Systems and Computers in Japan, Vol. 27, No. 5, 1996
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J78-D-II, No. 6, June 1995, pp. 946-953

# Capacity of Autocorrelation Associative Memory with Quantized Synaptic Weight

Kazushi Mimura, Masato Okada and Koji Kurata, *Members*

Faculty of Engineering Science, Osaka University, Toyonaka, Japan 560

## SUMMARY

The neural chips to speed up the process of a neural network have recently been developed actively. Since a conventional neural chip requires quantized synaptic weights, it is important to know the properties of the network with such weights. Devices for a 2-layer network with the same number of input and output elements have been well developed, and such a device can be used as an autocorrelation associative memory by feeding its output to its input.

This paper investigates theoretically the capacity of an autocorrelation associative memory when its synaptic weights are quantized with a finite number of bits. It also proposes an optimum quantization function. A system with a finite number of elements is computer simulated. The proposed method can be applied to the design of a neural chip for associative memory.

**Key words**: Associative memory; memory capacity; quantization noise; synaptic noise.

## 1. Introduction

Development of neural chips which speed up a neural network has been active recently. However, there has been a problem in that vast numbers of wirings are required between neurons in a large-scale neural network, e.g., the maximum $N$ of wires for a network consisting of $N$ neurons.

It has been proposed that space paralleling and high-speed light for such a large-scale wiring be used. A synaptic weight in an optical system is represented by the transmission speed of light. Since the synaptic weight is a real number, it is ideal that the transmission rate is continuously variable. A method has been proposed that uses phase modulation of a polarized light. This is not suitable for integrated circuits since the structure is too complex. There has also been a method which uses an optimal mask, but this requires a quantized synaptic weight since the mask can represent only transmission or nontransmission.

Quantization is not only important for an optical neural chip, but generally, for the synaptic weight of many electronic devices. Therefore, it is important to find the properties of networks with synaptic weight. Notably, devices for a 2-layer network have been well advanced [2, 3]. Such a device can easily be applied to an autocorrelation associative memory by feeding its output to its input. Therefore, this paper investigates the relationship between the memory capacity and the quantization of synaptic weight in this type of memory. It has been known that an associative memory can be analyzed by S/N analysis [4-6], but this cannot be applied to a synaptic weight. In this paper, an optimum quantization function is obtained by using Sompolinsky's theory [9] which is an expansion of statistical dynamics [7, 8].



ISSN0082-1666/96/0005-0055

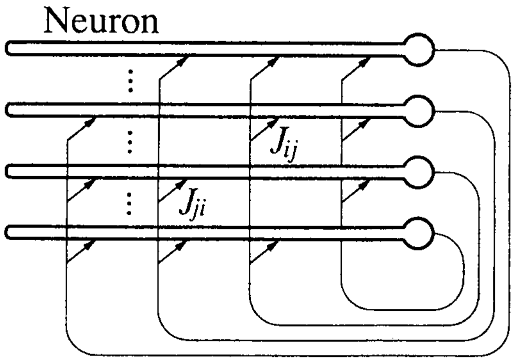


Fig. 1. Structure of associative memory.

## 2. Outline of Autocorrelation Associative Memory

Figure 1 shows an outline of the structure of an autocorrelation associative memory [1]. All the elements are connected to each other, and each element changes the state of the whole circuit. Let the state of the $i$-th element be $s_i$ ($= \pm 1$), and let the synaptic weight from the $j$-th element to the $i$-th element be $J_{ij} = J_{ji}$. No autosynapse is considered, i.e., $J_{ij} = 0$. The state of the circuit changes from $s$ to $s'$ in the following:

$$\mathrm{s}' = \mathrm{sgn}\,[\mathbf{J}\mathrm{s}] \tag{1}$$

$$\mathbf{J} = (J_{ij}),\ \mathrm{s} = (s_1, \ldots, s_N)^t,\ \mathrm{s}' = (s'_1, \ldots, s'_N)^t$$

where sgn is a sign function defined by

$$\mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

and t is a transposition, and $N$ is the number of elements.

