

The parallel dynamics of a dilute symmetric Hebb-rule network

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

1991 J. Phys. A: Math. Gen. 24 5427

(<http://iopscience.iop.org/0305-4470/24/22/025>)

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

Download details:

IP Address: 129.252.86.83

The article was downloaded on 01/06/2010 at 14:02

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

COMMENT

The parallel dynamics of a dilute symmetric Hebb-rule network

T L H Watkin and D Sherrington

Department of Theoretical Physics, Oxford University, 1 Keble Road, Oxford OX1 3NP, UK

Received 24 May 1991

Abstract. We find that the dynamics of a dilute network of symmetric Hebb-rule synapses with parallel evolution, the model considered by Patrick and Zagrebnov, are considerably more complicated than that paper predicted. The number of parameters is not constant but equal to the square of the number of time steps taken and they are generated by exact but increasingly complicated relations. A replica-symmetric calculation of long-term behaviour gives an estimate of $\alpha_c \approx 1.26$. The calculations also describe parallel evolution of fully connected spin glasses.

During the past few years there has been considerable interest in the dynamics of neural network models with very low asymmetric connectivity, due to the remarkable discovery [1] that the full evolution in time of these systems can be calculated exactly. This result is in strong contrast to that for fully connected networks which have, in general, extremely complicated dynamics [2, 3] governed by detailed correlations. Here we show (in contradiction to the predictions of [4]) that a dilute but symmetric network also has complicated (though different) dynamics, whose long-term behaviour may be examined using the techniques of thermodynamics.

To be more specific, [1] considered a dilute network of N Ising spins $\{s_i = \pm 1\}$ with $i = 1, \dots, N$, which stores p patterns $\{\xi_i^\mu = \pm 1\}$ with $i = 1, \dots, N$ and $\mu = 1, \dots, p$, where each ξ_i^μ is randomly chosen to be $+1$ or -1 with equal probability. Neurons i and j ($i \neq j$) are connected by bonds

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

where each c_{ij} is drawn independently from the distribution

$$\text{Prob}(c_{ij} = x) = \left(1 - \frac{C}{N}\right) \delta(x) + \frac{C}{N} \delta(x-1) \quad i \neq j \tag{2}$$

which implies that, on average, every neuron receives inputs from C other neurons and sends outputs to a different selection of C neurons.

The network evolves in parallel for τ timesteps, so that at each of the steps $t = 1, \dots, \tau$ every $s_i(t)$ is updated according to

$$\text{Prob}(s_i(t+1) = \pm 1) = \frac{1}{2} \left(1 \pm \tanh \left(\beta \sum_j J_{ij} s_j(t) \right) \right) \tag{3}$$

β is the inverse temperature.

To recall pattern 1 we could initialize the network by randomly and independently choosing each $s_i(t)$ to be $\pm \xi_i^1$ with probabilities g and $(1-g)$ respectively. Thus the initial overlap of the network with the pattern is $m(0) = 2g - 1$, where

$$m(t) = \left\langle \frac{1}{N} \sum_{i=1}^N \xi_i^1 s_i(t) \right\rangle. \quad (4)$$

For $t=0$ the brackets refer to the average over the initial setting of the network, but for $t>0$ it also includes the average over $\{c_{ij}\}$, $\{\xi_i^\mu\}$ and the thermal noise of (3)

For $C \ll \ln N$ the network has a local tree structure so that almost all pairs of neurons have entirely different sets of ancestors, and correlations in noise may be neglected. [1] showed that this gives the simple recursion relation:

$$m(t+1) = \int_{-\infty}^{+\infty} Dz \tanh[\beta(m(t) + z\sqrt{\alpha})] \quad (5)$$

where

$$Dz \equiv \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \quad (6)$$

and $\alpha \equiv p/C$. The limit of $m(t)$ as $t \rightarrow \infty$ can only be non-zero if $\alpha < (2/\pi) = 0.634$, which is the 'critical capacity', α_c , of the network; m goes continuously to zero as $\alpha \rightarrow \alpha_c$.

