Study of self-inhibited analogue neural networks using the self-consistent signal-to-noise analysis

1992 J. Phys. A: Math. Gen. 25 4799

(http://iopscience.iop.org/0305-4470/25/18/014)

J. Phys. A: Math. Gen. 25 (1992) 4799–4811. Printed in the UK

Study of self-inhibited analogue neural networks using the self-consistent signal-to-noise analysis

T Fukai† and M Shiino‡

† Department of Electronics, Tokai University, Kitakaname, Hiratsuka, Kanagawa, Japan
‡ Department of Applied Physics, Tokyo Institute of Technology, Ohokayama, Meguro-ku, Tokyo, Japan

Received 16 December 1991, in final form 5 May 1992

Abstract. We study an analogue neural network model whose response function is linear and consequently has no asymptotes in order to examine a possible new mechanism for regulating neuron activities by means of neural feedback circuits. Inhibitory self-coupling is introduced as an example of such feedback mechanisms, which for simplicity of the analysis is assumed to be piecewise linear. Our recently developed self-consistent signal-to-noise analysis is applied to explore the equilibrium properties of the model neural network. The analysis revealed that there exists a finite optimal value for the linear analogue gain that maximizes storage capacity for a given value of the self-coupling. The results of the analysis for such equilibrium properties as storage capacity are quite consistent with the results of computer simulations.

1. Introduction

Various attempts have been made to understand the retrieval properties of analogue neural networks since the models are of importance from the point of view of electronic implementation as well as relevance to physiological nerve systems (Hopfield 1984, Marcus and Westervelt 1989, Marcus *et al* 1990, Waugh *et al* 1990, Fukai and Shiino 1990a, Shiino and Fukai 1990, Kühn *et al* 1991, Fukai and Shiino 1991). Usually in analogue networks, the response function of a neuron is assumed to be of sigmoid-type with proper asymptotes, which introduces a necessary cut-off for the activity of a neuron. It seems, however, that the mechanism for regulating the activation of neurons has to be studied more seriously since a low firing rate (Miyashita and Chang 1988, Tsodyks and Feigel'man 1988, Amari 1989, Treves and Amit 1989, Treves 1990) in living nervous systems may imply the necessity of a very different scheme for the cut-off of neuron activities. Some authors, for instance, introduced a non-monotonic response function which may appear as a consequence of local feedback from surrounding inhibitory neurons to lower the avarage activity level of neurons (Morita *et al* 1990). One of the aims of the present paper is to examine the network properties of the models incorporating a new type of mechanism to regulate excessive neuron activities.

To this end, we introduce self-coupling into symmetric neural networks as a simple example of a local feedback mechanism for inhibiting the activity of each neuron. The self-coupling is assumed to act when the neuron activity exceeds a certain critical level. In neural networks with an appropriate local inhibition scheme, the response

0305-4470/92/184799+13$04.50

function may not necessarily have asymptotes. In the present paper, we assume for simplicity that the response function is linear, and, consequently, has no asymptotes. The model is found to yield an interesting feature, in contrast to the usual sigmoid-type analogue neural networks, in that there exists an optimal finite linear analogue gain which maximizes the storage capacity for a given value of the magnitude of the self-coupling.

The introduction of a linear response function with no asymptotes in the present model may give rise to the drawback that the energy function becomes unbounded from below in a certain case. We can, however, confine ourselves to discussing the properties of locally stable fixed points for the memory retrieval of a network attained by converging dynamical flows even if the energy function is globally unbounded from below. The conventional thermodynamic approach using the replica method (Amit *et al* 1985) will require a delicate argument of its validity in the case of the downwards unbounded energy function since the thermal equilibrium distribution may not be properly defined through the stochastic relaxation process which happens to yield run-away trajectories downwards infinitely.

Recently, by extending the so-called signal-to-noise analysis (McEliece *et al* 1987, Peretto 1988, Amari and Maginu 1988, Domany *et al* 1989) which does not require the existence of an energy function, we proposed a new method, which is straightforward and easy to handle to analyse the statistical properties of the equilibrium states of analogue neural networks. The validity of the method (hereafter referred to as self-consistent signal-to-noise analysis, SCSNA) is ensured for conventional symmetric analogue networks (Hopfield 1982, 1984, Amit *et al* 1987) since it yields the same set of equations for determining the equilibrium states as that derived by means of the replica method. The method is also available for the analysis of Ising-spin neural networks when the corresponding Thouless–Anderson–Palmer (TAP) equations (Thouless *et al* 1977, Mézard *et al* 1987) are known. Furthermore, as far as fixed-point equilibrium attractors are concerned, SCSNA can be applicable even to analogue neural networks with asymmetric coupling. Indeed SCSNA was successfully applied to an asymmetric analogue neural network to yield a storage capacity which is consistent with simulation results (Shiino and Fukai 1991). The second aim of the present paper is to examine whether SCSNA works for the present network model which possibly has a bottomless energy function.

The paper is organized as follows. In section 2, the model is presented together with its energy function. The retrieval states for a finite number of patterns is discussed in section 3. In section 4, SCSNA is applied to the derivation of equations for the order parameters when the network is loaded with an extensive number of patterns. In section 5, the equations obtained are numerically solved and the phase diagram is presented for various cases. Computer simulations are also conducted. We shall see that the results from SCSNA on the equilibrium properties of the model is in quite good agreement with those of the computer simulations. Section 6 is devoted to concluding remarks.