Data stored in the associative memory are in a dynamic equilibrium state as shown by Eq. (1). In other words, to memorize $p$ vectors, $\xi^1, \ldots, \xi^p$, is equivalent to making synaptic weight $J$ so that

$$\boldsymbol{\xi}^\mu = \mathrm{sgn}\,[J\boldsymbol{\xi}^\mu], \quad \mu = 1, \ldots, p \tag{2}$$

holds. A method to obtain such a synaptic weight is to define **J** as

$$J_{ij} = \frac{J}{N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \qquad i \neq j, \quad J > 0 \tag{3}$$

A constant $J/N$ is multiplied since the dynamics are not affected by multiplying a positive constant to all the elements as shown by Eq. (1). Also $\xi^\mu = (\xi_j^\mu)$ holds. This method is called "correlation learning." This method has an advantage in that no repeated learning is required, although the number of its vectors is fewer than other methods such as an error-correction learning.

If the number of vectors to be memorized exceeds a certain value, no equilibrium state can be obtained. The ratio of this criterion number and the number of the elements are called "memory capacity." Assuming

$$\xi_i^\mu = \begin{cases} 1 & \text{with prob. } \frac{1}{2} \\ -1 & \text{with prob. } \frac{1}{2} \end{cases} \tag{4}$$

the memory capacity for a one-half coding is given by Amit [7] as about 0.138.

The energy function $E$ of an autocorrelation associative memory is defined by

$$E = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij} s_i s_j \tag{5}$$

In an asynchronized autocorrelation associated memory in which each element changes a state every time, the energy function simply reduces. Hopfield [1] has suggested that an associated memory can be treated as a physical problem by introducing its energy.

## 3. Capacity of Associated Memory with Quantized Synaptic Weight

### 3.1. Outline of calculation method

The capacity of this type of memory reduces when noise is added to its synaptic weight. Such a noise is called "synoptic noise." Sompolinsky [9] has investigated theoretically the capacity of an associated memory when the synaptic noise is a random number having a Gaussian distribution with an arbitrary variance. By using his theory, let us obtain the memory capacity by converting it to a synaptic noise caused by a quantization of the synoptic weight.

### 3.2. Sompolinsky's theory

#### 3.2.1. Model with synaptic noise

A model of an associated memory having a synaptic weight with a random number $\delta_{ji}$ is represented by

$$J_{ij} = \frac{J}{N}\sum_{\mu=1}^{p} \xi_i^{\mu}\xi_j^{\mu}, +\delta_{ij} \quad i \neq j, \quad J > 0 \tag{6}$$

where it is assumed that the random numbers $\delta_{ij}$ have a Gaussian distribution of

$$\delta_{ij} \sim N\left(0, \frac{J^2}{N}\Delta^2\right) \tag{7}$$

and that this is independent of vector $\xi$, the size of the synaptic noise being $\Delta$; $J^2/N$ acts as a coefficient to scale the variance of the noise independent of the number of elements.

Let us consider a model in which each element behaves probabilistically and asynchronously so that methods in physics can be applied.

When an element changes its state, it is selected randomly or orderly. The new state $s_i$ of the $i$-th element is given by 1 for probability $P(u_i)$, and by 1 for probability 1 - $P(u_i)$, where

$$P(u_i) = \frac{1}{1 + \exp(-\frac{u_i}{T})}$$

$$u_i = \sum_j J_{ij}s_j$$

If temperature $T$ $(= 1/\beta)$ approaches zero in this model, the model coincides with the model given by Eq. (1), except for a synchronization of change of state of the elements, since the fluctuation of the probability disappears. Therefore, as far as the equilibrium is concerned, the two models are equivalent.

Let us consider the memorized vector $\xi^1$ and

$$m = \left[\ll \frac{1}{N}\sum_{i=1}^{N} \xi_i^1 < s_i >_T \gg\right] \tag{8}$$

which is the overlap of $\xi$ and an equilibrium state when it is recalled as an initial state, where $<\cdot>$ is a mean temperature, $<<\cdot>>$ is mean value of probability variable $\xi$, and $[\cdot]$ is the mean value of probability variable $\delta_{ij}$.

The overlap $m$ is 1 when the memorized vector is recalled in the same form; and $m$ is zero when a random vector is recalled. Therefore, whether a vector is memorized can be judged from the state of the overlap, 1 or 0. The capacity $\alpha_c$ of a memory can be obtained by $p_c/N$, where $p_c$ is the number of the maximum vectors which do not make $m = 0$. The first vector $\xi^1$ is chosen as the initial state to obtain $m$ in the above statement, but any vector can be chosen since the vector is selected randomly.

