Onset of 'super retrieval phase' and enhancement of the storage capacity in neural networks of nonmonotonic neurons

1993 J. Phys. A: Math. Gen. 26 L831

(http://iopscience.iop.org/0305-4470/26/17/014)

J. Phys. A: Math. Gen. 26 (1993) L831–L841. Printed in the UK

LETTER TO THE EDITOR

Onset of 'super retrieval phase' and enhancement of the storage capacity in neural networks of non-monotonic neurons

M Shiino† and T Fukai‡

† Department of Applied Physics, Tokyo Institute of Technology, Ohokayama, Meguro-ku, Tokyo, Japan
‡ Department of Electronics, Tokai University, Kitakaname 1117, Hiratsuka-shi, Kanagawa, Japan

Received 10 February 1993

Abstract. Analogue neural networks of associative memory with continuous time dynamics are studied for non-monotonic transfer functions using the method of self-consistent signal-to-noise analysis. The Hebb learning rule with unbiased random patterns is assumed for the synaptic couplings. A novel phenomenon is found to occur as a result of a phase transition concerning the property of the local field distribution. In retrieval states of the newly found phase which we refer to as the super retrieval phase, noise in the local fields vanishes and the memory retrieval without errors ensues even for an extensive number of memory patterns stored under the local learning rule. The storage capacity is obtained as a function of the parameter representing the degree of non-monotonicity of the transfer functions, with the result that an enhancement of the storage capacity can also occur.

Theory of associative memory neural networks with symmetric synaptic interactions [1–4] has made great progress amidst an intense effort by researchers to understand the collective behaviour of systems of interconnected neurons on the basis of statistical neurodynamics [5–10] and statistical mechanics of phase transitions in spin glasses [11–15]. Statistical mechanics makes full use of the existence of an energy function which not only plays the role of a Lyapunov function ensuring stability of the network system, but also allows an equilibrium distribution of the Gibbs type.

Analogue neural networks with deterministic continuous-time dynamics [4, 16–23] which are characterized by transfer functions representing input–output relations of graded-response neurons do not always satisfy the condition for the existence of an energy function and thus are, in general, expected to give rise to a variety of behaviours including rich dynamical phenomena. When a transfer function is monotonically increasing as in the sigmoidal one, which is often assumed in the study of associative memory models of neural networks, an energy function is allowed to exist for the case of symmetric synaptic interactions and enables one to analyse the storage capacity using a statistical mechanical approach [16, 18, 19]. For instance, in the case of the hyperbolic-tangent transfer function the behaviour of the storage capacity with change in the analogue gain was shown [16, 18, 19] to be qualitatively the same as that of a stochastic network of formal two-state neurons, i.e. Ising spin neurons, with change in the inverse temperature. If, on the other hand, the restriction of the monotonicity of the transfer functions is removed, one can easily see that the network has no longer a Lyapunov function even when it works as an associative memory.

One may be allowed to consider effective transfer functions that describe overall input–output relations of possible local clusters of neurons serving as functional units in the information processing of physiological nervous systems. Such effective transfer functions will take a variety of shapes and happen to exhibit non-monotonic behaviour.

The self-consistent signal-to-noise analysis (SCSNA) [22] we have recently developed is applicable to analogue networks with arbitrary transfer functions [23]. So the method is powerful particularly for networks without Lyapunov functions, such as the networks having non-monotonic transfer functions, since it substantially assumes nothing but the existence of fixed point type attractors for the dynamics of the networks. The SCSNA properly takes into account the correlation between components of the non-condensed patterns and the output of a neuron in the retrieval state with respect to the condensed pattern, and self-consistently splits the local field into three parts by means of a kind of renormalization procedure; signal, pure noise, and output proportional terms [22]. One of the characteristic features of the SCSNA is a clear cut explanation, based on the role of the output proportional term, for a non-Gaussian distribution of the local fields which manifests itself in a pronounced manner in the case of non-monotonic transfer functions.

The aim of the present letter is to report the results of the SCSNA applied to a continuous-time analogue network with a certain type of non-monotonic transfer function, which shows not only an enhancement of the storage capacity but also the remarkable phenomenon of the onset of 'super retrieval phase' associated with a non-Gaussian property of the local field distribution.

