Neural networks with biased bipolar synapses and biased patterns

1993 J. Phys. A: Math. Gen. 26 873

(http://iopscience.iop.org/0305-4470/26/4/014)

J. Phys. A: Math. Gen. 26 (1993) 873–881. Printed in the UK

Neural networks with biased bipolar synapses and biased patterns

R Raju Viswanathan

Centre for Artificial Intelligence and Robotics, Rajbhavan Circle, Bangalore 560001, India

Received 27 May 1992

Abstract. A neural network model with bipolar synaptic couplings is studied with a global constraint on the synapses which each neuron receives, corresponding to a bias of the couplings. When the network is used to store biased patterns, it is shown that the maximal storage capacity for any non-zero bias of the stored patterns corresponds to a particular type of the connectivity pattern of the network as encoded in the constraint. The 'space of interactions' approach yields the dependence of storage capacity on the distribution of the couplings for every pattern bias. It turns out that the optimal coupling bias is non-zero and independent of the bias in the patterns.

1. Introduction

Neural network models have become popular for associative memory in recent years. In this paper, we investigate a particular class of neural network models, namely those with synaptic connection weights which take values in the discrete set $\{+1,-1\}$, also referred to as Ising bonds. This class of models has the advantage of ease of hardware implementation. Such a model was first studied in [1] using the 'space of interactions' method; it was found that the replica symmetric solution was unstable. The nature of the replica symmetry breaking was elucidated in [2] and the maximal capacity was found there to be about $0.83N$, in agreement with the numerical evidence of [3,4]. Owing to the fact that replica symmetry in the space of interactions is broken, the maximal storage capacity corresponds to zero entropy in the space of interactions. In [5], the storage capacity for such networks used to store biased patterns was determined as a function of the bias; it was established that for models with broken replica symmetry, the zero entropy condition provides an accurate estimate of the storage capacity.

Here we answer the following question: What is the effect of the distribution of the signs of the synaptic couplings in the network on the storage capacity of biased patterns? For a network of N neurons in the large N limit, we study this question in the next section by imposing a constraint on the synaptic couplings which each neuron receives, namely that the sum of the synaptic couplings is proportional to $\sqrt{N}$, with the same constant of proportionality for all the neurons. This is a global constraint on the couplings which is a measure of the bias in the distribution of the signs of the couplings. A biological motivation for this question is to understand the significance of Dale's law† in the context of human neurobiology. While the model we study here is not a faithful representation of Dale's law, we nevertheless believe that our results support the general conjecture that there is an optimal distribution of the signs of the synaptic couplings which makes the storage of biased patterns efficient.

In the next section, the model is formulated and the saddle point equations in the space of interactions are obtained. The saddle point equations for the model are solved numerically; as for the random pattern case, replica symmetry is broken and so the equations are solved for zero entropy in the space of interactions. We find that there is a distribution of the couplings which maximizes the storage capacity of biased patterns. Moreover, this distribution appears to be 'universal' in the sense that it is independent of the bias in the stored patterns (as long as it is non-zero), and depends only slightly on the size of the basin of attraction of the stored patterns. This raises the interesting possibility that a given optimal set of synaptic couplings might be able to store differently biased patterns simultaneously. The peaks in the storage capacity are sharper for patterns with larger bias.

We end with a brief summary of our results.

2. Biased synapses and biased patterns

The model we shall consider is a discrete state asynchronously updated neural network model with states $s_i = \pm 1$ $(i = 1, \ldots, N)$ and the standard updating rule

$$s_i(t+1) = \mathrm{sgn}\left(\sum_{j \neq i} w_{ij} s_j(t)\right) \tag{1}$$

where the w_{ij} are the synaptic weights. It is assumed, as usual, that $w_{ii} = 0$. We shall constrain the connection weights w to be bipolar, i.e. they are only allowed to take the values ± 1.