In the work of [4], however, the distribution of the bonds is fundamentally different. The $\{c_{ij}\}$ are drawn from the distribution

$$\text{Prob}(c_{ij} = x) = \left(1 - \frac{C}{N}\right) \delta(x) + \frac{C}{N} \delta(x-1) \quad i < j \quad (7)$$

$$c_{ij} = c_{ji}.$$

Now all bonds are symmetric: the C neurons from which i receives inputs are also the C to which i sends outputs. The network is locally (over a few steps from any site) a Cayley tree, as in figure 1 where the neurons connected to i have been labelled by variable j . ([4] uses a trivially different definition of C and α : $C^{[4]} = C/2$ and $\alpha^{[4]} = \alpha/2$.)

The exact steady-state of this system with asynchronous dynamics (spins updated one at a time according to (3) instead of all at once), has already been derived [5]

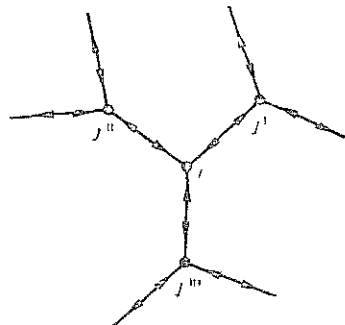


Figure 1. The Cayley tree of neural interactions. Neurons connected to i are labelled by j .

using a thermodynamic technique which requires only $C \ll N$ (a very much weaker constraint than $C \ll \ln N$ and one that renders the model susceptible to numerical simulation). It too shows a continuous transition to $m = 0$, but at a critical capacity of $\alpha_c = 1$ exactly.

To solve the parallel evolution problem Patrick and Zagrebnoy [4] assumed that noise on the branches of the Cayley tree connected to a given node, i , remains uncorrelated, which gives an evolution with just two order parameters. However, as noted in [1], dilute systems with symmetric bonds are considerably more complicated than the asymmetric versions. The noise on bond ij (figure 1) is equal to

$$\sum_{\mu \neq i} \xi_i^\mu \xi_j^\mu \tag{8}$$

which influences the effect $s_j(0)$ has upon $s_i(1)$, which in turn influences $s_j(2)$, interfering with the noise of the j th branch, to produce an effect at $s_i(3)$. Reference [4] also neglected, for $t \geq 3$, the effects of 'echoes' within the branches, $s_j(2)$ would be correlated with $s_j(0)$ even if neuron i were absent.

Correcting the analysis of [4], would be cumbersome and highly prone to errors. A cleaner and more reliable alternative is to use the generating function introduced by [3] and generalized in [6] into a Markov chain (an equivalent function was used in [7]):

$$Y(\{h(t)\}) = \left\langle \text{Tr}_{\{s_i(t)\}} \left[\prod_{i=1}^N \prod_{t=1}^{\tau} \frac{1}{2} \left(1 + \tanh \left(\beta \left(s_i(t) \sum_{j=1}^N J_{ij} s_j(t-1) \right) \right) \right) \right] \right. \\ \left. \times \exp \left(\frac{1}{N} \sum_i h(t) \sum_i \xi_i s_i(t) \right) \right\rangle_{\{J_{ij}\}, \{s_i(0)\}} \tag{9}$$

where the trace is over all N spins at all times. This function has the property that

$$\frac{d}{dh^i} Y|_{\{h^i=0\}} = m^i.$$

The average over $\{c_{ij}\}$ can now be performed in the manner of [6] and [7], requiring only that $C \ll N$. After τ steps the following τ^2 order parameters have been generated

$$m(t) \qquad \qquad \qquad t = 1, \dots, \tau \\ q(t_1, t_2) = \left\langle \frac{1}{N} \sum_i s_i(t_1-1) s_i(t_2-1) \right\rangle \qquad 0 < t_1 < t_2 < \tau \tag{10} \\ r(t_1, t_2) \qquad \qquad \qquad 0 < t_2 < t_1 < \tau.$$