Let the ratio of the number of memorized vectors $p$ to the number of elements $N$ be $\alpha$ as

$$\alpha = \frac{p}{N} \tag{9}$$

Let us evaluate the number of memorized vectors by using this ratio.

By applying the replica method [7, 8] to the model given by Eq. (6) which includes noise, the free energy per element [9] can be obtained as

$$\begin{aligned} f = & \frac{\alpha T}{2}\left(\log(1-C) + \frac{(1-\beta J)C}{1-C}\right) \\ & + \frac{1}{2}J(\alpha r C + m^2) - \frac{T}{4}\Delta^2 C^2 \\ & - T\int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \\ & \times \log\left\{2\cosh\beta J\left(z\sqrt{\alpha r + \Delta^2 q} + m\right)\right\} \end{aligned} \tag{10}$$

where

$$q \stackrel{\text{def}}{=} \left[\ll \frac{1}{N}\sum_{i=1}^{N} < s_i >_T^2 \gg\right], \tag{11}$$

$$r \stackrel{\text{def}}{=} \frac{1}{\alpha}\left[\ll \sum_{\mu=1}^{\alpha N}\left(\frac{1}{N}\sum_{i=1}^{N} < s_i >_T^2 \xi_i^{\mu}\right) \gg\right] - \frac{m^2}{\alpha} \tag{12}$$

$$C \stackrel{\text{def}}{=} \frac{J(1-q)}{T} \tag{13}$$

$q$ is called the "Edwards-Anderson orderly variable," and represent the correlation between the replicas; $r$ is the mean square of the overlap of the state of the element and the unrecalled pattern, and represents the variance of components (among the total of inputs from all the elements) which do not contribute to the recall; $C$ is a parameter for convenience.

By using the saddle-point equation of the free energy, self-constraint equations of $m$, $q$ and $r$ are obtained as follows:

$$m = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \times \tanh \beta J \left(z\sqrt{\alpha r + \Delta^2 q} + m\right) \quad (14)$$

$$q = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \times \tanh^2 \beta J \left(z\sqrt{\alpha r + \Delta^2 q} + m\right) \quad (15)$$

$$r = \frac{q}{(1-C)^2} \quad (16)$$

Since a model which behaves deterministically taking into account a thermal fluctuation corresponds to its original associative memory, it is necessary to choose the temperature $T$ as its limit ($T \to 0$, i.e., $q \to 1$). Then,

$$m = \mathrm{erf}\left(\frac{m}{\sqrt{2\left(\alpha r + \Delta^2\right)}}\right) \quad (17)$$

$$r = (1-C)^{-2} \quad (18)$$

$$C = \sqrt{\frac{2}{\pi\left(\alpha r + \Delta^2\right)}} \exp\left(-\frac{m^2}{2\left(\alpha r + \Delta^2\right)}\right) \quad (19)$$

hold. The overlap $m$ is obtained by numerically solving these three equations. Then the capacity of the associative memory is obtained from $m$; erf($x$) in Eq. (17) is an error integration function defined by

$$\mathrm{erf}(x) \stackrel{\mathrm{def}}{=} \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-\tau^2\right) d\tau \quad (20)$$

Equations (17) through (19) show that $m$ is the function of $\alpha$ and $\Delta$ alone as

$$m = m(\alpha, \Delta) \quad (21)$$

When a memorized vector can be recalled, $m$ is approximately 1. When a memorized vector cannot correctly be recalled, $m$ starts to reduce, and then $m$ suddenly becomes zero. Figure 2 shows the change of the overlap $m$ against the pattern ratio $\alpha$, obtained by a

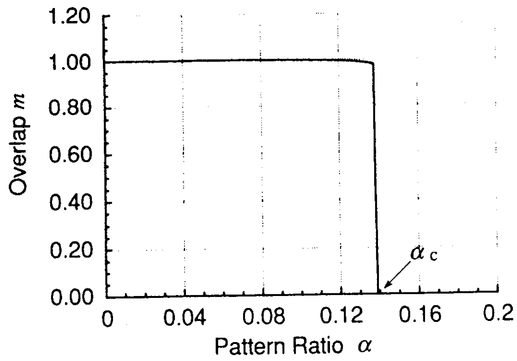


Fig. 2. Dependence on direction cosine to memorized vectors.