The fixed points of the model are those states which do not change under the updating rule (1); therefore such states satisfy [6]

$$s_i \sum_j \frac{w_{ij} s_j}{\sqrt{N}} > \kappa \tag{2}$$

where κ is a strictly positive number. The larger the value of κ, the bigger is the basin of attraction around the fixed point. We will assume that there are $p = \alpha N$ patterns or states $\{s_i^\mu\}$, $\mu = 1, \ldots, p$, to be stored. The condition for each pattern to be a fixed point is

$$\gamma_i^\mu = s_i^\mu \sum_j \frac{w_{ij} s_j^\mu}{\sqrt{N}} > \kappa \tag{3}$$

for $\mu = 1, \ldots, p$. We assume that the stored patterns are all biased with the same bias m. This means that the probability p_σ of choosing a bit with value σ from any one of the patterns is

$$p_\sigma = \frac{1+m}{2}\delta_{\sigma,1} + \frac{1-m}{2}\delta_{\sigma,-1}. \tag{4}$$

† This law is the empirical fact that the synaptic couplings of a single neuron in the nervous system are either all inhibitory or all excitatory.

The number m essentially measures the fractional excess of plus signs over minus signs in the patterns to be stored. We further assume that the synaptic couplings which each neuron receives are constrained to have a bias

$$\sum_j w_{ij} = r\sqrt{N}. \tag{5}$$

This is a global constraint on the input synapses of each neuron. The number r measures the excess of excitatory couplings over inhibitory ones; we shall call it the bias of the couplings.

The partition function in the space of interactions $[w]$ can be written as [6] (apart from an overall normalization)

$$Z = \int \left(\prod_{i,j} \mathrm{d}w_{ij} \right) \rho[w] \mathrm{e}^{-\beta H[w]} \tag{6}$$

where the Hamiltonian

$$H[w] = \sum_{i=1}^{N} \sum_{\mu=1}^{N} \theta(\kappa - \gamma_i^\mu[w]) \tag{7}$$

measures the number of sites that are not fixed points and β is a 'temperature' parameter. Here $\rho[w]$ is the density of states in the space of interactions; in our case

$$\rho[w] = (\delta(w_{ij} - 1) + \delta(w_{ij} + 1))\, \delta\left(\sum_j w_{ij} - r\sqrt{N} \right). \tag{8}$$

In the $\beta \to \infty$ limit, Z is simply the number of states with energy zero

$$\lim_{\beta\to\infty} Z = \lim_{\beta\to\infty} \sum_{E_n} \Omega(E_n) \mathrm{e}^{-\beta E_n} \to \Omega(0) \tag{9}$$

so that in this limit the entropy of the zero energy state is

$$S = \ln \Omega(0) \tag{10}$$

$$= \ln \int \left(\prod \mathrm{d}w_{ij} \right) \rho[w] \prod_{i=1}^{N} \prod_{\mu=1}^{p} \theta(\gamma_i^\mu - \kappa). \tag{11}$$

The maximal storage capacity is reached when the entropy in interaction space becomes zero, since this means that there is only one set of w's which maintain the patterns $\{s_i^\mu\}$ as fixed points. The quantity of interest is therefore the entropy, averaged over the distribution of the stored patterns $\{s_i^\mu\}$, since we are interested in the storage of general patterns governed by the distribution p_σ. The expression for the averaged entropy $\langle\langle S \rangle\rangle$ then becomes

$$\langle\langle S \rangle\rangle = \sum_i \langle\langle \ln \Omega_i \rangle\rangle \tag{12}$$

with

$$\Omega_i = \int \left(\prod \mathrm{d}w_{ij} \right) \rho[w] \prod_{\mu=1}^{p} \theta(\gamma_i^\mu - \kappa) \tag{13}$$

since the rows of the matrix w_{ij} are independent.

The replica trick can be used to evaluate the quantity $\langle\langle \ln \Omega_i \rangle\rangle$

$$\langle\langle \ln \Omega_i \rangle\rangle = \lim_{n \to 0} \frac{\langle\langle \Omega_i^n \rangle\rangle - 1}{n}. \tag{14}$$

Introducing replicas w_{ij}^a $(a = 1, \ldots, n)$, we have

$$\langle\langle \Omega_i^n \rangle\rangle = \langle\langle \prod_{a=1}^{N} \int \left(\prod \mathrm{d}w_{ij} \right) \rho[w] \prod_{\mu=1}^{p} \theta\left(s_i \sum_j \frac{w_{ij} s_j}{\sqrt{N}} - \kappa \right) \rangle\rangle. \tag{15}$$

