(t) = \sum_{i\neq j} J_{ij}\sigma_j(t)\,. \tag{6}$$

Recursion relations for the main and residual overlaps in a retrieval dynamics for the NN problem have been derived by Patrick and Zagrebnov (PZ) [21]. While the first and second steps in the recursion yield the Kinzel and Gardner–Derrida–Mottishaw results [11, 12], respectively, the third step and the conjectured general structure were of no use in giving a clue as to the long-time behaviour. It was instead argued that the latter follows from a set of plausible stationarity assumptions concerning the main and residual overlaps in the limit $N \to \infty$ and from a Gaussian distribution with a time-dependent variance, for the residual overlaps. The long-time behaviour that follows is that of MFT [3]. An alternative signal-to-noise ratio analysis has been carried out by Amari and Maginu (AM) [22]. Recent Monte Carlo simulations by Nishimori and Ozeki [23] seem to confirm their main assumption for retrieval dynamics of a Gaussian distributed residual overlap with a time-dependent variance.

The main purpose of this paper is to present a simpler and more general derivation of a deterministic (i.e. without external noise) discrete-time retrieval dynamics which enables one to identify the basic conditions which lead to the stationary MFT results in the long-time limit, in contrast to previous works [21, 22]. Our aim is to do this for a general distribution of random $\{n_i^\mu\}$. We show that our purpose can be achieved in terms of three macrovariables: the main overlap, the dispersion of the residual overlap and the correlation of the Gaussian component of the residual overlap at the next time step with the residual overlap itself. We show that the structure of the equation for the correlation depends on a strong-stationarity assumption for the Gaussian component. Alternative studies of a retrieval dynamics by Shukla [24] and by Coolen and Sherrington [25] have appeared recently.

A further aim of our work is to study the basins of attraction for the IRAM and the NN problem with the tri-modal distribution of random variables given by (4). We do not consider non-retrieval dynamics for which the assumption of a Gaussian distributed noise is presumably incorrect.

The outline of the paper is the following. In section 2 we present some formal relations for the low and high components of the overlaps. In section 3 we discuss the first step of the recursion, and in section 4 we consider the extension to large times and we show there the results for the attractors. Comparison with some previous works [21, 22] is made throughout the paper. We conclude in section 5 with a few remarks, and comment on the relationship between our work and Shukla's. A brief presentation of formal results for a stochastic noisy (finite-temperature) dynamics is also included.

2. Formal relations

The main variable of interest is the overlap (or magnetization) vector

$$\boldsymbol{m}_N(t) = \frac{1}{N}\sum_j \boldsymbol{n}_j \sigma_j(t) \tag{7}$$

with components $\boldsymbol{m}_N = \{m_N^\mu\}$ at time step t, in terms of which the local field in MFT may be written as

$$\begin{aligned} h_i(t) &= \boldsymbol{n}_i \cdot \boldsymbol{m}_{Ni}(t) \\ &= \sum_\mu n_i^\mu m_{Ni}^\mu(t) \end{aligned} \tag{8}$$

where $\boldsymbol{m}_{Ni}(t)$ is defined by the sum in (7) with the term $j = i$ excluded. Given the magnetization at $t = 0$, the magnetization at $t = 1$ follows from (5) and (7).

Following earlier work, we are interested in the time evolution of the magnetization with a single macroscopic component of order unity and $(p-1)$ remaining components of $O(1/\sqrt{p})$ in order to ensure a finite magnetization. We write, accordingly,

$$\boldsymbol{n}_i = \begin{pmatrix} \rho_i \\ \xi_i \end{pmatrix} \qquad \xi_i = \{\xi_i^\mu;\, \mu = 2, \ldots, p\} \tag{9}$$

and

$$\boldsymbol{m}_N(t) = \begin{pmatrix} m_N(t) \\ l_N(t) \end{pmatrix} \qquad l_N = \{l_N^\mu;\, \mu = 2, \ldots, p\}. \tag{10}$$

This will be useful for studying the evolution of the main overlap (the main component of the magnetization)

$$m_t = \lim_\alpha m_N(t) \tag{11}$$

in the so-called α-limit, or $\lim_\alpha$, in which $\alpha = p/N$ is kept finite while p and $N \to \infty$. Here

$$m_N(t) = \frac{1}{N}\sum_i \rho_i \sigma_i(t) \tag{12}$$

is the first component of $\boldsymbol{m}_N(t)$, which is of O(1). Since the remaining components are of O$(1/\sqrt{p})$, it is convenient to define the finite residual overlap

$$L_N^\mu(t) \equiv \sqrt{p}\, l_N^\mu(t) \tag{13}$$

in which

$$l_N^\mu(t) = \frac{1}{N}\sum_i \xi_i^\mu \sigma_i(t)\,. \tag{14}$$

The internal field (8) may then be written as

$$h_{Ni}(t) = \rho_i m_{Ni}(t) + \omega_{Ni}(t) \tag{15}$$

where

$$\omega_{Ni}(t) \equiv \boldsymbol{\xi}_i \cdot \boldsymbol{L}_{Ni}(t)/\sqrt{p} \qquad \boldsymbol{L}_{Ni} = \{L_{Ni}^\mu; \mu = 2, \ldots, p\} \tag{16}$$

is the noise term.