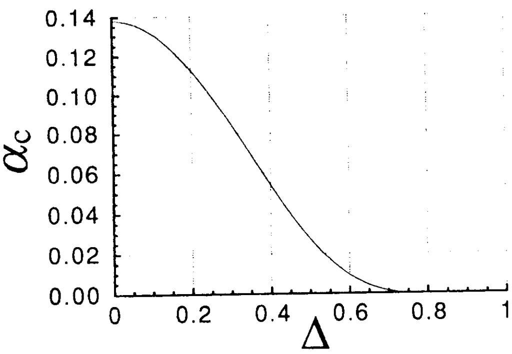


Fig. 3. Dependence on noise to capacity.

numerical calculation using Eqs. (17) through (19), assuming $\Delta = 0$ (no noise). The capacity of the associative memory $\alpha_c = \alpha_c(\Delta)$ is defined by the maximum value of the number of recalled vectors among the number of vectors intended to be memorized. From Fig. 2, the capacity of the memory can be obtained as

$$\frac{\partial m}{\partial \alpha}(\alpha_c, \Delta) = -\infty \quad (22)$$

The capacity of the memory $\alpha_c$ can be obtained by a numerical calculation when the size of a synaptic noise $\Delta$ is fixed. By repeating this procedure, the relationship between $\Delta$ and the memory capacity $\alpha_c$ is as shown in Fig. 3. This shows that $\alpha_c$ is a simple decreasing function.

### 3.2.2. Model of synaptic weight with function

Let us consider a circuit in which the synaptic weight is a function of weight obtained from the Hebb rule. That is,

$$J_{ij} = \frac{\sqrt{p}}{N} F\left(\frac{1}{\sqrt{p}}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu\right) \quad i \neq j \tag{23}$$

where $F$ is a nonlinear function to activate the synaptic weight. Coefficient $1\sqrt{p}$ keeps the variance of $\Sigma\xi_i^\mu\xi_j^\mu$ constant with respect to $p$. This model coincides with a model in which $F$ is linear and synaptic noise is not contained.

Let us set

$$x_{ij} = \frac{1}{\sqrt{p}}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \tag{24}$$

This shows that $x_{ij}$ has a normal distribution with a variance of 1, from the central limit theorem, since $\xi_i^\mu$ is independent. The size of a synaptic noise $\Delta$ (which corresponds to a noise due to a quantization of synaptic weight using a nonlinear function $F$) has been given by Sompolinsky [9] as

$$\Delta^2 = \alpha\left(\frac{\tilde{J}^2}{J^2} - 1\right) \tag{25}$$

where

$$\begin{aligned} \tilde{J}^2 &\stackrel{\text{def}}{=} \ll F^2(x) \gg \\ &= \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) F^2(x) \end{aligned} \tag{26}$$

$$\begin{aligned} J &\stackrel{\text{def}}{=} \ll xF(x) \gg \\ &= \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) xF(x) \end{aligned} \tag{27}$$

A simple calculation shows that $J$ represents the intensity memorizing as a coefficient $J$ in Eq. (3). As a special case where $F(x) = x$, a comparison with Eq. (3) shows that $J = 1$. To calculate these equations, the fact that $x$ in Eq. (24) has a normal distribution is used. From Eq. (25), $\Delta \propto \alpha$. By setting

$$\Delta = \Delta_0\sqrt{\alpha} \tag{28}$$

then

$$\Delta_0^2 = \frac{\tilde{J}^2}{J^2} - 1 \tag{29}$$

From Eq. (28), $m$ can be regarded as the functions of $\alpha$ and $\Delta_0$ alone, although $m$ is the function of $\alpha$ and $\Delta$. The memory capacity $\alpha$ is a simple reducing function of $\Delta_0$. Equations (26) and (27) show that the quantum function producing $\Delta = 0$ if $F(x) = 0$ in which $J$ is equal to $J$.