This expression can be evaluated in the saddle point approximation for $N \to \infty$, as is well known. Introducing an integral representation for the step function, averaging over the spins s_j^μ, and introducing delta functions with additional variables F_{ab} and P_a to enforce the constraint for the order parameter

$$q_{ab} = \frac{1}{N} \sum_{j=1}^{N} w_{ij}^a w_{ij}^b \tag{16}$$

and the constraint for the incoming synapses

$$r = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{ij}^a \tag{17}$$

respectively, the expression for $\langle\langle \Omega_i^n \rangle\rangle$ becomes

$$\langle\langle \Omega_i^n \rangle\rangle = \left\langle \int \left(\prod_a \frac{\mathrm{d}P_a}{2\pi} \right) \left(\prod_{a<b} \frac{\mathrm{d}q_{ab} \mathrm{d}F_{ab}}{2\pi/N} \right) \exp\left(N\alpha G_1(q_{ab}, r_a) + N G_2(P_a, F_{ab}) + \mathrm{i}N \sum_{a<b} q_{ab} F_{ab} \right) \right\rangle. \tag{18}$$

Here the brackets $\langle \cdots \rangle$ indicate that an average over s_i^μ remains to be performed, and $G_1(q_{ab}, r_a)$ is given by

$$\mathrm{e}^{N\alpha G_1(q_{ab}, r_a)} = \left[\int_\kappa^\infty \left(\prod_a \frac{\mathrm{d}\lambda^a}{2\pi} \right) \int_{-\infty}^{\infty} \left(\prod_a \mathrm{d}x^a \right) \exp\left(\mathrm{i} \sum_a x^a \lambda^a - \mathrm{i}mrs_i^\mu \sum_a x^a - \frac{1-m^2}{2} \left(\sum_a (x^a)^2 + 2 \sum_{a<b} q_{ab} x^a x^b \right) \right) \right]^p. \tag{19}$$

Using the ansatz of replica symmetry, $q_{ab} = q$, and taking the average over s_i^μ, the expression for G_1 reduces, in a standard manner, to

$$G_1(q, r) = n \int_{-\infty}^{\infty} \mathrm{D}z \left(\frac{1+m}{2} \ln H(\tau_-) + \frac{1-m}{2} \ln H(\tau_+) \right) \tag{20}$$

where Dz is the Gaussian measure $\mathrm{D}z \equiv \mathrm{e}^{-z^2/2}\mathrm{d}z/\sqrt{2\pi}$, and the function H is the complementary error function

$$H(\tau) = \int_\tau^\infty \mathrm{D}z. \tag{21}$$

The variables τ_- and τ_+ are defined by

$$\tau_- = (1-q)^{-1/2} \left(z\sqrt{q} + \frac{\kappa - mr}{\sqrt{1-m^2}} \right) \tag{22}$$

and

$$\tau_+ = (1-q)^{-1/2} \left(z\sqrt{q} + \frac{\kappa + mr}{\sqrt{1-m^2}} \right). \tag{23}$$

Taking into account the constraint that $w_{ij} = \pm 1$, G_2 is determined by

$$\mathrm{e}^{NG_2} = \left[\sum_{w^a = \pm 1} \exp \left(\mathrm{i} \sum_a P_a w^a - \mathrm{i} \sum_{a<b} F_{ab} w^a w^b \right) \right]^N. \tag{24}$$

Making the replica symmetric ansatz $F_{ab} = \mathrm{i}F$ and $P_a = \mathrm{i}P$, we need to evaluate the expression

$$\sum_{w^a = \pm 1} \exp \left(F \sum_{a<b} w^a w^b - P \sum_a w^a \right) \tag{25}$$

where the sum is over all possible distributions of the w's over the discrete set $\{+1, --1\}$. This expression evaluates, in the small n limit, to

$$\left(1 + n \ln 2 + \frac{n}{2} \int_{-\infty}^{\infty} \mathrm{D}u \, (\ln \cosh(u\sqrt{F} - P) + \ln \cosh(u\sqrt{F} + P)) \right) \mathrm{e}^{-nF/2}. \tag{26}$$

