J. Phys. A: Math. Gen. **31** (1998) 8607–8621. Printed in the UK PII: S0305-4470(98)94141-5

Phase diagram and storage capacity of sequence processing neural networks*

A Düring†, A C C Coolen‡ and D Sherrington†

† Department of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK
‡ Department of Mathematics, King's College, University of London, Strand, London WC2R 2LS, UK

Received 13 May 1998

Abstract. We solve the dynamics of Hopfield-type neural networks which store sequences of patterns, close to saturation. The asymmetry of the interaction matrix in such models leads to violation of detailed balance, ruling out an equilibrium statistical mechanical analysis. Using generating functional methods we derive exact closed equations for dynamical order parameters, namely the sequence overlap and correlation and response functions, in the thermodynamic limit. We calculate the time translation invariant solutions of these equations, describing stationary limit cycles, which leads to a phase diagram. The effective retarded self-interaction usually appearing in symmetric models is here found to vanish, which causes a significantly enlarged storage capacity of $\alpha_c \sim 0.269$, compared with $\alpha_c \sim 0.139$ for Hopfield networks storing static patterns. Our results are tested against extensive computer simulations and excellent agreement is found.

1. Introduction

The equilibrium properties of the Hopfield model [1], a globally coupled neural network, with the typically Hebbian prescription for the interaction strengths

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \tag{1}$$

(in which the ξ_i^μ represent components of patterns to be stored) have been successfully described in the regime close to saturation, where the number p of patterns stored scales as $p = \alpha N$, using replica methods [2, 3]. As an alternative approach, a path integral formalism developed in [4], was applied to the dynamics of the same system, and both approaches have indeed been shown to lead to identical phase diagrams [5]. Many modifications of the standard Hopfield model have been proposed, including models where the network does not statically recall individual patterns, but reproduces a *sequence* of stored patterns [1, 6–10]. The simplest way to induce this cyclical behaviour is by an asymmetric modification of the interaction matrix (1), in combination with a parallel execution of the neural dynamics. Numerical simulations show that in such (non-symmetric) models the number of patterns that can be stored successfully is significantly larger than that of models storing static patterns, with a storage capacity of $\alpha_c \approx 0.27$ [1, 7], compared with $\alpha_c \approx 0.14$ for the standard (symmetric) Hopfield model.

* This paper is dedicated to Professor Horner on the occasion of his 60th birthday, in recognition of his considerable services both to theoretical physics and to scientific publishing (for example as Editor of Zeitschrift für Physik).

0305-4470/98/438607+15$19.50

In this paper we study such a model, where a single sequence of extensive length is stored in a fully (but non-symmetrically) connected Ising spin neural network with parallel stochastic dynamics. The asymmetry of the interaction matrix and the resulting violation of detailed balance and associated fluctuation–dissipation theorems rule out equilibrium statistical mechanical methods of analysis, including conventional replica theory. Some time ago an approximate dynamical solution for this model was proposed [7], which provided results roughly in line with the numerical evidence available at the time. To our knowledge, an exact solution, however, has so far not yet been found.

In our present study we use the path integral methods of [4, 5, 11] to solve the dynamics close to saturation exactly in the thermodynamic limit for our fully connected sequence processing network, without having to resort to approximations. In the standard (symmetric) Hopfield network two effects limit the storage capacity: a Gaussian noise in the equivalent effective single-spin problem, which is non-local in time, and a retarded self-interaction. The magnitude of both depends on the load factor α. Our theory shows that for the present model the retarded self-interaction vanishes, similar to the situation in the non-symmetric Sherrington–Kirkpatrick (SK) model [12–14], which explains the extended storage capacity. Numerical simulations for large system sizes (up to 50 000 spins) are in excellent agreement with our analytical results, both with respect to the maximum storage capacity $\alpha_c \approx 0.269$ (at zero noise level) and with respect to the full phase diagram in the $\alpha - T$ plane.

2. Definitions

We study a system consisting of N Ising-type neurons $\sigma_i = \pm 1$ which evolve in time according to a stochastic alignment to local fields. The neurons change their states simultaneously, with probabilities

$$\mathrm{Prob}[\sigma_i(t+1) = -\sigma_i(t)] = \frac{1}{2}\left[1 - \tanh\left(\beta\sigma_i(t)\left[\sum_{j=1}^{N} J_{ij}\sigma_j(t) + \theta_i(t)\right]\right)\right] \tag{2}$$

where the entries of the interaction matrix $\boldsymbol{J}$ are given by

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^{\mu+1}\xi_j^{\mu} \tag{3}$$

(the pattern labels μ are understood to be taken modulo p). The non-negative parameter $\beta = T^{-1}$ controls the amount of noise in the dynamics, with $T = 0$ corresponding to deterministic evolution and with $T = \infty$ corresponding to purely random evolution. The variables $\theta_i(t)$ represent external fields. The p vectors $\boldsymbol{\xi}^{\mu} = (\xi_1^{\mu}, \ldots, \xi_N^{\mu}) \in \{-1, 1\}^N$ are randomly and independently drawn patterns. Our interest is in the saturation regime $p = \alpha N$. For discussions of the relation of such models to biological or artificial neural networks see e.g. [15–18]. The matrix $\boldsymbol{J}$ will generally be non-symmetric, so that (2) will not obey detailed balance. Hence we cannot use conventional equilibrium statistical mechanics to analyse the stationary behaviour: we will have to solve the dynamics. For the subsequent analysis, it will turn out to be useful to represent our expression for $\boldsymbol{J}$ in matrix notation as

$$\boldsymbol{J} = \frac{1}{N}(\boldsymbol{\xi}^T \boldsymbol{S}\boldsymbol{\xi}) \qquad S_{\mu\nu} = \delta_{\mu,(\nu+1) \bmod p}. \tag{4}$$

Here the $p \times N$ matrix $\boldsymbol{\xi}$ is defined as $\boldsymbol{\xi}_{\mu i} = \xi_i^{\mu}$. When $\boldsymbol{S}$ is replaced by the unity matrix $\mathbf{1}$, the definition (4) reverts to that of the standard Hopfield model.

To analyse the dynamics of the system we follow [4] and define a generating (or characteristic) functional $Z[\psi]$:

$$Z[\psi] = \sum_{\boldsymbol{\sigma}(0)\ldots\boldsymbol{\sigma}(t)} p[\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t)]\mathrm{e}^{-\mathrm{i}\sum_{s<t}\boldsymbol{\sigma}(s)\cdot\boldsymbol{\psi}(s)} \tag{5}$$

in which $\boldsymbol{\sigma}(s) = (\sigma_1(s),\ldots,\sigma_N(s))$ denotes the microscopic system state at time s, and with the usual notation $\boldsymbol{x}\cdot\boldsymbol{y} = \sum_i x_i y_i$. In the familiar way one can obtain from $Z[\psi]$ all averages of interest by differentiation, e.g.

$$m_i(s) = \langle\sigma_i(s)\rangle = \mathrm{i}\lim_{\psi\to 0}\frac{\partial Z[\psi]}{\partial\psi_i(s)} \tag{6}$$

$$G_{ij}(s,s') = \frac{\partial}{\partial\theta_j(s')}\langle\sigma_i(s)\rangle = \mathrm{i}\lim_{\psi\to 0}\frac{\partial^2 Z[\psi]}{\partial\psi_i(s)\partial\theta_j(s')} \tag{7}$$

$$C_{ij}(s,s') = \langle\sigma_i(s)\sigma_j(s')\rangle = -\lim_{\psi\to 0}\frac{\partial^2 Z[\psi]}{\partial\psi_i(s)\partial\psi_j(s')}. \tag{8}$$

The dynamics (2) is a Markov chain, so the path probabilities $p[\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t)]$ are simply given by products of the individual transition probabilities $W[\boldsymbol{\sigma}'|\boldsymbol{\sigma}]$ of the chain: $p[\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t)] = p[\boldsymbol{\sigma}(0)]\prod_{s=0}^{t-1} W[\boldsymbol{\sigma}(s+1)|\boldsymbol{\sigma}(s)]$. For the dynamics (2) these transition probabilities are given by

$$\begin{aligned} W[\boldsymbol{\sigma}(s+1)|\boldsymbol{\sigma}(s)] &= \prod_i \tfrac{1}{2}\left[1+\sigma_i(s+1)\tanh\left(\beta\left[\sum_j J_{ij}\sigma_j(t)+\theta_i(t)\right]\right)\right] \\ &= \prod_i \mathrm{e}^{\beta\sigma_i(s+1)[\sum_j J_{ij}\sigma_j(s)+\theta_i(s)]-\ln 2\cosh(\beta[\sum_j J_{ij}\sigma_j(s)+\theta_i(s)])}. \end{aligned}$$

