Recurrent Correlation Associative Memories

Tzi-Dar Chiueh, *Member, IEEE,* and Rodney M. Goodman, *Member, IEEE*

***Abstract*—This paper presents a model for a class of high-capacity associative memories. Since they are based on two-layer recurrent neural networks and their operations depend on the correlation measure, we call these associative memories recurrent correlation associative memories (RCAM's). The RCAM's are shown to be asymptotically stable in both synchronous and asynchronous (sequential) update modes as long as their weighting functions are continuous and monotone nondecreasing. In particular, a new high-capacity RCAM named the exponential correlation associative memory (ECAM) is proposed. The asymptotic storage capacity of the ECAM scales exponentially with the length of memory patterns, and it meets the ultimate upper bound for the capacity of associative memories. Furthermore, the asymptotic storage capacity of the ECAM with limited dynamic range in its exponentiation nodes is found to be proportional to that dynamic range. This paper also reports a 3 μm CMOS ECAM chip, which has been designed and fabricated. The prototype chip can store 32 24-bit memory patterns, and its speed is faster than one associative recall operation every 3 μs. An application of the ECAM chip to vector quantization is also described.**

I. Introduction

SINCE the seminal work of Hopfield [1], [2], there has been much interest in building associative memories using neural network approaches. The storage capacity of the Hopfield memory has been found, both empirically [1] and theoretically [3], to scale less than linearly (approximately $N/\log N$) with the number of components in memory patterns. Psaltis and Park [4], Dembo and Zeitouni [5], [6], and Sayeh and Han [7] all proposed new architectures that utilize nonlinear circuits and correlations between memory patterns and the input pattern. Previously, we also proposed a new associative memory model that adopts the exponentiation function [8], [9]. These models can all be implemented by a two-layer recurrent network: the first layer computes the correlations of the current-state pattern and all the memory patterns, followed by some nonlinear weighting function; the second layer calculates a weighted sum of all memory patterns and thresholds that sum to produce the next-state patterns. Since these recurrent neural network associative memories are based on the correlation measure, we call them recurrent correlation associative memories (RCAM's).

In Section II, a model for the RCAM is presented. Also, some known associative memories are shown to be instances of the RCAM model. Section III deals with the convergence property of the RCAM's. By defining a Liapunov ("energy") function and demonstrating that it never increases, the RCAM's are shown to be asymptotically stable in both synchronous and asynchronous update modes if their weighting functions are continuous and monotone nondecreasing. Section IV concentrates on a particular model called the exponential correlation associative memory (ECAM). The relationship between the storage capacity and the attraction radius of the ECAM is investigated. If all state patterns inside a hypersphere of some attraction radius centered at a memory pattern are to be attracted, in one iteration, to that memory pattern with very high probability, then as N (the number of components in the memory patterns) approaches infinity, the storage capacity (the maximum number of memory patterns) is proportional to c^N. The constant c is greater than 1 and it decreases as the attraction radius increases. More importantly, we find that under certain conditions the asymptotic storage capacity of the ECAM meets the ultimate upper bound for the capacity of associative memories derived by the sphere-packing arguments in [10]. We also find that the asymptotic storage capacity of the ECAM is proportional to the dynamic range of its exponentiation circuits if that dynamic range is limited. In Section V, we present the results of some simulation experiments of the ECAM, which confirm the theoretical findings about the asymptotic storage capacity of the ECAM, even though N is not excessively large. VLSI implementation of the ECAM and its application to an associative recall problem are discussed in Section VI.

II. A Model for Recurrent Correlation Associative Memories

Let $\boldsymbol{x}$ and $\boldsymbol{w}$ be two N-bit bipolar patterns whose components are either $+1$ or -1; then the correlation of $\boldsymbol{x}$ and $\boldsymbol{w}$ is denoted by

$$\langle \boldsymbol{x}, \boldsymbol{w} \rangle \equiv \sum_{j=1}^{N} x_j w_j. \tag{1}$$

