

Analytic solution of attractor neural networks on scale-free graphs

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2004 J. Phys. A: Math. Gen. 37 8789

(<http://iopscience.iop.org/0305-4470/37/37/002>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.64

The article was downloaded on 02/06/2010 at 19:06

Please note that [terms and conditions apply](#).

Analytic solution of attractor neural networks on scale-free graphs

I Pérez Castillo¹, B Wemmenhove², J P L Hatchett³, A C C Coolen³,
N S Skantzos⁴ and T Nikolettopoulos³

¹ Institute for Theoretical Physics, Celestijnenlaan 200D, Katholieke Universiteit Leuven, B-3001, Belgium

² Institute for Theoretical Physics, University of Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands

³ Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK

⁴ Departament de Física Fonamental, Facultat de Física, Universitat de Barcelona, 08028 Barcelona, Spain

E-mail: wemmenho@science.uva.nl, hatchett@math.kcl.ac.uk, isaac.perez@fys.kuleuven.ac.be, tcoolen@math.kcl.ac.uk, nikos@ffn.ub.es and theodore@math.kcl.ac.uk

Received 2 April 2004

Published 1 September 2004

Online at stacks.iop.org/JPhysA/37/8789

doi:10.1088/0305-4470/37/37/002

Abstract

We study the influence of network topology on retrieval properties of recurrent neural networks, using replica techniques for dilute systems. The theory is presented for a network with an arbitrary degree distribution $p(k)$ and applied to power-law distributions $p(k) \sim k^{-\gamma}$, i.e. to neural networks on scale-free graphs. A bifurcation analysis identifies phase boundaries between the paramagnetic phase and either a retrieval phase or a spin-glass phase. Using a population dynamics algorithm, the retrieval overlap and spin-glass order parameters may be calculated throughout the phase diagram. It is shown that there is an enhancement of the retrieval properties compared with a Poissonian random graph. We compare our findings with simulations.

PACS numbers: 75.10.Nr, 05.20.-y, 64.60.Cn

1. Introduction

The impressive ability of human and animal brains to recognize and manipulate complex patterns under real-world (i.e. noisy and often conflicting) conditions continues to appeal not only to biologists but also to physicists, computer scientists and engineers, albeit with the latter driven by different objectives and motivations. Hopfield [1] was one of the first to introduce a simple model to describe associative memory in recurrent neural networks successfully, based on the biologically motivated Hebbian rule for adapting the connections between the neurons (the 'synapses'). His model initiated a period of intense research activity. The success of these

early neural network models was mainly due to their analytic tractability, which was achieved upon sacrificing biological realism; all neurological connectivity structures were sacrificed by the first generation of (fully connected) models. However, the research area was thereby able to benefit significantly from recent advances in mean-field spin-glass (SG) theory [2, 3], and many new results were published in the second half of the 1980s; see e.g. [4–6] or [7].

A step towards increased biological realism was made by the introduction of diluted neural network models. Initially, in the thermodynamic limit, each neuron was on average connected to a vanishing fraction of the system, but this fraction contained an infinite number of nodes. These models were solvable by virtue of the specific nature of their architectures: one either chooses strictly symmetric dilution (so that detailed balance and hence equilibrium analysis are preserved, e.g. [9–11]), or strictly asymmetric dilution, which ensures that neuron states are statistically independent on finite times [8] (now the local fields are described by Gaussian distributions, leading to simple dynamic order parameter equations). In the early models, the bond statistics were uniform over the entire network, leading to thin tails in its degree distribution, whereas the connectivity of a real neuron is known to vary strongly within the brain [12]. In response to this, there have been several recent studies of recurrent neural network models with alternative connectivity distributions. Most evolve around numerical simulations of Hopfield-type models on graphs with power-law degree distributions [13–16]. Examples of recent analytic work on recurrent neural networks with finite connectivity can be found in [17, 18]; both deal with Poissonian graphs and apply the equilibrium statistical mechanical techniques of dilute disordered spin systems [19–21].

The objective of this paper is to extend and generalize the solution for finitely connected Poissonian neural networks [17, 18] to recurrent neural networks with arbitrary degree distribution, in the spirit of [22, 23] and within the replica-symmetric (RS) ansatz. We derive analytically phase diagrams for networks with Hebbian synapses and arbitrary degree distributions $p(k)$, and apply population dynamics algorithms to obtain the values of the order parameters in the three phases (namely paramagnetic (P), retrieval (R) and SG). This study thereby establishes a connection between the equilibrium statistical mechanics of neural networks and the theory of so-called ‘complex networks’. In line with biological reality, we find that recurrent neural networks with degree distributions with ‘fat tails’ are indeed superior to those with Poissonian degree distributions, in terms of the size of the recall region in the phase diagram.