The super retrieval phase refers to a phase where noise in the local fields of neurons in the retrieval states vanishes even with an extensive number of stored patterns. It has, to date, been believed that networks which store an extensive number of patterns through local learning rules cannot be free from a finite fraction of errors in their retrieval states, which increase as loading rate is increased [13]. In the non-local learning rule such as the pseudo-inverse rule, on the other hand, perfect memory recall with extensively many patterns has been shown to be possible as a result of the vanishing of noise in the local fields of neurons [24].

The novel phase found in the present network having non-monotonic transfer functions appears due to a phase transition in which the width of the local field distribution for the retrieval state implied by the SCSNA-order parameter equations approaches 0 at a certain critical value α_0 as the loading rate α is decreased. Below the critical value, the standard type of retrieval state with finite width of the local field distribution cannot exist any more and a different type of retrieval state with the vanishing noise comes into existence. The occurrence of such retrieval states implies that memory retrieval without errors ensues even in the case of the local learning rule of the Hebb type.

We begin by presenting the model network of non-monotonic analogue neurons together with a brief outline of the SCSNA, the original version of which was given in

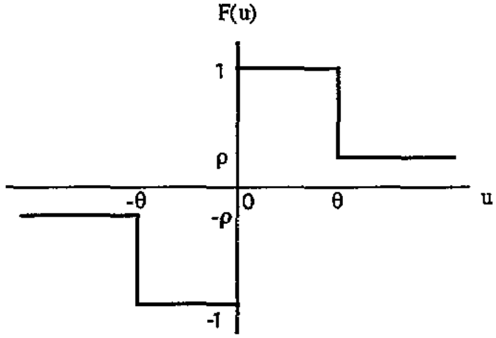

Figure 1. Transfer function representing input–output relation of a neuron.

[22]; its reformulated version elucidating the self-consistent character of the method has been detailed in [25].

The time evolution of the network of N neurons is assumed to be given by

$$\frac{\mathrm{d}}{\mathrm{d}t} u_i = -u_i + \sum_{j=1}^{N} J_{ij} F(u_j) \qquad i = 1, \ldots, N \tag{1}$$

with u_i representing membrane potential of neuron i and J_{ij} synaptic connections with the Hebb learning rule

$$J_{ij} = \frac{1}{N} \sum_{\substack{\mu=1 \\ J_{ii}=0}}^{p} \xi_i^{(\mu)} \xi_j^{(\mu)} \tag{2}$$

where $\{\xi_i^{(\mu)}\}$ ($\mu = 1, \ldots, p, i = 1, \ldots, N$) denote p sets of uncorrelated random patterns with $\Pr\{\xi_i^{(\mu)} = \pm 1\} = \frac{1}{2}$.

We deal with the transfer function F specified with the two parameters θ and ρ as shown in figure 1:

$$\begin{aligned} F(u) &= \operatorname{sgn} u \qquad & |u| < \theta \\ &\rho \operatorname{sgn} u \qquad & |u| > \theta \end{aligned} \tag{3}$$

which is non-monotonic when $\rho < 1$. The case of $\rho = 0$ has been studied in great detail elsewhere [25] to observe the effect of cutting off output activity of neurons on the behaviour of the storage capacity. Decreasing θ below $\theta = 1$ has been found to yield the onset of super retrieval phase for $\alpha < \alpha_0(\theta)$ as well as an appreciable enhancement of the storage capacity. In the present paper we want to confine ourselves to studying the effect of changing the other parameter ρ which can be considered to control the degree of non-monotonicity of the transfer function.

Assuming the existence of equilibrium solutions to (1), we will be concerned with a set of equations

$$x_i = F\left(\sum_j J_{ij} x_j \right) \qquad i = 1, \ldots, N \tag{4}$$

for output of neurons ($x_i \equiv F(u_i)$) to obtain the retrieval solution with respect to pattern 1 of the form

$$m^{(1)} = O(1) \qquad m^{(\mu)} = O(1/\sqrt{N}) \qquad \text{for } \mu \geqslant 2 \tag{5}$$

with $m^{(\mu)}$ standing for the order-parameter overlaps

$$m^{(\mu)} = \frac{1}{N} \sum_j \xi_j^{(\mu)} F(u_j) \qquad \mu = 1, \ldots, p. \tag{6}$$