To formally remove the coupling terms $\sigma_i(s+1)\sigma_j(s)$ we introduce the auxiliary variables $\boldsymbol{h}(s) = (h_1(s),\ldots,h_N(s))$, representing the local fields at each spin site at given times, by insertion of

$$1 = \int \mathrm{d}\boldsymbol{h}(s)\prod_i \delta\left[h_i(s) - \sum_j J_{ij}\sigma_j(s) - \theta_i(s)\right].$$

After writing the above δ-distributions in integral form, which generates conjugate field variables $\hat{\boldsymbol{h}}(s) = (\hat{h}_1(s),\ldots,\hat{h}_N(s))$, and upon introducting the more convenient notation $\{\mathrm{d}\boldsymbol{h}\,\mathrm{d}\hat{\boldsymbol{h}}\} = \prod_i\prod_{s<t}[\mathrm{d}h_i(s)\,\mathrm{d}\hat{h}_i(s)/2\pi]$, we can express (5) as

$$\begin{aligned} Z[\psi] = \sum_{\boldsymbol{\sigma}(0)\ldots\boldsymbol{\sigma}(t)} p(\boldsymbol{\sigma}(0))\int\{\mathrm{d}\boldsymbol{h}\,\mathrm{d}\hat{\boldsymbol{h}}\}\prod_{s<t}\mathrm{e}^{\beta\boldsymbol{\sigma}(s+1)\cdot\boldsymbol{h}(s)-\sum_i\ln 2\cosh[\beta h_i(s)]+\mathrm{i}\hat{\boldsymbol{h}}(s)\cdot[\boldsymbol{h}(s)-\boldsymbol{\theta}(s)]-\mathrm{i}\boldsymbol{\psi}(s)\cdot\boldsymbol{\sigma}(s)} \\ \times\mathrm{e}^{-\mathrm{i}N^{-1}\hat{\boldsymbol{h}}(s)\cdot(\boldsymbol{\xi}^T\boldsymbol{S}\boldsymbol{\xi})\boldsymbol{\sigma}(s)}. \end{aligned} \tag{9}$$

This expression describes the system dynamics (2), (3) in general form. To obtain quantitative information about particular regimes of operation, we have to make specific ansatze. Our ansatz will be one describing (possibly noisy) recall of the stored sequence of patterns. At each timestep exactly one stored pattern is assumed to be 'condensed', i.e. the overlap between that pattern (which without loss of generality can be labelled with the time index) and the system state is of $\mathcal{O}(1)$, whereas all other overlaps are of $\mathcal{O}(N^{-\frac{1}{2}})$. The cumulative impact of the overlaps of the non-condensed patterns will introduce an additional noise component into the system dynamics; the non-condensed patterns play the role of 'quenched disorder'. For $N\to\infty$ the mean-field physics of the problem should be self-averaging with respect to the realisation of the disorder, so we are allowed to average the generating functional (9) over the non-condensed patterns (such averages will be denoted

as $\overline{f[\{\boldsymbol{\xi}\}]}$). Since each pattern with $\mu \leqslant t$ will at some stage be condensed, in contrast with those patterns with $\mu > t$, we can simplify our calculation by averaging only over the latter. The resulting expressions will, for $N \to \infty$, turn out not to depend on the remaining patterns with $\mu \leqslant t$.

As in most dynamic mean-field calculations of disordered systems based on evaluating disorder-averaged generating functionals, we will consider the time t to be fixed, whereas we will take the limit $N \to \infty$. This restricts the predicting power of the theory to those processes that take place on finite timescales. In the present calculation we will find that a time-translation invariant state (representing motion on a stationary limit cycle) is indeed approached on finite timescales, so this restriction is not a problem.

3. Dynamic mean-field theory

In (9) only the term $\hat{\boldsymbol{h}}(s)\cdot(\boldsymbol{\xi}^T\boldsymbol{S}\boldsymbol{\xi})\boldsymbol{\sigma}(s)$ contains both condensed and non-condensed patterns. We isolate the non-condensed ones by introducing the variables $\boldsymbol{x}$ and $\boldsymbol{y}$:

$$1 = \int \mathrm{d}\boldsymbol{x} \prod_{s<t}\prod_{\mu\neq s} \delta\left[x_\mu(s) - \frac{1}{\sqrt{N}}\sum_i \xi_i^{\mu+1}\hat{h}_i(s)\right]$$

$$1 = \int \mathrm{d}\boldsymbol{y} \prod_{s<t}\prod_{\mu\neq s} \delta\left[y_\mu(s) - \frac{1}{\sqrt{N}}\sum_i \xi_i^{\mu}\sigma_i(s)\right].$$

Upon writing the above δ-distributions in integral form (which generates the additional integration variables $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$), we then arrive at the following expression for the disorder-averaged generating functional:

$$\begin{aligned}\overline{Z}[\psi] = \sum_{\boldsymbol{\sigma}(0)\ldots\boldsymbol{\sigma}(t)} p(\boldsymbol{\sigma}(0)) \int \{\mathrm{d}\boldsymbol{h}\,\mathrm{d}\hat{\boldsymbol{h}}\} \mathrm{e}^{\sum_{s<t}[\beta\boldsymbol{\sigma}(s+1)\cdot\boldsymbol{h}(s) - \sum_i \ln 2\cosh[\beta h_i(s)] + \mathrm{i}\hat{\boldsymbol{h}}(s)\cdot[\boldsymbol{h}(s)-\boldsymbol{\theta}(s)] - \mathrm{i}\psi(s)\cdot\boldsymbol{\sigma}(s)]} \\ \times \mathrm{e}^{-\mathrm{i}N^{-1}\sum_{s<t}[\hat{\boldsymbol{h}}(s)\cdot\boldsymbol{\xi}^{s+1}][\boldsymbol{\sigma}(s)\cdot\boldsymbol{\xi}^s]} \\ \times \int \frac{\mathrm{d}\boldsymbol{x}\,\mathrm{d}\hat{\boldsymbol{x}}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\hat{\boldsymbol{y}}}{(2\pi)^{2(p-1)t}} \mathrm{e}^{\mathrm{i}\sum_{s<t}\sum_{\mu\neq s}[\hat{x}_\mu(s)x_\mu(s) + \hat{y}_\mu(s)y_\mu(s) - x_\mu(s)y_\mu(s)]} \\ \times \overline{[\mathrm{e}^{-\mathrm{i}N^{-\frac{1}{2}}\sum_{s<t}\sum_{\mu\neq s}[\hat{x}_\mu(s)\hat{\boldsymbol{h}}(s)\cdot\boldsymbol{\xi}^{\mu+1} + \hat{y}_\mu(s)\boldsymbol{\sigma}(s)\cdot\boldsymbol{\xi}^\mu]}]. \end{aligned} \tag{10}$$

We can now carry out the disorder average in the last term, which is significantly simplified if in the exponent we use $\sum_{\mu\neq s}[\sum_i \ldots] = \sum_{\mu>t}[\sum_i \ldots] + \mathcal{O}(N)$. It gives

$$\begin{aligned}\overline{[\ldots]} &= \mathrm{e}^{\mathcal{O}(N^{\frac{1}{2}})}\overline{[\mathrm{e}^{-\mathrm{i}N^{-\frac{1}{2}}\sum_{s<t}\sum_{\mu>t}\sum_i \xi_i^\mu[\hat{x}_{\mu-1}(s)\hat{h}_i(s) + \hat{y}_\mu(s)\sigma_i(s)]}]} \\ &= \mathrm{e}^{\mathcal{O}(N^{\frac{1}{2}})}\prod_{\mu>t}\prod_i \cos\left[N^{-\frac{1}{2}}\sum_{s<t}[\hat{x}_{\mu-1}(s)\hat{h}_i(s) + \hat{y}_\mu(s)\sigma_i(s)]\right] \\ &= \mathrm{e}^{\mathcal{O}(N^{\frac{1}{2}})}\prod_{\mu>t}\mathrm{e}^{-\frac{1}{2N}\sum_{s,s'<t}\sum_i[\hat{x}_{\mu-1}(s)\hat{h}_i(s) + \hat{y}_\mu(s)\sigma_i(s)][\hat{x}_{\mu-1}(s')\hat{h}_i(s') + \hat{y}_\mu(s')\sigma_i(s')]}. \end{aligned} \tag{11}$$