2. Model definitions

Our model is a system of N Ising spin neurons $\sigma_i \in \{-1, 1\}$, with $i = 1, \dots, N$. The neurons are located on the nodes of a graph with arbitrary degree distribution $p(k) = N^{-1} \sum_i \delta_{k, k_i}$, where k_i denotes the number of neurons connected to neuron i . This system is assumed to be in thermodynamic equilibrium, described by the Hamiltonian

$$\mathcal{H} = - \sum_{i < j=1}^N \sigma_i J_{ij} \sigma_j - \sum_{\mu=1}^p h^\mu \sum_{i=1}^N \xi_i^\mu \sigma_i. \quad (1)$$

Here the $\{h^\mu\}$ represent generating fields and the vectors $(\xi_1^\mu, \dots, \xi_N^\mu)$ represent stored random N -bit patterns. We will abbreviate the bits to be stored at a given node i as $\xi_i = (\xi_i^1, \dots, \xi_i^p)$. Since the number of connections per neuron is finite, the number of patterns p must be of order $\mathcal{O}(N^0)$. The bonds J_{ij} depend on the patterns via

$$J_{ij} = \frac{c_{ij}}{\langle k \rangle} \phi(\xi_i \cdot \xi_j) \quad (2)$$

with $\xi_i \cdot \xi_j = \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu$ and $\langle k \rangle = \sum_{k \geq 0} p(k)k$. Special cases of interest are $\phi(x) = x$, namely Hebbian bonds, and $\phi(x) = \text{sign}(x)$, namely clipped Hebbian bonds.

The variables $c_{ij} \in \{0, 1\}$ specify our graph microscopically; we extend their definition to all pairs (i, j) by putting $c_{ij} = c_{ji}$ and $c_{ii} = 0$. It is known from complex network theory [24, 25] that a connectivity distribution $p(k)$ alone does not fully specify the statistics of a graph. Here we draw the matrix $\mathbf{c} = \{c_{ij}\}$, which represents quenched disorder for the spin system (1), randomly from the probability distribution

$$\mathcal{P}(\mathbf{c}) = \frac{\left[\prod_{i < j} P(c_{ij}) \delta_{c_{ij}, c_{ji}} \right] \left[\prod_i \delta_{k_i, \sum_{j \neq i} c_{ij}} \right]}{\sum_{\mathbf{c}'} \left[\prod_{i < j} P(c'_{ij}) \delta_{c'_{ij}, c'_{ji}} \right] \left[\prod_i \delta_{k_i, \sum_{j \neq i} c'_{ij}} \right]} \quad (3)$$

with the single-bond probabilities

$$P(c_{ij}) = \frac{\langle k \rangle}{N} \delta_{c_{ij}, 1} + \left(1 - \frac{\langle k \rangle}{N} \right) \delta_{c_{ij}, 0}. \quad (4)$$

These equations imply that the number of connections at each site i is constrained to the value k_i , but apart from this fact the connectivity matrix is maximally random. On average, each connection $c_{ij} = c_{ji}$ has probability $\langle k \rangle / N$ of being equal to one and probability $(1 - \langle k \rangle / N)$ of being equal to zero. After introducing integral representations for the delta-constraints, disorder averages $\langle\langle A(\mathbf{c}) \rangle\rangle_{\mathbf{c}}$ over the ensemble of graphs are given by [22, 23]

$$\langle\langle A(\mathbf{c}) \rangle\rangle_{\mathbf{c}} = \mathcal{N}^{-1} \sum_{\mathbf{c}} \left[\prod_{i < j} P(c_{ij}) \delta_{c_{ij}, c_{ji}} \right] \left[\prod_i \int \frac{d\psi_i}{2\pi} e^{i\psi_i (\sum_j c_{ij} - k_i)} \right] A(\mathbf{c}), \quad (5)$$

$$\mathcal{N} = \sum_{\mathbf{c}} \left[\prod_{i < j} P(c_{ij}) \delta_{c_{ij}, c_{ji}} \right] \left[\prod_i \int \frac{d\psi_i}{2\pi} e^{i\psi_i (\sum_j c_{ij} - k_i)} \right]. \quad (6)$$

Another statistical quantity to characterize graphs, beyond $p(k)$, is the degree–degree correlation: the joint probability $\omega(k_i, k_j)$ that a pair of nodes i and j are connected, and have connectivities k_i and k_j , respectively. For the present ensemble (3) one finds

$$\omega(k_i, k_j) = \frac{p(k_i) p(k_j) k_i k_j}{\langle k \rangle N}. \quad (7)$$

When comparing our theory with simulations in section 5, care has to be taken that the generated graphs are in accordance with (7).

3. Replica calculation of the free energy and order parameters

We calculate the free energy per spin and the relevant order parameters, using the replica techniques as developed for constrained connectivity graphs, along the lines of [22, 23]. Thus the asymptotic free energy per spin $f = -\lim_{N \rightarrow \infty} (\beta N)^{-1} \log \mathcal{Z}$ is written as

$$f = -\lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{Nn\beta} \log \langle\langle \mathcal{Z}^n \rangle\rangle_{\mathbf{c}}, \quad (8)$$

where \mathcal{Z}^n is the usual n -replicated partition function

$$\mathcal{Z}^n = \sum_{\sigma_1} \cdots \sum_{\sigma_N} \exp \left[\beta \sum_{\alpha=1}^n \sum_{i < j} J_{ij} \sigma_i^\alpha \sigma_j^\alpha + \beta \sum_{\mu=1}^P h^\mu \sum_{i=1}^N \sum_{\alpha=1}^n \xi_i^\mu \sigma_i^\alpha \right] \quad (9)$$