2. Neural network model with self-coupling

We provide a set of variables u_i and x_i ($-\infty < u_i,\ x_i < \infty,\ i = 1, \ldots, N$) representing a membrane potential and the output activity of neuron i, respectively. The response function f which defines the relation between the two quantities $x_i =$

$f(u_i)$ is usually taken to be a sigmoid function with suitable asymptotes, say, $f = \tanh(\beta u_i)$ with an analogue gain β. In the present paper, however, we assume linearity for the response function:

$$f(u) = \beta u \qquad (\beta > 0). \tag{1}$$

Now we need to introduce a mechanism to lower the activity of excessively activated neurons. In the present model, such a mechanism is provided by feedback through the self-coupling of the neurons, that is the output activity of a neuron is suppressed by the inhibition introduced through the self-coupling if its activity level exceeds a certain threshold Y_0. If h_i and J_{ij} denote the input stimulus for neuron i and the synaptic coupling between neurons i and $j(\neq i)$, respectively, the time evolution of u_i is given by

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = -u_i + h_i \tag{2}$$

$$h_i = \sum_{j\neq i} J_{ij} x_j - f_S(x_i) \tag{3}$$

with the self-coupling response function

$$f_S(x) = \begin{cases} J(x - Y_0) & x \geqslant Y_0 \\ J(x + Y_0) & x \leqslant -Y_0 \qquad (J > 0) \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

This equation implies that the self-coupling term acts as a regulator to adjust the neuron activity in the neighbourhood of the threshold activity $\pm Y_0$. Although SCSNA can deal with a certain type of asymmetric synaptic coupling (Shiino and Fukai 1991) we here assume, for the sake of simplicity, that the synaptic coupling is given by the standard Hebb rule, i.e.

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} (\xi_i^{(\mu)} - a)(\xi_j^{(\mu)} - a) \qquad i \neq j \tag{5}$$

where each component of p embedded biased random patterns $\{\xi_i^{(\mu)}\}$ $(\mu = 1, \ldots, p)$ takes on $+1$ and -1 with probabilities $(1 + a)/2$ and $(1 - a)/2$, respectively.

This model can be shown to have an energy function E:

$$E = -\frac{1}{2} \sum_{ij} J_{ij} x_i x_j + \frac{1}{2\beta} \sum_i x_i^2 + \frac{J}{2} \sum_i [(x_i - Y_0)^2 \theta(x_i - Y_0) + (x_i + Y_0)^2 \theta(-x_i - Y_0)]. \tag{6}$$

The condition that equation (6) has a lower bound is easily obtained as follows. We first note that the necessary and sufficient condition for the non-existence of such a diverging motion as $x_i(t) \to \pm\infty$, $|x_j(t)| < Y_0$ $(i = 1, \ldots, rN,\ j = rN + 1, \ldots, N;\ 0 \leqslant r \leqslant 1)$ is that the matrix

$$M_{lm} = -J_{lm} + \delta_{lm}((1/\beta) + J) \qquad l, m = 1, \ldots, rN \tag{7}$$

is positive definite. For the Hebb rule (5), this condition is ensured if inequality $J > -(1/\beta) + r(1-a^2)(1+2\sqrt{\alpha/r})$ holds. We now see that

$$J > -(1/\beta) + (1-a^2)(1+2\sqrt{\alpha}) \tag{8}$$

ensures the existence of a lower bound for equation (6). In this case, the network allows no run-away trajectories and hence will function normally as an associative memory.

The network state evolves so as to decrease the value of the energy function. Since the self-coupling (4) is a unique mechanism in the present system for stabilizing the neuron activity and functions only when the magnitude of output activity is larger than the threshold, it is naturally expected that $\xi_i^{(\mu)} x_i \gtrsim Y_0$ in the retrieval state corresponding to a memory pattern $\{\xi_i^{(\mu)}\}$. In particular for self-coupling $J \to \infty$, $x_i \to Y_0 \xi_i^{(\mu)}$ in the retrieval state. Therefore the state $\approx \{Y_0 \xi_i^{(\mu)}\}$ will be retrieved for certain values of J, β and $\alpha = p/N$.

We would like to note that the present approach using SCSNA developed for neural networks with infinitely many patterns will, in principle, be available to the system in which condition (8) is not obeyed so far as locally stable retrieval states are concerned.

3. Case of finite number of patterns

In the present section, we deal with the case in which the number of pattern p is finite to see how the network functions as a content-addressable memory and how the retrieval state is defined. For the sake of simplicity we assume that the patterns are random and unbiased ($a = 0$). Defining the overlap order parameters $g^{(\nu)}$ ($\nu = 1, \ldots, p$) as

$$g^{(\nu)} = \frac{\beta}{N} \sum_i \xi_i^{(\nu)} u_i \tag{9}$$

one can easily obtain values of u_i in equilibrium states of the network which will be classified into the following three types according to the cases $\beta u_i > Y_0$, $\beta u_i < -Y_0$ and $|\beta u_i| < Y_0$:

$$\begin{aligned} u_i &= \frac{\sum_\nu \xi_i^{(\nu)} g^{(\nu)} + JY_0}{1+\beta J} && \text{if } \beta \sum_\nu \xi_i^{(\nu)} g^{(\nu)} > Y_0 \\ u_i &= \frac{\sum_\nu \xi_i^{(\nu)} g^{(\nu)} - JY_0}{1+\beta J} && \text{if } \beta \sum_\nu \xi_i^{(\nu)} g^{(\nu)} < -Y_0 \\ u_i &= \sum_\nu \xi_i^{(\nu)} g^{(\nu)} && \text{if } \left|\beta \sum_\nu \xi_i^{(\nu)} g^{(\nu)}\right| < Y_0 . \end{aligned} \tag{10}$$