Now we can evaluate the entropy in the saddle point approximation for $N \to \infty$; the saddle point is determined by the stationary conditions

$$\frac{\partial G}{\partial P} = \frac{\partial G}{\partial F} = \frac{\partial G}{\partial q} \tag{27}$$

where

$$G = \alpha G_1 + G_2 - \frac{n(n-1)}{2} qF. \tag{28}$$

First, the condition $\partial G/\partial P = 0$ yields $P = 0$. Then for small n, we get (apart from a constant which is cancelled by a contribution coming from the normalization of Z)

$$G_2 \approx -\frac{nF}{2} + n\int_{-\infty}^{\infty} \mathrm{D}u \;\ln\cosh u\sqrt{F} \tag{29}$$

so that the saddle point condition $\partial G/\partial F = 0$ gives

$$\frac{2}{\sqrt{F}}\int_0^{\infty} \mathrm{D}u \tanh u\sqrt{F} = 1 - q. \tag{30}$$

The third saddle point condition $\partial G/\partial q = 0$ yields

$$\frac{\alpha}{\sqrt{2\pi}}\left[\frac{1+m}{2}\int_{-\infty}^{\infty} \mathrm{D}z\frac{1}{H(\tau_-)}\mathrm{e}^{-\tau_-^2/2}\left(\frac{\tau_-}{1-q}+\frac{z}{\sqrt{q(1-q)}}\right)\right.$$
$$\left.+\frac{1-m}{2}\int_{-\infty}^{\infty} \mathrm{D}z\frac{1}{H(\tau_+)}\mathrm{e}^{-\tau_+^2/2}\left(\frac{\tau_+}{1-q}+\frac{z}{\sqrt{q(1-q)}}\right)\right] = F \tag{31}$$

where τ_- and τ_+ were defined earlier.

Now the replica symmetric value $q \to 1$ corresponds to an unstable solution, and, in addition, the corresponding entropy is negative. As was shown in [2,5], the correct solution corresponding to maximal storage must be determined by the zero entropy solution, which is also a stable solution for this model [2,5].

The entropy averaged over the stored patterns is

$$\begin{aligned}\langle\langle S\rangle\rangle &= N\lim_{n\to 0}\frac{\langle\langle \Omega_i^n\rangle\rangle - 1}{n}\\ &= N^2\left[\alpha\int_{-\infty}^{\infty}\mathrm{D}z\left(\frac{1+m}{2}\ln H(\tau_-)+\frac{1-m}{2}\ln H(\tau_+)\right)\right.\\ &\quad\left.+F\frac{q-1}{2}+\int_{-\infty}^{\infty}\mathrm{D}u\;\ln\cosh u\sqrt{F}\right].\end{aligned} \tag{32}$$

The zero entropy equation is then

$$\alpha\int_{-\infty}^{\infty}\mathrm{D}z\left(\frac{1+m}{2}\ln H(\tau_-)+\frac{1-m}{2}\ln H(\tau_+)\right)$$
$$+F\frac{q-1}{2}+\int_{-\infty}^{\infty}\mathrm{D}u\;\ln\cosh u\sqrt{F} = 0. \tag{33}$$

Equations (30), (31), and (33) must be solved together to determine the maximal storage capacity $\alpha_{\max}$.

These equations can be solved numerically. We have solved them using a method of interval halving for various values of m and κ, to an accuracy of about 0.02 in α. Figure 1 shows the storage capacity as a function of the coupling bias for $\kappa = 0$; the pattern bias values m are 0.2, 0.4, 0.6 and 0.8, with larger values of m leading to higher peaks.

Figures 2, 3 and 4 show the capacity as a function of coupling bias for $m = 0.2$, $m = 0.6$ and $m = 0.8$ respectively. The three curves in each figure are for values of κ equal to 0, 1 and 2, with smaller peaks corresponding to larger values of κ.