and $\sigma_i \equiv (\sigma_i^1, \dots, \sigma_i^n)$ is the n -replicated spin at site i . Upon performing the trace over c_{ij} (i.e. the disorder average) one obtains

$$\begin{aligned} \langle\langle \mathcal{Z}^n \rangle\rangle_c &= \frac{1}{\mathcal{N}} \prod_{i=1}^N \left[\sum_{\{\sigma_i\}} \int \frac{d\psi_i}{2\pi} e^{-i\psi_i k_i} \exp \left[\beta \sum_{\mu=1}^p h^\mu \sum_{i=1}^N \sum_{\alpha=1}^n \xi_i^\mu \sigma_i^\alpha \right] \right. \\ &\quad \left. \times \exp \left[\frac{\langle k \rangle}{2N} \sum_{i,j=1}^N (e^{i(\psi_i + \psi_j) + \frac{\beta}{\langle k \rangle} \phi(\xi_i \cdot \xi_j) \sigma_i \cdot \sigma_j} - 1) + \mathcal{O}(N^0) \right] \right]. \end{aligned} \tag{10}$$

In order to arrive at an effective single-site problem, one needs the concept of sublattices $I_\xi = \{i | \xi_i = \xi\}$ [26]. The sublattices divide the whole network into classes of sites i sharing the same pattern vector ξ_i . They are a necessary ingredient in the present analysis due to the explicit site-dependence of the interaction values through the values of the pattern vectors. This distinguishes neural network models from models with bond-disorder, like spin glasses. One now defines the following order parameter functions:

$$P_\xi(\sigma) = \frac{1}{|I_\xi|} \sum_{i \in I_\xi} e^{i\psi_i} \delta_{\sigma, \sigma_i}. \tag{11}$$

These are reminiscent of the replicated spin probability distributions within sublattices, as in [17], but here include extra phase factors $e^{i\psi_i}$ whose effect is to replace k_i by $k_i - 1$ in expressions of the type (5). After applying saddle-point integration techniques in the subsequent calculation, one finds as a result

$$P_\xi(\sigma) = \frac{1}{|I_\xi|} \sum_{i \in I_\xi} \langle\langle \delta_{\sigma, \sigma_i} e^{i\psi_i} \rangle\rangle_c \tag{12}$$

with $\langle \dots \rangle$ denoting a thermal average over the n -replicated spins. Indeed this involves a phase-factor in an expression of the type (5), such that k_i is replaced by $k_i - 1$ in the disorder average. Thus the physical meaning of the order parameter (11) is the distribution of a replicated *cavity* spin (a spin from which one connection is removed) in sublattice ξ (see [22]). After some straightforward manipulations (factorization over sites, integration over the variables ψ_i , etc, see also [23]) we find $f = \lim_{n \rightarrow 0} \text{extr}_{\{P, \hat{P}\}} f[\{P, \hat{P}\}]$, where

$$\begin{aligned} f[\{P, \hat{P}\}] &= \frac{\langle k \rangle}{\beta n} \sum_{\sigma} \langle \hat{P}_\xi(\sigma) P_\xi(\sigma) \rangle_\xi - \frac{1}{\beta n} \sum_k p_k \left\langle \log \left[\sum_{\sigma} e^{\beta \mathbf{h} \cdot \xi \sum_{\alpha=1}^n \sigma_\alpha \hat{P}_\xi^k(\sigma)} \right] \right\rangle_\xi \\ &\quad - \frac{\langle k \rangle}{2\beta n} \sum_{\sigma, \sigma'} \langle\langle P_\xi(\sigma) P_{\xi'}(\sigma') e^{(\beta/\langle k \rangle) \phi(\xi \cdot \xi') (\sigma \cdot \sigma')} \rangle\rangle_{\xi, \xi'} - \frac{\langle k \rangle}{2\beta n}. \end{aligned} \tag{13}$$

This involves the sublattice averages $\langle f(\xi) \rangle_\xi = \sum_{\xi} p(\xi) f(\xi)$, with $p_\xi = \lim_{N \rightarrow \infty} |I_\xi|/N$. To find the extremum of the variational free energy (13), we vary the latter with respect to $\{P_\xi(\sigma), \hat{P}_\xi(\sigma)\}$, leading to the saddle-point equations

$$\hat{P}_\xi(\sigma) = \sum_{\sigma'} \langle P_{\xi'}(\sigma') e^{(\beta/\langle k \rangle) \phi(\xi \cdot \xi') (\sigma \cdot \sigma')} \rangle_{\xi'}, \tag{14}$$

$$P_\xi(\sigma) = \sum_k \frac{k p_k}{\langle k \rangle} \frac{e^{\beta \mathbf{h} \cdot \xi \sum_{\alpha=1}^n \sigma_\alpha \hat{P}_\xi^{k-1}(\sigma)}}{\sum_{\sigma} e^{\beta \mathbf{h} \cdot \xi \sum_{\alpha=1}^n \sigma_\alpha \hat{P}_\xi^k(\sigma)}}. \tag{15}$$