Note that $\langle \boldsymbol{x}, \boldsymbol{w} \rangle = N - 2d_{\text{Hamming}}(\boldsymbol{x}, \boldsymbol{w})$. Now, let $\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \cdots, \boldsymbol{u}^{(M)}$ be the M N-bit bipolar ($+1$ or -1) memory patterns to be stored in an RCAM. Also, let $\boldsymbol{x}$ be the N-bit bipolar current-state pattern and $\boldsymbol{x}'$ be the N-bit bipolar next-state pattern; then the evolution equation (motion equation) of the RCAM is defined as

$$\boldsymbol{x}' = \operatorname{sgn}\left\{ \sum_{k=1}^{M} f_k(\langle \boldsymbol{u}^{(k)}, \boldsymbol{x} \rangle) \cdot \boldsymbol{u}^{(k)} \right\} \tag{2}$$

where the f_k's are called weighting functions. Fig. 1 illustrates the architecture of the RCAM's. Matrix $\boldsymbol{U}$ is an $M \times N$ matrix made up of the M memory patterns $\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \cdots, \boldsymbol{u}^{(M)}$ and $y_k = \langle \boldsymbol{u}^{(k)}, \boldsymbol{x} \rangle$. Let us now describe how some known neural network associative memories can be expressed as instances of the RCAM.

A. Correlation-Matrix Associative Memory

This model is essentially the same as the Hopfield memory except that it does not have a feedback connection in the original form and the diagonal of the connection weight matrix is

Manuscript received May 9, 1990; revised November 9, 1990.

T.-D. Chiueh was with the Department of Electrical Engineering, California Institute of Technology. He is now with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan 10764.

R. M. Goodman is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 9041928.

1045-9227/91/0300-0275$01.00 © 1991 IEEE

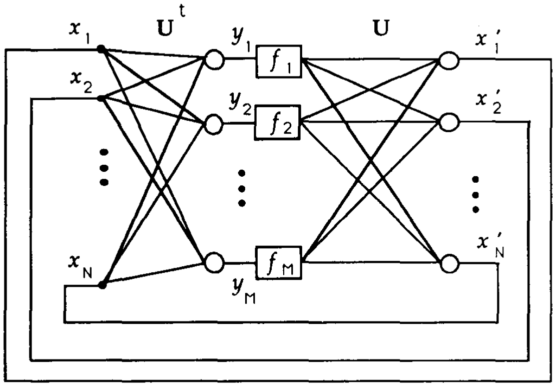

Fig. 1. Architecture of the recurrent correlation associative memories. Matrix U is an $M \times N$ matrix made up of M N-bit bipolar ($+1$ or -1) memory patterns, $u^{(k)}$, $k = 1, 2, \cdots, M$. x and x' are the current-state and the next-state patterns, respectively. The f_k's are the weighting functions.

not zeroed [11], [12]. The connection weight matrix is given by

$$T_{ij} = \sum_{k=1}^{M} u_i^{(k)} \cdot u_j^{(k)}.$$

It is easily shown that the correlation-matrix associative memory is an instance of the RCAM with all f_k's equal to $f(\cdot)$ and

$$f(t) \equiv t.$$

B. High-Order Correlation Associative Memory

In these types of associative memories [4], the evolution equation is (2) with all f_k's equal to $f(\cdot)$ and

$$f(t) \equiv (N + t)^q$$

where q is an integer greater than 1. The storage capacity of the high-order correlation associative memory is asymptotically proportional to N^q.

C. Potential-Function Correlation Associative Memory

Dembo and Zeitouni [5], [6] and Sayeh and Han [7] independently introduced similar models that utilize a potential-type function. Originally, they are continuous-time systems with real-valued patterns. Nonetheless, it is straightforward to express these models in discrete-time formulation with bipolar patterns. The evolution equation then takes the form of (2) with all f_k's equal to $f(\cdot)$ and

$$f(t) = (N - t)^{-L}$$

where $L \geq 1$. The storage capacity of this model grows exponentially with the number of components in memory patterns [6]. The primary disadvantage of this model is that hardware implementation of the potential function can be cumbersome.

D. Exponential Correlation Associative Memory

We have introduced the exponential correlation associative memory (ECAM), which is an instance of the RCAM with all f_k's equal to an exponentiation function [8], [9], i.e.,

$$f_k(t) \equiv a^t$$

where $a > 1$. The storage capacity of the ECAM will be explored in Section IV.

F. Other Recurrent Correlation Associative Memories

In principle, as soon as one comes up with a weighting function $f(\cdot)$, one builds an RCAM. However, the important thing is to find a weighting function that is easy to implement and that produces an RCAM that is asymptotically stable and has a large storage capacity. In the next section, we will present a condition on the weighting function $f(\cdot)$ that is sufficient for the asymptotic stability of the corresponding RCAM.