Substituting equation (10) into equation (9) in turn yields the self-consistent equation for $g^{(\nu)}$:

$$g^{(\nu)} = \frac{\beta}{N} \sum_i \xi_i^{(\nu)} \left[\theta(\beta \sum_\mu \xi_i^{(\mu)} g^{(\mu)} - Y_0) \frac{\sum_\lambda \xi_i^{(\lambda)} g^{(\lambda)} + JY_0}{1+\beta J} \right.$$

$$+ \theta(-\beta\sum_\mu \xi_i^{(\mu)} g^{(\mu)} - Y_0)\frac{\sum_\lambda \xi_i^{(\lambda)} g^{(\lambda)} - JY_0}{1+\beta J}$$

$$+ \theta(Y_0 - \beta\sum_\mu \xi_i^{(\mu)} g^{(\mu)})\theta(Y_0 + \beta\sum_\eta \xi_i^{(\eta)} g^{(\eta)})\sum_\lambda \xi_i^{(\lambda)} g^{(\lambda)}\Bigg]. \tag{11}$$

The sum over i will be carried out by introducing the concept of a 'sublattice' due to the finiteness of p (Hemmen 1982, Riedel *et al* 1988, Shiino *et al* 1989, Shiino 1990, Fukai and Shiino 1990b, Domany *et al* 1991). Assuming a p-dimensional vector $\boldsymbol{\xi}$ of the hypercube $\boldsymbol{H}^p = \{-1, 1\}^p$ to represent the sublattice $\{i|\xi_i^{(\nu)} = \xi^{(\nu)}, \nu = 1, \ldots, p\}$, one can rewrite equation (11) as

$$g^{(\nu)} = \beta\sum_{\boldsymbol{\xi}} \in \boldsymbol{H}^p r(\boldsymbol{\xi})\xi^{(\nu)}\Bigg[\theta(\beta\sum_\mu \xi^{(\mu)} g^{(\mu)} - Y_0)\frac{\sum_\lambda \xi^{(\lambda)} g^{(\lambda)} + JY_0}{1+\beta J}$$

$$+ \theta(-\beta\sum_\mu \xi^{(\mu)} g^{(\mu)} - Y_0)\frac{\sum_\lambda \xi^{(\lambda)} g^{(\lambda)} - JY_0}{1+\beta J}$$

$$+ \theta(Y_0 - \beta\sum_\mu \xi^{(\mu)} g^{(\mu)})\theta(Y_0 + \beta\sum_\eta \xi^{(\eta)} g^{(\eta)})\sum_\lambda \xi^{(\lambda)} g^{(\lambda)}\Bigg] \tag{12}$$

where $r(\boldsymbol{\xi})$ is the ratio of the size of sublattice specified by $\boldsymbol{\xi}$ to the total number of neurons N and is taken to be 2^{-p} for each $\boldsymbol{\xi}$ due to the assumption that the patterns are random and unbiased.

We now consider the $p = 2$ case to elaborate on the solution to equation (12). Setting $X(\xi^{(1)}, \xi^{(2)}) = \sum_{\nu=1}^{2} \xi^{(\nu)} g^{(\nu)}$ in equation (12) and noting $X(-1,-1) = -X(1,1)$ etc, one obtains

$$\beta^{-1}X = \theta(\beta X - Y_0)\frac{X + JY_0}{1+\beta J} + \theta(-\beta X - Y_0)\frac{X - JY_0}{1+\beta J} + \theta(Y_0 - \beta X)\theta(Y_0 + \beta X)X$$

with X representing any of $X(1,1)$, $X(1,-1)$, $X(-1,1)$ and $X(-1,-1)$. The solution to this equation other than the trivial one $X = 0$ is given by

$$X = \frac{\pm\beta JY_0}{1+\beta J - \beta} \tag{13}$$

under the condition

$$\frac{\beta^2 J}{1+\beta J - \beta} > 1. \tag{14}$$

Noting $g^{(1)} = (X(1,1) + X(1,-1))/2$ and $g^{(2)} = (X(1,1) - X(1,-1))/2$, one has

$$g^{(1)} = \frac{\pm\beta JY_0}{1+\beta J - \beta} \qquad g^{(1)} = 0$$
$$g^{(2)} = 0 \qquad g^{(2)} = \frac{\pm\beta JY_0}{1+\beta J - \beta} \tag{15}$$

$$|g^{(1)}| = |g^{(2)}| = \frac{1}{2}\frac{\beta JY_0}{1+\beta J - \beta} \tag{16}$$

$$g^{(1)} = g^{(2)} = 0. \tag{17}$$

It is easily seen from equation (15) and inequality (14) that the solution $g^{(1)} = \pm\beta J Y_0/(1+\beta J-\beta)$, $g^{(2)} = 0$ implies $\beta u(1,1) > Y_0$, $\beta u(1,-1) > Y_0$, $\beta u(-1,1) < -Y_0$ and $\beta u(-1,-1) < -Y_0$ when $u(\xi^{(1)},\xi^{(2)})$ stands for the membrane potential of neurons belonging to the sublattice labelled by $\boldsymbol{\xi}$. In other words, this solution represents the retrieval state of the pattern $\{\xi_i^{(1)}\}$ in which the output of neuron i satisfies $\beta u_i \xi_i^{(1)} > Y_0$.