Since the leading order N in the exponent of (11) does not involve components of $\{\boldsymbol{x}, \hat{\boldsymbol{x}}, \boldsymbol{y}, \hat{\boldsymbol{y}}\}$ with pattern index $\mu \leqslant t$, the latter can be integrated out in expression (10). We now isolate the various relevant macroscopic observables occurring in (11) by inserting integrals over appropriate δ-functions:

$$1 = \int \frac{\mathrm{d}\boldsymbol{m}\,\mathrm{d}\hat{\boldsymbol{m}}}{(2\pi/N)^t} \mathrm{e}^{\mathrm{i}N\sum_{s<t}\hat{m}(s)[m(s) - \frac{1}{N}\sum_i \xi_i^s\sigma_i(s)]}$$

$$1=\int\frac{\mathrm{d}\boldsymbol{k}\,\mathrm{d}\hat{\boldsymbol{k}}}{(2\pi/N)^t}\mathrm{e}^{\mathrm{i}N\sum_{s<t}\hat{k}(s)[k(s)-\frac{1}{N}\sum_i\xi_i^{s+1}\hat{h}_i(s)]}$$

$$1=\int\frac{\mathrm{d}\boldsymbol{q}\,\mathrm{d}\hat{\boldsymbol{q}}}{(2\pi/N)^{t^2}}\mathrm{e}^{\mathrm{i}N\sum_{s,s'<t}\hat{q}(s,s')[q(s,s')-\frac{1}{N}\sum_i\sigma_i(s)\sigma_i(s')]}$$

$$1=\int\frac{\mathrm{d}\boldsymbol{Q}\,\mathrm{d}\hat{\boldsymbol{Q}}}{(2\pi/N)^{t^2}}\mathrm{e}^{\mathrm{i}N\sum_{s,s'<t}\hat{Q}(s,s')[Q(s,s')-\frac{1}{N}\sum_i\hat{h}_i(s)\hat{h}_i(s')]}$$

$$1=\int\frac{\mathrm{d}\boldsymbol{K}\,\mathrm{d}\hat{\boldsymbol{K}}}{(2\pi/N)^{t^2}}\mathrm{e}^{\mathrm{i}N\sum_{s,s'<t}\hat{K}(s,s')[K(s,s')-\frac{1}{N}\sum_i\sigma_i(s)\hat{h}_i(s')]}.$$

Combination of (11) with (10) will then give us an expression for $\overline{Z}[\psi]$ which will factorize over sites if we choose a factorized initial distribution $p(\boldsymbol{\sigma}(0))=\prod_i p_i(\sigma_i(0))$, resulting in an integral which for $N\to\infty$ will be dominated by saddle points:

$$\overline{Z}[\psi]=\int\mathrm{d}\boldsymbol{m}\,\mathrm{d}\hat{\boldsymbol{m}}\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\hat{\boldsymbol{k}}\,\mathrm{d}\boldsymbol{q}\,\mathrm{d}\hat{\boldsymbol{q}}\,\mathrm{d}\boldsymbol{Q}\,\mathrm{d}\hat{\boldsymbol{Q}}\,\mathrm{d}\boldsymbol{K}\,\mathrm{d}\hat{\boldsymbol{K}}\,\mathrm{e}^{N\{\Psi[\ldots]+\Phi[\ldots]+\Omega[\ldots]\}+\mathcal{O}(N^{\frac{1}{2}})}\tag{12}$$

in which the functions $\Psi[\ldots]$, $\Phi[\ldots]$ and $\Omega[\ldots]$ are given by:

$$\Psi[\boldsymbol{m},\boldsymbol{k},\hat{\boldsymbol{m}},\hat{\boldsymbol{k}},\boldsymbol{q},\boldsymbol{Q},\boldsymbol{K},\hat{\boldsymbol{q}},\hat{\boldsymbol{Q}},\hat{\boldsymbol{K}}]=\mathrm{i}\sum_{s<t}[\hat{m}(s)m(s)+\hat{k}(s)k(s)-m(s)k(s)]$$
$$+\mathrm{i}\sum_{s,s'<t}[\hat{q}(s,s')q(s,s')+\hat{Q}(s,s')Q(s,s')+\hat{K}(s,s')K(s,s')]\tag{13}$$

$$\Phi[\boldsymbol{m},\boldsymbol{k},\hat{\boldsymbol{q}},\hat{\boldsymbol{Q}},\hat{\boldsymbol{K}}]=\frac{1}{N}\sum_i\ln\Bigg\{\sum_{\sigma(0)\ldots\sigma(t)}p_i(\sigma(0))\int\{\mathrm{d}h\,\mathrm{d}\hat{h}\}\mathrm{e}^{\sum_{s<t}[\beta\sigma(s+1)h(s)-\ln 2\cosh[\beta h(s)]]}$$
$$\times\mathrm{e}^{-\mathrm{i}\sum_{s,s'<t}[\hat{q}(s,s')\sigma(s)\sigma(s')+\hat{Q}(s,s')\hat{h}(s)\hat{h}(s')+\hat{K}(s,s')\sigma(s)\hat{h}(s')]}$$
$$\times\mathrm{e}^{\mathrm{i}\sum_{s<t}\hat{h}(s)[h(s)-\theta_i(s)-\hat{k}(s)\xi_i^{s+1}]-\mathrm{i}\sum_{s<t}\sigma(s)[\hat{m}(s)\xi_i^s+\psi_i(s)]}\Bigg\}\tag{14}$$

$$\Omega[\boldsymbol{q},\boldsymbol{Q},\boldsymbol{K}]=\frac{1}{N}\ln\int\frac{\mathrm{d}\boldsymbol{x}\,\mathrm{d}\hat{\boldsymbol{x}}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\hat{\boldsymbol{y}}}{(2\pi)^{2(p-t)t}}\mathrm{e}^{\mathrm{i}\sum_{\mu>t}\sum_{s<t}[\hat{x}_\mu(s)x_\mu(s)+\hat{y}_\mu(s)y_\mu(s)-x_\mu(s)y_\mu(s)]}$$
$$\times\mathrm{e}^{-\frac{1}{2}\sum_{\mu>t}\sum_{s,s'<t}[\hat{x}_\mu(s)Q(s,s')\hat{x}_\mu(s')+\hat{x}_{\mu-1}(s)K(s',s)\hat{y}_\mu(s')+\hat{y}_\mu(s)K(s,s')\hat{x}_{\mu-1}(s')+\hat{y}_\mu(s)q(s,s')\hat{y}_\mu(s')]}$$
$$=\frac{1}{N}\ln\int\frac{\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{v}}{(2\pi)^{(p-t)t}}\mathrm{e}^{\mathrm{i}\sum_{\mu>t}\sum_{s<t}u_{\mu+1}(s)v_\mu(s)}$$
$$\times\mathrm{e}^{-\frac{1}{2}\sum_{\mu>t}\sum_{s,s'<t}[u_\mu(s)Q(s,s')u_\mu(s')+u_\mu(s)K(s',s)v_\mu(s')+v_\mu(s)K(s,s')u_\mu(s')+v_\mu(s)q(s,s')v_\mu(s')]}$$

(15)

with the shorthand $\{\mathrm{d}h\,\mathrm{d}\hat{h}\}=\prod_{s<t}[\mathrm{d}h(s)\,\mathrm{d}\hat{h}(s)/2\pi]$. The final expression (15) for Ω was obtained by integrating out the variables $(\boldsymbol{x},\boldsymbol{y})$, followed by a simple pattern index shift transformation.