III. The Convergence Property of the RCAM's

Since the RCAM's are based on a recurrent network structure, understanding their asymptotic behavior is important. Hopfield [1] proved that his model is asymptotically stable when running in the asynchronous update mode (when only one neuron in the output layer updates itself at a time). At first, he introduced a Liapunov ("energy") function of the system, and went on to demonstrate that the Liapunov function either decreases or stays the same after each iteration. Moreover, he showed that the energy function has a lower bound and that the system cannot stay at the same energy level forever. These facts imply that the Hopfield memory will eventually reach a stable state with minimum "energy" level. However, if the Hopfield memory is running in the synchronous update mode (all neurons in the output layer update themselves at the same time), it may not converge to a fixed point, but may instead become oscillatory between two states [13]. In this section, we prove that the first four RCAM's in the previous section are all asymptotically stable in both synchronous and asynchronous update modes.

To begin with, let us introduce a lemma.

Lemma 1: Let $f(t)$ be continuous and monotone nondecreasing over $[-N, N]$; then the RCAM with the following evolution equation,

$$x' = \text{sgn}\left\{\sum_{k=1}^{M} f(\langle u^{(k)}, x\rangle) \cdot u^{(k)}\right\}$$

is asymptotically stable in both synchronous and asynchronous (sequential) update modes.

Proof: see Appendix I.

Theorem 2: The first four RCAM's in the previous section are all asymptotically stable in both synchronous and asynchronous (sequential) update modes.

Proof: First of all, all four weighting functions in the previous section are continuous. Also, for any $t_1 > t_2$, we have

$$t_1 \geq t_2$$

$$(N + t_1)^q \geq (N + t_2)^q$$

$$(N - t_1)^{-L} \geq (N - t_2)^{-L}$$

$$a^{t_1} \geq a^{t_2}$$

where $q > 1$, $L \geq 1$, and $a > 1$. Consequently, Lemma 1 can be applied and the theorem proved. ■

The significance of Lemma 1 is that it ensures that one can employ any continuous, monotone nondecreasing weighting function and the resulting RCAM will be asymptotically stable in both synchronous and asynchronous update modes. This proves to be very helpful when it comes to hardware implementation of RCAM's, because any physical device exhibits some deviation from its ideal characteristic. Accordingly, as long as the real response of the nonlinear circuits is continuous and monotone nondecreasing, the RCAM will always be asymptotically stable, although its performance in storage capacity and error-correction ability may become poorer. However, Lemma 1 gives only a sufficient condition for an RCAM to be stable; it says nothing about necessary conditions.

IV. The Capacity and the Attraction Radius of the ECAM

The ECAM seems to be the one RCAM that is most amenable to VLSI implementation; therefore this section is devoted to an exploration of the storage capacity and the attraction radius of the ECAM. Our definition of the storage capacity is somewhat similar to that given by McEliece *et al.* [3]. Suppose we choose $M = M(N)$ N-bit memory patterns at random, program an ECAM with those M patterns, and initialize that ECAM with an input pattern $r (r \equiv \rho N$, and $0 \leq \rho < 1/2)$ bits away from the nearest memory pattern. We then ask, as $N \to \infty$, what the greatest rate of growth of $M(N)$ is so that after one iteration the bit-error probability (the probability that a bit in the next-state pattern is different from the corresponding bit in the nearest memory pattern) is less than $(4\pi T)^{-1/2} e^{-T}$, where T is a fixed and large number. By adjusting T, one can make a tradeoff between the bit-error probability and the storage capacity of an ECAM.

To begin with, assume that all M N-bit memory patterns $\boldsymbol{u}^{(k)}$, $k = 1, 2, \cdots, M$, are randomly chosen; in other words, each bit in any of the M memory patterns is the outcome of a Bernoulli trial ($+1$ or -1). Let us now present the theorem about the storage capacity of the ECAM.

Theorem 3: Suppose an ECAM is loaded with

$$M(N) = \begin{cases} [a^4/(4T)]\, 2^{N(1 - \mathcal{H}(\rho'))} + 1 \\ \qquad \text{if } \rho' \geq (1 + a^2)^{-1} \\ [a^4/(4T)]\, 2^N [(1 + a^{-2}) a^{2\rho'}]^{-N} + 1 \\ \qquad \text{if } \rho' < (1 + a^2)^{-1} \end{cases} \tag{3}$$

N-bit memory patterns, where $\rho' = \rho + (1/N)$, $0 \leq \rho < 1/2$, and $\mathcal{H}(\rho')$ is the binary information entropy of ρ'. If the current-state pattern $\boldsymbol{x}$ is ρN bits away from the nearest memory pattern, then as $N \to \infty$, the bit-error probability (P_e) is less than $(4\pi T)^{-1/2} e^{-T}$.