As usual [20, 21, 27], one adopts the replica symmetric (RS) ansätze

$$P_{\xi}(\sigma) = \int dh W_{\xi}(h) \frac{e^{\beta h \sum_{\alpha=1}^n \sigma_{\alpha}}}{[2 \cosh(\beta h)]^n}, \quad (16)$$

$$\widehat{P}_{\xi}(\sigma) = \int du Q_{\xi}(u) \frac{e^{\beta u \sum_{\alpha=1}^n \sigma_{\alpha}}}{[2 \cosh(\beta u)]^n}, \quad (17)$$

which are invariant under permutations of replica indices. Inserting these into (15), one finds the replica symmetric form of the saddle-point equations

$$Q_{\xi}(u) = \left\langle \int dh W_{\xi'}(h) \delta \left[u - \frac{1}{\beta} \tanh^{-1} \left[\tanh(\beta h) \tanh \left(\frac{\beta \phi(\xi \cdot \xi')}{\langle k \rangle} \right) \right] \right] \right\rangle_{\xi'} \quad (18)$$

$$W_{\xi}(h) = \sum_k \frac{k p_k}{\langle k \rangle} \int \left[\prod_{\ell=1}^{k-1} du_{\ell} Q_{\xi}(u_{\ell}) \right] \delta \left[h - \sum_{\ell=1}^{k-1} u_{\ell} - \sum_{\mu=1}^p h^{\mu} \xi^{\mu} \right], \quad (19)$$

where $W_{\xi}(h)$ and $Q_{\xi}(u)$ are probability distributions for the effective cavity fields and propagated fields (messages) in sublattice ξ , respectively [23]. Note that one could eliminate equation (18) by insertion into (19). Substitution of (16) and (17) into (13) gives us the RS free energy per spin:

$$\begin{aligned} \beta f = \langle k \rangle & \left\langle \int dh du W_{\xi}(h) Q_{\xi}(u) \log [1 + \tanh(\beta u) \tanh(\beta h)] \right\rangle_{\xi} \\ & - \frac{\langle k \rangle}{2} \left\langle \left\langle \int dh dh' W_{\xi}(h) W_{\xi'}(h') \log \left[1 + \tanh(\beta h) \tanh(\beta h') \tanh \left[\frac{\beta \phi(\xi \cdot \xi')}{\langle k \rangle} \right] \right] \right\rangle_{\xi, \xi'} \right. \\ & - \sum_k p_k \left\langle \int \left[\prod_{\ell=1}^k du_{\ell} Q_{\xi}(u_{\ell}) \right] \log \left(\frac{2 \cosh(\beta \sum_{\ell=1}^k u_{\ell} + \beta \sum_{\mu=1}^p h^{\mu} \xi^{\mu})}{\prod_{\ell=1}^k 2 \cosh(\beta u_{\ell})} \right) \right\rangle_{\xi} \\ & - \frac{\langle k \rangle}{2} \left\langle \left\langle \log \cosh \left[\frac{\beta \phi(\xi \cdot \xi')}{\langle k \rangle} \right] \right\rangle_{\xi, \xi'} \right\rangle - \langle k \rangle \log 2. \end{aligned} \quad (20)$$

We may finally use the generating fields h_{μ} to find explicit expressions for the disorder-averaged pattern recall overlaps $m^{\mu} = \lim_{N \rightarrow \infty} N^{-1} \sum_i \xi_i^{\mu} \langle \langle \sigma_i \rangle \rangle_c = -(\partial f / \partial h^{\mu})|_{h=0}$:

$$m^{\mu} = \sum_k p_k \left\langle \xi^{\mu} \int \left[\prod_{\ell=1}^k du_{\ell} Q_{\xi}(u_{\ell}) \right] \tanh \left(\beta \sum_{\ell=1}^k u_{\ell} \right) \right\rangle_{\xi}. \quad (21)$$

In a similar manner, one may derive an expression for the disorder-averaged RS SG order parameter $q = \lim_{N \rightarrow \infty} N^{-1} \sum_i \langle \langle \sigma_i^2 \rangle \rangle_c$, upon adding a term of the form $\lambda \sum_{\alpha < \beta} \sigma_{\alpha} \sigma_{\beta}$ to the replicated Hamiltonian in (9). The result is

$$q = \sum_k p_k \left\langle \int \left[\prod_{\ell=1}^k du_{\ell} Q_{\xi}(u_{\ell}) \right] \tanh^2 \left(\beta \sum_{\ell=1}^k u_{\ell} \right) \right\rangle_{\xi}. \quad (22)$$

These expressions have a transparent interpretation, given that within the cavity formalism the local magnetization at a site i is indeed given by $m_i = \tanh(\beta \sum_{\ell=1}^{k_i} u_{\ell})$.