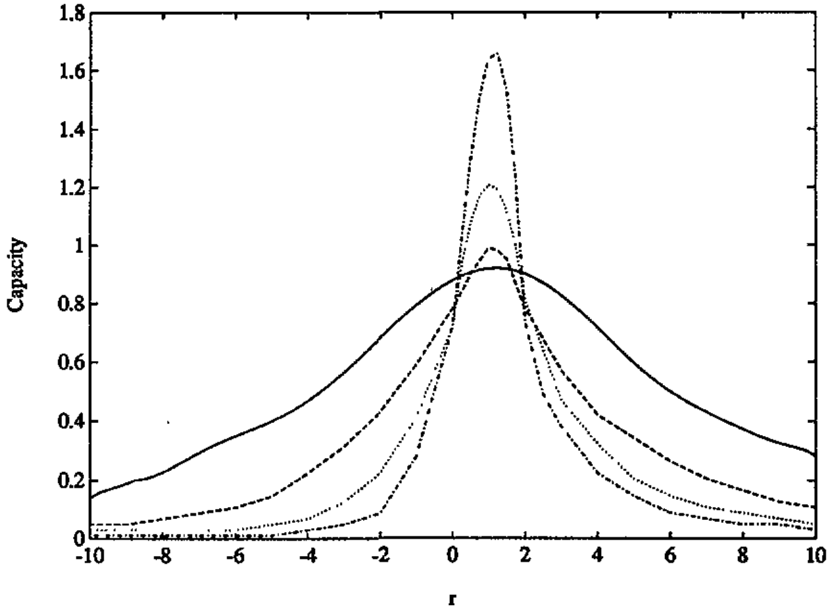

Figure 1. The storage capacity as a function of r for $\kappa = 0$ and $m = 0.2, 0.4, 0.6$ and 0.8; the higher peaks are for larger values of m.

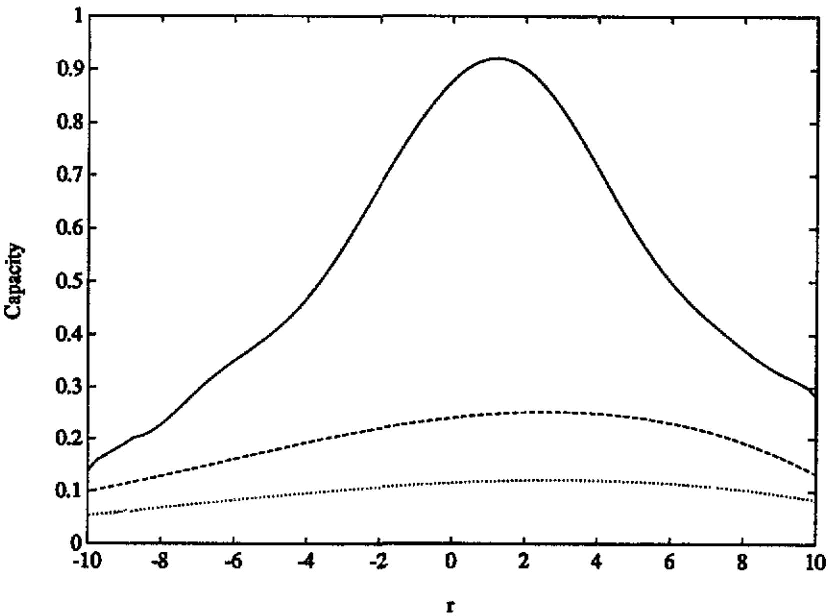

Figure 2. The storage capacity as a function of r for $m = 0.2$ and $\kappa = 0, 1$ and 2; the upper curves are for the smaller values of κ.

The interesting fact that emerges from these solutions is that for every value of κ, the storage capacity is optimal for a value of bias in the couplings which is