Proof: Only an outline of the proof is given here; details can be found in Appendix II. Since the current-state pattern $\boldsymbol{x}$ is assumed to be ρN bits away from the nearest memory pattern, say $\boldsymbol{u}^{(l)}$, the evolution equation then takes the form

$$\begin{aligned} \boldsymbol{x}' &= \operatorname{sgn} \left\{ \sum_{k=1}^{M} a^{\langle u^{(k)}, x \rangle} \boldsymbol{u}^{(k)} \right\} \\ &= \operatorname{sgn} \left\{ a^{N(1-2\rho)} \boldsymbol{u}^{(l)} + \sum_{k=1, k \neq 1}^{M} a^{\langle u^{(k)}, x \rangle} \boldsymbol{u}^{(k)} \right\}. \end{aligned}$$

Considering only the ith component of $\boldsymbol{x}'$ and letting $u_i^{(l)} = -1$ without loss of generality yields

$$x_i' = \operatorname{sgn} \left\{ -a^{N(1-2\rho)} + \sum_{k=1, k \neq 1}^{M} a^{\langle u^{(k)}, x \rangle} u_i^{(k)} \right\}. \tag{4}$$

Note that the second term of the argument of the sgn function in (4) is a sum of $(M - 1)$ independent, identically distributed (i.i.d.) random variables.

Now define

$$\omega_k \equiv a^{\langle u^{(k)}, x \rangle} u_i^{(k)}, \qquad k = 1, 2, \cdots, M$$

and let

$$\omega \equiv \sum_{k=1, k \neq l}^{M} \omega_k$$

$$v \equiv \sum_{k=1}^{M} \omega_k = -a^{N(1-2\rho)} + \omega.$$

After some lengthy derivation (see Appendix II), we have the following results in order:

$$E[\omega] \ll a^{N(1-2\rho)}$$

and

$$\operatorname{Var}[\omega] < a^{2N(1-2\rho)}/(2T).$$

As $N, M \to \infty$, the central limit theorem [14] can be applied, which leads to

$$\begin{aligned} P_e = \operatorname{Prob}[v > 0] &= \operatorname{Prob}[\omega > a^{N(1-2\rho)}] \\ &< (4\pi T)^{-1/2} e^{-T}. \end{aligned} \tag{5}$$

■

Therefore, we conclude that the ECAM has a storage capacity that scales *exponentially* with N—the number of bits in the memory patterns. In other words, the ECAM can store c^N memory patterns—all N-bit wide—and still be capable of some error correction. The base constant c actually depends on two parameters, a and ρ. Refer to Fig. 2 and see how c decreases with smaller a and larger ρ. Also note that c is never less than 1. More importantly, in the case where $\rho' \geq (1 + a^2)^{-1}$, one has, as $N \to \infty$,

$$\begin{aligned} [\log_2 M(N)]/N &= 1 - \mathcal{H}(\rho') + [4 \log_2(a) \\ &\quad - \log_2(4T)]/N \\ &\simeq 1 - \mathcal{H}(\rho') \\ &\simeq 1 - \mathcal{H}(\rho). \end{aligned}$$

Hence, when $\rho' \geq (1 + a^2)^{-1}$ the asymptotic storage capacity of the ECAM meets the ultimate upper bound for the capacity of associative memories [10] (the $a = \infty$ curve in Fig. 2).

This exponential capacity is very attractive; however, the dynamic range required of the exponentiation circuit also grows exponentially with N. In any real implementation, this requirement is very difficult to meet, if not impossible. In a typical CMOS VLSI process, a transistor operating in the subthreshold region working as an exponentiation circuit has a dynamic range of approximately 10^5 to 10^7 [15]. Thus, we need to study how the storage capacity of the ECAM changes if the dynamic range of its exponentiation circuits is limited.

