PHYSICAL REVIEW E VOLUME 57, NUMBER 2 FEBRUARY 1998

Random and systematic dilutions of synaptic connections in a neural network with a nonmonotonic response function

Masato Okada
Kawato Dynamic Brain Project, Japan Science and Technology Corporation, 2-2 Hikaridai Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Tomoki Fukai
Department of Electronics, Tokai University, Kitakaname 259-12, Hiratsuka, Kanagawa 259-12, Japan
and Laboratory for Neural Modeling, FRP, Institute of Physical and Chemical Research (RIKEN),
Hirosawa 2-1, Wako, Saitama 350-01, Japan

Masatoshi Shiino
Department of Applied Physics, Tokyo Institute of Technology, Ohokayama, Meguroku, Tokyo 152, Japan

(Received 6 June 1997)

It has been observed that the dilution of synaptic connections in neural networks has relevance to biology and applicability to engineering. From this viewpoint, the effects of synaptic dilution on the retrieval performance of an associative memory model with a nonmonotonic response function are investigated through the self-consistent signal-to-noise analysis. Compared with a fully connected neural network, for which a nonmonotonic response function is known to achieve a large enhancement of storage capacity and the occurrence of the superretrieval phase leads to an errorless memory retrieval, the nonmonotonic neural network with a random synaptic dilution undergoes a considerable decrease in storage capacity. It is shown, however, that by employing a systematic dilution technique characterized by a nonlinear learning rule, in which larger connections are retained, it is possible to significantly reverse the undesirable rapid reduction in storage capacity. It is also proved that the superretrieval phase is structurally unstable against the dilution of synapses. [S1063-651X(98)04801-6]

PACS number(s): 87.10.+e, 89.70.+c, 05.90.+m

INTRODUCTION

Single neurons display monotonic increases in firing rates as input currents are increased. However, there are many sources of inhibition in biological neural networks, and the effective neural processing units, which may consist of a small number of excitatory and inhibitory neurons, possibly exhibit more complicated response profiles to external stimuli. To explore the potential abilities of such neural systems in biological information processing, nonmonotonic response functions were introduced to associative memory neural networks [1,2]. In the models, an output of each processing unit, or ‘‘neuron,’’ decreases (increases) for relatively large (small) values of membrane potential rather than showing a sigmoidal response profile.

Several noteworthy features were found in the nonmonotonic neural networks, for which the existence of a Lyapunov function ensuring network stability is not guaranteed. When standard Hebbian learning is used to define a connection matrix with random memory patterns, the storage capacity increases to about three times [2,3] that of the case where a step function [4] or a sigmoidal response function is used [4–6]. Such an enhancement also occurs for sparsely coded memory patterns [7]. If synaptic connections are optimally learned, the storage capacity is enhanced by a similar numeral factor [8] compared with an optimally learned monotonic neural network [9]. Furthermore, in a certain equilibrium state termed ‘‘superretrieval phase,’’ the inputs (or local fields) of individual neurons do not suffer from noise components [2,10] even under extensive memory loading. In the monotonic neural networks, the noise arising from the interference by nonretrieved memory patterns prevents errorless memory retrieval. The occurrence of the superretrieval phase is robust against the introduction of biased memory patterns, asymmetric synaptic connections memorizing the presynaptic and postsynaptic activity, and the profile of transfer functions including positive-valued ones [11]. However, the extent of the superretrieval phase in the phase diagrams depends significantly on the profiles of nonmonotonic response functions [12]. Nonmonotonic neurons, which are responsible for a considerable reduction in the number of spurious states, are also useful for the retrieval of temporal sequences [13] and practical applications to real-world pattern recognition by an electrically implemented neural networks [14].

Since information about a memory pattern is retained by many synaptic connections, the performance of associative memory models is robust against the damage to neural circuits. In fact, the problems of synaptic dilution and nonlinear learning rules such as clipping of synapses have been studied for monotonic response functions in stochastic Ising spin networks, including the zero-temperature case [15–17]. If synaptic connections are eliminated at random while the connections’ symmetry is maintained, the storage capacity drops almost linearly with the number of eliminated synapses, implying that a neural network can still work as an efficient associative memory [15]. Diluting synapses asymmetrically was investigated analytically only in the case of extreme dilution [16,18].

It is tempting to assume that the behavior of nonmonotonic neural networks with synaptic dilution resembles that of monotonic neural networks, and accordingly their superior abilities would be expected to endure despite such damage. In this paper, we analytically and numerically examine how the dilution of synaptic connections affects the performance of a nonmonotonic neural network. We employ the self-consistent signal-to-noise analysis (SCSNA) to investigate the equilibrium properties of the neural network. Contrary to the above expectation, the results show that the excellent abilities exhibited by nonmonotonic transfer functions suffer considerably from a random dilution compared with the performance of a randomly diluted monotonic neural network. In particular, the superretrieval phase is proved to be structurally unstable for any type of synapse dilution. Consequently, we show how the deterioration of the retrieval abilities of nonmonotonic neural network can be countered by retaining large synaptic connections in the dilution.

SCSNA for diluted neural networks

The SCSNA starts from the fixed-point equations for the dynamics of an N-neuron network:

$$h_i=\sum_{j\neq i}^{N} J_{ij}F(h_j), \quad i=1,\ldots,N, \tag{1}$$

or more conveniently

$$x_i=F\left(\sum_{j\neq i}^{N} J_{ij}x_j\right), \quad i=1,\ldots,N, \tag{2}$$

where h_i and $x_i=F(h_i)$ are the local field of the ith neuron at equilibrium and output activity of neuron i, respectively. J_{ij} stands for the synaptic connection from neuron j to i. In this study, the response function is assumed to be given by

$$F(x)=\begin{cases} 1, & 0<x<\theta \\ -1, & -\theta<x<0 \\ 0, & |x|>\theta \end{cases} \tag{3}$$

for which the properties of a fully connected nonmonotonic neural network were extensively studied [2]. The parameter θ can be regarded as a cutoff parameter for the membrane potential. Random memory patterns are generated according to the probability distribution

$$P(\xi_i^\mu=\pm 1)=\tfrac{1}{2}. \tag{4}$$

In the following, we derive a set of equations for order parameters when the synaptic connections are randomly diluted. This analysis is valid for a generic response function other than that given in Eq. (3).