This argument can be easily extended to the cases with $p > 2$ in which the retrieval state corresponding to pattern μ is found to exist only when the condition (14) is satisfied, i.e. $1 - J < \beta^{-1} < 1$, and to be represented by $g^{(\nu)} = \pm\delta^{\mu\nu}\beta J Y_0/(1+\beta J-\beta)$, $\nu = 1,\ldots,p$.

4. Self-consistent signal-to-noise analysis

We now proceed to deal with SCSNA of the present network which is loaded with infinitely many patterns. For the time being, we assume an arbitrary input–output response function f to show the general framework of SCSNA. SCSNA is based on the systematic splitting of the signal and noise components of the local field h_i appearing in the fixed-point equation obtained from equation (2):

$$x_i = f(h_i) = f\left(\sum_{\rho=1}^{p}(\xi_i^{(\rho)} - a)m^{(\rho)} - \alpha(1-a^2)x_i - f_S(x_i)\right) \qquad (18)$$

where $m^{(\mu)} = (1/N)\sum_i(\xi_i^{(\mu)} - a)x_i$ are modified pattern overlaps which yield order parameters to the present system. Consider the retrieval state corresponding to pattern $\{\xi_i^{(1)}\}$, that is $m^{(1)} \sim \mathrm{O}(1)$ and $m^{(\rho)} \sim \mathrm{O}(1/\sqrt{N})$ $(\rho > 1)$. SCSNA should be distinguished from conventional signal-to-noise analysis (McEliece *et al* 1987, Peretto 1988, Amari and Maginu 1988, Domany *et al* 1989) in which $\sum_{\rho>1}(\xi_i^{(\rho)} - a)m^{(\rho)}$ is viewed as a Gaussian noise with mean 0. Such a naive treatment of the noise component proves to be incorrect although it seemingly recovers the result of the replica calculation by Amit *et al* (1985). In SCSNA, one can properly deal with the previously mentioned term to extract a systematic contribution which is proportional to output x_i.

Fixed-point equation (18) implies that x_i can be formally solved as a function of $\sum_{\rho=1}^{p}(\xi_i^{(\rho)} - a)m^{(\rho)}$.

$$x_i = \mathcal{F}\Big(\sum_{\rho=1}^{p}(\xi_i^{(\rho)} - a)m^{(\rho)}\Big). \qquad (19)$$

To work with SCSNA, we first evaluate the pattern overlap for noise-generating patterns $\{\xi_i^{(\rho)}\}$ $(\rho > 1)$ which can be written as

$$m^{\rho} = \frac{1}{N}\sum_i(\xi_i^{(\rho)} - a)\mathcal{F}\left((\xi_i^{(\rho)} - a)m^{(\rho)} + \sum_{\mu\neq\rho}(\xi_i^{(\mu)} - a)m^{(\mu)}\right). \qquad (20)$$

By expanding the right-hand side of this expression with respect to $m^{(\rho)}$, we can obtain

$$m^{\rho} = \frac{1}{1-C'}\frac{1}{N}\sum_i(\xi_i^{(\rho)} - a)\mathcal{F}\left(\sum_{\mu\neq\rho}(\xi_i^{(\mu)} - a)m^{(\mu)}\right) \qquad (21)$$

where

$$C' = \frac{1}{N}\sum_i (\xi_i^{(\rho)} - a)^2 \mathcal{F}'\Big(\sum_{\mu\neq\rho}(\xi_i^{(\mu)} - a)m^{(\mu)}\Big)$$
$$= (1 - a^2)\Big\langle\!\Big\langle \mathcal{F}'\Big(\sum_{\mu\neq\rho}(\xi_i^{(\mu)} - a)m^{(\mu)}\Big)\Big\rangle\!\Big\rangle \tag{22}$$

with $\langle\!\langle\cdots\rangle\!\rangle$ denoting averaging over biased memory patterns.

Now these formal expressions for the pattern overlaps are substituted into the noise-generating term $\Omega \equiv \sum_{\mu>1}(\xi_i^{(\mu)} - a)m^{(\mu)}$ to yield

$$\Omega = \frac{1}{1-(1-a^2)C}\frac{1}{N}\sum_{\rho>1}(\xi_i^{(\rho)} - a)\sum_{j=1}^{N}(\xi_j^{(\rho)} - a)\mathcal{F}\Big(\sum_{\mu\neq\rho}(\xi_i^{(\mu)} - a)m^{(\mu)}\Big) \tag{23}$$

where $C = \langle\!\langle\mathcal{F}'\rangle\!\rangle$. To extract a systematic part from Ω, we split the j-summation in (10) into $j \neq i$ and $j = i$ terms. Since the former is a sum of almost uncorrelated random variables with mean 0, it will be assumed to yield a noise z

$$z = \frac{1}{1-(1-a^2)C}\frac{1}{N}\sum_{\rho>1}(\xi_i^{(\rho)} - a)\sum_{j\neq i}(\xi_j^{(\rho)} - a)\mathcal{F}\Big(\sum_{\mu\neq\rho}(\xi_i^{(\mu)} - a)m^{(\mu)}\Big) \tag{24}$$

which obeys a Gaussian distribution with mean 0 and variance

$$\sigma^2 = \langle\!\langle z^2\rangle\!\rangle = \frac{\alpha(1-a^2)q}{[1-(1-a^2)C]^2}. \tag{25}$$