Suppose the dynamic range (D) of the exponentiation circuits is fixed and

$$D \equiv a^N.$$

Then as N increases, a will decrease and M will no longer scale exponentially with N. We now concentrate on the case where N

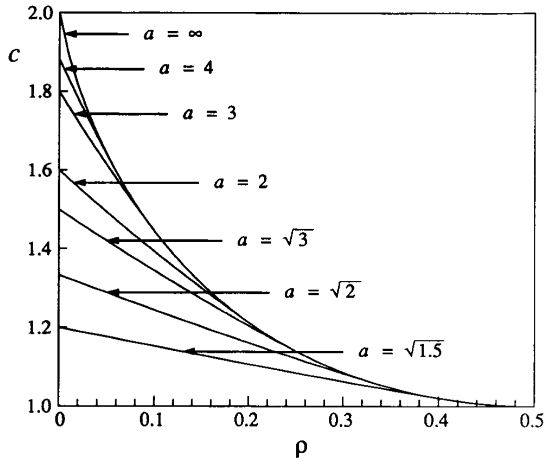

Fig. 2. Relationship of the base constant c and the two parameters, a and ρ. The ECAM has an exponential storage capacity that is proportional to c^N.

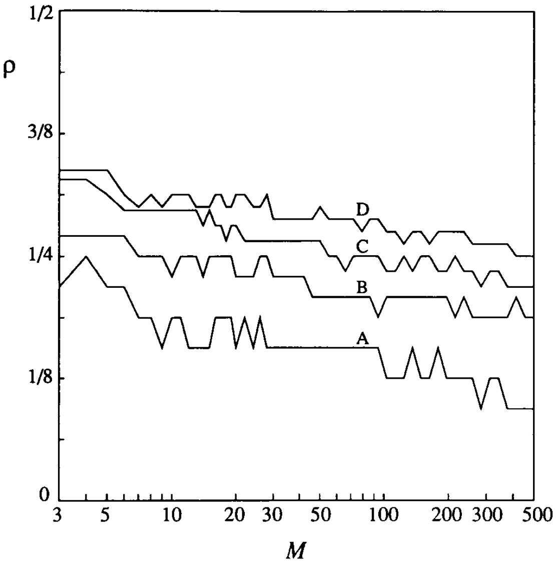

Fig. 3. Normalized attraction radius (ρ, $\rho = r/N$) versus number of stored memory patterns (M) in the ECAM, Curve A: $N = 32$. Curve B: $N = 48$. Curve C: $N = 64$. Curve D: $N = 80$.

approaches infinity. Since N is very large and D is fixed, a will be near 1. Let

$$a \equiv 1 + \mu$$

where μ is a small positive number; then

$$\log D = N \log a = N \log (1 + \mu) \simeq N\mu.$$

As a approaches 1, ρ' will be less than $(1 + a^2)^{-1}$ in practically all cases (remember that $\rho < 1/2$ and $\rho' = \rho + 1/N$); therefore, only the second formula in (3) need be considered. It follows that with fixed D and as N approaches infinity,

$$\begin{aligned} M(N) &= [a^4/(4T)]\, 2^N[(1 + a^{-2})a^{2\rho'}]^{-N} \\ &\simeq [a^4/(4T)]\, 2^N [(2 - 2\mu)(1 + 2\rho'\mu)]^{-N} \\ &\simeq [a^4/(4T)]\, \left(1 + [N\mu(1 - 2\rho')]/N\right)^N \\ &\simeq [a^4/(4T)]\, e^{N\mu(1-2\rho)} \\ &\simeq [a^4/(4T)]\, D^{1-2\rho}. \end{aligned} \tag{6}$$

From the above equation, one sees that the asymptotic storage capacity of the ECAM is proportional to the dynamic range (D) when the required attraction radius (ρ) is 0. However, as the attraction radius is increased, the storage capacity decreases. These findings are not at all discouraging since they say that the ECAM can be only as good as one of its components—the exponentiation circuit.

V. Simulation Results

Simulations have been conducted in order to confirm the theoretical results about the storage capacity of the ECAM. We let $a = 2$ and randomly choose ten sets of M N-bit memory patterns. For each set of M memory patterns, program an ECAM with these M patterns. For each ECAM, 100 initial-state patterns are generated by randomly picking a memory pattern and flipping d bits. They are then fed to the ECAM and the ECAM is allowed to run until it becomes stable. The resulting fixed point is then compared with the original memory pattern, and the run is called a *success* if they match, a *failure* otherwise. The number of successes out of 1000 runs is then collected. If this number is greater than 998, we say that loaded with M memory patterns, the ECAM can tolerate d errors. The largest d for a fixed M is called the *attraction radius* (r).