Random dilution

The synaptic connections are given by

$$J_{ij}=\frac{\sqrt{p}\,c_{ij}}{Nc}\,T_{ij}, \tag{5}$$

where T_{ij} is a conventional Hebbian connection matrix

$$T_{ij}=\frac{1}{\sqrt{p}}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \tag{6}$$

and c_{ij} takes its value as 0 or 1 with probabilities

$$P(c_{ij}=1)=1-P(c_{ij}=0)=c \quad (0\leqslant c\leqslant 1), \tag{7}$$

$$c_{ij}=c_{ji} \tag{8}$$

for symmetric dilution. For asymmetric dilution, c_{ij} and c_{ji} are independently determined by the same probability distribution. Moreover we can analyze a more general case, where c_{ij} and c_{ji} have an arbitrary correlation such that

$$\mathrm{Cov}(c_{ij},c_{ji})=k_c\mathrm{Var}(c_{ij})=k_c c(1-c). \tag{9}$$

In this general case, the symmetric and asymmetric dilutions correspond to $k_c=1$ and $k_c=0$, respectively. Defining the loading level as $\alpha=p/N$, and noting that Eq. (5) is written as

$$J_{ij}=\frac{1}{N}\sum_{\mu=1}^{\alpha N} \xi_i^\mu \xi_j^\mu+\frac{c_{ij}-c}{Nc}\sum_{\mu=1}^{\alpha N} \xi_i^\mu \xi_j^\mu, \tag{10}$$

we can rewrite the local field h_i for neuron i as

$$h_i=\sum_{\mu=1}^{\alpha N} \xi_i^\mu m_\mu+\frac{1}{Nc}\sum_{\mu=1}^{\alpha N}\sum_{j\neq i}^{N} (c_{ij}-c)\xi_i^\mu \xi_j^\mu x_j-\alpha x_i, \tag{11}$$

where m_μ is the overlap between the stored pattern $\boldsymbol{\xi}^\mu$ and the equilibrium state $\boldsymbol{x}$,

$$m_\mu=\frac{1}{N}\sum_{i=1}^{N} \xi_i^\mu x_i. \tag{12}$$

By substituting Eq. (11) into Eq. (2), we can easily see that x_i can be formally represented as a function of $\Sigma_{\mu=1}^{\alpha N}\xi_i^\mu m_\mu+(1/Nc)\Sigma_{\mu=1}^{\alpha N}\Sigma_{j\neq i}^{N}(c_{ij}-c)\xi_i^\mu\xi_j^\mu x_j$. Thus we obtain

$$x_i=\tilde{F}\left(\sum_{\mu=1}^{\alpha N} \xi_i^\mu m_\mu+\frac{1}{Nc}\sum_{\mu=1}^{\alpha N}\sum_{j\neq i}^{N} (c_{ij}-c)\xi_i^\mu \xi_j^\mu x_j\right), \tag{13}$$

with a certain function $\tilde{F}(x)$ to be determined later.

Let $\{\xi_i^1\}$ be the target pattern to be retrieved. Therefore, we can assume that $m_1=O(1)$ and $m_\mu=O(1/\sqrt{N})(\mu>1)$. Then we can use the Taylor series expansion to obtain

$$\begin{aligned} m_\mu&=\frac{1}{N}\sum_{i=1}^{N} \xi_i^\mu \tilde{F}\left(\sum_{\nu=1}^{\alpha N} \xi_i^\nu m_\nu+\frac{1}{Nc}\sum_{\nu=1}^{\alpha N}\sum_{j\neq i}^{N} (c_{ij}-c)\xi_i^\nu \xi_j^\nu x_j\right) \\ &=\frac{1}{N}\sum_{i=1}^{N} \xi_i^\mu x_i^{(\mu)}+Um_\mu+\frac{1}{Nc}\sum_{i=1}^{N} x_i'^{(\mu)} \\ &\quad\times\frac{1}{N}\sum_{j\neq i}^{N} (c_{ij}-c)\xi_i^\mu \xi_j^\mu x_j \end{aligned} \tag{14}$$

for $\mu>1$, where

$$x_i^{(\mu)}=\tilde{F}\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_i^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{j\neq i}^{N}(c_{ij}-c)\xi_i^\nu\xi_j^\nu x_j\right), \tag{15}$$

$$x_i'^{(\mu)}=\tilde{F}'\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_i^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{j\neq i}^{N}(c_{ij}-c)\xi_i^\nu\xi_j^\nu x_j\right), \tag{16}$$

$$U=\frac{1}{N}\sum_{i=1}^{N}x_i'^{(\mu)}. \tag{17}$$

If the averages over $\xi_i^\mu(\mu>1)$ and c_{ji} are taken in the right-hand side (rhs) of Eq. (14), the last term vanishes. To show this, the ξ_i^μ and c_{ji} dependences of x_j are extracted from x_j before averaging is performed:

$$x_j=x_j^{(\mu)(c_{ji})}+\left[\xi_j^\mu m_\mu+\frac{1}{Nc}\sum_{k\neq i,j}^{N}(c_{jk}-c)\xi_j^\mu\xi_k^\mu x_k+\frac{1}{Nc}\sum_{\nu\neq\mu}(c_{ji}-c)\xi_j^\nu\xi_i^\nu x_i+\frac{1}{Nc}(c_{ji}-c)\xi_j^\mu\xi_i^\mu x_i\right]x_j'^{(\mu)(c_{ji})}, \tag{18}$$

where

$$x_j^{(\mu)(c_{ji})}=\tilde{F}\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_j^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{k\neq i,j}^{N}(c_{jk}-c)\xi_j^\nu\xi_k^\nu x_k\right), \tag{19}$$

$$x_j'^{(\mu)(c_{ji})}=\tilde{F}'\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_j^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{k\neq i,j}^{N}(c_{jk}-c)\xi_j^\nu\xi_k^\nu x_k\right). \tag{20}$$