4. Phase diagram and order parameters

The P phase, where (21) and (22) are zero, has $W_{\xi}(x) = Q_{\xi}(x) = \delta(x)$ for all ξ . The recall phase (R) is defined by $m^{\mu} \neq 0$ for some μ . In the SG phase, $q > 0$ but $m^{\mu} = 0$ for all μ . Transitions away from the P phase are expected to be second order, allowing us to find the $P \rightarrow (R, SG)$ transitions via a simple continuous bifurcation analysis. In contrast, locating the $SG \rightarrow R$ transition requires knowledge of the (nontrivial) functions $W_{\xi}(h)$ and $Q_{\xi}(u)$ in the R or SG regimes; to find these transitions we will solve (18) and (19) numerically using a population dynamics algorithm.

Following [17, 18] we apply a bifurcation analysis to compute the second-order transition lines away from the P phase. By assuming the effective fields to be small, i.e. $\int dh W_{\xi}(h)h^{\ell} = \mathcal{O}(\epsilon^{\ell})$ for all ξ , and expanding equations (18) and (19) up to order ϵ^2 , one finds the following bifurcation conditions for transitions away from P:

$$P \rightarrow R: \int dh W_{\xi}(h)h = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \left\langle \tanh \left[\frac{\beta\phi(\xi \cdot \xi')}{\langle k \rangle} \right] \int dh W_{\xi'}(h)h \right\rangle_{\xi'}, \quad (23)$$

$$P \rightarrow SG: \int dh W_{\xi}(h)h^2 = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \left\langle \tanh^2 \left[\frac{\beta\phi(\xi \cdot \xi')}{\langle k \rangle} \right] \int dh W_{\xi'}(h)h^2 \right\rangle_{\xi'}. \quad (24)$$

For $p_{\xi} = 2^{-p}$ (random patterns) one can find the eigenvalues of the relevant (symmetric) matrices

$$M_{\xi\xi'} = 2^{-p} \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \tanh \left[\frac{\beta\phi(\xi \cdot \xi')}{\langle k \rangle} \right] \quad (25)$$

and

$$Q_{\xi\xi'} = 2^{-p} \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \tanh^2 \left[\frac{\beta\phi(\xi \cdot \xi')}{\langle k \rangle} \right] \quad (26)$$

by exploiting the fact that they only depend on the inner products $\xi \cdot \xi'$ and the eigenvectors are simultaneously the eigenvectors of a set of symmetry operators on the inner product. The conditions for a second order transition become (see [17])

$$P \rightarrow R: \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \frac{2^{-p}}{p} \sum_{n=0}^p \binom{p}{n} (p-2n) \tanh \left[\frac{\beta\phi(p-2n)}{\langle k \rangle} \right] = 1, \quad (27)$$

$$P \rightarrow SG: \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} 2^{-p} \sum_{n=0}^p \binom{p}{n} \tanh^2 \left[\frac{\beta\phi(p-2n)}{\langle k \rangle} \right] = 1. \quad (28)$$

For a Poissonian degree distribution, one has $\langle k^2 \rangle = \langle k \rangle^2 + \langle k \rangle$, and we recover the results in [17]. For non-Poissonian distributions, equation (27) predicts an enlargement of the R phase when the degree distribution has fat tails, e.g. for power-law distributions $p(k) \sim k^{-\gamma}$. As noted by [23], if the second moment of the degree distribution is not finite (e.g. for power-law distributions with $\gamma \leq 3$), so is the left-hand side of (27), implying that there is always a R phase in the thermodynamic limit, for any temperature and any storage ratio. An appropriate rescaling of the temperature $\beta \rightarrow \beta\langle k^2 \rangle / \langle k \rangle^2$ reveals that the R transition comes before the SG transition in this regime (i.e. there is no SG phase), as long as $\langle k \rangle$ does not diverge (as long as, e.g., $\gamma > 2$ for power-law distributions). For a single pattern, the R phase boundary is given,

in accordance with [23], by

$$\beta_c = -\frac{\langle k \rangle}{2} \log \left(1 - \frac{2\langle k \rangle}{\langle k^2 \rangle} \right) \quad (29)$$

(our expression differs slightly, due to our rescaling of the bonds by a factor $\langle k \rangle^{-1}$).

To find our order parameters, we solve equations (18) and (19) numerically, using a population dynamics algorithm [28, 29]. We have two distributions for each of our 2^p sublattices, so the required CPU time grows exponentially with p . However, one may exploit sublattice symmetries, especially when the system is condensed in a single pattern retrieval state [30]. If the first pattern is condensed, numerical and analytical evidence (as yet short of a proof) suggests the solution of (18) and (19) to have $W_{\xi}(h) = W_{\xi^1}(h)$. Moreover, for random patterns one expects the symmetry $W_{\xi^1}(h) = W(\xi^1 h)$. Insertion of this ansatz (and a similar one for Q_{ξ}) into (18) and (19) leads us to

$$W(h) = \sum_k \frac{k p_k}{\langle k \rangle} \int \left[\prod_{\ell=1}^{k-1} du_{\ell} Q(u_{\ell}) \right] \delta \left[h - \sum_{\ell=1}^{k-1} u_{\ell} \right], \quad (30)$$

$$Q(u) = \frac{1}{2^{p-1}} \sum_{n=0}^{p-1} \binom{p-1}{n} \int dh W(h) \delta \left[u - \frac{1}{\beta} \tanh^{-1} \left[\tanh(\beta h) \tanh \left(\frac{\beta \phi(p-2n)}{\langle k \rangle} \right) \right] \right]. \quad (31)$$