One can deduce the physical meaning of the various dynamic order parameters introduced along the way in the usual manner by (repeated) derivation of the definition (5) with respect to the external fields $\theta_i(s)$ and $\psi_i(s)$, in combination with usage of the normalization identity $Z[\mathbf{0}]=1$. Evaluation of a function $f[\ldots]$ at the dominating (physical) saddle-point of the extensive exponent in (12) will be indicated by $f|_{\mathrm{saddle}}$. The external fields occur in the function Φ only (not in Ψ or Ω). The resulting identities can be summarized in a compact form upon introduction of an effective single-site measure $\langle\ldots\rangle_i$, defined as

$$\langle f[\{\sigma,h,\hat{h}\}]\rangle_i=\frac{\sum_{\sigma(0)\ldots\sigma(t)}\int\{\mathrm{d}h\,\mathrm{d}\hat{h}\}W_i[\{\sigma,h,\hat{h}\}]f[\{\sigma,h,\hat{h}\}]}{\sum_{\sigma(0)\ldots\sigma(t)}\int\{\mathrm{d}h\,\mathrm{d}\hat{h}\}W_i[\{\sigma,h,\hat{h}\}]}$$

with

$$W_i[\{\sigma,h,\hat{h}\}] = p_i(\sigma(0))[\mathrm{e}^{\sum_{s<l}[\beta\sigma(s+1)h(s)-\ln 2\cosh[\beta h(s)]+\mathrm{i}\hat{h}(s)[h(s)-\theta_i(s)-\hat{k}(s)\xi_i^{s+1}]-\mathrm{i}\sigma(s)\hat{m}(s)\xi_i^s]} \times \mathrm{e}^{-\mathrm{i}\sum_{s,s'<l}[\hat{q}(s,s')\sigma(s)\sigma(s')+\hat{Q}(s,s')\hat{h}(s)\hat{h}(s')+\hat{K}(s,s')\sigma(s)\hat{h}(s')]}]|_{\mathrm{saddle}}. \tag{16}$$

In particular we now find, in leading order in N:

$$\overline{\langle\sigma_i(s)\rangle} = \mathrm{i}\lim_{\psi\to 0}\frac{\partial(N\Phi)}{\partial\psi_i(s)}\bigg|_{\mathrm{saddle}} = \langle\sigma(s)\rangle_i \tag{17}$$

$$0 = \frac{\partial\overline{Z}[0]}{\partial\theta_i(s)} = \lim_{\psi\to 0}\frac{\partial(N\Phi)}{\partial\theta_i(s)}\bigg|_{\mathrm{saddle}} = -\mathrm{i}\langle\hat{h}(s)\rangle_i \tag{18}$$

$$\begin{aligned}\overline{\langle\sigma_i(s)\sigma_j(s)\rangle} &= -\lim_{\psi\to 0}\frac{\partial^2(N\Phi)}{\partial\psi_i(s)\partial\psi_j(s')}\bigg|_{\mathrm{saddle}} - \lim_{\psi\to 0}\left[\frac{\partial(N\Phi)}{\partial\psi_i(s)}\frac{\partial(N\Phi)}{\partial\psi_j(s')}\right]\bigg|_{\mathrm{saddle}} \\ &= \delta_{ij}\langle\sigma(s)\sigma(s')\rangle_i + [1-\delta_{ij}]\langle\sigma(s)\rangle_i\langle\sigma(s')\rangle_i\end{aligned} \tag{19}$$

$$\begin{aligned}\frac{\partial\overline{\langle\sigma_i(s)\rangle}}{\partial\theta_j(s')} &= \mathrm{i}\lim_{\psi\to 0}\frac{\partial^2(N\Phi)}{\partial\psi_i(s)\partial\theta_j(s')}\bigg|_{\mathrm{saddle}} + \mathrm{i}\lim_{\psi\to 0}\left[\frac{\partial(N\Phi)}{\partial\psi_i(s)}\frac{\partial(N\Phi)}{\partial\theta_j(s')}\right]\bigg|_{\mathrm{saddle}} \\ &= -\mathrm{i}\delta_{ij}\langle\sigma(s)\hat{h}(s')\rangle_i\end{aligned} \tag{20}$$

$$\begin{aligned}0 = \frac{\partial^2\overline{Z}[0]}{\partial\theta_i(s)\partial\theta_j(s')} &= \lim_{\psi\to 0}\frac{\partial^2(N\Phi)}{\partial\theta_i(s)\partial\theta_j(s')}\bigg|_{\mathrm{saddle}} + \lim_{\psi\to 0}\left[\frac{\partial(N\Phi)}{\partial\theta_i(s)}\frac{\partial(N\Phi)}{\partial\theta_j(s')}\right]\bigg|_{\mathrm{saddle}} \\ &= -\delta_{ij}\langle\hat{h}(s)\hat{h}(s')\rangle_i.\end{aligned} \tag{21}$$

Note that we have already used identity (18) to simplify (20) and (21).

4. Derivation of saddle-point equations

In the limit $N\to\infty$, the integral (12) will be dominated by the dominating (physical) saddle point of the extensive exponent $\Psi+\Phi+\Omega$. We are now in a position to derive the saddle-point equations by differentiation with respect to our integration variables $\{\boldsymbol{m},\boldsymbol{k},\hat{\boldsymbol{m}},\hat{\boldsymbol{k}},\boldsymbol{q},\boldsymbol{Q},\boldsymbol{K},\hat{\boldsymbol{q}},\hat{\boldsymbol{Q}},\hat{\boldsymbol{K}}\}$. These equations will involve the average sequence overlap $m(s)$ (which measures the quality of the sequence recall) and the average single-site correlation and response functions $C(s,s')$ and $G(s,s')$:

$$m(s) = \lim_{N\to\infty}\frac{1}{N}\sum_i\overline{\langle\sigma_i(s)\rangle}\xi_i^s \tag{22}$$

$$C(s,s') = \lim_{N\to\infty}\frac{1}{N}\sum_i\overline{\langle\sigma_i(s)\sigma_i(s')\rangle} \tag{23}$$

$$G(s,s') = \lim_{N\to\infty}\frac{1}{N}\sum_i\frac{\partial\overline{\langle\sigma_i(s)\rangle}}{\partial\theta_i(s')}. \tag{24}$$

Straightforward differentiation, followed by usage of the identities (17)–(21) wherever possible, then leads us to the following saddle-point equations:

$$\text{for all } s,s': \qquad k(s) = \hat{m}(s) = Q(s,s') = 0 \tag{25}$$

$$\text{for all } s: \qquad \hat{k}(s) = m(s) = \lim_{N\to\infty}\frac{1}{N}\sum_i\langle\sigma(s)\rangle_i\xi_i^s \tag{26}$$

$$\text{for all } s,s': \qquad q(s,s') = C(s,s') = \lim_{N\to\infty}\frac{1}{N}\sum_i\langle\sigma(s)\sigma(s')\rangle_i \tag{27}$$

$$\text{for all } s,s': \qquad K(s,s') = iG(s,s') = \lim_{N\to\infty}\frac{1}{N}\sum_i \langle \sigma(s)\hat{h}(s')\rangle_i \tag{28}$$

$$\text{for all } s,s': \qquad \hat{q}(s,s') = \left.\frac{\mathrm{i}\partial\Omega}{\partial q(s,s')}\right|_{\text{saddle}} \tag{29}$$

$$\text{for all } s,s': \qquad \hat{Q}(s,s') = \left.\frac{\mathrm{i}\partial\Omega}{\partial Q(s,s')}\right|_{\text{saddle}} \tag{30}$$

$$\text{for all } s,s': \qquad \hat{K}(s,s') = \left.\frac{\mathrm{i}\partial\Omega}{\partial K(s,s')}\right|_{\text{saddle}}. \tag{31}$$

The effective single-site measure (16) simplifies considerably due to (25), (26), and since the function Ω depends on the trio $\{\boldsymbol{q},\boldsymbol{Q},\boldsymbol{K}\}$ only (see (15)), our saddle-point equations can be reduced to a problem involving only the key physical observables $m(s)$, $C(s,s')$ and $G(s,s')$.