The SCSNA extracts pure noise part in the local field h_i of neuron i which is written as

$$h_i \equiv \sum_j J_{ij} x_j = \xi_i^{(1)} m^{(1)} + \xi_i^{(\mu)} m^{(\mu)} + \sum_{\nu \neq \mu, 1} \xi_i^{(\nu)} m^{(\nu)} - \alpha x_i \qquad \mu \neq 1 \tag{7}$$

by assuming the decomposition [25]

$$\sum_{\nu \neq \mu, 1} \xi_i^{(\nu)} m^{(\nu)} = \overline{z_{i\mu}} + \gamma x_i \qquad \mu \geqslant 2 \tag{8}$$

where $\overline{z_{i\mu}}$ represents pure noise. The spirit of the SCSNA is in the self-consistent determination of $\overline{z_{i\mu}}$ and γ [22, 25]. We note that the so-called conventional treatment implies $\gamma = \alpha$, giving rise to 0 mean of the naive noise

$$\sum_{\nu \geqslant 2} \sum_{j \neq i} \frac{1}{N} \xi_i^{(\nu)} \xi_j^{(\nu)} x_j.$$

Equation (4) then takes the form

$$x_i = F(\xi_i^{(1)} m^{(1)} + \xi_i^{(\mu)} m^{(\mu)} + \overline{z_{i,\mu}} + \Gamma x_i), \tag{9}$$

with

$$\Gamma = \gamma - \alpha \tag{10}$$

in which we refer to Γx_i as the output proportional term. After some manipulations for the renormalization procedure presented in [22, 25], it follows that γ is given as

$$\gamma = \frac{\alpha}{1 - U} \tag{11}$$

and the noise $\overline{z_{i\mu}}$ is almost independent of μ and i, obeying an identical Gaussian distribution with mean U and variance σ^2, where U and σ are self-consistently determined by [22, 25]

$$U = \left\langle \left\langle \frac{\partial x_i}{\partial z_{i\mu}} \right\rangle \right\rangle \tag{12}$$

$$\sigma^2 = \frac{\alpha}{(1 - U)^2} \langle\langle x_i^2 \rangle\rangle \tag{13}$$

with x_i being given by the solution of (9) and $\langle\langle \ \rangle\rangle$ representing average over $\xi_i^{(1)}$ and Gaussian noise $\overline{z_{i\mu}}$. Equations (12) and (13) together with the one for the pattern overlap $m^{(1)}$

$$m^{(1)} = \langle\langle \xi_i^{(1)} x_i \rangle\rangle \tag{14}$$

constitute a fundamental set of equations for the SCSNA.

Since the transfer function (3) is odd ($F(-u)=-F(u)$), taking average over $\xi_i^{(1)}$ turns out to be equivalent to simply setting $\xi_i^{(1)}=1$ in the above equations. Further setting $\sigma^2=\alpha r$ together with $z=\tilde{z}/\sigma$ one can rewrite the set of equations in a more familiar form [22, 25]:

$$m=\int_{-\infty}^{\infty} dz \frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}} Y(z) \tag{15a}$$

$$(1-U)^2 r=\int_{-\infty}^{\infty} dz \frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}} Y(z)^2 \tag{15b}$$

$$U\sqrt{\alpha r}=\int_{-\infty}^{\infty} dz \frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}} zY(z) \tag{15c}$$

where the renormalized output $Y(z)$ is given by solving the equation:

$$Y(z)=F(m+\sqrt{\alpha r}\,z+\Gamma Y(z)) \tag{15d}$$

with $\Gamma=\alpha(U/1-U)$. Since the above equation with F given by (3) admits more than one solution, we have to resort to use of the Maxwell rule to pick up the available solution. Note that the application of the rule can be ensured in the case of monotonic transfer functions within the context of the saddle point method of replica symmetric theory [16].

The resultant Y can take several different types of shape as a function of z according to the value of Γ, which is determined by solving (15). We depict one of them obtained for the case of $-(2\theta/1-\rho)<\Gamma<0$ which covers a wide region of parameter space of interest:

$$Y(z)=-\rho(z<z_1) \qquad Y(z)=-1(z_1<z<z_2) \qquad Y(z)=\frac{-m-\sqrt{\alpha r}z}{\Gamma}(z_2<z<z_3)$$

$$Y(z)=1(z_3<z<z_4) \qquad Y(z)=\rho(z_4<z)$$

where

$$z_1=\frac{-\theta+\frac{1+\rho}{2}\Gamma-m}{\sqrt{\alpha r}} \qquad z_2=\frac{\Gamma-m}{\sqrt{\alpha r}}$$

$$z_3=\frac{-\Gamma-m}{\sqrt{\alpha r}} \qquad z_4=\frac{\theta-\frac{1+\rho}{2}\Gamma-m}{\sqrt{\alpha r}}. \tag{16}$$