Here we defined $q = \langle\!\langle\mathcal{F}^2\rangle\!\rangle$ which serves as the Edwards–Anderson order parameter for the present system. On the other hand, the $j = i$ term yields a non-vanishing systematic part:

$$\frac{\alpha(1-a^2)}{1-(1-a^2)C}\mathcal{F}. \tag{26}$$

Noting that local field is given by $h_i = (\xi_i^{(1)} - a)m^{(1)} + \Omega - \alpha(1-a^2)x_i - f_S(x_i)$, we now obtain from equations (18), (19) and (26) the following equation that determines the output response $Y \equiv \mathcal{F}$ implicitly as a function of noise z:

$$Y = f\Big((\xi - a)m + z + \frac{\alpha(1-a^2)^2 C}{1-(1-a^2)C}Y - f_S(Y)\Big) \tag{27}$$

where we have renamed $\xi_i^{(1)}$ and $m^{(1)}$ as ξ and m, respectively. Now the separation of the signal and noise components is properly done since equation (27) completely determines the dependence of the response on signal $(\xi - a)m$ and noise z.

This expression reveals that local field h_i obeys a non-Gaussian distribution due to the nonlinearity involved in equation (27) although noise z has been assumed to obey a Gaussian distribution. The appearance of the systematic part proportional to output Y in the local field of equation (27) is characteristic of analogue neural networks and is not observed in stochastic neural networks like the Boltzmann machine. This is because that part is cancelled out by the Onsager reaction term appearing in the TAP equation for the Boltzmann machine. The extraction of the systematic part from noise component in the local field h_i has been performed in order to define the effective output Y in a self-consistent manner, as seen from equations (20)–(27). For this reason, the present method may be called SCSNA.

Using the definitions of m, q and C and assuming that the average $\langle\langle\cdots\rangle\rangle$ can be replaced by the average over noise z and condensed pattern ξ, one obtains what is usually called saddle-point equations in the neural network theory:

$$m = \left\langle \int_R \mathrm{D}_\sigma z (\xi - a) Y(z) \right\rangle \tag{28}$$

$$q = \left\langle \int_R \mathrm{D}_\sigma z Y(z)^2 \right\rangle \tag{29}$$

$$\sigma C = \left\langle \int_R \mathrm{D}_\sigma z \frac{z}{\sigma} Y(z) \right\rangle \tag{30}$$

$$\mathrm{D}_\sigma z = \frac{\mathrm{d}z}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right)$$

where $\langle\cdots\rangle$ stands for the average over ξ and σ is determined by equation (25). Solving this set of equations for m, q and C, we obtain phase boundaries for the analogue neural networks: retrieval phase ($m \neq 0,\ q \neq 0$), spin glass phase ($m = 0,\ q \neq 0$) and disordered phase ($m = 0,\ q = 0$).

5. Phase boundaries

The SCSNA explored in the previous section can be applied straightforwardly to the present neural network to obtain

$$m = \left\langle \int_R \mathrm{D}z (\xi - a) Y(z) \right\rangle \tag{31}$$

$$q = \left\langle \int_R \mathrm{D}z Y(z)^2 \right\rangle \tag{32}$$

$$\sqrt{\frac{q}{\alpha}} \frac{v}{1 - a^2} = \left\langle \int_R \mathrm{D}z\, z Y(z) \right\rangle \tag{33}$$

$$Y = \beta((\xi - a)m + \sigma z + vY - f_S(Y)) \tag{34}$$

$$\sigma = \sqrt{\frac{q}{\alpha}}(v + \alpha(1 - a^2)) \tag{35}$$

$$\mathrm{D}z = \frac{\mathrm{d}z}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}z^2\right)$$

from (28)–(30) after a change of variable $v = \alpha(1 - a^2)^2 C/(1 - (1 - a^2)C)$ and a rescaling $z/\sigma \to z$ of the Gaussian integration variable. In general equation (34) allows more than one solution. For such a case, we assume that the available solution can be selected by applying the Maxwell rule as will be shown below. This recipe was justified in the case of symmetric analogue neural networks with energy functions bounded from below where the saddle-point method in the replica calculation ensures the validity of the Maxwell rule. We list all the possible cases to be considered, although only case A occurs in the numerical analysis of equations (31)–(33) within the parameter regions which we will study later. Note that we have assumed $\beta,\ J > 0$.

Case A ($\beta^{-1} - v > 0$) *and case B* ($\beta^{-1} - v < -J$). The solution of equation (34) is uniquely determined for any values of z and is given by

$$\begin{aligned} Y &= \frac{\sigma z + (\xi - a)m + JY_0}{\beta^{-1} - v + J} \qquad z > z_{\xi+} \\ Y &= \frac{\sigma z + (\xi - a)m - JY_0}{\beta^{-1} - v + J} \qquad z < z_{\xi-} \\ Y &= \frac{\sigma z + (\xi - a)m}{\beta^{-1} - v} \qquad \text{otherwise} \\ z_{\xi\pm} &= \frac{\pm(\beta^{-1} - v)Y_0 - (\xi - a)m}{\sigma} \end{aligned} \tag{36}$$

for case A. The inequality signs should be reversed for case B.