The values of the order parameters may be expressed as

$$m(t) = g \langle s(t) s(0) \rangle_{z^+} + (g-1) \langle s(t) s(0) \rangle_{z^-} \\ ir(t_1, t_2) = \alpha g \langle x^{t_2} s(t_2) s(t_1-1) \rangle_{z^+} + \alpha (1-g) \langle x^{t_2} s(t_2) s(t_1-1) \rangle_{z^-} \tag{11} \\ q(t_1, t_2) = g \langle s(t_1-1) s(t_2-1) \rangle_{z^+} + (1-g) \langle s(t_1-1) s(t_2-1) \rangle_{z^-}$$

where the averages are with respect to 'one-particle' partition functions:

$$z^\pm = \frac{1}{2} \text{Tr}_{s_i(t)} \int_{-\infty}^{\infty} \prod_{t=1}^{\tau} d\lambda^t \int_{-\infty}^{\infty} \frac{dx^t}{2\pi} \frac{1}{2} (1 + \tanh(\beta \lambda^t)) \\ \times \exp \left(i \sum_{t=1}^{\tau} x^t \lambda^t \mp i \sum_{t=1}^{\tau} x^t s(t) s(0) m(t-1) - \alpha \sum_{t_1 < t_2} q(t_1, t_2) x^{t_1} x^{t_2} s(t_1) s(t_2) \right. \\ \left. - i \sum_{t_1 > t_2} r(t_1, t_2) x^{t_1} s(t_1) s(t_2-1) - \frac{\alpha}{2} \sum_i (x^i)^2 \right). \tag{12}$$

For simplicity we state the results of these equations for zero temperature ($\beta \rightarrow \infty$). After one step there is only one parameter

$$m(1) = 2 \operatorname{erf}(m(0)/\sqrt{\alpha}) \quad (13)$$

where

$$\operatorname{erf}(x) = \int_0^x Dz.$$

The new parameters after two steps are

$$\begin{aligned} q(1, 2) &= m(0)m(1) \\ r(2, 1) &= \sqrt{\frac{2\alpha}{\pi}} \exp\left(-\frac{(m(0))^2}{2\alpha}\right) \\ m(2) &= \sum_{\sigma=\pm} (1 + \sigma m(0)) \operatorname{erf}\left(\frac{m(1) + \sigma r(2, 1)}{\sqrt{\alpha}}\right). \end{aligned} \quad (14)$$

So far these results are in accordance with those of [1] (with a trivial difference in the definition of the error function). However, for $t > 3$ the results differ. After three time steps the new order parameters are

$$q(1, 3) = \sum_{\sigma=\pm} (1 + \sigma m(0)) \operatorname{erf}\left(\frac{r(2, 1) + \sigma m(1)}{\sqrt{\alpha}}\right) \quad (15)$$

$$r(3, 2) = \sqrt{\frac{\alpha}{2\pi}} \sum_{\sigma=\pm} (1 + \sigma m(0)) \exp\left(-\frac{(r(2, 1) + \sigma m(1))^2}{2\alpha}\right) \quad (16)$$

$$r(3, 1) = 0 \quad (17)$$

$$\begin{aligned} m(3) &= \operatorname{Tr}_{\sigma_1, \sigma_2} \frac{\sigma_2 + m(0)}{2\sqrt{2\pi\alpha}} \left[\int_{-\sigma_1 - \sigma_2 m(2)}^{\infty} d\lambda - \int_{-\infty}^{-\sigma_1 - \sigma_2 m(2)} d\lambda \right] \\ &\quad \times \exp\left(-\frac{\lambda^2}{2\alpha}\right) \left[\frac{1}{2} + \sigma_1 \operatorname{erf}\left(\frac{q(1, 3)\lambda + \sigma_2 m(0)}{\sqrt{\alpha(1 + (q(1, 3))^2)}}\right) \right]. \end{aligned} \quad (18)$$

$q(2, 3)$ is also non-zero, but does not enter the formula for $m(3)$. Equations (13)–(17) are corrected for non-zero temperature by the substitutions:

$$\begin{aligned} \exp\left(-\frac{x^2}{2\alpha}\right) &\rightarrow \int_{-\infty}^{\infty} \frac{d\lambda}{\alpha} \tanh(\beta(\lambda + x)) \lambda \exp\left[-\frac{\lambda^2}{2\alpha}\right] \\ \operatorname{erf}\left(\frac{x}{\sqrt{\alpha}}\right) &\rightarrow \int_{-\infty}^{\infty} \frac{d\lambda}{2\sqrt{2\pi\alpha}} \tanh(\beta(\lambda + x)) \exp\left(-\frac{\lambda^2}{2\alpha}\right) \end{aligned} \quad (19)$$

with a similar complication of the formula for $m(3)$.