The retrieval state is given by the solution of (15) with $m\neq 0$ which satisfies both the condition that the corresponding fixed point of the dynamics (1) be stable and a certain requirement concerning the degree of the pattern retrieval in associative memory. The requirement will be specified by $g \approx 1$ with the tolerance overlap g being defined by [25]

$$g = \frac{1}{N} \sum_i \xi_i \operatorname{sgn} u_i \tag{17}$$

This quantity measures quality of retrieval better than the order-parameter overlap m in the case of networks with transfer functions which differ much from the sigmoidal ones. In other words, it should be understood that pattern retrieval with a general type of transfer function will require the definition of the overlap as given by (17).

We show in figure 2 the plot of m satisfying $g \approx 1$ against α (thick line) which was obtained by numerically solving (15) under the condition $-(2\theta/1-\rho) < \Gamma < 0$ for $\theta = 0.4$ and $\rho = 0.2$. Note that here the standard type of retrieval solution with $m \neq 0$ and $r \neq 0$ is allowed to exist only for a certain interval of α unlike the common case of the Hopfield model with sigmoidal transfer functions. We denote its upper and lower bounds by $\tilde{\alpha}_c$ and α_0, respectively, although the existence of the latter is not always ensured but depends on the values of θ and ρ.

If the retrieval solutions obtained turn out stable attractors of (1) for α up to $\alpha = \tilde{\alpha}_c$, the $\tilde{\alpha}_c$ should coincide with the storage capacity of the network. However, it is not the case for the network with the parameters depicted, as is shown in the same figure which also displays some $m(\alpha)$ points obtained from the results of successful retrieval in numerical simulations with $N = 200$–300 for various values of α. Whereas the data points obtained are seen to fit well the theoretical $m(\alpha)$ curve, few data points showing successful retrieval were collected for $\alpha > \approx 0.38$, because starting with any initial conditions almost failed to make the network state settle into the expected retrieval state with $g \approx 1$. This implies that the theoretically obtained retrieval solutions with α ranging from $\tilde{\alpha}_c$ down to a certain value, 0.38 in this case, lose their

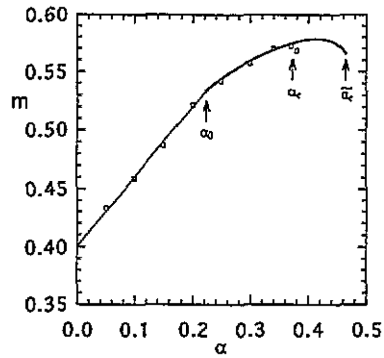

Figure 2. Plots of the order-parameter overlap m against the loading rate α obtained by the SCSNA (solid lines) and by numerical simulations (square) with $N = 200$–300 for $\theta = 0.4$ and $\rho = 0.2$. The thick line which has the end points at $\alpha = \alpha_0$ and $\alpha = \tilde{\alpha}_c$ represents the standard type of retrieval states with $r \neq 0$ given by the solutions to (15). The thin line representing the relation $m = \theta + (1+\rho)\alpha/2$ gives the super retrieval states with $r = 0+$ (see text). It terminates at the end point ($\alpha = \alpha_0$) of the thick line. Squares denote the results for successful retrieval with $g \approx 1$, showing excellent agreement with the result of the SCSNA. The storage capacity α_c is determined to be the upper bound for α yielding successful retrieval.