Inserting (31) into (30) then gives the relatively simple expression

$$W(h) = \sum_k \frac{k p_k}{\langle k \rangle} \prod_{\ell=1}^{k-1} \left[\frac{1}{2^{p-1}} \sum_{n_{\ell}=0}^{p-1} \binom{p-1}{n_{\ell}} \int dh_{\ell} W(h_{\ell}) \right] \\ \times \delta \left\{ h - \frac{1}{\beta} \sum_{\ell=1}^{k-1} \tanh^{-1} \left[\tanh(\beta h_{\ell}) \tanh \left[\frac{\beta \phi(p-2n_{\ell})}{\langle k \rangle} \right] \right] \right\}. \quad (32)$$

5. Comparison with simulations

In figure 1 we present the resulting RS phase diagrams in the (α, T) -plane, where $\alpha = p/\langle k \rangle$ for Hopfield-type networks with power-law degree distribution $p(k) \sim k^{-\gamma}$. The distributions $p(k)$ are characterized by $\gamma = 3.1$ and 4 respectively, and by a variable K which defines a lower cutoff ($p(k) = 0$ for $k < K$). Notice that the phase diagrams are calculated for low values of the average connectivity $\langle k \rangle$ (see caption to figure 1), even though in biological networks $\langle k \rangle$ is of the order 100–1000. Our choice is made for numerical convenience only, although the differences between small and large values of $\langle k \rangle$ will be only quantitative. At high T one finds the P phase. At sufficiently low T one finds a R phase (small α) or a SG phase (large α). The value $\gamma = 3.1$ is close to the critical value $\gamma_c = 3$ below which there is no P phase, yet here the phase diagram is found to be qualitatively similar to that corresponding to Poissonian graphs [17] (with re-scaled values of T and α). In the $\gamma = 4$ phase diagram we also indicate the location of the R \rightarrow SG transition, resulting from a population dynamics calculation. The P \rightarrow R and P \rightarrow SG boundaries were found by solving the bifurcation equations (27) and (28) numerically. For small values of the order parameters, i.e. large T or α , finding accurate

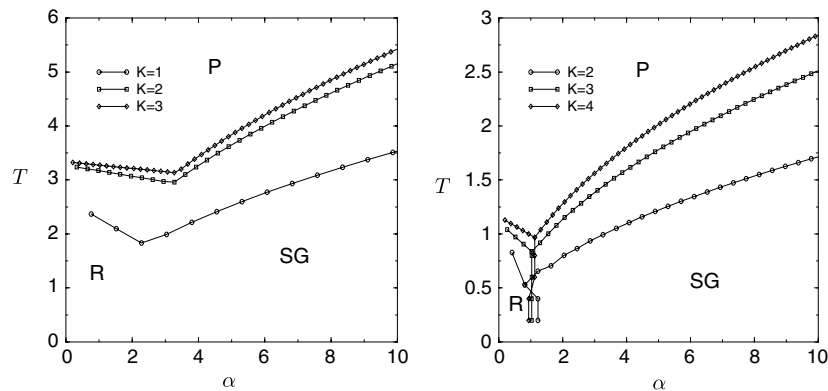


Figure 1. RS phase diagrams for $p(k) \sim k^{-\gamma}$ with $p(k) = 0$ for $k < K$ and $\gamma = 3.1$ (left) and 4 (right). Phase boundaries found by bifurcation analysis, separating P from R and SG phases. For $\gamma = 4$ the $R \rightarrow SG$ phase boundary is calculated for T values which are multiples of 0.2, and is found via population dynamics (with a condensed ansatz). For $\gamma = 3.1$, average connectivity values corresponding to $K = 1$ (circles), 2 (squares) and 3 (diamonds) are $\langle k \rangle = 1.318, 3.055$ and 4.898 respectively. For $\gamma = 4$, average connectivity values corresponding to $K = 2$ (circles), 3 (squares) and 4 (diamonds) are $\langle k \rangle = 2.454, 3.887$ and 5.352 respectively.

numerical values for the order parameters becomes increasingly difficult. Hence, for large fluctuations in the connectivity (i.e. large values of $\langle k^2 \rangle$), one cannot expect accurate results for the $R \rightarrow SG$ transition on the basis of a population dynamics algorithm. This is why we have omitted the $R \rightarrow SG$ line in the $\gamma = 3.1$ phase diagram.

The overall picture supports the findings in numerical studies [13–16], namely that the more fat-tailed the connectivity distribution, the larger the R regime in the phase diagram. The networks are more robust against noise (the critical temperature is enlarged) and a larger number of patterns can in principle be retrieved. The reason, pointed out in the above-mentioned literature, is that the ‘hubs’ in the network (i.e. nodes with a high degree of connectivity) are capable of maintaining order in the network in more noisy circumstances, and induce a nonvanishing contribution to the retrieval order parameter. Fat-tailed distributions (relatively small γ) obviously result in more and larger hubs. The literature however also points out that, although the R regime is enlarged for some scale-free networks, the value of the retrieval overlap is smaller in a large part of the R regime, than in the case of a Poissonian degree distribution.