Equations of a similar form to (11) and (12) describe the dynamics of a fully connected Hebb network [3] at zero temperature, but with many more order parameters related in even more complicated ways.

Figure 2 plots $m(1)$, $m(2)$ and $m(3)$ against the initial overlap $m(0)$ for $\alpha = 0.3$ and zero temperature. $\tilde{m}(3)$ is the result for the third overlap if correlations between the branches of the Cayley tree are neglected. The 'experimental' values are the averages

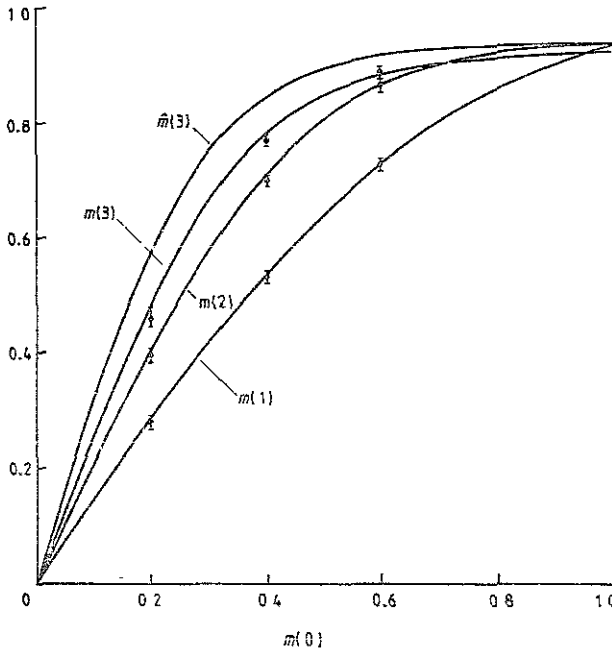


Figure 2. A comparison of theory and the results of a numerical simulation. The overlap of the network with the pattern at timesteps 1, 2 and 3 is shown against the initial overlap $m(0)$. $\hat{m}(3)$ is the prediction for $m(3)$ neglecting inter-branch correlations. Experimental points are the results of five completely independent simulations with $C = 50$ and $N = 5000$ and the error bars mark the standard deviation.

of five completely independent simulations with $C = 50$ and $N = 5000$, which provide very good support for our theoretical predictions. As the number of steps increases, however, the order parameters are related in ever more complicated ways and no simple, general recursion relation exists.

To derive the critical capacity of this network we resort to a thermodynamic calculation using the hamiltonian appropriate to parallel updating [8].

$$H = -\frac{1}{\beta} \sum_{i=1}^N \ln \left(2 \cosh \left(\beta \left(\sum_{j=1}^C J_{ij} s_j \right) \right) \right) \tag{20}$$

giving a partition function

$$Z = \text{Tr}_{\{s_i\}} \exp(-\beta H).$$

The free energy averaged over the network disorder,

$$F_a = \langle \ln Z \rangle_{J_{ij}}$$

is obtained using the method of 'replicas' (reviewed in [9]), which introduces n replicas of the system (labelled by $\gamma = 1, \dots, n$). Taking N and C to be large, with $C \ll N$, and with the conventional assumption of 'replica symmetry', gives the average free

energy F as a function of six order parameters.

$$F_{av}(m, \bar{m}, q, Q, \delta, \varepsilon)$$

$$= m\bar{m} - \beta(q-1) \left(Q - \frac{\alpha}{2} \right) + \frac{\alpha\beta}{2} (\delta^2 - \varepsilon^2) - \frac{1}{\beta} \int_{-\infty}^{\infty} D\gamma Dz \\ \times \ln \left[\sum_{\sigma=\pm} \exp(\sigma\beta(\bar{m} + z\sqrt{Q})) 2 \cosh \left(\beta \left(\frac{\sigma\alpha\beta(\varepsilon - \delta)}{2} + \gamma\kappa - \frac{z\alpha\varepsilon}{2\sqrt{Q}} - m \right) \right) \right] \quad (21)$$

where

$$\kappa^2 \equiv \frac{\alpha}{2} \left(q - \frac{\varepsilon^2 \alpha}{2Q} \right).$$

Only two of these parameters have an obvious physical interpretation.

$$m = \langle \xi s_\gamma \rangle$$

is the average overlap of any replica with the pattern, and

$$q = \langle s_\gamma s_{\gamma'} \rangle \quad \gamma \neq \gamma'$$

is the overlap between any pair of replicas. As the storage capacity $\alpha \rightarrow 0$, the free energy tends to that found in [8] for the far simpler case of low-loading networks; only the parameters \bar{m} and m are of significance in this limit, and these have been shown to be identical.