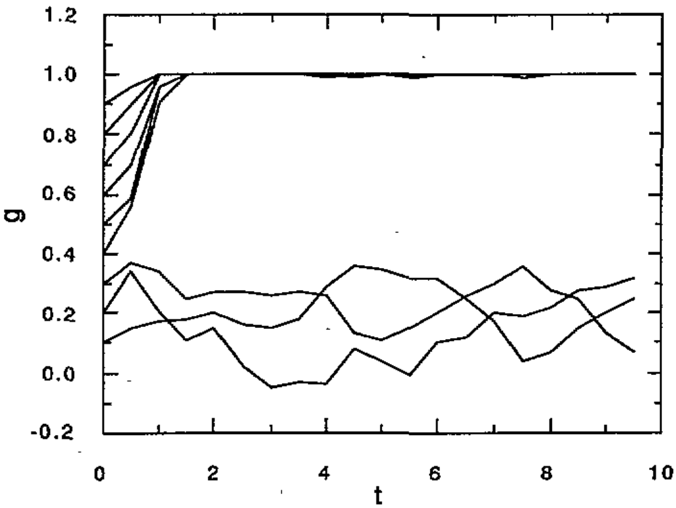

Figure 3. Time evolution of the tolerance overlap g for several initial conditions for $\theta=0.4$, $\rho=-0.2$, and $\alpha=0.3$.

stability. The occurrence of instability in general includes that of oscillatory instability, since the present network has no Lyapunov function. Then the storage capacity α_c has to be given by the value of α corresponding to the stability limit. One has in general the storage capacity lower than $\tilde{\alpha}_c$ ($\alpha_c \leqslant \tilde{\alpha}_c$). Note however that the value of α_c itself still gets enhanced relative to the commonly known value ≈ 0.14 of the Hopfield model with sigmoidal transfer functions.

We present in figure 3 the time course of the tolerance overlap g in the retrieval process obtained in numerical simulations with $N=300$ for $\theta=0.4$, $\rho=-0.2$ and $\alpha=0.3$. It can be seen that the network with the given parameters, if started with an initial condition ensuring an appropriately large overlap, settles into the retrieval state specified with $g\approx 1$ after some time. This implies that the present system works properly as an associative memory.

Let us turn to the problem of the appearance of the lower bound α_0 in figure 2. The critical point α_0 comes into existence in such a way that with loading rate α approaching α_0 from above, r tends to 0, and below α_0 no standard type of retrieval solution can exist. Recall that $\sigma^2=\alpha r$ represents variance of the noise in the local fields. Interestingly, what is happening to the system below α_0 is the disappearance of noise in the local fields of neurons. The clue to this singular phenomenon caused by the anomalous behaviour of $r\to 0+$ of the standard type of retrieval solution can be obtained by examining (15) and (16) in the limit $r\to 0+$.

We first note that in this limit $|U|\to\infty$ and hence $\Gamma\to -\alpha$. Since in addition we know by solving numerically (15) that while $z_1, z_2, z_3 \to -\infty$, z_4 remains finite, it follows

$$m\to\theta+\frac{(1+\rho)\alpha}{2} \qquad (r\to 0). \tag{18}$$

In figure 2 the straight line $m=\theta+(1+\rho)\alpha/2$ is also drawn (thin line) showing that the curve representing the standard type of retrieval solution m of (15) just disappears upon crossing it at $\alpha=\alpha_0$. Note then that the critical value α_0 turns out to satisfy the equations which are obtained by taking the limit $r\to 0+$ on (15) together with (18):

$$\begin{aligned} &\alpha\left\{\frac{1+\rho^2}{2}+(1-\rho^2)N[0,z_4]\right\}=(1-\rho)^2\frac{e^{-z_4^2}}{2\pi} \\ &\frac{1+\rho}{2}+(1-\rho)N[0,z_4]=\theta+\frac{\alpha(1+\rho)}{2} \end{aligned} \tag{19}$$

where

$$N[X,Y]=\int_X^Y dz\,\frac{1}{\sqrt{2\pi}}e^{-z^2/2}.$$

Regarding the behaviours of the networks for $\alpha\leqslant\alpha_0$, we will have no other choice than admitting that $m=\theta+(1+\rho)\alpha/2$ holds true with $r=0+$, because the $r\neq 0$ solution of the SCSNA cannot exist any longer. The validity of the proposition has been confirmed by numerical simulations on the networks of an appropriately large number of neurons. Indeed, the results of numerical simulations depicted in figure 2 display excellent agreement with the claimed relation between m and α below α_0, which lends credit to the establishment of $r=0+$.