In figure 2 we show as a function of temperature the recall overlap m^1 (for states where only pattern 1 is condensed) as obtained from a population dynamics calculation, together with the measurements of m^1 in numerical simulations. Both simulation and population dynamics had $N = 10^4$ corresponding to system size and population size, respectively, and all data are averages over 10 runs.

To compare our theory with simulations we must first generate a scale-free network. A convenient and fast way to produce it, is to use the celebrated Barabasi–Albert algorithm for network growth [31, 32], which results in an exponent $\gamma = 3.0$. The left-hand figure refers to $p(k) \sim [k(k+1)(k+2)]^{-1}$, corresponding to the degree distribution resulting from this algorithm. It should be noted, however, that the degree–degree correlation generated by the latter algorithm differs from the one (7) in our present model (see e.g. [24]). Consequently, a different algorithm, similar to the one described in [25], had to be used here, and we could not test our results directly by comparison to other simulation results in the literature. First, a set

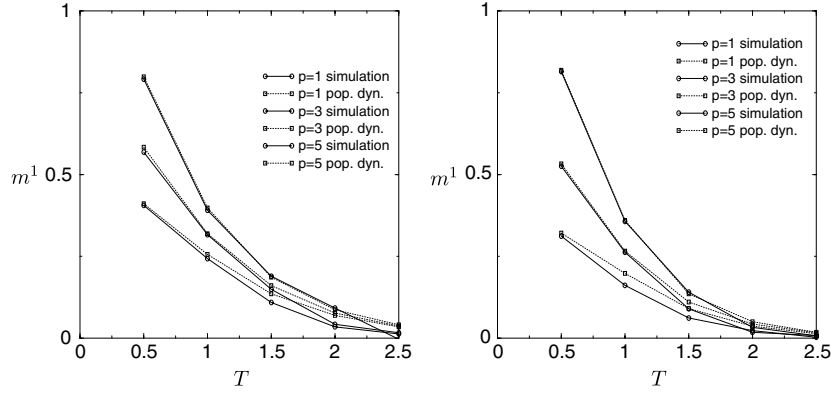


Figure 2. Condensed recall overlap m^1 as calculated from our RS theory via population dynamics (squares) versus numerical simulations (circles), for a graph with $p(k) \sim [k(k+1)(k+2)]^{-1}$ (left), and a graph with $p(k) \sim k^{-3}$ (right), both with $p(k) = 0$ for $k < 3$. The corresponding average connectivities are $\langle k \rangle = 6$ (left) and $\langle k \rangle = 5.123$ (right). All data shown are averages over 10 runs, in both simulations ($N = 10^4$ spins) and in population dynamics (with populations of size $N = 10^4$).

$\{k_1, \dots, k_N\}$ is generated in accordance with the distribution $p(k)$. One then chooses randomly two sites i and j with probabilities $p(i) = k_i / \sum_j k_j \sim p(k_i)k_i$. Unless these two sites coincide or already share a bond, they are connected. If site i already has k_i connections, it is excluded from the process to speed up the algorithm. This process is repeated until all connectivities have the correct value. We also used this algorithm to generate power-law distributed graphs with $\gamma = 3$ (right-hand figure). For both architectures, the overlaps m^1 have been plotted for networks with $p = 1, 3$ and 5 . At low temperature, the results of the simulation agree well with the population dynamics results. For the values of p used in figure 2 one has pattern retrieval at sufficiently low temperatures. In fact, our theory claims that for $p(k) \sim k^{-\gamma}$ with $\gamma \leq 3$ one will have retrieval at any T . For large values of p and T , however, the overlaps, although indeed nonzero, become smaller and hence our numerical accuracy decreases. Moreover, the equilibration times in the simulations grow rapidly as p increases. The discrepancies at high temperatures and large p are, we believe, due to finite-size effects. In figure 3 we show that the agreement between theory and simulations indeed improves for larger system sizes.

6. Conclusions

We have solved attractor neural network models on random graphs with arbitrary connectivity distributions $p(k)$, using the replica method within the RS ansatz, in the spirit of [17, 22, 23]. The RS order parameters are the effective cavity field distributions $W_\xi(h)$ in each sublattice, or equivalently, the distributions of messages $Q_\xi(u)$. Second-order phase transitions from the P phase to a R or SG phase could be derived explicitly, given the assumption that these transitions are second order and provided the second moment $\langle k^2 \rangle$ of the connectivity degree distribution of the graph is finite. The overlap and SG order parameters in each phase can in principle be calculated via a population dynamics algorithm. The latter is limited by numerical accuracy when the values of these order parameters are small (as for large T and α). We find that the retrieval region in the phase diagram is larger for fat-tailed degree distributions than for those with exponential decay (e.g. Poissonian), but it is not clear whether this can be exploited in associative memories since it goes at the cost of the magnitude of the retrieval