In Fig. 3, the normalized attraction radius (ρ, $\rho = r/N$) is plotted against the number of memory patterns (M) for various N. Note that if a horizontal line is drawn across the plot, it will intersect the four curves in the figure at points with equidistant intervals. Since the four curves correspond to the cases where N increases linearly, the previous observation implies that for a fixed ρ the storage capacity of the ECAM scales exponentially with N, confirming Theorem 3. Next we fix N at 32 and vary the dynamic range of the exponentiation circuits. Fig. 4 illustrates how the relationship between the attraction radius (r) and the number of loaded memory patterns (M) changes for different dynamic ranges when $N = 32$. As one can easily see, the curves intersect the vertical axis ($r = 1$) at four points, each of which is approximately twice as large as the point before. Since the dynamic ranges of these four curves double successively, the storage capacity of the ECAM is thus proportional to the dynamic range of the exponentiation circuits for fixed N. Furthermore, if one draws a vertical line at larger r, it again intersects the four curves at points equidistantly apart, although with a smaller interval than the previous case. Therefore, we conclude that the previous result about the storage capacity of the ECAM with fixed-dynamic-range exponentiation circuits (i.e., (6)) is valid.

VI. VLSI Implementation of the ECAM

In the previous sections, we have introduced a model for the recurrent correlation associative memories. We addressed, in particular, the case where the weighting functions are exponential, namely, the ECAM. The evolution equation of the ECAM

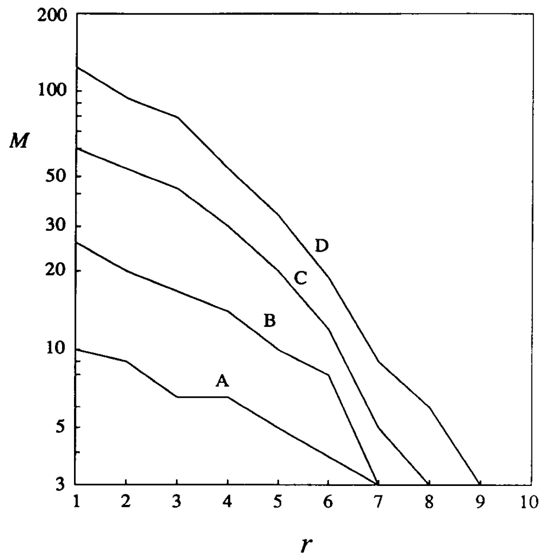

Fig. 4. Number of stored memory patterns (M) versus attraction radius (r) of the ECAM with $N = 32$ and fixed dynamic range, D. Curve A: $D = 2^4$. Curve B: $D = 2^5$. Curve C: $D = 2^6$. Curve D: $D = 2^7$.

is given by

$$x' = \text{sgn}\left\{ \sum_{k=1}^{M} a^{\langle u^{(k)}, x \rangle} u^{(k)} \right\} \tag{7}$$

where a is a constant greater than 1.

The ECAM chip we have designed is *programmable*; that is, one can change the set of memory patterns stored in an ECAM chip at will. To perform an associative recall, one first loads a set of memory patterns onto the chip. The chip is then switched to the associative recall mode, and an input pattern is presented to the ECAM chip. The ECAM chip then computes the next-state pattern according to (7) and presents it at the output port of the chip. No clock signal is necessary since after the internal circuits have settled the components of the next-state pattern appear in parallel at the output port. Feedback is easily incorporated by connecting the output port to the input port. Details on the circuits of the ECAM chip have previously been presented [16]. Here, only a brief description of the chip is given.

A. Design of the ECAM Circuits

From the evolution equation of ECAM, one notices that there are essentially three operations that need to be carried out:

- $\langle u^{(k)}, x \rangle$: correlation computation;
- $\Sigma_{k=1}^{M} a^{\langle u^{(k)}, x \rangle} u^{(k)}$: exponentiation, multiplication, and summation;
- sgn $(\cdot)$: thresholding.

For easy VLSI implementation, we designed a basic ECAM cell (see Fig. 5) that realizes all the aforementioned computations. An ECAM that holds M N-bit memory patterns can be constructed by replicating the basic ECAM cell M times in the vertical direction and N times in the horizontal direction.