The disappearance of noise in the local fields below α_0 can be more directly confirmed by observing the local field distribution $P(h)$, which is expected to take a non-Gaussian form due to the presence of the output proportional term ΓY in the renormalized local field. Figures 4(a) and 4(b), respectively, show the profiles of $P(h)$ obtained by numerical simulations on the networks with $\theta=0.4$ and $\alpha=0.2$ for $N=300$ in the cases of $r=0+$ state ($\rho=0.2$) and $r\neq 0$ state ($\rho=0.5$). Focusing one's attention only to either half of the $P(h)$ because of the expected symmetry $P(h)=P(-h)$, one sees a pair of peaks separated from each other by a gap which is not so clearly visible in figure 4(b). The gap in the local field distribution can easily be understood, based on the SCSNA, to arise in connection with solving (15d) for $Y(z)$

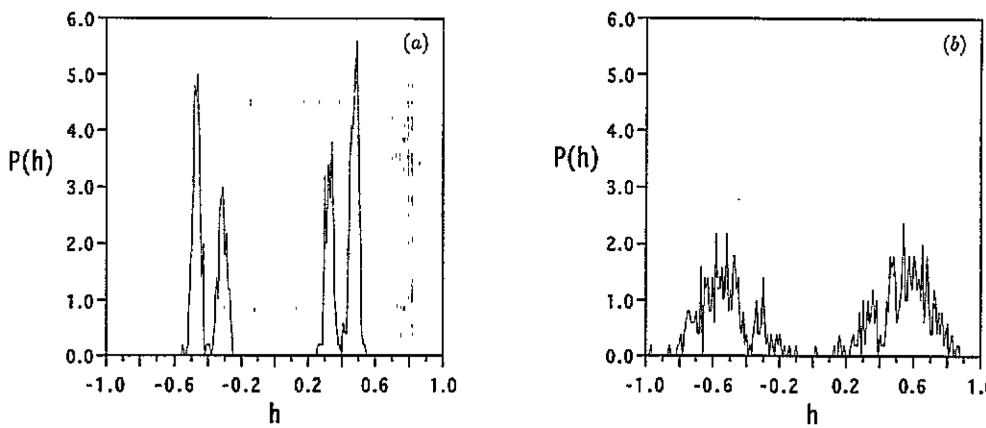

Figure 4. Local field distributions $P(h)$ of neurons in the retrieval states obtained by numerical simulations with $N=300$ for $\theta=0.4$ and $\alpha=0.2$. (a) $\rho=0.2$ (super retrieval state); (b) $\rho=0.5$. The non-Gaussian distributions ($h>0$) are seen to exhibit a pair of peaks separated from each other by the gap whose size is $(1-\rho)|\Gamma|$ according to the theory. The value for Γ is (a) $-\alpha=-0.2$ and (b) -0.0976.

with use of the Maxwell rule: the local fields are forbidden to take their values in the interval of width $(1-\rho)|\Gamma|$, $\theta-(1-\rho)/2\,|\Gamma|<|h|<\theta+(1-\rho)/2|\Gamma|$. One also notices the sharp contrast between the two profiles of $P(h)$ with the same loading rate expecially in terms of the width of each peak of the distributions: the one for $\rho=0.2$, indeed, becomes extremely narrowed in accordance with $r=0+$, though it is seen to remain finite due probably to finite size effect. The profiles of $P(h)$ obtained by numerical simulations even with $N=200$–500 for various values of the parameters ρ, θ, and α are in good agreement with the theoretical results based on the SCSNA [25], which are though not shown here.

The appearance of the gap in the local field distribution showing a clear evidence for its non-Gaussian property will differ, in character, from such a gap as inferred by the order-parameter equations (15) with $\Gamma>0$ for the sigmoidal transfer functions with sufficiently high analogue gain, which are considered to exhibit replica symmetry–breaking instability leading to possible breakdown of the self-averaging property assumed in the SCSNA. In fact, the problem of the gap for high analogue gain with $\Gamma>0$ has been discussed recently by Kühn and Bös in the light of replica symmetry breaking [26]. The gap which occurs with $\Gamma<0$ in the present network having a non-monotonic transfer function and is related to the negative slope of the transfer function with

$$\left|\frac{\mathrm{d}}{\mathrm{d}Y}F^{-1}(Y)\right|<\left|\Gamma\right|$$

however, should be free from the matter of replica symmetry breaking because of the good agreement with the results of numerical simulations supporting the presence of the gap in the local field distribution.