Next we apply these formal relations to work out explicitly the main overlap and discuss the probability distribution for the noise, in the first step.

3. First step in the recursion

Let the initial configuration $\{\sigma_j(t=0); j = 1, \ldots, N\}$ be such that the main overlap

$$\lim_\alpha m_N(t=0) = m_0 \tag{17}$$

is finite, of O(1), and

$$\langle L_0^\mu \rangle = 0 \qquad \langle (L_0^\mu)^2 \rangle = l_0^2 \tag{18}$$

where the initial *residue* defined by

$$L_0^\mu \equiv \lim_\alpha L_N^\mu(0) \tag{19}$$

is a random variable in μ of a further unspecified distribution.

Here we relax the usual strong initial condition that $\sigma_j(0)$ be uncorrelated with the $\{\xi_i^\mu\}$ which yields a variance of the residual overlap in the $\lim_\alpha$ equal to α, in place of l_0^2. We assume instead the weaker condition that the residue L_0^μ be independent of the $\{\xi_i^\mu\}$. This ensures that the averages of the first two moments of these quantities factorize into a product of averages.

3.1. Main overlap

If the noise converges to a sequence ω_0 of IIDRV (a point to be justified below) then the $\lim_\alpha$ of the main overlap, equation (12), in the first time step becomes, with the law of large numbers,

$$\begin{aligned} m_1 &= \langle \rho\sigma(0) \rangle_{\rho,\omega} \\ &= \langle \rho\, \mathrm{sgn}(h_0) \rangle_{\rho,\omega} \end{aligned} \tag{20}$$

making use of (5), in which

$$h_0 = \rho m_0 + \omega_0 . \tag{21}$$

Here

$$h_0 = \lim_{\alpha} h_N(t=0) \tag{22}$$

is the limiting initial internal field and

$$\omega_0 = \lim_{\alpha} \omega_N(t=0) \tag{23}$$

while the brackets denote the average over both the probability distribution of the initial noise and over ρ.

To justify the convergence of the noise term we follow PZ and note that $\omega_{Ni}(0)$, given by (16), is a sum of IIDRV to which the central limit theorem (CLT) can be applied so that

$$\lim_{\alpha} \frac{\omega_{Ni}(0) - \langle \omega_{Ni}(0)\rangle}{\sqrt{\mathrm{Var}[\omega_{Ni}(0)]}} \doteq z \tag{24}$$

in which the left-hand-side is distributed ($\doteq$) as a Gaussian random variable z with a mean of zero and unit variance. From (16)

$$\begin{aligned}\langle \omega_0 \rangle &= \lim_{\alpha} \langle \omega_{Ni}(0) \rangle \\ &= \lim_{\alpha} \sum_{\mu} \langle L^{\mu}_{Ni}(0)\xi^{\mu}_i \rangle / \sqrt{p}\end{aligned} \tag{25}$$

and this vanishes, (i.e. $\langle \omega_0 \rangle = 0$) due to our assumption of independence of L^{μ}_0 with $\{\xi^{\mu}_i\}$ and the distribution in (4). This is also in accordance with other authors [21, 22]. Similarly, the variance

$$\mathrm{Var}(\omega_0) = \lim_{\alpha} \frac{1}{p} \sum_{\mu} \mathrm{Var}\big[L^{\mu}_{Ni}(0)\xi^{\mu}_i\big] \tag{26}$$

becomes

$$\begin{aligned}\mathrm{Var}(\omega_0) &= \lim_{\alpha} \frac{1}{p} \sum_{\mu} \left(L^{\mu}_0\right)^2 \\ &= l_0^2\end{aligned} \tag{27}$$

from (18). Thus with the first two moments of $\omega_{Ni}(0)$ in the $\lim_{\alpha}$ we have for the initial noise that

$$\omega_0 \doteq l_0 z \tag{28}$$

where z is a Gaussian random variable with a mean of zero and unit variance. This should be compared with the distribution of width $\sqrt{\alpha}$ of other authors [21, 22].