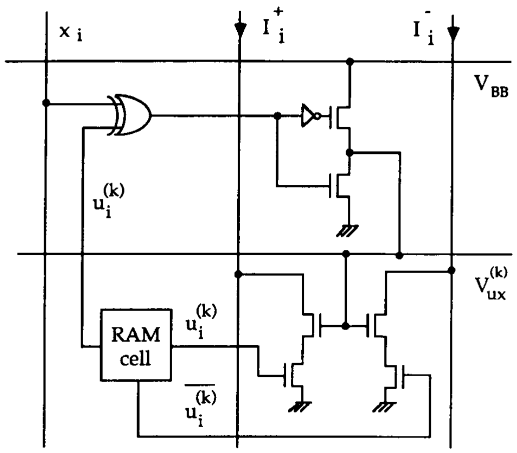

Fig. 5. Circuit diagram of the basic ECAM cell.

A voltage-divider circuit consisting of NMOS transistors working as controlled resistors (linear resistors or open circuits) computes the correlation between the input pattern x and a memory pattern $u^{(k)}$. The output voltage ($V_{ux}^{(k)}$) is proportional to the number of positions at which x and $u^{(k)}$ match. The maximum output voltage is controlled by an externally supplied voltage V_{BB}.

The exponentiation function is implemented by an NMOS transistor whose gate voltage is set to $V_{ux}^{(k)}$. If V_{BB} is set to be around the threshold voltage (V_{TH}), the NMOS transistor is in the subthreshold region, where its drain current depends exponentially on its gate-to-source voltage [17]. Since the multiplier $u_i^{(k)}$ is either $+1$ or -1, the multiplication is done by forming two branches, each made up of a pass transistor in series with an exponentiation transistor. One of the two pass transistors is controlled by $u_i^{(k)}$, the other by the complement of $u_i^{(k)}$. Summation of M terms in the evolution equation is done by current summing. The final results are two currents, I_i^+ and I_i^-.

The thresholding is done by comparing these two currents, which can be implemented by the top portion of a differential amplifier. The result of that comparison determines the sign of the ith bit of the next-state pattern, x_i'.

B. The ECAM Chip and Test Results

The complete ECAM chip includes 32 × 24 ECAM cells, read/write circuit, sense amplifiers, row decoders, and I/O multiplexers. It is then fabricated on a 3 μm CMOS process; the total chip area is 47 mm^2. Since all circuits in the ECAM chip, including the exponentiation transistors, would function normally when they are scaled down, one can fabricate higher capacity ECAM chips using more advanced technologies, e.g., 1 μm CMOS technology.

The ECAM chip has been fully tested and the results show that the chip performs almost as well as a computer simulation of the ECAM. Fig. 6 illustrates the testing results of the ECAM chip. The number of successful associative recalls in 1000 trials is plotted against the number of errors in input patterns for the following four cases: 1) a simulation with $a = 2$; 2) The ECAM chip with $V_{BB} = 5$ V; 3) $V_{BB} = 2$ V; and 4) $V_{BB} = 1$ V. As the number of errors increases, the number of successes decreases. Also, one notices that the simulated ECAM is by far the best case, which is expected because the ECAM chip is only an approximation of the ECAM model and thus will definitely do worse.

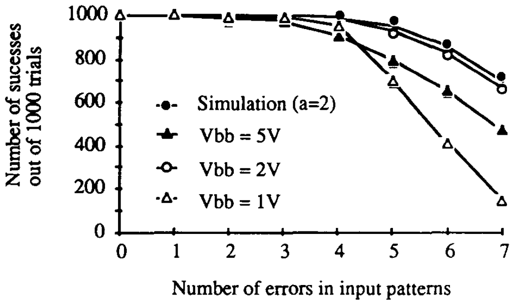

Fig. 6. Error-correcting ability of the ECAM chip with different V_{BB} and an ECAM simulation with $a = 2$.

The best performance of the ECAM chip occurs when V_{BB} is set to 2 V, which is about twice the threshold voltage. This phenomenon arises from two contradicting effects brought about by increasing V_{BB}. On the one hand, increasing V_{BB} increases the dynamic range of the exponentiation transistors in the ECAM chip, thus improving the error-correcting ability of the ECAM chip. On the other hand, as V_{BB} increases beyond the threshold voltage, the exponentiation transistors leave the subthreshold region and may enter saturation, where the drain current is approximately proportional to the *square* of the gate-to-source voltage. Since the second-order correlation associative memory in general possesses a smaller storage capacity than the ECAM, one would expect that with a fixed number of memory patterns, the ECAM should do better than the second-order correlation associative memory. To sum up, two contradicting effects are going on as V_{BB} is raised. One tends to enhance the performance of the ECAM chip, while the other tends to degrade it. A compromise between these two effects is reached, and the best performance is achieved when $V_{BB} = 2$ V.