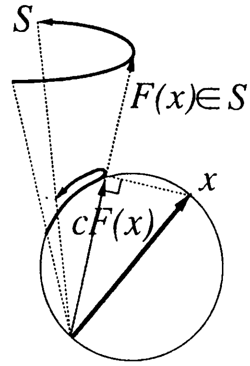


Fig. 4. Geometrical illustration of optimal quantum function.

Let us consider geometrically a quantum function which minimizes a quantum noise as shown in Fig. 4. Let us define the inner product and the norm of a function space as

$$\begin{aligned} (f, g) &\stackrel{\text{def}}{=} \int \frac{dx}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) f(x)g(x) \\ \|f\| &\stackrel{\text{def}}{=} \sqrt{(f, f)} \end{aligned}$$

The weight of the inner produce represents the distribution of the synaptic weight. Using this, $\Delta_0$ is represented by

$$\Delta_0^2 = \frac{\|F\|^2}{(x, F)^2} - 1 \tag{30}$$

Let the set of quantum functions $F$ be $S$. Then, $F$ which gives

$$\min_{F\in S} \Delta_0^2$$

is the final value. Let us define a quantity $D$ represented by

$$\begin{aligned} D^2 &\stackrel{\text{def}}{=} \min_{c>0} \|cF - x\|^2 \\ &= \min_{c>0}\left(c^2\|F\|^2 - 2c(x, F) + \|x\|^2\right) \\ &= -\frac{(x, F)^2}{\|F\|^2} + \|x\|^2 \end{aligned} \tag{31}$$

where $D > 0$. Equation (31) represents the square errors of quantum function $F$ and function $x$. As Eq. (1) shows, the dynamics of recall of a memory is independent of the positive constant of the whole synaptic weight. Therefore, $F$ is multiplied by $c$ so that this becomes as close as possible to function $x$, where $c > 0$. Figure 4 shows that $c$ which minimizes $\|cF - x\|^2$ is given by taking $cF$ on the foot of a perpendicular from $x$ to $F$. (Note, a 2-dimensional plane is illustrated for simplicity.) Therefore, the quantum function $F$ which minimizes $\Delta$ is obtained by projecting a set of quantum functions $S$ onto a hyperspace having a diameter of function $x$, and by choosing the value of $F$ which is closest to $x$ in the projection. This is represented by

$$\min_{F \in S} D$$

### 3.3. Model with quantized synaptic weight

The capacity of a memory can be obtained by knowing the intensity of relative noise. When the quantum function is applied to a synaptic weight, the relative intensity of a noise can be calculated, and this can be used to calculate the capacity. The relative intensity of the noise $\Delta$ is obtained by giving a step function (a quantum function) to this nonlinear function. Let us consider a nonlinear function $F$ shown in Fig. 5. If the quantum function has 8 bits, $n = 2^{k-1} - 1$. If

$$F(x) = \begin{cases} a_1, & 0 \leqq x < b_1 \\ \vdots & \\ a_i, & b_{i-1} \leqq x < b_i \\ \vdots & \\ a_n, & b_{n-1} \leqq x < b_n \\ 1, & b_n \leqq x \end{cases} \qquad (32)$$

are substituted into Eqs. (26), (27) and (29),

$$\Delta_0^2 = \frac{\sum_{k=0}^{n} a_{k+1}^2 \left\{ \mathrm{erf}\left(\frac{b_{k+1}}{\sqrt{2}}\right) - \mathrm{erf}\left(\frac{b_k}{\sqrt{2}}\right) \right\}}{\left\{ \sum_{k=0}^{n} (a_{k+1} - a_k) \exp\left(-\frac{b_k^2}{2}\right) \right\}^2} \cdot \frac{\pi}{2} - 1 \qquad (33)$$

is obtained, where

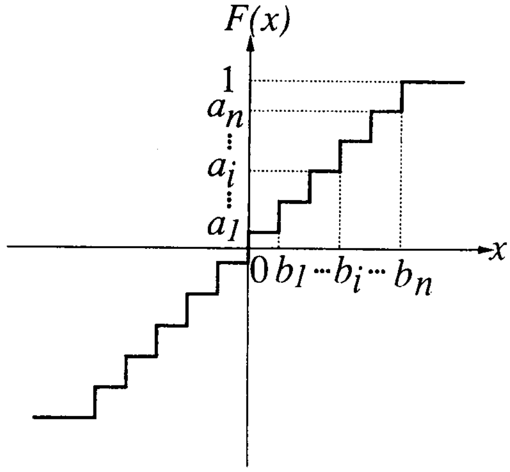


Fig. 5. General quantum function.