Substituting Eq. (18) into the r.h.s. of Eq. (14) and averaging the resultant expressions over ξ_i^μ and c_{ji} shows that the last term vanishes. Thus, m_μ can be expressed as

$$\begin{aligned}m_\mu&=\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu x_i^{(\mu)}+Um_\mu\\&=\frac{1}{N(1-U)}\sum_{i=1}^{N}\xi_i^\mu x_i^{(\mu)}\end{aligned} \tag{21}$$

for $\mu>1$. Similarly, by using Eq. (18), we can show that the second term in Eq. (11) is expressed as

$$\frac{1}{Nc}\sum_{\mu=1}^{\alpha N}\sum_{j\neq i}^{N}(c_{ij}-c)\xi_i^\mu\xi_j^\mu x_j^{(\mu),(c_{ji})}+k_c\frac{\alpha(1-c)}{c}Ux_i. \tag{22}$$

Equations (21) and (22) give the following expression for the local field:

$$\begin{aligned}h_i={}&\xi_i^1 m_1-\alpha x_i+k_c\frac{\alpha(1-c)}{c}Ux_i+\frac{1}{N(1-U)}\sum_{\mu=2}^{\alpha N}\tilde{F}\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_i^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{k\neq i}^{N}(c_{ik}-c)\xi_i^\nu\xi_k^\nu x_k\right)\\&+\frac{1}{Nc}\sum_{\mu=1}^{\alpha N}\xi_i^\mu\sum_{j\neq i}^{N}(c_{ij}-c)\xi_j^\mu\tilde{F}\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_j^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{k\neq i,j}^{N}(c_{jk}-c)\xi_j^\nu\xi_k^\nu x_k\right)\\&+\frac{1}{N(1-U)}\sum_{\mu=2}^{\alpha N}\xi_i^\mu\sum_{j\neq i}^{N}\xi_j^\mu\tilde{F}\left(\sum_{\nu\neq\mu}^{\alpha N}\xi_j^\nu m_\nu+\frac{1}{Nc}\sum_{\nu\neq\mu}^{\alpha N}\sum_{k\neq j}^{N}(c_{jk}-c)\xi_j^\nu\xi_k^\nu x_k\right).\end{aligned} \tag{23}$$

Note that the second term in Eq. (23) arises from the ith term of the summation over neurons, which results from substituting Eq. (21) into the "naive noise" term $\Sigma_{\mu=2}^{\alpha N}\xi_i^\mu m_\mu$. In the last term in Eq. (23), we expand $\tilde{F}$ with respect to $(c_{ji}-c)$. Then the lowest order term yields a noise term that has a vanishing mean. The higher order terms can be dropped from Eq. (23), since they only yield vanishing means, variances, and cross-talk correlations with the third term, which is another noise term. Thus, we finally obtain

$$h_i=\xi_i^1 m_1+\sigma z_i+\left[\frac{1}{1-U}+k_c\frac{1-c}{c}\right]\alpha Ux_i, \tag{24}$$

$$\sigma^2=\frac{\alpha q}{(1-U)^2}+\frac{\alpha(1-c)}{c}q, \tag{25}$$

where the noise terms in Eq. (23) are replaced by the Gaussian noise σz_i, with z_i obeying the normal distribution $\mathcal{N}(0,1)$.

Finally, the results of the SCSNA for the symmetric dilution are summarized by the following order-parameter equations:

$$m=\left\langle\int_{-\infty}^{\infty}Dz\,\xi Y(z;\xi)\right\rangle_\xi, \tag{26}$$

$$q=\left\langle\int_{-\infty}^{\infty}Dz\,Y(z;\xi)^2\right\rangle_\xi, \tag{27}$$

$$U=\left\langle\int_{-\infty}^{\infty}Dz\,zY(z;\xi)\right\rangle_\xi, \tag{28}$$

$$Dz\equiv dz\,\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right),$$

where $\langle\cdots\rangle_\xi$ implies averaging over the target pattern. The effective response function $Y(z;\xi)$, or $\tilde{F}(x)$, is obtained implicitly from $F(x)$ by solving

$$Y(z;\xi)=F\left(\xi m+\sigma z+\left[\frac{1}{1-U}+k_c\,\frac{1-c}{c}\right]\alpha U Y(z;\xi)\right). \tag{29}$$

It is remarked that the order parameter U is well defined for nondifferential response functions in the final expression given in Eq. (28). The order-parameter equations for the asymmetric dilution can be obtained simply by omitting the c-dependent term in Eq. (29).

Sompolinsky showed that adding symmetric noise to synaptic connections

$$J_{ij}=\frac{1}{N}\sum_{\mu=1}^{\alpha N}\xi_i^\mu\xi_j^\mu+\delta_{ij}, \tag{30}$$

$$\delta_{ij}\sim\mathcal{N}\left(0,\frac{\Delta^2}{N}\right),\quad \delta_{ij}=\delta_{ji}, \tag{31}$$

is equivalent to diluting the connections symmetrically [15]. We note that such an equivalence can also be proved by the SCSNA in the general cases covered by Eq. (9), including the asymmetric dilution case to which the conventional replica calculation is not applicable. In the general case, the variances Δ^2 of the synaptic noise equivalent to the dilutions are found to be

$$\Delta^2=\frac{\alpha(1-c)}{c}, \tag{32}$$

$$k=k_c,\quad \text{where}\quad \mathrm{Cov}(\delta_{ij},\delta_{ji})=k\,\frac{\Delta^2}{N}, \tag{33}$$

from Eqs. (24) and (25).