Note that the average over the noise of the sign of the local field can now be calculated and this yields

$$\lim_{\alpha} \langle \mathrm{sgn}[h_N(0)] \rangle_{\omega} = \mathrm{erf}(x_0 \rho/\sqrt{2}) \tag{29}$$

in which $x_0 \equiv m_0/l_0$ and

$$\mathrm{erf}(x/\sqrt{2}) \equiv \sqrt{2/\pi} \int_0^x \mathrm{d}z\, \mathrm{e}^{-z^2/2} . \tag{30}$$

We are now able to write down the $\lim_{\alpha}$ of the main overlap in the first step as

$$m_1 = H(x_0) \equiv \langle \rho\, \mathrm{erf}(x_0 \rho/\sqrt{2}) \rangle_{\rho} \tag{31}$$

with the remaining average over the distribution of ρ. This is the generalization of Kinzel's result for the overlap for all time, which yields a critical $\alpha_c \sim 0.64$ for retrieval with a finite overlap [11]. This first-step result is exact only for the dilute network [13] and is far from the statistical mechanical result of $\alpha_c \sim 0.14$ for the fully connected network [3].

3.2. Residual overlap

To continue with the first step we now consider the evolution of the residual overlap which involves the high components of the magnetization. In the noise term (16) it is convenient to separate a part for a given μ

$$\omega_N(0) = L_N^\mu(0)\xi^\mu/\sqrt{p} + v_N(0) \tag{32}$$

where, for simplicity, the site dependence has been suppressed and in which the remainder noise

$$v_N(0) \equiv \sum_{\nu\neq\mu} L_N^\nu(0)\xi^\nu/\sqrt{p} \tag{33}$$

is distributed in the $\lim_\alpha$

$$\lim_\alpha v_N(0) \doteq l_0 z \tag{34}$$

as a Gaussian random variable of mean zero and variance l_0^2. This follows from the CLT applied to the sequence $\{L_0^\nu\xi^\nu\}$, $\nu \neq \mu$.

Next we consider the local field (15) and use (32) to write

$$h_N(0) = \hat{h}(0)l_0 + L_N^\mu\xi^\mu(0)/\sqrt{p} \tag{35}$$

where

$$\hat{h}(0)l_0 \equiv \rho m_N(0) + v_N(0)\,. \tag{36}$$

Using the fact that, from (13), $L_N^\mu(0)$ of order unity is much smaller than $\sqrt{p}$, we expand

$$\mathrm{sgn}[h_N(0)] \simeq \mathrm{sgn}\big[\hat{h}(0)\big] + 2\frac{L_N^\mu(0)\xi^\mu}{l_0\sqrt{p}}\delta\big[\hat{h}(0)\big] \tag{37}$$

to first order in $1/\sqrt{p}$ to be used in the expression for the residue

$$L_N^\mu(1) = \sum_i \lambda_i \qquad \lambda_i \equiv \frac{\sqrt{p}}{N}\xi_i^\mu\,\mathrm{sgn}[h_{Ni}(0)] \tag{38}$$

in the first step. It should be noted that $\{\lambda_i\}$ is not a set of independent random variables, as pointed out in a somewhat different context by PZ [21]. Indeed, ξ_i^μ is correlated with $\mathrm{sgn}[h_{Nk}(0)]$, if $i \neq k$, through a term proportional to $L_{Nk}^\mu(0)/\sqrt{p}$. Note, however, that this is a weak correlation, implying that the CLT may still be used for the distribution of $L_N^\mu(1)$.

The content of the theorem in this context is that

$$\lim_\alpha \frac{\sum_i \lambda_i - N\langle\lambda_i\rangle}{\sqrt{N\,\mathrm{Var}(\lambda_i)}} \doteq z \tag{39}$$

in which the left-hand side is again distributed as a Gaussian random variable of mean zero and unit variance, and

$$\mathrm{Var}(\lambda_i) = \alpha/N \tag{40}$$

follows from (38).

Moreover, we may write

$$N\langle\lambda_i\rangle = \sqrt{p}\big\langle\xi_i^\mu\langle\mathrm{sgn}[h_{Ni}(0)]\rangle_z\big\rangle_{\xi\rho} \tag{41}$$

where the averages are to be taken over the distributions of ξ, ρ and z. In the $\lim_\alpha$, the interesting limit for our purpose, equation (41) yields

$$\lim_\alpha(N\langle\lambda_i\rangle) = \frac{L_0^\mu}{l_0}G(x_0) \tag{42}$$

where L_0^μ has been defined in (19) and

$$G(x_0) \equiv \sqrt{2/\pi}\big\langle\exp\left[-(x_0\rho)^2/2\right]\big\rangle_\rho \tag{43}$$

in which $x_0 = m_0/l_0$. On the other hand, the variance of L_N^μ becomes α, through equation (40).

We are now able to state the following theorem.