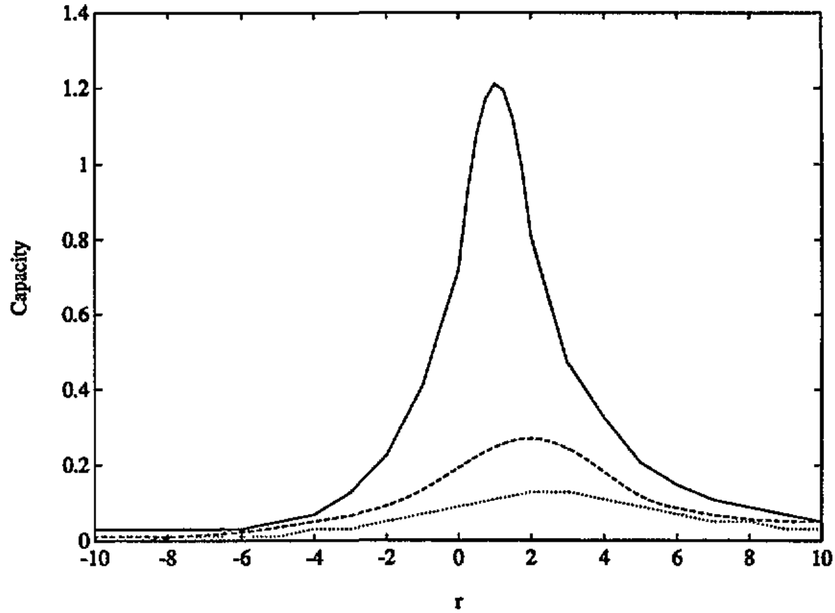

Figure 3. The storage capacity as a function of r for $m = 0.6$ and $\kappa = 0$, 1 and 2; the upper curves are for the smaller values of κ.

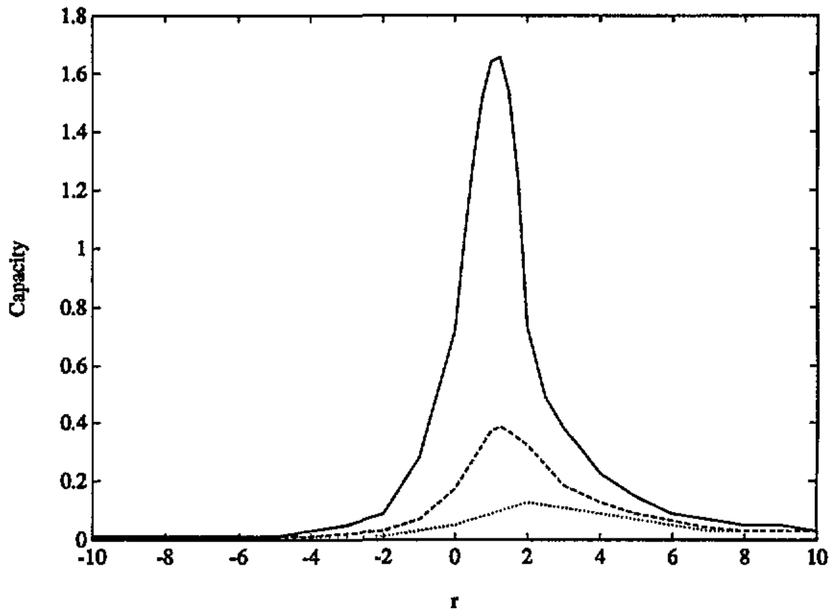

Figure 4. The storage capacity as a function of r for $m = 0.8$ and $\kappa = 0, 1$ and 2; the upper curves are for the smaller values of κ.

independent of the bias in the patterns, as long as m is non-zero. For random patterns ($m = 0$) of course the maximum capacity $\alpha_{\max} = 0.83$ is independent of the coupling bias r, as is evident from equations (31) and (33), since r drops out of equations (22) and (23) when $m = 0$. For $\kappa = 0$, it can be seen from figure 1 that $r = 1$ is an optimal value for the coupling bias. The optimal value of r increases slightly to about 2 as κ increases to 2, as can be seen from the smallest peaks in figures 2, 3 and 4.

The peaks in the storage capacity are sharper for patterns which are more severely biased, indicating that only a narrow range of coupling bias values store these patterns efficiently. Further, the capacity is independent of the sign of the pattern bias m since the relevant equations (31) and (33) are invariant under $m \to -m$.

3. Conclusions

We conclude by summarizing our results. We have investigated the dependence of the storage capacity on the distribution of the signs of the couplings in a bipolar neural network model. We find that there is a value of the coupling bias which maximizes the storage capacity independently of the bias of the stored patterns. The peak in the storage capacity is higher and narrower for more strongly biased patterns. It seems plausible that the existence of such an optimal distribution of couplings, for biased patterns, persists in more general networks as well. This issue is under further investigation and the results will be reported elsewhere.

References

[1] Gardner E and Derrida B 1988 *J. Phys. A: Math. Gen.* **21** 271
[2] Krauth W and Mezard M 1989 *J. Physique* **50** 3057
[3] Gardner E and Derrida B 1989 *J. Phys. A: Math. Gen.* **22** 1983
[4] Krauth W and Opper M 1989 *J. Phys. A: Math. Gen.* **22** L519
[5] Gutfreund H and Stein Y 1990 *J. Phys. A: Math. Gen.* **23** 2613
[6] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257