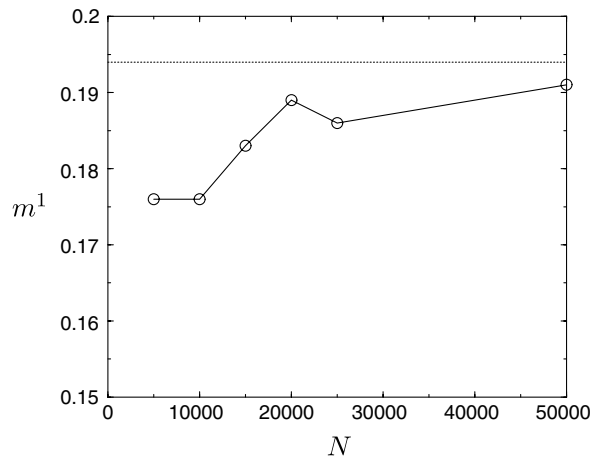


Figure 3. Average retrieval overlap m^1 over 10 simulations as a function of system size N , for connectivity degree distribution $p(k) \sim k^{-3}$ (with $p(k) = 0$ for $k < 3$, at $T = 1$ and $p = 5$). The dotted line corresponds to the value predicted by our RS population dynamics.

overlaps. The possible occurrence of replica symmetry breaking is beyond the scope of this paper. Within our numerical accuracy, we can conclude that upon comparing the results of our replica symmetric theory (including population dynamics) to numerical simulations, for degree distributions $p(k) \sim [k(k+1)(k+2)]^{-1}$ and $p(k) \sim k^{-\gamma}$, we obtain satisfactory agreement.

Acknowledgments

This study was initiated during an informal Finite Connectivity Workshop at King's College London in November 2003. TN, IPC, NS and BW acknowledge financial support from the State Scholarships Foundation (Greece), the Fund for Scientific Research (Flanders, Belgium), the ESF SPHINX programme and the FOM Foundation (Fundamenteel Onderzoek der Materie, The Netherlands), respectively.

References

- [1] Hopfield J J 1982 *Proc. Nat. Acad. Sci. USA* **79** 2554
- [2] Mezard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
- [3] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing* (Oxford: Oxford University Press)
- [4] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. A* **32** 1007
- [5] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530
- [6] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257
- [7] Domany E, van Hemmen J L and Schulten K (ed) 1991 *Models of Neural Networks I* (Berlin: Springer)
- [8] Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett.* **4** 167
- [9] Watkin T L H and Sherrington D 1991 *Europhys. Lett.* **14** 791
- [10] Watkin T L H and Sherrington D 1991 *J. Phys. A: Math. Gen.* **24** 5427
- [11] Canning A and Naef J P 1992 *J. Physique I* **2** 1791
- [12] Eguiluz V M, Chialvo D R, Cecchi G, Baliki M and Apkarian A V 2003 *Preprint cond-mat/0309092*
- [13] Torres J J, Munoz M A, Marro J and Garrido P L 2003 *Preprint cond-mat/0310205*
- [14] Torres J J, Marro J, Garrido P L, Cortes J M, Ramos F and Munoz M A 2003 *Preprint cond-mat/0312230*
- [15] Mcgraw P N and Menzinger M 2003 *Phys. Rev. E* **68** 047102
- [16] Stauffer D *et al* 2003 *Eur. Phys. J. B* **32** 395

-
- [17] Wemmenhove B and Coolen A C C 2003 *J. Phys. A: Math. Gen.* **36** 9617
 - [18] Pérez Castillo I and Skantzos N S 2003 *Preprint* cond-mat/0309655
 - [19] Viana L and Bray A J 1985 *J. Phys. C: Solid State Phys.* **18** 3037
 - [20] Kanter I and Sompolinsky H 1987 *Phys. Rev. Lett.* **58** 164
 - [21] Mezard M and Parisi G 1987 *Europhys. Lett.* **3** 1067
 - [22] Wong K Y and Sherrington D 1987 *J. Phys. A: Math. Gen.* **20** L793
 - [23] Leone M, Vázquez A, Vespignani A and Zecchina R 2002 *Eur. Phys. J. B* **28** 191
 - [24] Gronlund A, Sneppen K and Minnhagen P 2004 *Preprint* cond-mat/0401537
 - [25] Farkas I, Derényi I, Palla G and Vicsek T 2004 *Preprint* cond-mat/0401640
 - [26] van Hemmen J L and Kühn R 1986 *Phys. Rev. Lett.* **57** 913
 - [27] Monasson R 1998 *J. Phys. A: Math. Gen.* **31** 513
 - [28] Mezard M and Parisi G 2001 *Eur. Phys. J. B* **20** 217
 - [29] Berg J and Sellito M 2001 *Phys. Rev. E* **65** 016115
 - [30] Wemmenhove B and Skantzos N S 2004 *J. Phys. A: Math. Gen.* **37** 7843
 - [31] Barabási A L and Albert R 1999 *Science* **286** 509
 - [32] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47