Theorem. In the $\lim_\alpha$, the residue L_1^μ in the first time step, defined as

$$L_1^\mu \equiv \lim_\alpha L_N^\mu(1) \tag{44}$$

is given by the stochastic equation

$$L_1^\mu \doteq \frac{L_0^\mu}{l_0}G(x_0) + \sqrt{\alpha}z_1 \tag{45}$$

where z_1 is a Gaussian random variable with mean zero and unit variance in the first time step.

A similar but more restricted relationship has been derived by PZ [21] and AM [22]. To complete the proof of (45) through the CLT we must ensure the condition in Liapunov's theorem for uniform convergence to a Gaussian. This condition requires that

$$a_\delta \equiv \sum_i M_{\delta+2}^i \Big/ \Big[\sum_i M_2^i\Big]^{\delta+2} \tag{46}$$

for some $0 < \delta \leqslant 1$ and vanishes in the limit of validity of (45). Here

$$M_n^i \equiv \big\langle(\lambda_i - \langle\lambda_i\rangle)^n\big\rangle . \tag{47}$$

Indeed, we find that $a_\delta = \mathrm{O}(1/N)$ with a coefficient which can be explicitly calculated for $\delta = 1$.

3.3. Correlation of the residue

Defining

$$l_t^2 \equiv \lim_\alpha \frac{1}{p}\sum_\mu \left[L_N^\mu(t)\right]^2 = \big\langle(L_t^\mu)^2\big\rangle \tag{48}$$

for any time step $t \geqslant 0$, we note that the variance of L_1^μ becomes, with equation (45),

$$l_1^2 = \alpha + 2\sqrt{\alpha}c_0G(x_0) + [G(x_0)]^2 . \tag{49}$$

Here

$$c_0 \equiv \big\langle z_1 L_0^\mu\big\rangle / l_0 \tag{50}$$

is the correlation between the Gaussian component of the residue in the first time step and the initial residue. This is, in general, non-zero, except in the perceptron layered NN problem [19].

Postulating that a relationship like (49), relating the dispersion of the residue in the first time step to initial quantities, applies between two consecutive time steps at almost all later times and anticipating the following section, we define for the second step the correlation

$$c_1 \equiv \langle z_2 L_1^\mu \rangle / l_1 \tag{51}$$

to be determined in the first time step, where z_2 is a Gaussian random variable with mean zero and unit variance in the second step. Note that

$$c_1^2 \leqslant \langle (z_2)^2 \rangle \langle (L_1^\mu)^2 \rangle / l_1^2 \equiv 1 \tag{52}$$

yields a bound for c_1.

We do not know how the Gaussian component of the residue at a given time step, z_t, is related to that at a previous time. *If* they were the same (a point to be discussed next) the average in (51), making use of (45), becomes

$$c_1 = [\sqrt{\alpha} + c_0 G(x_0)]/l_1 \,. \tag{53}$$

We now have (31), (49) and (53) for m_1, l_1 and c_1, completely describing the first step of the recursion. Although we expect the first two to be exact, the latter rests on the assumption on z_2.

4. Attractors

Here we are interested in the asymptotic behaviour of the model in which the equilibrium situation is reached. With that purpose we postulate, with no proof, the *strong stationarity* of z_t for long times in which

$$z_t \doteq z \tag{54}$$

is distributed as a Gaussian random variable with mean zero and unit variance. If L_t^μ is at most weakly dependent on $\{\xi_i^\mu\}$ for almost all t so that the conditions for (27) apply in the $\lim_\alpha$, then the noise would become

$$\omega_t \doteq l_t z \tag{55}$$

that is, a Gaussian random variable with mean zero and variance l_t^2. Although we cannot prove this rigorously, we checked by means of an expansion like that in (37) that the average of L_t^μ with ξ_i^μ vanishes to order $1/\sqrt{p}$ and that the average

$$C_{\mu\nu} = \langle L_N^\mu(t)\xi^\mu L_N^\nu(t)\xi^\nu \rangle \tag{56}$$

also needed in the proof vanishes as $1/p$ in the $\lim_\alpha$.

Having the exact equation for the main overlap and the residue at the first time step, we assume they have the right structure for almost all t. With the strong stationarity of z_t, equation (53) yields the corresponding recursion relation between c_t and c_{t+1}, in which