In order to calculate the remaining Gaussian integral in Ω we have to define matrices operating in the product space of vectors labelled by both time indices s and pattern indices μ. Note: in the case of the standard symmetric Hopfield model, where integration variables with different pattern labels μ immediately decouple, this would not have been necessary. We define a matrix $\boldsymbol{\Gamma} = \boldsymbol{S}\otimes\boldsymbol{R}$ as having matrix elements $\Gamma_{\mu\mu'}(s,s') = S_{\mu\mu'}R(s,s')$, where $\mu,\mu' = 1,\ldots,p$ and where $s,s' = 0,\ldots,t-1$. It will operate as follows: if $\boldsymbol{y} = \boldsymbol{\Gamma}\boldsymbol{x}$ then $y_\mu(s) = \sum_{\mu'>t}\sum_{s'<t} S_{\mu\mu'}R(s,s')x_{\mu'}(s')$ for each (μ,s). Note that in evaluating Ω for $N\to\infty$, and due to t remaining finite, we can safely drop the restriction that $\mu > t$, and instead have $\mu = 1,\ldots,p$. The above definition allows us to write

$$\begin{aligned}
\lim_{N\to\infty}\Omega[\boldsymbol{C},\boldsymbol{Q},\mathrm{i}\boldsymbol{G}] &= \lim_{N\to\infty}\frac{1}{N}\ln\int\frac{\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{v}}{(2\pi)^{pt}}\mathrm{e}^{-\frac{1}{2}\boldsymbol{u}\cdot[\mathbf{1}\otimes\boldsymbol{Q}]\boldsymbol{u}-\frac{1}{2}\boldsymbol{v}\cdot[\mathbf{1}\otimes\boldsymbol{C}]\boldsymbol{v}+\mathrm{i}\boldsymbol{v}\cdot[\boldsymbol{S}\otimes\mathbf{1}-\boldsymbol{G}\otimes\mathbf{1}]\boldsymbol{u}} \\
&= \lim_{N\to\infty}\frac{1}{N}\ln\left\{\det{}^{-\frac{1}{2}}[\mathbf{1}\otimes\boldsymbol{C}]\int\frac{\mathrm{d}\boldsymbol{u}}{(2\pi)^{pt/2}}\right. \\
&\quad\left.\times\mathrm{e}^{-\frac{1}{2}\boldsymbol{u}\cdot\{\mathbf{1}\otimes\boldsymbol{Q}+[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]^\dagger[\mathbf{1}\otimes\boldsymbol{C}]^{-1}[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]\}\boldsymbol{u}}\right\} \\
&= -\lim_{N\to\infty}\frac{1}{2N}\{\ln\det[\mathbf{1}\otimes\boldsymbol{C}] \\
&\quad+\ln\det\{\mathbf{1}\otimes\boldsymbol{Q}+[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]^\dagger[\mathbf{1}\otimes\boldsymbol{C}]^{-1}[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]\}\}.
\end{aligned}\tag{32}$$

We use (32) to work out the saddle-point equations (29)–(31). The first of the three equations comes out trivially:

$$\begin{aligned}
\hat{q}(s,s') &= \mathrm{i}\frac{\partial}{\partial C(s,s')}\Omega[\boldsymbol{C},\boldsymbol{0},\mathrm{i}\boldsymbol{G}] \\
&= -\lim_{N\to\infty}\frac{\mathrm{i}}{2N}\frac{\partial}{\partial C(s,s')}\ln\det\{[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]^\dagger[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]\} = 0.
\end{aligned}\tag{33}$$

In order to work out the remaining two equations we use the general matrix identity $\ln\det[\boldsymbol{M}+\boldsymbol{Q}] = \ln\det\boldsymbol{M} + \mathrm{Tr}[\boldsymbol{M}^{-1}\boldsymbol{Q}] + \mathcal{O}(\boldsymbol{Q}^2)$, as well as the specific properties of the $p\times p$ matrix $S_{\mu\nu} = \delta_{\mu,\nu+1}$. In particular we will be using its unitarity, $\boldsymbol{S}^\dagger\boldsymbol{S} = \mathbf{1}$, the identity $[(\boldsymbol{S}^\dagger)^m\boldsymbol{S}^n]_{\mu\mu} = \delta_{mn}$, and its p eigenvalues s_μ being given by $s_\mu = \mathrm{e}^{-2\pi\mathrm{i}\mu/p}$. Equation (30) now reduces to

$$\hat{Q}(s,s') = \mathrm{i}\lim_{\boldsymbol{Q}\to\boldsymbol{0}}\frac{\partial}{\partial Q(s,s')}\Omega[\boldsymbol{C},\boldsymbol{Q},\mathrm{i}\boldsymbol{G}]$$

$$= -\lim_{N\to\infty}\frac{\mathrm{i}}{2N}\sum_{\mu\leqslant p}\{[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]^{\dagger}[\mathbf{1}\otimes\boldsymbol{C}]^{-1}[\boldsymbol{S}\otimes\mathbf{1}-\mathbf{1}\otimes\boldsymbol{G}]\}^{-1}_{\mu\mu}(s',s)$$

$$= -\lim_{N\to\infty}\frac{\mathrm{i}}{2N}\sum_{\mu\leqslant p}\{[\mathbf{1}\otimes\mathbf{1}-\boldsymbol{S}^{\dagger}\otimes\boldsymbol{G}]^{-1}[\mathbf{1}\otimes\boldsymbol{C}][\mathbf{1}\otimes\mathbf{1}-\boldsymbol{S}\otimes\boldsymbol{G}^{\dagger}]^{-1}\}_{\mu\mu}(s',s)$$

$$= -\frac{1}{2}\alpha\mathrm{i}\sum_{n,m\geqslant 0}\lim_{p\to\infty}\frac{1}{p}\sum_{\mu\leqslant p}\{[\boldsymbol{S}^{\dagger}\otimes\boldsymbol{G}]^{n}[\mathbf{1}\otimes\boldsymbol{C}][\boldsymbol{S}\otimes\boldsymbol{G}^{\dagger}]^{m}\}_{\mu\mu}(s',s)$$

giving:

$$\hat{\boldsymbol{Q}} = -\tfrac{1}{2}\alpha\mathrm{i}\sum_{n\geqslant 0}(\boldsymbol{G}^{\dagger})^{n}\boldsymbol{C}(\boldsymbol{G})^{n}. \tag{34}$$

Finally we turn to equation (31):

$$\begin{aligned}\hat{K}(s,s') &= \frac{\partial}{\partial G(s,s')}\Omega[\boldsymbol{C},\mathbf{0},\mathrm{i}\boldsymbol{G}]\\ &= -\frac{1}{2}\alpha\frac{\partial}{\partial G(s,s')}\lim_{p\to\infty}\frac{1}{p}\{\ln\det[\mathbf{1}\otimes\mathbf{1}-\boldsymbol{S}^{\dagger}\otimes\boldsymbol{G}]^{\dagger}+\ln\det[\mathbf{1}\otimes\mathbf{1}-\boldsymbol{S}^{\dagger}\otimes\boldsymbol{G}]\}\\ &= -\frac{1}{2}\alpha\frac{\partial}{\partial G(s,s')}\lim_{p\to\infty}\frac{1}{p}\sum_{\mu\leqslant p}\{\ln\det[\mathbf{1}-\mathrm{e}^{-2\pi\mathrm{i}\mu/p}\boldsymbol{G}^{\dagger}]+\ln\det[\mathbf{1}-\mathrm{e}^{2\pi\mathrm{i}\mu/p}\boldsymbol{G}]\}\\ &= -\frac{1}{2}\alpha\frac{\partial}{\partial G(s,s')}\operatorname{Tr}\sum_{n>0}\frac{1}{n}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\{\mathrm{e}^{-n\mathrm{i}\omega}(\boldsymbol{G}^{\dagger})^{n}+\mathrm{e}^{n\mathrm{i}\omega}(\boldsymbol{G})^{n}\}=0.\end{aligned} \tag{35}$$