Systematic dilution

The systematic dilution of synaptic connections can be achieved by introducing synaptic noise with an appropriate nonlinear function $f(x)$ [15]:

$$J_{ij}=\frac{\sqrt{p}}{N}f(T_{ij}). \tag{34}$$

Note that T_{ij} obeys the normal distribution $\mathcal{N}(0,1)$ for $p=\alpha N\to\infty$ [in the finite loading case, i.e., $p\sim O(1)$, another treatment is needed [15,19,20]]. If we define $f(x)$ as

$$f(x)=\begin{cases}x, & |x|>\delta\\ 0, & |x|\leqslant\delta,\end{cases} \tag{35}$$

the connections in the range $[-\delta,\delta]$ are eliminated (we call this case "bottom-cut-off dilution") [21]. On the other hand, if

$$f(x)=\begin{cases}x, & |x|<\delta\\ 0, & |x|\geqslant\delta,\end{cases} \tag{36}$$

the connections in the ranges $(-\infty,-\delta]$, $[\delta,\infty)$ are eliminated ("top-cut-off dilution"). The cutting rate is easily expressed in terms of δ: e.g.,

$$R_c=\int_{-\delta}^{\delta}Dx, \tag{37}$$

for the bottom-cut-off dilution. We note that for the above two cases

$$\int_{-\infty}^{\infty}Dx f(x)=0, \tag{38}$$

which is assumed in the following analysis.

As a first step, we apply a naive signal-to-noise (S/N) analysis to the nonlinear learning rule Eq. (34) in order to estimate the signal and noise terms. Assuming that the state $\{x_i\}$ with components $x_i=\xi_i^1$ is stable and that $F(u)=\mathrm{sgn}(u)$, the local field h_i in the equilibrium $\{x_i\}$ is

$$h_i=\sum_{j\neq i}^{N}J_{ij}\xi_j^1=\frac{1}{N}\sum_{j\neq i}^{N}\xi_i^1\xi_j^1\xi_j^1 f'(T_{ij}^{(1)})+\frac{\sqrt{p}}{N}\sum_{j\neq i}^{N}f(T_{ij}^{(1)})\xi_j^1, \tag{39}$$

$$T_{ij}^{(1)}=\frac{1}{\sqrt{p}}\sum_{\mu=2}^{p}\xi_i^\mu\xi_j^\mu. \tag{40}$$

The first term in Eq. (39),

$$\frac{1}{N}\sum_{j\neq i}^{N}\xi_i^1\xi_j^1\xi_j^1 f'(T_{ij}^{(1)})=J\xi_i^1, \tag{41}$$

$$J\equiv\int Dx f'(x)=\int Dx\,x f(x), \tag{42}$$

is the signal, while the second one is the noise, of which the mean and variance are

$$E\left[\frac{\sqrt{p}}{N}\sum_{j\neq i}^{N}f(T_{ij}^{(1)})\xi_j^1\right]=0, \tag{43}$$

$$E\left[\left(\frac{\sqrt{p}}{N}\sum_{j\neq i}^{N}f(T_{ij}^{(1)})\xi_j^1\right)^2\right]=\alpha\tilde{J}^2, \tag{44}$$

$$\tilde{J}^2\equiv\int Dx(f(x))^2, \tag{45}$$

respectively, where $E[\cdots]$ implies averaging over all of the random memory patterns $\{\xi_i^\mu\}$.

According to this naive S/N analysis, we rewrite the connections as

$$J_{ij}=\frac{J}{N}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu+\left[\frac{\sqrt{p}}{N}f(T_{ij})-\frac{J}{N}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu\right]$$
$$=\frac{\sqrt{p}}{N}\{JT_{ij}+[f(T_{ij})-JT_{ij}]\}. \tag{46}$$

The following derivation suggests that the residual overlap m^μ for the first term in Eq. (46) is enhanced by a factor of $1/(1-U)$, while any enhancement of the last part is canceled because of the subtraction. It also implies that the last part corresponds to the synaptic noise.

For the SCSNA of the nonlinear learning rule, we obtain

$$h_i=J\sum_{\mu=1}^{p}\xi_i^\mu m_\mu-\alpha Jx_i+\frac{\sqrt{p}}{N}\sum_{j\neq i}[f(T_{ij})-JT_{ij}]x_j. \tag{47}$$

In the last term in Eq. (47), the output activity $x_j=F(h_j)$ generally depends on the connection T_{ji}. We define $x_j^{(T_{ji})}$ by subtracting T_{ji} dependences from x_j at the leading order:

$$x_j=x_j^{(T_{ji})}+x_j'^{(T_{ji})}\frac{\sqrt{p}}{N}[f(T_{ji})-T_{ji}]x_i, \tag{48}$$

where

$$x_j^{(T_{ji})}=F\left(h_j-\frac{\sqrt{p}}{N}[f(T_{ji})-T_{ji}]x_i\right), \tag{49}$$

$$x_j'^{(T_{ji})}=F'\left(h_j-\frac{\sqrt{p}}{N}[f(T_{ji})-T_{ji}]x_i\right). \tag{50}$$

Substituting Eq. (48) into Eq. (47) gives

$$h_i=J\sum_{\mu=1}^{p}\xi_i^\mu m_\mu-\alpha Jx_i+\frac{\sqrt{p}}{N}\sum_{j\neq i}[f(T_{ij})-JT_{ij}]x_j^{(T_{ji})}$$
$$+\frac{\alpha}{N}x_i\sum_{j\neq i}[f(T_{ij})-JT_{ij}][f(T_{ji})-JT_{ji}]x_j'^{(T_{ji})}. \tag{51}$$

Then Eq. (51) and the relation $x_i=F(h_i)$ indicate that the output activity can be formally expressed as

$$x_i=\widetilde{F}\left(J\sum_{\mu=1}^{p}\xi_i^\mu m_\mu+\frac{\sqrt{p}}{N}\sum_{j\neq i}[f(T_{ij})-JT_{ij}]x_j^{(T_{ji})}\right). \tag{52}$$