$$c_t = \langle z L_t^\mu \rangle / l_t \,. \tag{57}$$

We thus have the stochastic equation

$$L_{t+1}^\mu \doteq \frac{L_t^\mu}{l_t} G(x_t) + \sqrt{\alpha} z \tag{58}$$

and the deterministic equations

$$l_{t+1}^2 = \alpha + 2\sqrt{\alpha} c_t G(x_t) + [G(x_t)]^2 \tag{59}$$

$$c_{t+1} = [\sqrt{\alpha} + c_t G(x_t)]/l_{t+1} \tag{60}$$

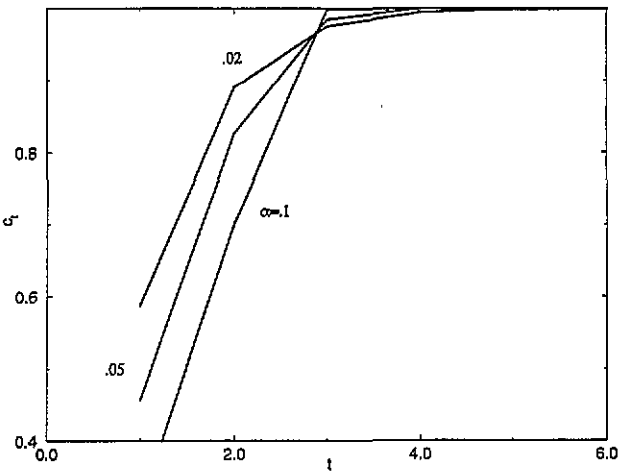

Figure 1. Convergence of the parameter c_t to asymptotic behaviour $c^* = 1$ for $\alpha = 0.1, 0.05$ and 0.02, when $u = 0$ (i.e. $b = 1$), with initial main overlap $m_0 = 0.55$ and initial dispersion $l_0 = 0.5$. The approach of c_t to $c^* = 1$ indicates the onset of the equilibrium solution discussed in the text.

for almost all t, in extension of (45), (49) and (53) where $x_t = m_t/l_t$ and

$$m_{t+1} = H(x_t) \tag{61}$$

is the recursion relation for the main overlap, with $H(x)$ given by (31).

In generalizing our equations for the first time step we proceeded in a different way to PZ [21], who generalized their first step equation to the long-time limit introducing an appropriate variance for the residual overlap. Here we have one more parameter , c_t, whose long-time recursion relation depends on the strong stationarity of z_t, except in the perceptron case, where c_t is identically zero and (59) becomes the exact equation for the dispersion of the residue [19, 26]. Equations (58)–(61) also generalize the results of AM [22].

We are interested in the stationary states of the equations. The stable states represent configurations in which the macroscopic properties take their equilibrium values and the basins of attraction provide an idea of the domain of stability of these solutions.

The formal results, so far, are for a general distribution of $\{n_i^\mu\}$. Using equation (4) the numerical solution of the equations with various initial values for the parameters reveal that the fixed-point solution $c^* = c_{t+1} = c_t = 1$ is the only stable solution reached after a few time steps, as shown in figure 1 for various values of α when $u \equiv 1 - b = 0$, and initial values $m_0 = 0.55$ and $l_0 = 0.5$. The curve for $\alpha = 0.1$ corresponds to initial values slightly inside the ferromagnetic (retrieval) phase shown in figures 2 and 3, while the other two curves correspond to points well within that phase. Note that the convergence to the asymptotic behaviour is slower for larger α. A similar behaviour follows for increasing values of u, except that the convergence to asymptotic behaviour is slower. When $c^* = 1$ is reached equations (59) and (60) become the same and, together with (61), they coincide with the equilibrium solutions for the main overlap and the dispersion of the residue at zero temperature (no noise) of our earlier work [10]. Thus, the long-time behaviour of our equations yields the expected equilibrium mean-field-theory results that generalize the work of Amit *et al* [3] within replica symmetry. This is a strong indication of the correctness of our assumptions.

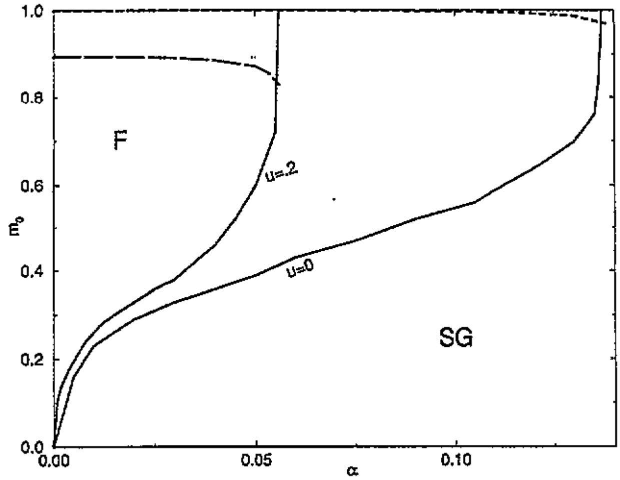

Figure 2. Boundaries of the basins of attraction (full curves) of either the ferromagnetic (F) or spin-glass (SG) phase for two values of u (equation (4)) as functions of the initial main overlap m_0 and $\alpha = p/N$, for $l_0 = 0.5$ and $c_0 = 0$. The fixed-point values m^*, for $u = 0$ and 0.2, are given by the upper and lower broken curves, respectively.