Case C ($0 > \beta^{-1} - v > -J$). In a certain range of z, multiple solutions are allowed for equation (34) as shown in figure 1 where the solutions are given by nodes of the two curves representing both sides of the equation. Then the available solution will be given by the one (marked by black dots) determined by the Maxwell rule which states that the solution corresponding to the nodes associated with a larger area enclosed by the two curves should be more stable than the others. Therefore we adopt a solution with negative Y when $z < z_* \equiv -(\xi - a)/\sigma$ and the one with positive Y when $z > z_*$. The value of Y jumps from $-Y_*$ to Y_* ($Y_* \equiv JY_0/(\beta^{-1} - v + J) > 0$) at $z = z_*$. Thus

$$\begin{aligned} Y &= \frac{\sigma z + (\xi - a)m + JY_0}{\beta^{-1} - v + J} \qquad z > z_* \\ Y &= \frac{\sigma z + (\xi - a)m - JY_0}{\beta^{-1} - v + J} \qquad z < z_*. \end{aligned} \tag{37}$$

The phase boundary between para and spin glass phases is easily obtained by linear analysis around $m = q = 0$. Noting that a solution with vanishing q can be allowed only in case A, and $z_{\xi\pm} \approx \pm JY_0/\sigma \to \pm\infty$, we can easily perform the Gaussian integrations involved in equations (32) and (33) to obtain

$$1 = \frac{[v + \alpha(1 - a^2)]^2}{\alpha(\beta^{-1} - v)^2} \tag{38}$$

$$\frac{v}{1 - a^2} = \frac{v + \alpha(1 - a^2)}{\beta^{-1} - v}. \tag{39}$$

From equations (38) and (39), we derive the equation that yields the disordered spin glass phase boundary:

$$\beta^{-1} = (1 - a^2)(1 + 2\sqrt{\alpha}) \tag{40}$$

which is identical to that for the analogue neural network with a sigmoid response function if β^{-1} is interpreted as the analogue gain in the sigmoid. This coincidence is reasonable since the boundary between the para and spin glass phases is determined only by the behaviour of the response function near the origin.

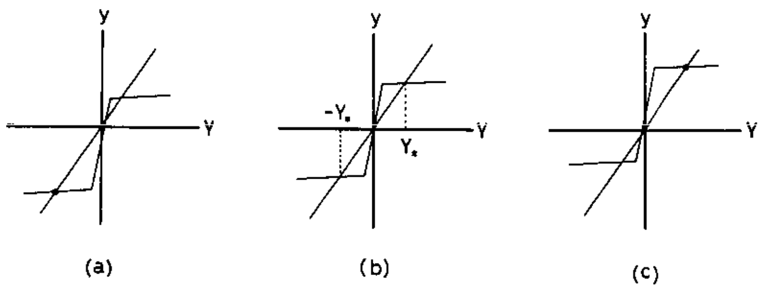

Figure 1. Diagrammatical explanation of the Maxwell rule for selecting a suitable solution of equation (34). The straight line stands for $y = \beta^{-1}Y$ and the piecewise linear curve for $y = vY - f_S(Y) + (\xi - a)m + \sigma z$. The figures are for the cases that (*a*) $z < z_*(\equiv -(\xi - a)m/\sigma)$, (*b*) $z = z_*$ and (*c*) $z > z_*$, respectively. In (*a*) and (*b*), as marked by black dots, the solutions corresponding to the nodes associated with a larger area which is enclosed by the two curves should be adopted. The value of Y jumps from $-Y_*$ to $Y_*(\equiv JY_0/[\beta^{-1} - v + J])$ at $z = z_*$.

It is necessary to solve equations (31)–(33) numerically to study the retrieval phase of the system. We deal with the case of unbiased random patterns, i.e. $a = 0$ for the sake of simplicity. Note that we can set $Y_0 = 1$ in equations (31)–(33) without any loss of generality by rescaling the variables as $Y/Y_0 \to Y$, $m/Y_0 \to m$, $\sigma/Y_0 \to \sigma$, $q/Y_0^2 \to q$. The critical storage capacity is shown in figure 2(*a*) as a function of β^{-1}, which may play a similar role to the temperature parameter in the Boltzmann machine, for various values of self-coupling J. The storage capacity of the present neural network increases with an increasing J. It, however, does not exceed 0.138, the maximal storage value corresponding to the vanishing temperature limit of the Boltzmann machine or to the infinitely large gain limit of the analogue neural network with a sigmoid response function. It is also seen that no retrieval solution exists for $\beta^{-1} > 1$, just as neither the Boltzmann machine nor the analogue network with sigmoid response exhibits a retrieval phase when the temperature or the inverse analogue gain is greater than unity. A characteristic feature of the present neural network is that there exists a finite optimal value for the analogue gain β that maximizes the storage capacity at a given value of J. Note that the results are consistent with the condition (14), that is $1 - J < \beta^{-1} < 1$, which will be interpreted as the phase boundary at the limit $\alpha \to 0$.

It is also noted that condition (8) is satisfied in the whole region of the retrieval phase for all values of J. Therefore no run-away trajectories appear in the dynamical evolution of the present neural network model if the number of stored patterns is less than the critical storage level $\alpha_c N$.