The order parameters for any temperature and α may be found from the stationary points of F , but the search in six dimensions is difficult numerically. As $\beta \rightarrow \infty$ the equation remains physical since we expect both $\beta(1-q)(Q - (\alpha/2))$ and $\beta(\varepsilon - \delta)$ to remain of order 1. In this limit, where the critical α is expected to lie, $q \rightarrow 1$ and $Q \rightarrow (\alpha/2)$, but even here the free energy landscape is difficult to compute and rich in spurious minima. Stationary points, however, seem to have $\beta(\varepsilon - \delta)$ small, and using the ansatz that it is zero gives an analytic result for $\alpha_c = (4/\pi) = 1.24$, double the replica-symmetric answer for asynchronous updating, which we take as a good estimate of the true value. It is in reasonable agreement with computer simulations using $C = 50$ and $N = 5000$.

A mapping exists between this problem and that of a fully-connected spin glass with symmetric bonds ($J_{ij} = J_{ji}$) distributed as for the SK model [10]:

$$N \langle J_{ij} \rangle = J_0 \\ N [\langle J_{ij}^2 \rangle - \langle J_{ij} \rangle^2] = J^2 \quad i < j \quad (22)$$

with the same updating rule (3), but operating at inverse temperature β^{SK} . Equations (13)–(21), with the change of variables $\alpha \rightarrow (J^2/J_0^2)$ and $\beta \rightarrow (J_0\beta^{\text{SK}})$, describe this spin glass exactly, in the $N \rightarrow \infty$ limit. m now represents the overall magnetization of the system. Equations similar to (11) and (12) were in fact derived for the spin glass in [3], but restricted to zero temperature and with the very considerable simplification of $J_0 = 0$, which makes many of the order parameters zero and simplifies the algebra allowing further steps to be taken more easily. $J_0 = 0$ is not appropriate for a neural network because final states have no correlation with stored patterns, nor did [3] analyse the equilibrium behaviour of the system. The same mapping between a fully connected spin glass and a dilute Hebb network was found in [5] for the case of asynchronous updating.

This model is worthy of further analysis. Even at zero temperature limit cycles (in which the whole network alternates between two states) may exist for the Hebb network, and do for the spin glass. What meaning can we then attach to $q \rightarrow 1$, all the replicas becoming exactly correlated? Nor is it clear if just six parameters can represent the equilibrium behaviour of a very large number of dynamical order parameters, nor whether replica symmetry is broken [9]. A thorough investigation of the equilibrium behaviour of parallel updating systems is currently in preparation.

In summary, even in a highly dilute network symmetric bonds cause non-trivial correlations to develop. In the large N limit these can be treated by analysis of Markov chain relations, yielding results which are exact for $c \ll N$ but become increasingly tedious to evaluate. Long-term behaviour may be predicted by thermodynamics, with an estimate of critical capacity of $\alpha_c = 1.24$.

Note added in proof. After seeing our results Dr Patrick and Dr Zagrebnoy were able to derive our result for $m(3)$ by extending and correcting their previous method. Their calculation will be presented elsewhere.

References

- [1] Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett* **4** 167
- [2] Horner H, Bormann D, Frick M, Kinzelbach H and Schmidt A 1989 *Z. Phys.* **B 76** 381
- [3] Gardner E, Derrida B and Mottishaw P 1987 *J. Physique* **48** 741
- [4] Patrick A E and Zagrebnoy V A 1990 *J. Phys. A: Math. Gen.* **23** L1323
- [5] Watkin T L H and Sherrington D 1991 *Europhys. Lett.* **14** 791
- [6] Watkin T L H and Sherrington D (preprint)
- [7] Kree R and Zippelius A 1991 *Models of Neural Networks* ed E Domany, J L van Hemmen and K Schulten (Berlin: Springer)
- [8] Peretto P 1984 *Biol. Cybern.* **50** 51
- [9] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)