Let $\{\xi_i^1\}$ be the target pattern. We substitute Eq. (52) into the definition of the pattern overlap and expand the resultant expression by $\xi_i^\mu m_\mu$ $(\mu>1)$, which has the order of $O(1/\sqrt{N})$. This leads to

$$m_\mu=\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu x_i^{(\mu)}+\frac{1}{N}\sum_{i=1}^{N}\left[Jm_\mu+\frac{1}{N}\sum_{j\neq i}^{N}\xi_j^\mu x_j^{(T_{ji})}f'(T_{ij}^{(\mu)})\right.$$
$$\left.-\frac{J}{N}\sum_{j\neq i}^{N}\xi_j^\mu x_j^{(T_{ji})}\right]x_i'^{(\mu)}$$
$$=\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu x_i^{(\mu)}+Jm_\mu U=\frac{1}{N(1-JU)}\sum_{i=1}^{N}\xi_i^\mu x_i^{(\mu)} \tag{53}$$

for $\mu>1$, where

$$x_i^{(\mu)}=\widetilde{F}\left(J\sum_{\nu\neq\mu}^{p}\xi_i^\nu m_\nu+\frac{\sqrt{p}}{N}\sum_{j\neq i}[f(T_{ij}^{(\mu)})-JT_{ij}^{(\mu)}]x_j^{(T_{ji})}\right), \tag{54}$$

$$x_i'^{(\mu)}=\widetilde{F}'\left(J\sum_{\nu\neq\mu}^{p}\xi_i^\nu m_\nu+\frac{\sqrt{p}}{N}\sum_{j\neq i}[f(T_{ij}^{(\mu)})-JT_{ij}^{(\mu)}]x_j^{(T_{ji})}\right), \tag{55}$$

$$U=\frac{1}{N}\sum_{i=1}^{N}x_2'^{(\mu)}, \tag{56}$$

$$T_{ij}^{(\mu)}=T_{ij}-\frac{1}{\sqrt{p}}\xi_i^\mu\xi_j^\mu. \tag{57}$$

Substituting Eq. (53) into Eq. (51) and averaging Eq. (51) over ξ_i^μ $(\mu>1)$ yield

$$h_i=J\xi_i^1 m_1+\frac{\sqrt{p}}{N}\sum_{j\neq i}^{N}[f(T_{ij})-JT_{ij}]x_j^{(T_{ji})}-\alpha Jx_i$$
$$+\frac{\alpha}{N}x_i\sum_{j\neq i}[f(T_{ij})-JT_{ij}][f(T_{ji})-JT_{ji}]x_j'^{(T_{ji})}$$
$$+\frac{1}{N(1-JU)}\sum_{\mu=2}^{p}\xi_i^\mu\sum_{j=1}^{N}\xi_j^\mu x_j^{(\mu)}$$
$$=J\xi_i^1 m_1+\frac{\sqrt{p}}{N}\sum_{j\neq i}^{N}[f(T_{ij})-JT_{ij}]x_j^{(T_{ji})}$$
$$+\left[\frac{J}{1-JU}+(\widetilde{J}^{\,2}-J^{\,2})\right]\alpha Ux_i$$
$$+\frac{1}{N(1-JU)}\sum_{\mu=2}^{p}\xi_i^\mu\sum_{j\neq i}^{N}\xi_j^\mu x_j^{(\mu)}. \tag{58}$$

The second and last terms in the rhs of Eq. (58) yield the Gaussian noise which has a vanishing mean. The variance of the noise term is given by

$$\sigma^2=\left[\frac{J^2}{(1-JU)^2}+\widetilde{J}^{\,2}-J^2\right]\alpha q. \tag{59}$$

Thus, after rewriting $\xi_i^1 \rightarrow \xi_i$ and $m_1 \rightarrow m$, we finally obtain the following implicit relation for the effective response function:

$$Y(z;\xi)=F\left(\xi m+\sigma z+\left[\frac{J}{1-UJ}+\tilde{J}^2-J^2\right]\alpha U Y(\xi;z)\right), \tag{60}$$

where m, q, and U are given by Eqs. (26)–(28).

Reduction in storage capacity

The equilibrium properties of the partially connected nonmonotonic networks can be obtained by numerically solving the equations for the order parameters. Accordingly, we use an additional rule, which is similar to the Maxwell rule in the thermodynamics, to solve the effective response function in terms of z. We do not repeat the detailed analysis here since it can be found in Ref. [2]. The order-parameter equations are studied for two values of θ, i.e., $\theta=1$ and $\theta=2$, to compare the effects of synapse dilution between different extents of nonmonotonicity. For the latter value, the response function can be regarded approximately as a step function owing to the fact that the distribution range of the local fields is practically limited. The choice $\theta=1$ gives the smallest possible value, since the storage capacity obtained by the SCSNA for the fully connected nonmonotonic network tends to be larger than the actual values obtained by simulations for $\theta<1$ [10]. This discrepancy seems to imply that the SCSNA describes unstable fixed points of the nonmonotonic network for these values of θ.

When the synaptic connections are randomly eliminated, the symmetric and asymmetric dilutions yield almost identical results by the SCSNA. Therefore, the following only presents results for asymmetric random dilution.

Figure 1(a) shows the storage capacity α_c of the nonmonotonic network for $\theta=1$ and $\theta=2$ as functions of the cutting rate $R_c \equiv 1-c$. In both cases, α_c decreases as R_c is increased. However, the reduction rate of α_c is small for $\theta=2$, i.e., when the response function is regarded as a step function. On the other hand, when $\theta=1$ and the degree of nonmonotonicity is high, α_c decreases rapidly. For instance, α_c becomes the half value of the fully connected case at $R_c=0.3$ for $\theta=1$, and at $R_c=0.6$ for $\theta=2$.

The values of the tolerance pattern overlap g defined in terms of the local field h_i as

$$g=\frac{1}{N}\sum_{i=1}^{N}\xi_i^1 \operatorname{sgn}(h_i) \tag{61}$$

can also be easily obtained by the SCSNA and the results are shown in Fig. 1(b) as functions of R_c for the two values of θ, when the network is maximally loaded with $\alpha_c N$ patterns. The reduction rate of g is also larger for $\theta=1$, but the difference between the reductions for the two θ values is not very large. Since the tolerance overlap is always close to unity unless $R_c \approx 1$, the quality of retrieved patterns is not significantly influenced by the dilution of connections in both cases.