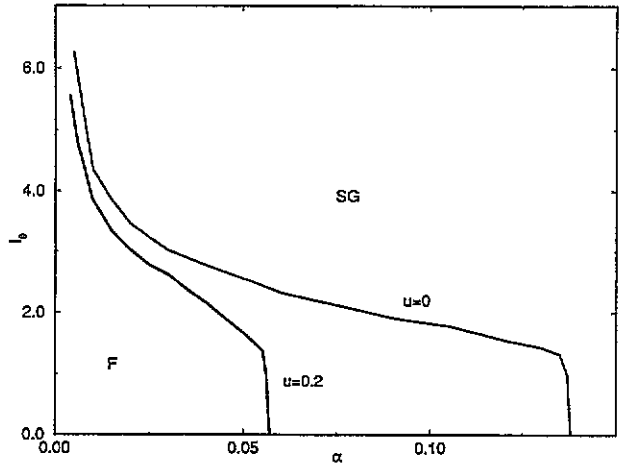

Figure 3. Boundaries of the basins of attraction, as in figure 2, for the initial dispersion of the residue, l_0, for $m_0 = 1$ and $c_0 = 0$.

Numerical results for the basins of attractions are shown in figures 2–4 for two values of $u = 1 - b$. There are two distinct regions which constitute the domains of attraction of the ferromagnetic (F) and spin-glass (SG) regions on either side of the full curves. The former is characterized by a stable fixed-point solution $m^* = m_{t+1} = m_t$ with $m^* \neq 0$, while $m^* = 0$ for the latter. Figure 2 shows the basins of attraction as functions of the initial overlap m_0, for given c_0 and l_0. The ferromagnetic (retrieval) region cannot be reached unless m_0 is larger than a critical value $m_c(\alpha)$ lying on either one of the curves for a given α, even if $\alpha < \alpha_c$, the critical storage capacity. Here $\alpha_c = 0.138$ and 0.056 for $u = 0$ and 0.2, respectively, are the values where the phase boundaries become vertical. The final stable values for $m^* > 0$ are also shown (the upper dotted curves). If $m_0 < m_c(\alpha)$, the fixed point

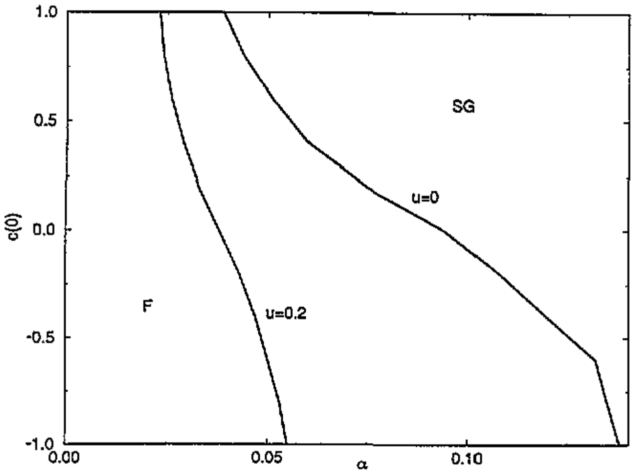

Figure 4. Boundaries of the basins of attraction, as in figure 2, for the initial correlation c_0, for $m_0 = 1 = l_0$.

$m^* = 0$ (not shown in the figure) is reached instead.

The basins of attraction as functions of the initial dispersion of the residue l_0, for given c_0 and m_0, are shown in figure 3. If l_0 is larger than a critical value $l_c(\alpha)$ for a given α, there is a spin-glass-like phase, even if $\alpha < \alpha_c$, as one would expect. Similarly, if the initial correlation $c_0 > c_c(\alpha)$, one also ends up with a spin-glass-like phase, even if $\alpha < \alpha_c$, as can be seen from the basins of attraction in figure 4.

Another quantity of interest is the relaxation time τ to the limiting m^*, such that

$$|m_t - m^*| \sim \exp(-t/\tau). \tag{62}$$

As noted by Meir and Domany [26] in the context of layered feed-forward networks, there is a 'critical slowing down' with a divergent relaxation time as $\alpha \to \alpha_c$, even though the transition to ferromagnetism (retrieval) is of first order. This is due to the merging of the stable and unstable branches for m^* as $\alpha \to \alpha_c$, so that the fixed point at α_c is only marginally stable. We find that this also takes place in our case when (59)–(61) are solved and $|m_t - m_{t+1}|$ is fitted to (62) for increasing t. The numerical solution of our equations in the vicinity of the fixed point yields for

$$\tau = (\alpha_c - \alpha)^{-\nu} \tag{63}$$

$\nu = 0.51 \pm 0.03$, as shown in figure 5 when $u = 0$, which is the same as the layered-network result.