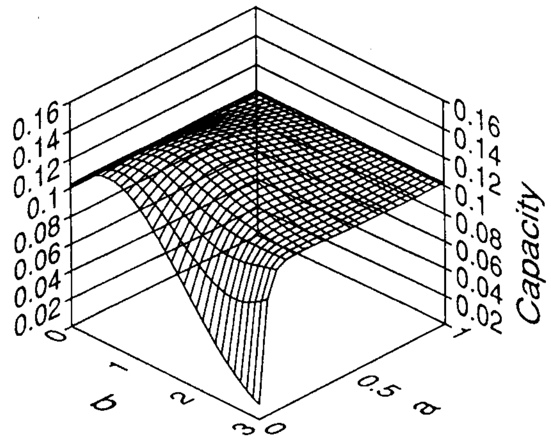


Fig. 6. Dependence on 4-level quantum function to capacity.

$$a_0 = b_0 = 0, \ a_{n+1} = 1, \ b_{n+1} = \infty$$

Since the memory capacity is a simple reducing function of $\Delta_0$, the quantum function which maximizes the memory capacity is determined by obtaining parameters $a_1, ..., a_n, b_1, ..., b_n$ which minimizes $\Delta_0$.

### 3.4. Theoretical results

The optimum quantum function and its memory capacity were calculated by using Eq. (33), when the synaptic weight is quantized in 2 levels (1 bit representation), 4 levels (2 bits), and 8 levels (3 bits), $n$ being 0, 1 and 3, respectively. Figure 6 shows the relationship between the quantum function and the mem-

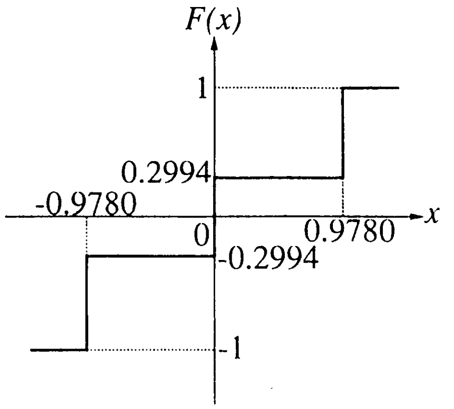


Fig. 7. Optimal 4-level quantum function.

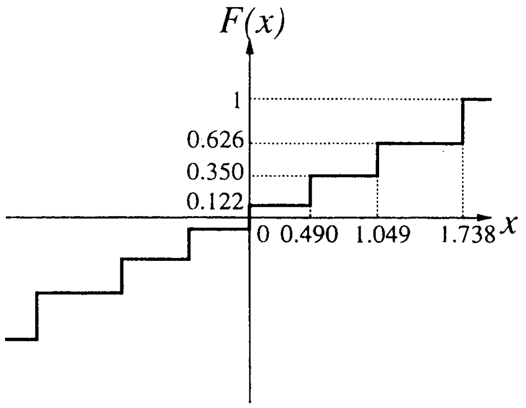


Fig. 8. Optimal 8-level quantum function.

ory capacity, for four levels and $n = 1$. In this example, $a_1$ and $b_1$ are the parameters to make the quantum function optimum.

Similarly, the optimum quantum function for two levels becomes a code function (which is 1 at $x > 0$, and -1 at $x < 0$). Figures 7 and 8 show the optimum quantum functions for four levels (2 bits) and eight levels (3 bits) respectively. Figure 9 shows the relationship between the number of bits in the quantization and the memory capacity.

Figure 9 shows that the memory does not change significantly even if a very coarse quantization is used. Figure 6 shows that the memory capacity changes little around the optimum parameter, i.e., an approximation of $F(x) = x$ gives the same result.