We refer to the noiseless state with $\alpha\leqslant\alpha_0$ as $r=0+$ state or super retrieval state to distinguish it from the normal retrieval state representing the standard type of retrieval state with $r\neq 0$. The onset of the super retrieval state will be surprising, because it can be easily proved that $g=1$ exactly follows ensuring memory retrieval without errors in spite of the presence of an extensive number of stored patterns with the Hebb learning rule. In the vanishing of the width of the local field distribution characterizing the new phase, the presence of an output proportional term in the renormalized local fields of a neuron, which arises from the so-called renormalization procedure of the SCSNA, plays a key role together with a prescription of the Maxwell rule appled to the case of non-monotonic transfer functions of a certain type. Indeed, the relation $m=\theta+(1+\rho)\alpha/2$ obtained under the limit $r\to 0+$ is a direct consequence of both $\Gamma\to-\alpha$ and the occurrence of the discontinuous jump at $z=z_4$ of $Y(z)$ resulting from the use of the Maxwell rule. Note that in the case of a finite number of stored patterns ($p<\infty$) with $\Gamma=0$ and $r=0$, one has $m=\theta$ when $\rho<\theta$. Furthermore, simply setting $\Gamma=-\alpha$ and $r=0$ in (15*a*) without using the Maxwell rule only yields $\theta/(1-\alpha)$ for $\rho<\theta$, apart from the stability problem. For this reason one must distinguish between $r=0$ and $r=0+$. This makes a sharp contrast with the case of stochastic networks with the pseudo-inverse rule where $r=0$ is just a solution to the order-parameter equations describing the retrieval states [24]. Furthermore, it is noted that in our case setting $r=0+$ cannot be consistent with all of the three equations (15*a*), (15*b*), and (15*c*), the first two of which concern the order-parameter overlap m and spin glasss order-parameter q respectively, and can be shown to hold [25].

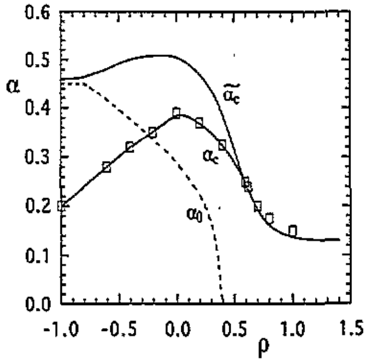

Figure 5. Phase diagram on the $\rho-\alpha$ plane for $\theta=0.4$ displaying the super retrieval phase ($0<\alpha<\alpha_0$) and the enhancement of the storage capacity. The storage capacity α_c represented by the solid curve with square is given as the upper bound for α ensuring successful memory recall in numerical simulations.

We may say that the phenomenon of the vanishing noise occurs as a result of the condensation of the so-called naive noise in the local fields for α below α_0, since Γ does not vanish but takes a finite value of $-\alpha$ while noise z vanishes ($\sigma=0+$). The naive noise which is generated by the non-condensed patterns has so far been treated as Gaussian noise with mean 0 in a conventional signal-to-noise analysis-type approach to the problem of determining the storage capacity. Our present results have revealed that such treatment for the naive noise indeed is invalid because of the correlation between components of the non-condensed patterns and the output of a neuron in the retrieval state.

The phase diagram on the ρ–α plane showing the dependence of α_c, $\tilde{\alpha}_c$, and α_0 on ρ for fixed θ ($=0.4$) is given in figure 5. The storage capacity α_c has been determined by the numerical simulations as the stability limit of the SCSNA-solution and α_0 by solving (19). Note that $\rho=1$ represents the Hopfield model with $F(u)=\mathrm{sgn}\,(u)$ yielding $\alpha_c=\tilde{\alpha}_c\approx 0.14$. As the parameter ρ controlling the degree of non-monotonicity of the transfer function is decreased, $\tilde{\alpha}_c$ is seen to increase until it attains a broad maximum of ≈ 0.51. The relation $\alpha_c=\tilde{\alpha}_c$, however, holds only down to $\rho\approx 0.6$ and thereafter α_c starts to deviate from $\tilde{\alpha}_c$ due to the occurrence of instability of the SCSNA solution. The α_c takes its maximum at around $\rho\approx 0$ and decreases with a further decrease in ρ. Concerning the behaviour of α_0, we see α_0 start to increase from $\alpha_0=0$ at $\rho=\theta$ ($=0.4$) to get intercepted by the decreasing curve of $\alpha_c(\rho)$, as ρ is decreased. The super retrieval phase is then represented by the region below both the $\alpha_c(\rho)$ curve ($\rho<\approx -0.25$) and the $\alpha_0(\rho)$ curve ($-0.25<\rho\leqslant 0.4$), where the relation $m=\theta+(1+\rho)\alpha/2$ holds. The condition $\rho\leqslant\theta$ for the occurrence of the super retrieval phase is likely to be related with the fact that under such condition the network with a finite number of patterns ($\alpha=0$) yields $m=\theta$. In this connection, it is worth noting that the super retrieval phase can exist under the condition $\theta<1$ [25].