Similarly, α_c is shown in Fig. 2 as a function of R_c for the two types defined by Eqs. (35) and (36) of the systematic synapse dilution. The cutoff parameter is fixed at $\theta=1$ in both cases. We see that the storage capacity for the bottom-cut-off dilution remains much larger than that for the random dilution. Indeed, α_c becomes half its value for the fully connected case at a cutting rate as large as $R_c=0.75$. On the other hand, the reduction in α_c occurs much more rapidly for the top-cut-off dilution than it does for the random dilution, as expected.

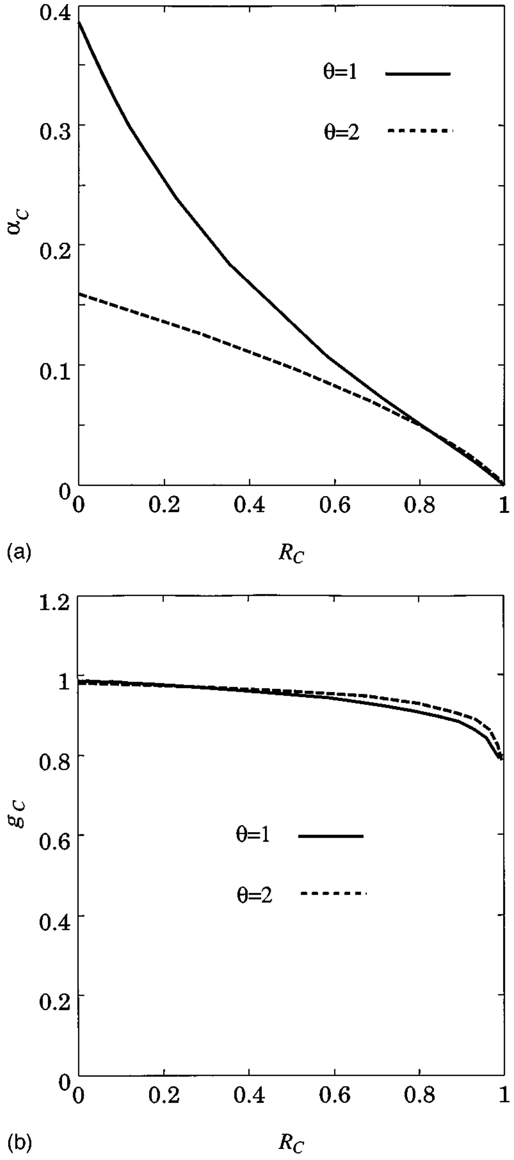

FIG. 1. SCSNA results for (a) storage capacity and (b) tolerance pattern overlap as a function of cutting rate R_c when synaptic connections are diluted randomly and asymmetrically. Solid and dashed curves are for nonmonotonic response function with cut-off activity $\theta=1$ and $\theta=2$, respectively. Random symmetric dilution gives results that are almost identical to those presented here.

To confirm the results of the SCSNA, numerical simulations of the retrieval dynamics of the nonmonotonic neural network are conducted for two cases, i.e., the random and bottom-cut-off dilutions. The results are shown in Figs. 3(a) and 3(b) for α_c, m, and the tolerance overlap g, respectively. α_c is obtained by evaluating the loading rates at which the network succeeds to retrieve memory patterns with approximately 50% probability. A trial is regarded as

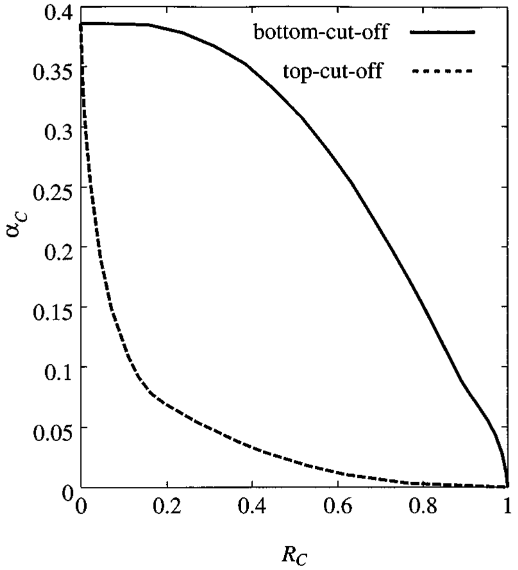

FIG. 2. Storage capacity obtained by the SCSNA for bottom-cut-off (solid curve) and top-cut-off (dashed curve) systematic dilutions.

successful if g is more than 0.96 after the network evolves into a stationary state. It is found that the results of the SCSNA for α_c and g are in good agreement with those of the simulations. Relatively large discrepancies in values of m between the results of the analytical and numerical studies are due to the fact that they vary significantly with the values of α near α_c. Taking this fact into account, we can conclude that the discrepancies are at an acceptable level.

Instability of superretrieval phase

So far, the superretrieval phase has been found only for associative memory models with nonmonotonic response functions. In this phase, the noise from uncondensed memory patterns in the local fields disappears, and consequently an evoked activity pattern coincides with a memory pattern without error. In the framework of the SCSNA, the occurrence of the superretrieval phase is indicated by the disappearance of the variance σ^2 of the Gaussian noise in the solution to the order-parameter equations. From Eq. (25), we see that $|U|\rightarrow\infty$ implies that $\sigma^2\rightarrow 0$ if the variance Δ^2, given in Eq. (32) and which arises from the random synapse dilution, is zero. This is indeed the case for fully connected nonmonotonic neural networks [2,10,12]. However, for $\Delta^2\neq 0$, σ^2 never disappears for any retrieval state since $q\approx 1$ in the state. This implies that the superretrieval phase is unstable against random dilution.

Although σ^2 cannot be zero due to the noise from the synapse dilution, the noise r from uncondensed patterns may still vanish. To examine this possibility, we solve the order-parameter equations for $\Delta^2=0$ (the fully connected case) and $\Delta^2=0.01$ and investigate whether the parameter r disappears at a certain value of α. We fix θ at 0.7, a value that ensures the existence of the superretrieval phase for $\Delta^2=0$.