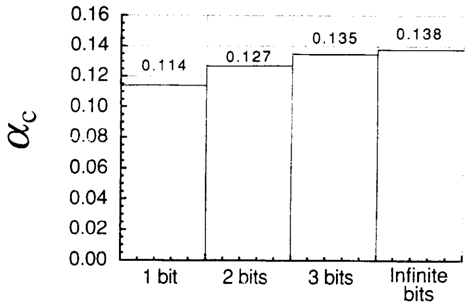


Fig. 9. Dependence on amount of information synapse to capacity.

## 4. Computer Simulation

### 4.1. Method

The theoretical results of proposed method are not accurate, since it assumes that the number of elements $N$ is infinite. Therefore, an optimum quantum function and its memory capacity were obtained by a computer simulation. The recall capacity of the memory is defined by a rate of the number of vectors which can recall the memorized vectors to the number of the memory elements. However, it is difficult to obtain the memory capacity by using Eq. (22), since the Monte Carlo simulation produces results containing some errors. In this paper, therefore, the memory capacity is obtained by using the data stored in the circuit.

When the number of memory is infinite, the value at a discontinuous point of the quantum function does not cause any problem, since the number of the vectors to be memorized increases greatly so that the distribution of the synaptic weight approaches a continuous point. When the number of elements are finite, however, some quantum functions have their synaptic weight on a discontinuous point of the quantum function. As shown in Fig. 5, when the synaptic weight is zero, the quantum function is independent of other parameters $a_1, ..., a_n, b_1, ..., b_n$. Therefore the value of the quantum function becomes important. If the synaptic function is zero, this paper assumes that the quantum function is $a_1$ or $-a_1$ with a probability of 1/2.

The computer simulation was carried out with $N = 32$ (the number of elements). Each vector element having a value of 1 or -1 is half-coded by using the M-series random numbers. Then $p$ chosen vectors are memorized. A vector is chosen from the memorized vectors. Taking this as an initial state, the state of the vector is shifted asynchronously. The overlap of the state, after each element has changed its state 15 times, and its initial state, is obtained, that is, $d = d(p)$. Note, it has been confirmed that each element almost reaches a balanced state. The overlap of vectors $\mathbf{x} = (x_i)$ and $\mathbf{y} = (y_i)$ is defined by

$$d \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=0}^{N} x_i y_i \tag{34}$$

An overlap is obtained by calculating the mean of 3200 values obtained with a different random number. The data stored in a memory circuit $I(p)$ are given by

$$I(p) = \frac{Np}{2} \{(1-d)\log(1-d) + (1+d)\log(1+d)\} \tag{35}$$

Then, the memory capacity is given by

$$\alpha_c = \frac{1}{N^2} \max_p I(p) \tag{36}$$

This definition coincides with the definition stated by Eq. (22), if an overlap $m$ for $N \to \infty$ is approximated by

$$m(\alpha) = \begin{cases} 1 & \alpha \leqq \alpha_c \\ 0 & \alpha > \alpha_c \end{cases}$$

In this paper, an optimum quantum function was obtained by using this equation and by calculating the memory capacity with the discrete synaptic weight.

#### 4.2. Results of simulation

Figure 10 shows the relationship between the number of the 4-level quantum functions ($n = 1$) and the capacity of the memory, the number of elements being $N = 32$.

In the simulation, $a_1$ around 0.3 and $b_1$ around 1.0 give the maximum memory capacity. These values are close to the theoretical values in Fig. 7 which are $a_1 = 0.2994$ and $b_1 = 0.9780$. This shows that even a simulation with only 32 elements produces values close to the theory. If more simulations are carried out, values closer to the theory would be obtained. In other words, in a case using a finite number of elements, its quantum function can be obtained theoretically. The optimum quantum function is dependent of the distribution of the synaptic weights. However, its memory capacity changes little compared with a case where a quantization is rough.

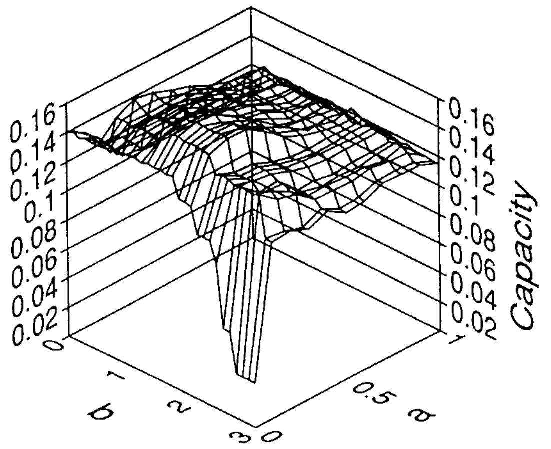


Fig. 10. Dependence on 4-level quantum function to capacity of a system with finite cells.