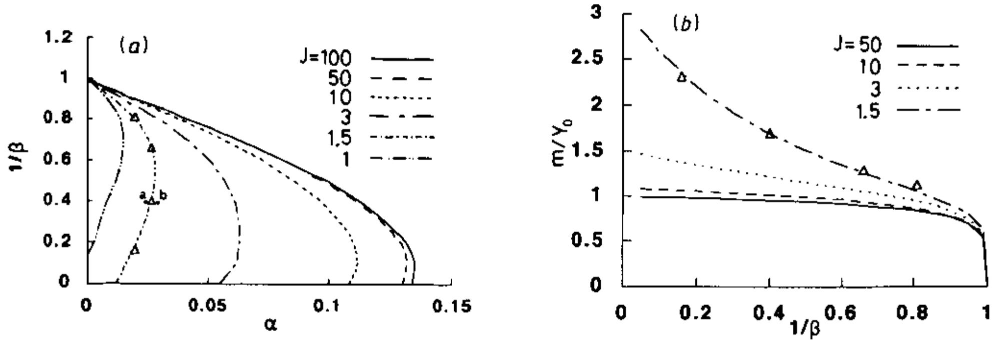

Figure 2. The results of the SCSNA for (a) critical storage capacities and (b) pattern overlaps at the critical storage levels shown as functions of the inverse analogue gain β^{-1} for various values of self-coupling J. Triangular points plot the simulation results for $J = 1.5$ on the network with $N = 3000$, which are in reasonable agreement with the SCSNA results. For the parameters α and β which are marked by the dots 'a' and 'b', typical features of the time evolution of the pattern overlap are shown respectively in figures 3(a) and 3(b).

The retrieval-phase solutions of equations (31)–(33) for the overlap m/Y_0 are shown in figure 2(b). It is noted that the values of the overlap obtained by SCSNA are quite consistent with those estimated in section 2 for the case of finite number of patterns (see equation (15) and the paragraph below it). For instance, the overlaps for the three cases that $J = 1.5$ and $\beta = 0.2, 0.5$ and 0.8 are obtained respectively as $m/Y_0 = 2.21$, 1.51 and 1.07 in SCSNA, while they are estimated as 2.14, 1.50 and 1.15 from the finite-pattern result $m/Y_0 = J/(\beta^{-1} + J - 1)$ (for $a = 0$, $m^{(\mu)} = \beta^{-1}g^{(\mu)}$).

Since the output of each neuron is not bounded by unity, m/Y_0 takes an arbitrarily large number. Therefore the quantity, which is a natural order parameter of the present system, is inappropriate for measuring the retrieval quality. Then it will be useful to introduce quantity Q_r which represents the quality of retrival: $Q_r \equiv$ ratio of the number of neurons with $x_i\xi_i^{(1)}/Y_0 > 1$ to total number N. Once the output $Y(z)$ is obtained, Q_r can be easily calculated by $Q_r = \int \mathrm{D}z\theta(\xi Y(z)/Y_0 - 1)$. It is found that Q_r does not sensitively depend on J and becomes almost unity in the retrieval phase.

We have conducted a numerical simulations on a neural network with 3000 neurons to observe the dynamical behaviour of the present model and to compare it with the result obtained using SCSNA. In figure 3(a), we show examples of the time evolution of the pattern overlap started with various initial values for the case allowing retrieval states. It is seen that the initial pattern overlap $m(0)/Y_0$ should be larger than a certain threshold value in order for the network to retrieve a memory pattern when the storage level is lower than the corresponding critical storage capacity (Amari and Maginu 1988). As the storage level increases, so does the threshold. When, on the other hand, the storage level is higher than the critical storage capacity, retrieval hardly takes place for any values of the initial pattern overlap (figure 3(b)). In the phase diagram shown in figure 2(a), the two cases in figures 3(a) and 3(b) are marked by the dots 'a' and 'b', respectively, with both points satisfying condition (8). The storage capacity which was estimated by the simulations as the most probable value of α at which the system ceases to possess a threshold value for some cases with $J = 1.5$ is plotted by triangles in figure 2(a). The simulation results are reasonably