5. The effective single-spin problem

Let us summarize the present stage of our calculation. Most macroscopic integration variables are found to vanish in the relevant physical saddle point: $k(s)=\hat{m}(s)=Q(s,s')=\hat{q}(s,s')=\hat{K}(s,s')=0$. The remaining ones can all be expressed in terms of three macroscopic observables, namely the overlaps $m(s)$ and the single-site correlation and response functions $C(s,s')$ and $G(s,s')$, as defined in (22)–(24), by using the four equations (26)–(34). We are thus left with a set of closed equations (26)–(28) from which to solve $\{m(s),C(s,s'),G(s,s')\}$. These equations are defined in terms of an effective single-spin problem. At this stage it is natural to choose the remaining external fields $\theta_i(s)$ to be so-called 'staggered' ones, i.e. $\theta_i(s)=\theta(s)\xi_i^{s+1}$. If used as symmetry-breaking perturbations, such fields will exactly single out macroscopic solutions of the type we introduced as an ansatz. This choice also removes the formal need to break symmetries via initial conditions, so that we may now choose $p_i(\sigma(0))=p(\sigma(0))$. As a consequence we find that the single-site measure (16) becomes site independent, since the remaining site dependence due to pattern components ξ_i^{μ} can be eliminated via a gauge transformation whereby $\sigma(s)\to\sigma(s)\xi_i^{s}$ and $h(s)\to h(s)\xi_i^{s+1}$. The resulting single-spin problem involves the following measure (which is properly normalized, as can be verified by explicit evaluation of $\langle 1\rangle_{\star}$):

$$\begin{aligned}\langle f[\{\sigma\}]\rangle_{\star} = \sum_{\sigma(0)\ldots\sigma(t)}\int\{\mathrm{d}h\,\mathrm{d}\hat{h}\}p(\sigma(0))f[\{\sigma\}]\mathrm{e}^{\sum_{s<t}[\beta\sigma(s+1)h(s)-\ln 2\cosh[\beta h(s)]]}\\ \times\mathrm{e}^{\mathrm{i}\sum_{s<t}\hat{h}(s)[h(s)-\theta(s)-m(s)]-\frac{1}{2}\alpha\sum_{s,s'<t}R(s,s')\hat{h}(s)\hat{h}(s')}\end{aligned} \tag{36}$$

with $R(s,s')=\sum_{n\geqslant 0}[(\boldsymbol{G}^{\dagger})^{n}\boldsymbol{C}(\boldsymbol{G})^{n}](s,s')$. This measure describes an effective single spin $\sigma(s)$ with a stochastic alignment to local fields given by $h(s)=m(s)+\theta(s)+\alpha^{\frac{1}{2}}\phi(s)$,

in which the term $\phi(s)$ represents a zero-average Gaussian random field with (non-zero) temporal correlations $\langle\phi(s)\phi(s')\rangle = R(s,s')$. Note that, as a consequence of (35), there is no term representing a retarded self-interaction, in contrast with the standard (symmetric) Hopfield model. This is the mathematical explanation of the higher storage capacity in the present sequence processing model. The asymmetry of the interaction matrix prevents the build-up of a microscopic memory, similar to the situation in the non-symmetric SK-model [12–14]. The equations from which to solve our remaining order parameters can be written as

$$m(s) = \langle\sigma(s)\rangle_\star \tag{37}$$

$$C(s,s') = \langle\sigma(s)\sigma(s')\rangle_\star \tag{38}$$

$$G(s,s') = \frac{\partial}{\partial\theta(s')}\langle\sigma(s)\rangle_\star. \tag{39}$$

Since the measure (36) factorizes with respect to spin variables at different times, we can immediately perform the spin summations in (37)–(39) (which would not have been possible for the standard Hopfield model). After a simple rescaling of fields and conjugate fields we then arrive at

$$m(s) = \int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}}\tanh\beta[m(s-1)+\theta(s-1)+\alpha^{\frac{1}{2}}v(s-1)] \tag{40}$$

$$\begin{aligned}C(s,s') = \delta_{s,s'} + [1-\delta_{s,s'}]\int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}}\\ \times\tanh\beta[m(s-1)+\theta(s-1)+\alpha^{\frac{1}{2}}v(s-1)]\\ \times\tanh\beta[m(s'-1)+\theta(s'-1)+\alpha^{\frac{1}{2}}v(s'-1)]\end{aligned} \tag{41}$$

$$\begin{aligned}G(s,s') = \beta\delta_{s,s'+1}\Bigg\{1-\int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}}\\ \times\tanh^2\beta[m(s-1)+\theta(s-1)+\alpha^{\frac{1}{2}}v(s-1)]\Bigg\}\end{aligned} \tag{42}$$

with $R(s,s') = \sum_{n\geqslant 0}[(\boldsymbol{G}^\dagger)^n\boldsymbol{C}(\boldsymbol{G})^n](s,s')$. The response function is found to be non-zero only if field perturbation and spin measurement are temporally separated by exactly one iteration step. Thus anomalous response cannot occur, and macroscopic stationarity should be achieved on finite timescales.

6. The stationary state

We now choose stationary external fields $\theta_i(s) = \theta\xi_i^{s+1}$, giving $\theta(s) = \theta$ in terms of the single-spin problem, and inspect time-translation invariant solutions of our macroscopic equations (40)–(42), which will describe motion on a macroscopic limit cycle:

$$m(s) = m \qquad C(s,s') = C(s-s') \qquad G(s,s') = G(s-s'). \tag{43}$$

In order to do this we shift the initial time in (40)–(42) from $t_0 = 0$ to $t_0 = -\infty$, and the final time to $t = \infty$. According to (43) the matrices $\boldsymbol{C}$ and $\boldsymbol{G}$ become Toeplitz matrices and commute, which implies that the matrix $\boldsymbol{R}$ simplifies to

$$\boldsymbol{R} = \boldsymbol{C}[\boldsymbol{1}-\boldsymbol{G}^\dagger\boldsymbol{G}]^{-1} \qquad R(s,s') = R(s-s') \tag{44}$$

and that we may thus write the stationary version of (40)–(42) as

$$m = \int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}}\tanh\beta[m+\theta+\alpha^{\frac{1}{2}}v(0)]$$

$$C(\tau \neq 0) = \int \{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}}\tanh\beta[m+\theta+\alpha^{\frac{1}{2}}v(\tau)]\tanh\beta[m+\theta+\alpha^{\frac{1}{2}}v(0)]$$

$$G(\tau) = \beta\delta_{\tau,1}\left\{1-\int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}}\tanh^2\beta[m+\theta+\alpha^{\frac{1}{2}}v(0)]\right\}.$$

We separate in $C(\tau)$ and $R(\tau)$ the persistent from the non-persistent parts, i.e.

$$C(\tau) = q+\tilde{C}(\tau) \qquad R(\tau) = r+\tilde{R}(\tau) \qquad \lim_{\tau\to\pm\infty}\tilde{C}(\tau) = \lim_{\tau\to\pm\infty}\tilde{R}(\tau) = 0.$$

The persistent part r of $R(\tau)$ can be expressed in terms of the persistent part q of $C(\tau)$, by combining (44) with the above expression for $G(\tau)$. This separation of persistent parts induces a frozen random field into the above order parameter equations, which can subsequently be absorbed into the local fields:

$$\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\boldsymbol{R}\boldsymbol{w}} = \mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}r[\sum_s w(s)]^2-\frac{1}{2}\boldsymbol{w}\cdot\tilde{\boldsymbol{R}}\boldsymbol{w}} = \int \mathrm{D}z\,\mathrm{e}^{\mathrm{i}\sum_s w(s)[v(s)-z\sqrt{r}]-\frac{1}{2}\boldsymbol{w}\cdot\tilde{\boldsymbol{R}}\boldsymbol{w}}$$

(with the familiar abbreviation $\mathrm{D}z = (2\pi)^{-\frac{1}{2}}\mathrm{e}^{-\frac{1}{2}z^2}$). Upon rewriting $G(\tau) = \beta\delta_{\tau,1}[1-\tilde{q}]$ and $r = q\rho$, we arrive at the following expressions for our persistent observables:

$$m = \int \mathrm{D}z\int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\tilde{\boldsymbol{R}}\boldsymbol{w}}\tanh\beta[m+\theta+z\sqrt{\alpha q\rho}+\alpha^{\frac{1}{2}}v(0)]$$

$$q = \lim_{\tau\to\infty}\int \mathrm{D}z\int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\tilde{\boldsymbol{R}}\boldsymbol{w}}\tanh\beta[m+\theta+z\sqrt{\alpha q\rho}+\alpha^{\frac{1}{2}}v(\tau)]$$
$$\times\tanh\beta[m+\theta+z\sqrt{\alpha q\rho}+\alpha^{\frac{1}{2}}v(0)]$$

$$\tilde{q} = \int \mathrm{D}z\int\{\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{w}\}\mathrm{e}^{\mathrm{i}\boldsymbol{v}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{w}\cdot\tilde{\boldsymbol{R}}\boldsymbol{w}}\tanh^2\beta[m+\theta+z\sqrt{\alpha q\rho}+\alpha^{\frac{1}{2}}v(0)]$$