In conclusion, we have shown that the continuous-time alalogue network with the non-monotonic transfer function (3) yields remarkable features of the enhancement of the storage capacity and appearance of the super retrieval phase, both of which contribute to improving the network performances of associative memory with the Hebb learning rule. Those phenomena found here will be generic and be expected for a certain class of transfer functions. In fact, using the systematic method of the SCSNA we can show that in most cases networks with non-monotonic transfer functions give rise to an enhancement of the storage capacity more or less to the same extent as in the present case under no self-coupling condition. Such an enhancement of the storage capacity was first noted by Morita *et al* [27] in their study based on numerical simulations. We have also studied elsewhere that introduction of an appropriate amount of self-couplings can lead to a further enhancement of the storage capacity [25]. With regard to the newly discovered phase, i.e. super retrieval phase, now that the mechanism underlying the vanishing of noise in the local fields has turned out to be closely related with the appearance of jumps in the renormalized output $Y(z)$ which arises from use of the Maxwell rule, one can easily expect the novel phase to occur for transfer functions with sufficiently steep negative slopes. It should be noted that the appearance of jumps in the original transfer function $F(u)$ as in the present model is not necessarily required for the onset of the super retrieval phase [25].

This work is partially supported by Grant-in-Aid for General Scientific Research (02460289) and Grant-in-Aid for Encouragement of Young Scientist (03740205) from the Ministry of Education.

References

[1] Amari S 1972 *IEEE Trans.* **C-21** 1197
[2] Little W A 1974 *Math. Biosci.* **19** 101
[3] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[4] Hopfield J. J. 1984 *Proc. Natl Acad. Sci. USA* **81** 3088
[5] Weisbuch G and Fogelmann-Soulie F 1985 *J. Physique Lett.* **46** L-623
[6] McEliece R J, Posner E C, Rodemich E R and Venkatesh S S 1987 *IEEE Trans.* **IT-33** 461
[7] Gardner E, Derrida B and Mottishaw P 1987 *J. Physique* **48** 741
[8] Amari S and Maginu K 1988 *Neural Networks* **1** 63
[9] Domany E, Kinzel W and Meir R 1989 *J. Phys. A: Math. Gen.* **22** 2081
[10] Patrick A E and Zagrebnov V A 1991 *J. Phys. A: Math. Gen.* **24** 3413
[11] Peretto P 1984 *Biol. Cybern.* **50** 51
[12] Kinzel W 1985 *Z. Phys.* B **60** 205
[13] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530; 1987 *Ann. Phys.* **173** 30
[14] Gardner E 1986 *J. Phys. A: Math. Gen.* **19** L1047
[15] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257
[16] Shiino M and Fukai T 1990 *J. Phys. A: Math. Gen.* **23** L1009
[17] Fukai T and Shiino M 1990 *Phys. Rev.* A **42** 7459
[18] Kühn R 1990 *Statistical Mechanics of Neural Networks* ed Garrido Springer Lecture Notes in Physics 398 (Heidelberg: Springer)
[19] Kühn R, Bös S and van Hemmen J L 1991 *Phys. Rev.* A **43** 2084
[20] Marcus C M, Waugh F R and Westervelt R M 1990 *Phys. Rev.* A **41** 3355
[21] Waugh F R, Marcus C. M. and Westervelt R M 1991 *Phys. Rev.* A **43** 3131
[22] Shiino M and Fukai T 1992 *J. Phys. A: Math. Gen.* **25** L375
[23] Fukai T and Shiino M 1992 *J. Phys. A: Math. Gen.* **25** 4799
[24] Kanter I and Sompolinsky H 1987 *Phys. Rev.* A **35** 380
[25] Shiino M and Fukai T 1992 Tokyo Institute of Technology Tokai University *preprint*
[26] Kühn R and Bös S 1992 Universitat Heidelberg *preprint* Nr 683
[27] Morita M, Yoshizawa S and Nakano K 1990 *Trans. IEICE* **J73-DII** 232