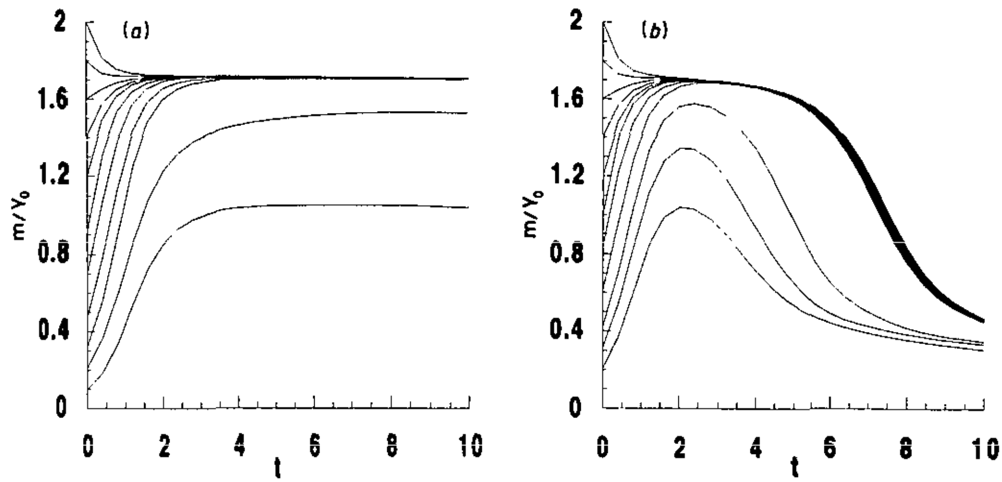

Figure 3. Examples of simulations of the retrieval process of the present network system with $N = 3000$ when (a) $\alpha \lesssim \alpha_c$ and (b) $\alpha \gtrsim \alpha_c$. The parameters used are $a = 0$, $\beta^{-1} = 0.4$ and $J = 1.5$. The storage levels, (a) $\alpha = 0.025$ and (b) $\alpha = 0.3$, are a little less and larger, respectively, than the critical storage capacity $\alpha_c = 0.027$ which is obtained by SCSNA for $\beta^{-1} = 0.4$ and $J = 1.5$. The two cases are marked by the dots 'a' and 'b' in the phase diagram shown in figure 2(a). For the initial states with $m(0)/Y_0 < 1$, $x_i(0)/Y_0$ is either equal to $\xi_i^{(1)}$ or $-\xi_i^{(1)}$ and the probability of finding $x_i(0)/Y_0 = \xi_i^{(1)}$ is $(1 + m(0)/Y_0)/2$, while for those with $m(0)/Y_0 > 1$, $x_i(0) = \xi_i^{(1)} m(0)/Y_0$ for arbitrary i. As in the conventional Hopfield network, the system appears to possess a threshold value for the initial pattern overlap $m(0)/Y_0$ which yields pattern retrieval.

consistent with those of SCSNA. In particular we emphasize that the values of m and Q_r in the retrieval states obtained by the simulation are in good agreement with those obtained by SCSNA.

6. Concluding remarks

We have presented an analogue neural network model in which the response function is a simple linear function of the membrane potential unlike the conventional sigmoid model. Self-coupling was introduced as an example of a regulation mechanism which relies on local feedback neuronal circuits in order to inhibit neuron activities exceeding a threshold level. The model constructed this way with symmetric synaptic coupling defines a system with an energy function which has no lower bound in certain parameter regions. To confirm the validity of the recently developed SCSNA (self-consistent signal-to-noise analysis), which avoids replica calculations and thus is very simple, we calculated the storage capacity or the phase boundaries for the present network system. The results were compared with those obtained by computer simulations on large-size neural networks and exhibited satisfactory agreement.

The SCSNA results show that the storage capacity increases for larger self-coupling but never exceeds the well known value $\alpha = 0.138$ for a conventional Hopfield neural network. It is noted that a finite optimal value for the linear analogue gain that maximizes the storage capacity at a given finite value of the self-coupling exists. This will suggest, in general, that the analogue gain should be controlled within a suitable range if neural networks utilize local feedback circuits to regulate neuron activities.

Although the present neural network system could possibly exhibit run-away solutions due to an energy function unbounded from below, all the retrieval states were found to satisfy the network stability condition (8). We note that, as far as the equilibrium properties of the network are concerned, the present system will be equivalent to an analogue neural network with a transfer function $f_{\text{eff}}(u)$, where $f_{\text{eff}}^{-1}(x) = \beta^{-1}x + f_S(x)$, which will work normally under the same condition (8). In view of this fact, the availability of the SCSNA should be reasonable in the present system. It may be of interest to study whether SCSNA is still valid when dealing with stable retrieval states of network systems without stability conditions such as equation (8).

Acknowledgments

This work is supported by Grant-in-Aid for Encouragement of Young Scientists (no 03740205) and Grant-in-Aid for General Scientific Research (no 02460289). The authors express their sincere thanks to Kazuo Akashi and the Asahi Corporation for technical support on the numerical computation.

References

Amari S 1989 *Neural Networks* **2** 451
Amari S and Maginu K 1988 *Neural Networks* **1** 63
Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530
—— 1987 *Phys. Rev.* A **35** 2283
Domany E, Hemmen J L and Schulten K (eds) 1991 *Models of Neural Networks* (Berlin: Springer)
Domany E, Kinzel W and Meir R 1989 *J. Phys. A: Math. Gen.* **22** 2081
Fukai T and Shiino M 1990a *Phys. Rev.* A **42** 7459
—— 1990b *Phys. Rev. Lett.* **64** 1465
—— 1992 *J. Phys. A: Math. Gen.* **25** 2873
Hemmen J L 1982 *Phys. Rev. Lett.* **49** 409
Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
—— 1984 *Proc. Natl Acad. Sci. USA* **81** 3088
Kühn R, Bös S and Hemmen J L 1991 *Phys. Rev.* A **43** 2084
Marcus C M, Waugh F R and Westervelt R M 1990 *Phys. Rev.* A **41** 3355
Marcus C M and Westervelt R M 1989 *Phys. Rev.* A **40** 501
McEliece R J, Posner E C, Rodemich E R and Venkatesh S S 1987 *IEEE Trans. Inform. Technol.* **IT-33** 461
Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
Miyashita Y and Chang H 1988 *Nature* **331** 68
Morita M, Yoshizawa S and Nakano K 1990 *Trans. IEICE* **J73-DII** 232
Shiino M, Nishimori H and Ono M 1989 *J. Phys. Soc. Japan* **58** 763
Peretto P 1988 *J. Physique* **49** 711
Riedel U, Kühn R and Hemmen J L 1988 *Phys. Rev.* A **38** 1105
Shiino M 1990 *J. Stat. Phys.* **59** 1051
Shiino M and Fukai T 1990 *J. Phys. A: Math. Gen.* **23** L1009
—— 1992 *J. Phys. A: Math. Gen.* **25** L375
Thouless D J, Anderson P W and Palmer R G 1977 *Philos. Mag.* **35** 593
Treves A 1990 *J. Phys. A: Math. Gen.* **23** 2631
Treves A and Amit D J 1989 *J. Phys. A: Math. Gen.* **22** 2205
Tsodyks M V and Feigel'man M V 1988 *Europhys. Lett.* **6** 101
Waugh F R, Marcus C M and Westervelt R M 1990 *Phys. Rev. Lett.* **64** 1986