$$\rho = [1-\beta^2(1-\tilde{q})^2]^{-1}.$$

We only need to know the joint probability distribution of the pair $(v(\tau), v(0))$ in the limit $\tau\to\infty$ to work out the remaining integrals over $\boldsymbol{v}$ and $\boldsymbol{w}$. This distribution is clearly a zero-average Gaussian one, so finding the second-order moments suffices. Integration over $\boldsymbol{w}$ gives

$$\langle v(\tau)v(0)\rangle = [\det\tilde{\boldsymbol{R}}]^{-\frac{1}{2}}\int\prod_s\left[\frac{\mathrm{d}v(s)}{\sqrt{2\pi}}\right]\mathrm{e}^{-\frac{1}{2}\boldsymbol{v}\cdot\tilde{\boldsymbol{R}}^{-1}\boldsymbol{v}}v(\tau)v(0) = \tilde{R}(\tau)$$

from which we conclude that $\langle v(0)^2\rangle = \tilde{R}(0)$, and that $\lim_{\tau\to\infty}\langle v(\tau)v(0)\rangle = 0$. The variance $\tilde{R}(0) = R(0)-r$ immediately follows from (44):

$$\tilde{R}(0) = \frac{1-q}{1-\beta^2(1-\tilde{q})^2} = (1-q)\rho.$$

All remaining integrals are now expressed in terms of persistent observables only:

$$m = \int \mathrm{D}z\int \mathrm{D}x\,\tanh\beta\left[m+\theta+z\sqrt{\alpha q\rho}+x\sqrt{\alpha(1-q)\rho}\right]$$

$$q = \int \mathrm{D}z\left[\int \mathrm{D}x\,\tanh\beta\left[m+\theta+z\sqrt{\alpha q\rho}+x\sqrt{\alpha(1-q)\rho}\right]\right]^2$$

$$\tilde{q} = \int \mathrm{D}z\int \mathrm{D}x\,\tanh^2\beta\left[m+\theta+z\sqrt{\alpha q\rho}+x\sqrt{\alpha(1-q)\rho}\right].$$

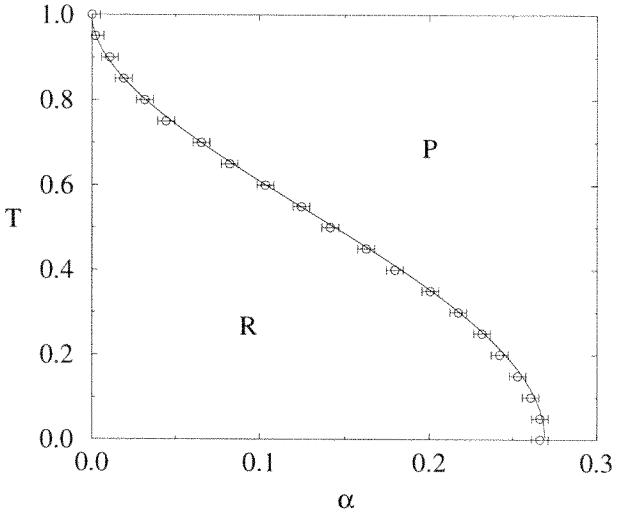

Figure 1. Phase diagram of the sequence processing network, in which one finds two phases: a recall phase (R), characterized by $\{m \neq 0, q > 0, \tilde{q} > 0\}$, and a paramagnetic phase (P), characterized by $\{m = 0, q = 0, \tilde{q} > 0\}$. Full curve: the theoretical prediction for the phase boundary. Markers: simulation results for systems of $N = 10\,000$ spins measured after 2500 iteration steps. The precision in terms of α is at least $\Delta\alpha = 0.005$ (indicated by error bars); the values for T are exact.

If we finally combine the two Gaussian variables in the equations for m and $\tilde{q}$ into a single Gaussian variable we arrive at our final result:

$$\rho = [1 - \beta^2(1 - \tilde{q})^2]^{-1} \tag{45}$$

$$m = \int \mathrm{D}z \, \tanh\beta\left[m + \theta + z\sqrt{\alpha\rho}\right] \tag{46}$$

$$\tilde{q} = \int \mathrm{D}z \, \tanh^2\beta\left[m + \theta + z\sqrt{\alpha\rho}\right] \tag{47}$$

$$q = \int \mathrm{D}z\left[\int \mathrm{D}x \, \tanh\beta\left[m + \theta + z\sqrt{\alpha q\rho} + x\sqrt{\alpha(1-q)\rho}\right]\right]^2. \tag{48}$$

Note that the trio (45)–(47) form itself a closed set, from the solution of which the persistent correlation q simply follows.

7. Phase diagram and storage capacity

We have solved the coupled equations (45)–(47) numerically for $\theta = 0$† in order to determine the region in the α–T plane where solutions with $m \neq 0$, which describe pattern sequence recall, exist‡. The boundary of this region determines the storage capacity of the system. This theoretical result was tested against numerical simulations of the present model, carried out at the spin level (2). We show the results in figure 1. One finds that, for $T > 0$ and

† The alternative choice $\theta \neq 0$ would have described the less interesting scenario where the $m \neq 0$ state would not be sustained autonomously (if at all), but where at each timestep and at each site a very specific external field $\theta_i(s) = \theta\xi_i^{s+1}$ would have actively pushed the system towards the pattern sequence.

‡ With the proviso that the present (standard) version of dynamic field theory applies only on timescales which are small compared with the system size N.

$\alpha < \infty$, the equations (45)–(47) admit only two types of solutions: a recall solution (R) characterized by $\{m \neq 0, q > 0, \tilde{q} > 0\}$, and a paramagnetic solution (P) characterized by $\{m = 0, q = 0, \tilde{q} > 0\}$. The absence of the analogon of a spin-glass phase will be discussed in more detail below. The (first order) phase boundary R $\to$ P as obtained theoretically (solid line) shows an excellent agreement with the computer simulations (markers), as performed for systems of size $N = 10\,000$ (using bisection). The maximum storage capacity α_c is obtained in the zero noise limit $T \to 0$ (or $\beta \to \infty$). For $\beta \to \infty$, where $\tilde{q} \to 1$ and $q \to 1$, the saddle-point equations can be simplified in the usual manner, using identities such as

$$\lim_{\beta\to\infty} \int \mathrm{D}z \, \tanh\beta\left[m + z\sqrt{\alpha\rho}\right] = \mathrm{erf}\left[\frac{m}{\sqrt{2\alpha\rho}}\right]$$

$$\lim_{\beta\to\infty} \beta(1-\tilde{q}) = \frac{\partial}{\partial m} \lim_{\beta\to\infty} \int \mathrm{D}z \, \tanh\beta\left[m + z\sqrt{\alpha\rho}\right] = \sqrt{\frac{2}{\pi\alpha\rho}} \exp\left[-\frac{m^2}{2\alpha\rho}\right].$$

With the definition $x = m/\sqrt{2\alpha\rho}$, from which the overlap m follows according to $m = \mathrm{erf}(x)$, we can combine our saddle-point equations for $\beta \to \infty$ into the single transcendental equation

$$x\sqrt{2\alpha} = \pm\sqrt{\mathrm{erf}^2(x) - \frac{4x^2}{\pi}\exp(-2x^2)}. \tag{49}$$

This equation is identical to that obtained in the $T = 0$ limit for the layered model of [19], and for the present model we thus obtain the same maximum storage capacity, which is defined as the largest value of α for which (49) has non-trivial solutions, of $\alpha_c \approx 0.269$. Note, however, that this equivalence does not extend beyond the $T = 0$ limit. To also verify this latter result with numerical simulations, taking into account the possibility of finite size effects, we measured the maximum storage capacity in zero temperature simulations for different system sizes, ranging from $N = 2500$ to $N = 50\,000$. This resulted in figure 2. The numerical data are again perfectly consistent with the result $\alpha_c \approx 0.269$ of our $N = \infty$ theory.