The results are shown in Fig. 4. For the fully connected nonmonotonic network, r disappears very rapidly when α approaches a critical value (≈ 0.125) from above. This indicates that the superretrieval phase appears for α less than the critical value. However, when $\Delta^2=0.01$, r does not show a singular behavior for any value of α. Rather, the results suggest that $r\propto e^{k\alpha}$ $(k>0)$ for this value of Δ^2. This implies that the superretrieval phase never occurs when a small but finite number of connections are eliminated.

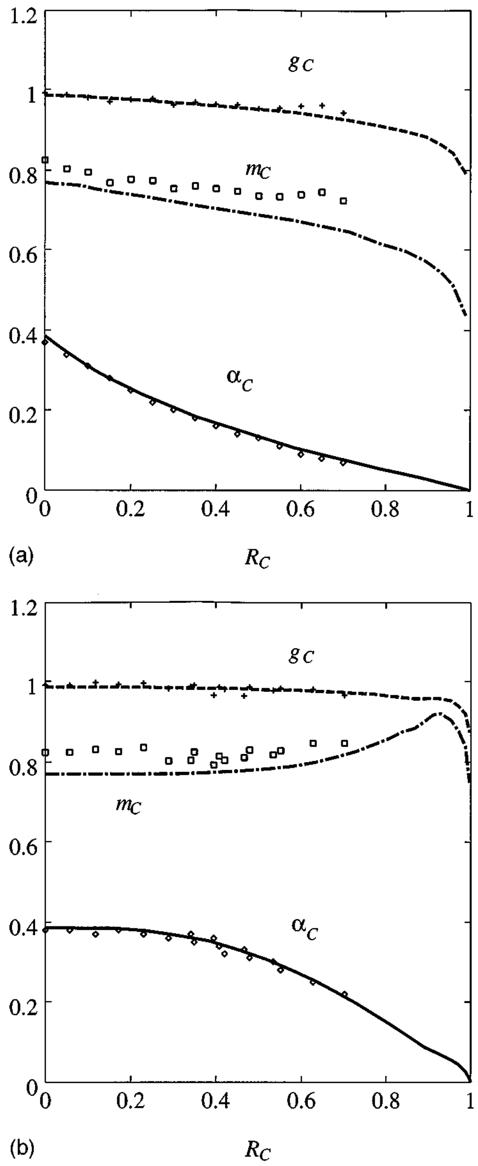

FIG. 3. Comparison between the SCSNA (curves) and simulation (plots) results for (a) random synaptic dilution and (b) bottom-cut-off systematic dilution. In both figures, solid, dash-dotted, and dashed curves show storage capacity, pattern overlap, and tolerance pattern overlap, respectively, as a function of R_c.

The disappearance of the superretrieval phase can also be seen by numerical simulations. To this end, the distribution $P(h)$ of the local fields in the retrieval states is calculated for the random dilution by numerical simulations. The results are shown in Figs. 5(a) and 5(b) when the cutting rate R_c is 0% and 1%, respectively. The SCSNA predicts that $P(h)$ has four peaks represented by the delta functions at $|h|=\theta\pm\alpha/2$ when $N\rightarrow\infty$ [2]. These sharp peaks are clearly seen in Fig. 5(a) for the superretrieval states of the fully connected network. However, in Fig. 5(b) the peaks are significantly smeared by the dilution of only a small number of the synaptic connections.

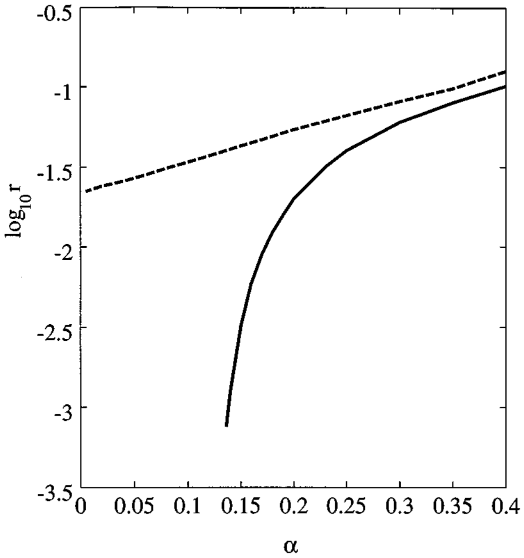

FIG. 4. Behavior of order parameter r is obtained by the SC-SNA as a function of loading rate α for a fully connected nonmonotonic neural network (solid curve) and randomly diluted one (dashed curve). In a fully connected network, $r \to 0$ as $\alpha \to \alpha_0 \approx 0.125$ from above, which implies the appearance of a superretrieval phase for $\alpha < \alpha_0$. Such singular behavior is not seen for diluted network.

By using Eq. (59), the disappearance of the superretrieval phase can be shown for a much wider class of synapse dilution. We define $f(x)$ as

$$f(x)=\begin{cases} x & x \in \mathbf{S} \\ 0 & x \in \overline{\mathbf{S}}, \end{cases} \tag{62}$$

where $\mathbf{S}$ is an arbitrary subset of $\mathbf{R}$ that satisfies Eq. (38). With this $f(x)$, a synaptic connection is eliminated if its value does not lie in $\mathbf{S}$. We show that the necessary condition, $J^2=\tilde{J}^2$, for the occurrence of the superretrieval phase is satisfied only if $\mathbf{S}=\mathbf{R}$. This is easily shown by noting that

$$J=\int_{\mathbf{S}} Dx\, x f(x)=\int_{\mathbf{S}} Dx f(x)^2=\tilde{J}^{\,2}=\int_{\mathbf{S}} Dx\, x^2 \leqslant 1. \tag{63}$$

Therefore,

$$J^2 \leqslant J=\tilde{J}^{\,2}. \tag{64}$$

The equality in Eq. (64) holds only when $J=1$ namely, when $\mathbf{S}=\mathbf{R}$ and neurons are fully connected. Although the above definition of $f(x)$ does not include all types of synapse dilution (for instance, one may cut half the connections randomly and the other half systematically), the superretrieval phase is presumably unstable to any type of synapse dilution done at a finite cutting rate.