5. Discussion and further results

We formulated a discrete-time retrieval dynamics in the absence of external noise and discussed the conditions that have to be met in order to set up the recursion relations for the first time step. We obtained the main overlap and residue (the $\lim_\alpha$ for the residual overlap), the dispersion and the correlation of the Gaussian component of the residue with the initial residue. We proceeded differently than PZ [21] and AM [22] in order to derive a retrieval dynamics based on weaker initial conditions and also to be able to derive results for a generalized distribution of random variables.

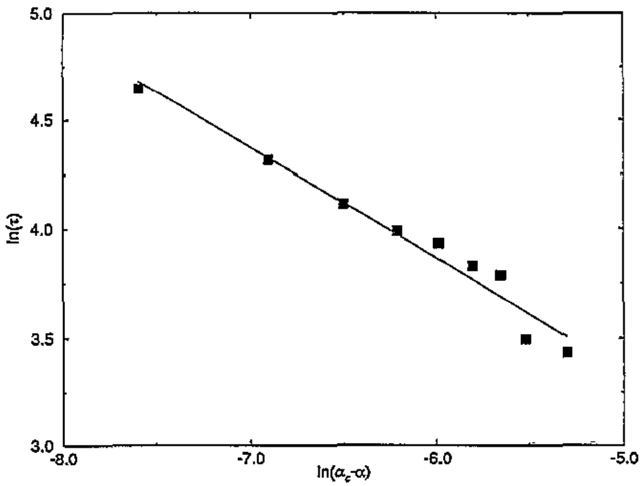

Figure 5. Logarithmic plot of the relaxation time near saturation (equation (63)) for $u = 0$. The best fit of the solution of (59)–(61) to (62) gives $\nu = 0.51 \pm 0.03$, when $u = 0$.

Since it seems very difficult to obtain the general structure of the equations rigorously for all times from a step-by-step procedure, as recognized by PZ [21], we proceeded assuming that our equations for the main overlap and stochastic residue have the right structure for almost all times *if* the Gaussian component of the residue satisfies a strong-stationarity hypothesis which seems reasonable for long times. In distinction to [21, 22], our equations contain an additional macrovariable: the correlation c_t between the Gaussian component of the residue in the next time and the residue itself. When the fixed-point equations are solved for our retrieval dynamics we recover the MFT results of our earlier work for a trimodal distribution of random variables which generalize the replica symmetric work of Amit *et al*. In this work we have not considered non-retrieval dynamics to account for remanence memory effects. In distinction to the recent retrieval dynamics formulated by Coolen and Sherrington [25], who used the replica method to calculate explicitly the intrinsic noise distribution which yields the replica-symmetric equilibrium MFT results, we obtain these without using the replica method.

At this point, a comparison of our work with that of Shukla [24] for the NN problem is in order. Following PZ, we aimed at a microscopic derivation of the equations, which goes beyond a signal-to-noise ratio analysis [22] for the main overlap, the residue, the dispersion of the residue and the correlation of the Gaussian component as macrovariables of the system. These are all quantities that can be calculated. In contrast, reference [24] presents a macroscopic description for the main overlap in terms of the dynamical energy per bit (or energy per spin) in the thermodynamic limit, $e(\alpha, t)$. This is very difficult to obtain but a signal-to-noise ratio argument was used to suggest an energy-conserving parallel dynamics for the NN problem in which $e(\alpha, t) = -0.5$ for $\alpha \leqslant 0.14$ and *all* times after the initial time step. Thus, $e(\alpha) = e(\alpha, t)$ takes the fixed-point value proposed earlier by Kohring [16]. In the low storage limit for the NN case ($b = 1$) our equations yield (at large times and when $c_t \to 1$) the result

$$m_{t+1} = \operatorname{erf}\left(\frac{m_t}{\sqrt{2}l_t}\right) \qquad l_{t+1}^2 = \alpha + 1 - m_t^2 \tag{64}$$

to *leading* order in α. These equations coincide with Shukla's equations (3.12) and (3.14).

For higher values of α the two dynamical theories differ. Our results are those of replica-symmetric MFT [3], for *any* α, whereas the numerical simulations in [24] are done for a limiting α between the replica-symmetric results and the value found in previous numerical simulations that agree very well with first-step replica-symmetry breaking [27].

The dynamics discussed in this work, in the $\lim_\alpha$, can be extended to other NN problems. The case considered in this work is the retrieval dynamics of a network trained with simple patterns. An interesting extension, which is being studied, is the generalization dynamics for a network trained with examples.