## 5. Conclusions

When a synaptic weight is represented by 1 bit (two levels), the memory capacity is about 0.114, and when that is represented by 2 bits (four levels), the memory capacity is about 0.125. Therefore, it is obvious that the doubling of elements increases the memory capacity. The main problem is to obtain an optimum memory capacity with a multiple-bits representation. If there are an infinite number of bits, its quantum function becomes a linear function $F(x) = x$. When the number of bits is finite, a quantum function which makes the memory capacity optimum soon approaches linear, but the memory capacity does not increase significantly compared with the case of an infinite number of bits. It is considered that this is due to the scattered data in the circuit.

In this paper, the relationship between a quantized synaptic weight and a memory capacity has been investigated taking a correlation learning with 1/2 coding alone. It is necessary to investigate in the future the effect of a quantized noise of a synaptic weight (which is determined by other learning methods, such as the error-correction learning) on the circuit, and the effect of space coding.

**Acknowledgement.** This work has been supported by the Inamori Fund, and Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, General Research (No. 02402035) and Priority Research (No. 05267229, Higher-Order Information Processing in the Brain). The authors which to express their thanks to these organizations.

## REFERENCES

1. J. J. Hopfield. Neural network and physical system with emergent collective computational abilities. Proc. Natl. Acad. Sci. (USA), **79**, pp. 2554-2558 (1982).
2. I. Ogura and K. Kasahara. Applications of VSTEP to optical connection and optical-signal processing. Technical Reports of Institute of TV, Japan, **14**[7], pp. 25-30, IPU'90-5 (Jan. 1990).
3. S. Tai and K. Hisama. Neural computer and optical computer. Collected Papers of Specialized Lectures, IECE Japan, 40-46 (Jan. 1991).
4. S. Amari and K. Maginu. Statistical neurodynamics of associative memory. Neural Networks, **1**, pp. 63-73 (1988).
5. M. Shiino and T. Fukai. Self-consistent signal-to-noise analysis and its application to analogue neural network with asymmetric connections. J. Phys. A: Math. Gen., **25**, L375-L381 (1992).
6. M. Okada. A Hierarchy of Macrodynamical Equations of Associative Memory. To appear In: Neural Networks.
7. D. J. Amit, H. Gutfreund and H. Sompolinsky. Statistical Mechanics of Neural Network near Saturation. Annals of Physics, **173**, 30-67 (1987).
8. D. J. Amit. The Properties of Models in Simple Neural Networks. In: Heidelberg Colloquium on Glassy Dynamics, edited by J. L van Hemman and I. Morgenstern (Springer-Verlag, 1986, pp. 430-484).
9. H. Sompolinsky. The Theory of Neural Networks: The Hebb Rule and Beyond. In: Heidelberg Colloquium on Glassy Dynamics, edited by J. L. van Hemmen and I. Morgenstern (Springer-Verlag, 1968, pp. 485-528).

## AUTHORS (from left to right)



**Kazushi Mimura** graduated from Biological Engineering Department, Faculty of Engineering Science, Osaka University in 1992, and received his Master's degree from the same university in 1994, and at present, he is in the doctoral program.

**Masato Okada** graduated from Physics Department, Faculty of Engineering Science, Osaka City University in 1985, and the first half of the Postgraduate program at Osaka University in 1991. He joined Mitsubishi Electric Co. Ltd. in 1991. He became a Research Associate (Bio-Engineering) on the Faculty of Engineering Science, Osaka University in 1993. He is engaged in research and development on semiconductor laser, and neural networks.

**Koji Kurata** graduated in 1981 from Computer Department, University of Tokyo, where he received his Master's degree in 1983. He was appointed Research Associate (Computer) at the same university and Lecturer (Bio-Engineering) at Osaka University in 1990. He is engaged in research on pattern formation theory in response diffusion systems, neural networks, and skin-map formation theory.