CONCLUSIONS

This paper has shown that the enhanced storage capacity of the associative memory model with a nonmonotonic response function is considerably reduced by random dilutions of synaptic connections. The reduction rate is larger than that for the conventional associative memory models with a monotonic response function. This result contrasts with the results of a similar study that suggested an oscillator neural network of associative memory was more robust against dilution than the standard Hopfield model [22]. To retain as large a storage capacity as possible in dilution of synaptic connections, we proposed bottom-cut-off dilution, in which the synaptic connections are systematically eliminated according to the order of their magnitudes. Although other attempts to consider robustness with respect to the dilution exist [18,23,21], this method of dilution is relatively simple and significantly minimizes the loss of the nonmonotonic neural network's excellent retrieval abilities. In both cases, however, the superretrieval phase, in which noise from uncondensed memory patterns disappears in the local fields of

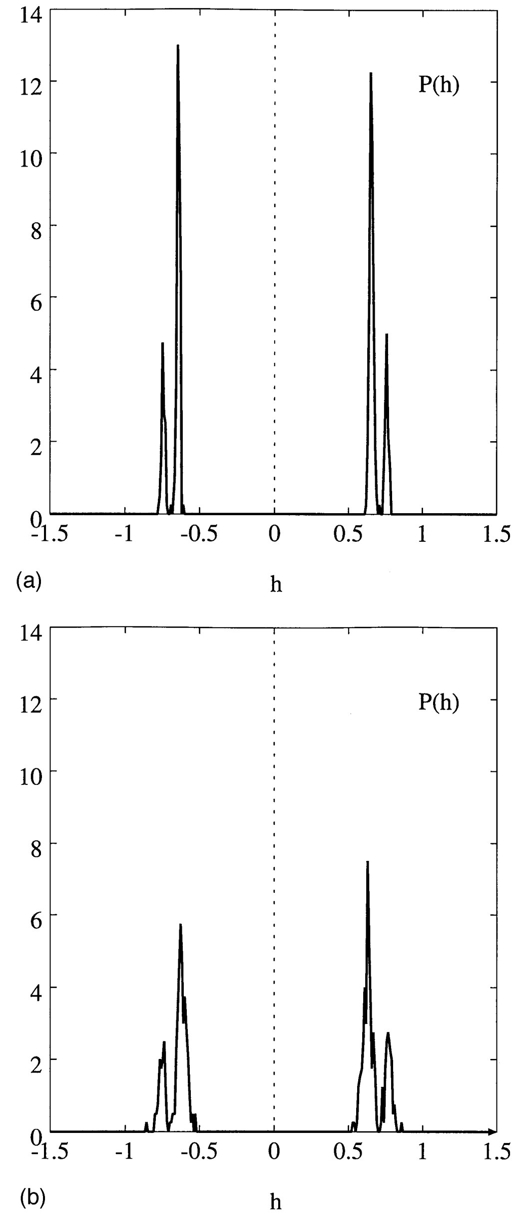

FIG. 5. Numerically obtained distributions of local fields h_i at equilibrium states for (a) fully connected and (b) 1%-diluted nonmonotonic neural networks. Values of parameters were fixed at $\theta=0.7$ and $\alpha=0.1$, for which the fully connected neural network operates in the superretrieval phase.

neurons, does not appear if a finite number of connections are eliminated. These results clearly show that enhancement of the storage capacity of non-monotonic neural networks is not achieved by disappearance of noise, which has already been suggested by the fact that the shape of response functions has a large influence on the extent of the superretrieval phase but not on the maximum values of the storage capacity [12].

[1] M. Morita, Neural Networks **6**, 115 (1993).

[2] M. Shiino and T. Fukai, J. Phys. A **26**, L831 (1993).

[3] S. Yoshizawa, M. Morita, and S. Amari, Neural Networks **6**, 167 (1993).

[4] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. **55**, 1530 (1985).

[5] M. Shiino and T. Fukai, J. Phys. A **23**, L1009 (1990).

[6] R. Kühn, S. Bös, and J. L. van Hemmen, Phys. Rev. A **43**, 2084 (1991).

[7] M. Okada, Neural Networks **9**, 1429 (1996).

[8] G. Boffetta, R. Monasson, and R. Zecchina, J. Phys. A **26**, L507 (1993).

[9] E. Gardner, J. Phys. A **21**, 257 (1988).

[10] M. Shiino and T. Fukai, Phys. Rev. E **48**, 867 (1993).

[11] M. Yoshioka and M. Shiino, Phys. Rev. E **55**, 7401 (1997).

[12] T. Fukai, J. Kim, and M. Shiino, Neural Networks **8**, 391 (1995).

[13] M. Morita, Neural Networks **9**, 1477 (1996).

[14] T. Asai, H. Yokotsuka, and T. Fukai, IEEE Trans. Neural Netw. **7**, 182 (1996).

[15] H. Sompolinsky, Phys. Rev. A **34**, 2571 (1986).

[16] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).

[17] H. Yanai, Y. Sawada, and S. Yoshizawa, Network **2**, 295 (1991).

[18] S. Mertens, J. Phys. A **24**, 337 (1991).

[19] J. L. van Hemmen, D. Grensing, A. Huber, and R. Kühn, J. Stat. Phys. **50**, 231 (1988).

[20] J. L. van Hemmen, D. Grensing, A. Huber, and R. Kühn, J. Stat. Phys. **50**, 259 (1988).

[21] P. Kuhlmann, R. Garcés, and H. Eissfeller, J. Phys. A **25**, L593 (1992).

[22] T. Aoyagi and K. Kitano, Phys. Rev. E **55**, 7424 (1997).

[23] A. Komoda, R. Serneels, K. Y. Wong, and M. Bouten, J. Phys. A **24**, L743 (1991).