The noiseless (zero-temperature) dynamics can be extended [21,24] by adding noisy terms $\phi_i, i = 1, \ldots, N$ to the internal field so that

$$\sigma_i(t+1) = \mathrm{sgn}[h_i(t) + \phi_i] \tag{65}$$

in which $\{\phi_i\}$ are IIDRV with a probability distribution

$$P(\phi \leqslant x) = \tfrac{1}{2}[1 + \tanh(\beta x)] \tag{66}$$

where β is the inverse temperature. The conditional probability that the spin takes a value $\sigma_i(t)$ given the internal field h_i at the previous step is then

$$P[\sigma_i(t)|h_i(t-1)] = \exp[\beta h_i \sigma_i(t)]/[2\cosh(\beta h_i)]\,. \tag{67}$$

The calculation of the main overlap requires the average over ϕ

$$\langle\sigma(t)\rangle_\phi = \tanh[\beta A(t)] \tag{68}$$

in which

$$A(t) \equiv \rho m_t + l_t z \tag{69}$$

where z is a Gaussian random variable with mean zero and unit variance. This yields

$$m_{t+1} = \langle\rho\tanh[\beta A(t)]\rangle_{\rho,z} \tag{70}$$

for the main overlap and

$$L^\mu_{t+1} \doteq \sqrt{\alpha} z_{t+1} + \frac{L^\mu_t}{l_t} G_\beta \tag{71}$$

for the stochastic residue where

$$G_\beta \equiv \langle z\tanh[\beta A(t)]\rangle_{\rho z}\,. \tag{72}$$

We also obtain, with the strong-stationarity assumption, the variance of the residue (71)

$$l^2_{t+1} = \alpha + \sqrt{\alpha} c_t G_\beta + G^2_\beta \tag{73}$$

and its correlation with the Gaussian component at the next time

$$c_{t+1} = [\sqrt{\alpha} + c_t G_\beta]/l_{t+1}\,. \tag{74}$$

When c_t converges to the stable fixed-point solution $c^* = 1$ this equation and (70) reproduce the finite-temperature equations of our earlier work in MFT.

Acknowledgments

We thank Alba Theumann for calling our attention to the work by Coolen and Sherrington [25]. This work was supported in part by CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnológico) and FINEP (Financiadora de Estudos e Projetos), Brazil.

References

[1] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[2] Amit D J 1989 *Modeling Brain Function* (Cambridge: Cambridge University Press)
[3] Amit D J, Gutfreund H and Sompolinsky H 1987 *Ann. Phys.* **173** 30
[4] Fisher K and Zippelius A 1985 *J. Phys. C: Solid State Phys.* **18** L1139 (corrigendum 1986 *J. Phys. C: Solid State Phys.* **19** L265)
[5] Harris A B, Caflisch R G and Banavar J R 1987 *Phys. Rev.* B **35** 4929
[6] Fisch R 1989 *Phys. Rev.* B **39** 873; 1990 *Phys. Rev.* B **41** 11 705
[7] Jayaprakash C and Kirkpatrick S 1980 *Phys. Rev.* B **21** 4072
[8] Derrida B and Vannimenus J 1980 *J. Phys. C: Solid State Phys.* **13** 3261
[9] Harris R, Plischke M and Zuckermann M J 1973 *Phys. Rev. Lett.* **31** 160
[10] Dominguez D R C and Theumann W K 1993 *Phys. Rev.* B **48** 6234
[11] Kinzel W 1985 *Z. Phys.* B **60** 205
[12] Gardner E, Derrida B and Mottishaw P 1987 *J. Phys.* **48** 741
[13] Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett.* **4** 167
[14] Horner H, Bohrmann D, Frick M, Kinzenbach H and Schmidt A 1989 *Z. Phys.* B **76** 381
[15] Kepler T B and Abbott L F 1988 *J. Physique* **49** 1657
[16] Kohring G A 1989 *Europhys. Lett.* **8** 697
[17] Opper M, Kleinz J, Köhler H and Kinzel W 1989 *J. Phys. A: Math. Gen.* **22** L407
[18] Zagrebnov V A and Chyvrov A S 1989 *Sov. Phys.–JETP* **68** 153
[19] Patrick A E and Zagrebnov V A 1990 *J. Physique* **51** 1129
[20] Patrick A E and Zagrebnov V A 1991 *J. Phys. A: Math. Gen.* **24** 59
[21] Patrick A E and Zagrebnov V A 1991 *J. Stat. Phys.* **63** 3413
[22] Amari S and Maginu K 1988 *Neural Networks* **1** 63
Amari S 1990 *Proc. IEEE* **78** 1443
[23] Nishimori H and Ozeki T 1993 *J. Phys. A: Math. Gen.* **26** 859
[24] Shukla P 1993 *J. Stat. Phys.* **71** 705
[25] Coolen A C C and Sherrington D 1993 *Phys. Rev. Lett.* **71** 3886
[26] Meir R and Domany E 1987 *Phys. Rev. Lett.* **59** 359
[27] Crisanti A, Amit D J and Gutfreund H 1986 *Europhys. Lett.* **2** 337