Finally we turn to the non-recall phases, still for zero external field, where $m = 0$ and where the remaining order parameters $q \in [0, 1]$ and $\tilde{q} \in [0, 1]$ follow from solving the coupled equations

$$\tilde{q} = \int \mathrm{D}z \, \tanh^2\left[\beta z\sqrt{\frac{\alpha}{1-\beta^2(1-\tilde{q})^2}}\right] \tag{50}$$

$$q = \int \mathrm{D}z \left\{\int \mathrm{D}x \, \tanh\beta\left[z\sqrt{\frac{\alpha q}{1-\beta^2(1-\tilde{q})^2}} + x\sqrt{\frac{\alpha(1-q)}{1-\beta^2(1-\tilde{q})^2}}\right]\right\}^2. \tag{51}$$

The first of these equations (50) determines $\tilde{q}$, which is related to the response function via $G(\tau) = \beta(1-\tilde{q})\delta_{\tau,1}$. Its solution is unique. For finite temperature one finds that $\tilde{q}$ is always non-zero, approaching zero only asymptotically as $\tilde{q} = \alpha\beta^2 + \mathcal{O}\beta^4$ for $T \to \infty$. The persistent correlation q subsequently follows from solving (51). This second equation always admits the paramagnetic solution $q = 0$. Careful numerical and analytical inspection reveals that for $T > 0$ and $\alpha < \infty$ it admits no solutions with $q > 0$, which would have been the analogon of a spin-glass state. Only in the limits $T \to 0$ and $\alpha \to \infty$, where $\beta(1-\tilde{q}) \to (1+\frac{1}{2}\pi\alpha)^{-\frac{1}{2}}$ and $\beta(1-\tilde{q}) \to 1$ respectively, and where equation (51) converts into

$$q = \int \mathrm{D}z \, \mathrm{erf}^2\left[\frac{z\sqrt{q}}{\sqrt{2(1-q)}}\right]$$

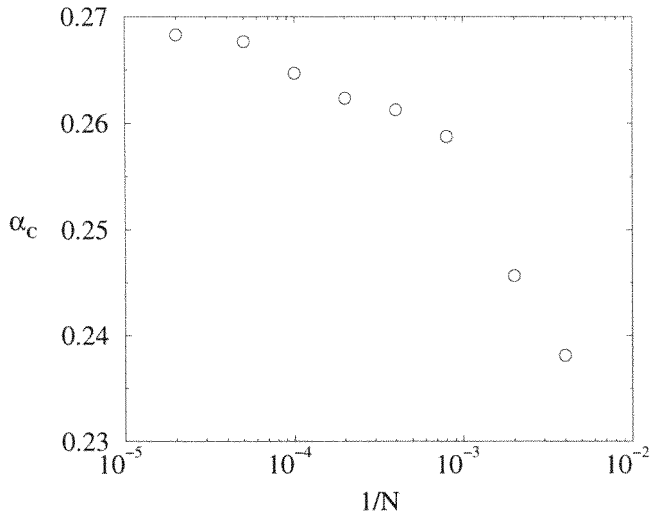

Figure 2. Results of determining the maximum sequence storage capacity α_c at $T = 0$ via numerical simulation of networks with different sizes N. The values for α_c have been determined with a precision of at least $\Delta\alpha = 0.001$ where possible. Note that the $N = \infty$ theory predicts $\alpha_c \approx 0.269$.

does one find a non-trivial solution, namely $q = 1$. This implies that in the phase diagram of figure 1 the phase beyond the boundary of the recall region is a paramagnetic state, with only a transition to a spin-glass type frozen state precisely at $T = 0$. This type of behaviour is very similar to that observed in non-symmetric spin-glass models [12–14].

8. Discussion

In this paper we have used path integral methods to solve in the thermodynamic limit the dynamics of a non-symmetric neural network model, designed to store and recall sequences of stored patterns, close to saturation. For about a decade this model has been known from numerical simulations to have a significantly enlarged storage capacity (by about a factor two) compared with the more familiar symmetric Hopfield network [1–3], which stores static patterns and obeys detailed balance. So far the sequence processing model had not yet been solved, and thus the enlarged storage capacity had not yet been explained, mainly due to the complication that the absence of detailed balance rules out the more traditional equilibrium statistical mechanical methods, including replica theory. In contrast, even in the regime of interest where the number of patterns in the sequence scales as $p = \alpha N$, and thus where the dynamical methods of simple mean-field models cannot be used, the powerful path integral methods of [4, 11, 5] do still apply; they allow us perform the disorder average in a dynamical framework, and thereby to calculate the system's phase diagram without having to resort to additional approximations.

In the standard (symmetric) Hopfield network two effects limit the storage capacity: a Gaussian noise in the equivalent effective single-spin problem, which is non-local in time, and a retarded self-interaction. The magnitude of both contributions depends on the load factor α. For the present model we find, in contrast, that the retarded self-interaction vanishes, which explains the extended storage capacity. Numerical simulations for large system sizes (up to $N = 50\,000$ spins) are in excellent agreement with our analytical results,

both with respect to the maximum storage capacity $\alpha_c \approx 0.269$ (at zero noise level) and with respect to the full phase diagram in the $\alpha - T$ plane. In the limit of zero noise level we find that the equation from which to solve the order parameter which describes the quality of the sequence recall reduces to that of the layered model of [19]. Our order parameter equations and their solutions also turn out to be very similar to those found for various versions of the non-symmetric SK spin-glass model, as studied in [12–14]. In particular, common features are the absence of a retarded self-interaction in the effective single-spin problem, and the absence of a spin-glass-type phase for non-zero temperatures.

We have restricted ourselves to calculating limit cycle solutions of the macroscopic dynamical equations which describe pure sequence reproduction, in view of our main goal: finding the maximum storage capacity α_c. Clearly there will be many alternative solutions, e.g. those describing motion towards a limit cycle (transients), and so-called non-equilibrium mixture state solutions (generated upon choosing non-staggered symmetry breaking fields) whereby at any time the microscopic state vector has a non-zero overlap with two or more of the stored patterns. A full classification of all possible solutions of the macroscopic dynamical equations is, if at all possible, beyond the scope of this study.

As a next step one could apply the present formalism to networks which store more than one periodic pattern sequence. By varying the scaling with N of both the sequence length and of the number of sequences, one should expect a transition between the behaviour similar to that of the symmetric Hopfield model (with an effective retarded self-interaction) and the behaviour observed in the present model (without such a retarded self-interaction). This will be the subject of a future study.

Acknowledgments

AD would like to thank the Rhodes Trust for a Scholarship. ACCC and DS would like to thank Heinz Horner for valuable discussions on the dynamics of spin glasses and neural networks, including the use of generating functional analysis to study them.

References

[1] Hopfield J 1982 *Proc. Natl Acad. Sci.* **79** 2554
[2] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530
[3] Amit D J, Gutfreund H and Sompolinsky H 1987 *Ann. Phys.* **173** 30
[4] de Dominicis G 1978 *Phys. Rev.* B **18** 4913
[5] Rieger H, Schreckenberg M and Zittartz J 1988 *J. Phys. A: Math. Gen.* **21** L263
[6] Sompolinsky H and Kanter I 1986 *Phys. Rev. Lett.* **57** 2861
[7] Amari S 1988 *Associative Memory and Its Statistical Neurodynamical Analysis* (Berlin: Springer) p 85
[8] Bauer U and Krey U 1990 *Z. Phys.* B **79** 461
[9] Kühn R and van Hemmen J L 1991 *Models of Neural Networks I* ed E Domany, J L van Hemmen and K Schulten (Berlin: Springer) p 213
[10] Coolen A C C and Sherrington D 1992 *J. Phys. A: Math. Gen.* **25** 5493
[11] Sommers H J 1987 *Phys. Rev. Lett.* **58** 1268
[12] Crisanti A and Sompolinsky H 1987 *Phys. Rev.* A **36** 4922
[13] Crisanti A and Sompolinsky H 1987 *Phys. Rev.* A **37** 4865
[14] Rieger H, Schreckenberg M and Zittartz J 1989 *Z. Phys.* B **74** 527
[15] Amit D J 1989 *Modeling Brain Function* (Cambridge: Cambridge University Press)
[16] Hertz J, Krogh A and Palmer R G 1991 *Introduction to the Theory of Neural Computation* (Reading, MA: Addison-Wesley)
[17] Peretto P 1992 *An Introduction to the Modeling of Neural Networks* (Cambridge: Cambridge University Press)

[18] Sherrington D 1993 *Mathematical Approaches to Neural Networks* ed J G Taylor (Amsterdam: Elsevier Science) p 261
[19] Domany E, Kinzel W and Meir R 1989 *J. Phys. A: Math. Gen.* **22** 2081