In the case when $V_{BB} = 2$ V, the drain current versus gate-to-source voltage characteristic of the exponentiation transistors is actually a hybrid of a square function and an exponentiation function. At the bottom it is exponential, and it gradually flattens out to a square function once the gate-to-source voltage becomes larger than the threshold voltage. Since the overall characteristic is still continuous and monotone increasing, the ECAM chip operating at $V_{BB} = 2$ V is asymptotically stable.

C. A Vector-Quantization Example

We have tested the speed of the ECAM chip using binary-image vector quantization as an example problem. Vector quantization is a means of data compression (source coding) on information to be transmitted or stored, e.g., speech waveforms or images [18]. In principle, given a set of code words and an input, a vector quantizer should find the nearest code word to the input. Then only the index of the nearest code word is transmitted or stored instead of the information itself. Usually, the number of possible code words is much smaller than the number of possible information patterns, thereby reducing the required transmission/storage bandwidth.

Each pixel in the test images is either black or white. At first, input images were partitioned into 4 × 4 blocks, and each block was vector-quantized by the ECAM chip. A set of 32 code words are chosen, and they correspond to all-white, all-black, horizontal-edge, vertical-edge, and diagonal-edge blocks. Note that the choice of these code words was totally heuristic and not optimized in any way since the objective of this experiment was to apply the ECAM chip to solve a real problem and to measure the speed of the chip. The ECAM chip was programmed with these code words, and 4 × 4 blocks from a binary image were fed to the ECAM chip one at a time. The nearest code word to each input block then appeared at the output of the ECAM chip. The indices of those code words could then be transmitted or stored, achieving a compression ratio of 16/5. A reconstructed image was formed by replacing each block by the nearest code word. However, there are times when the output pattern of the ECAM chip is not a code word; in this case an all-white block is generated instead. Fig. 7 illustrates an original binary image and its ECAM-chip-reconstructed image. It is obvious that the reconstructed binary image is not as good as the original; yet this is the price paid for reduced bandwidth. In addition, any real application would optimize the code words for less distortion.

Working on the above task, the ECAM chip performed one associative recall operation on a 4 × 4 block in less than 3 μs (this includes the communication time between the ECAM chip and the controlling computer). This projects to about 49 ms for a 512 × 512 binary image, or more than 20 images per second—fast enough for real-time applications. If one simulates the ECAM on a serial digital computer, it would take approximately 3072 simple instructions (multiply or add instructions) plus other complex operations for the computer to perform one associative recall operation. Therefore, in terms of associative recall operations the ECAM chip runs faster than a 1024 MIPS serial digital computer.

VII. Conclusion

In this paper, we have proposed a model for a group of associative memories called the recurrent correlation associative memories. We also proved that these RCAM's are asymptotically stable as long as their weighting functions are continuous and monotone nondecreasing. In particular, a new high-capacity RCAM called the exponential correlation associative memory (ECAM) was presented. We have also shown that the asymptotic storage capacity of the ECAM scales *exponentially* with the length of memory patterns. It was also found that under certain conditions the asymptotic storage capacity of the ECAM meets the ultimate upper bound for the capacity of associative memories. Nevertheless, in order to store M N-bit memory patterns, one needs $M \times N$ connection weights, M exponentiation nodes, and N hard-limiter neurons. Hence, to store an exponential number of memory patterns, exponential hardware complexity is required. We believe that this is not discouraging since it means only that one can store as much information in the ECAM as one's hardware allows. Moreover, the asymptotic storage capacity of the ECAM with fixed dynamic range in its exponentiation nodes is found to be proportional to that dynamic range.

Simulation results confirming the theoretical findings about the attraction radius and the storage capacity of the ECAM were also presented. A VLSI chip based on the ECAM model was fabricated and tested. The ECAM chip was shown to perform almost as well as the computer simulation of the ECAM. The speed of the chip was measured by employing it to do vector quantization on binary images. It was found that the ECAM chip can process binary images in real time, i.e., about 20–30 images every second.

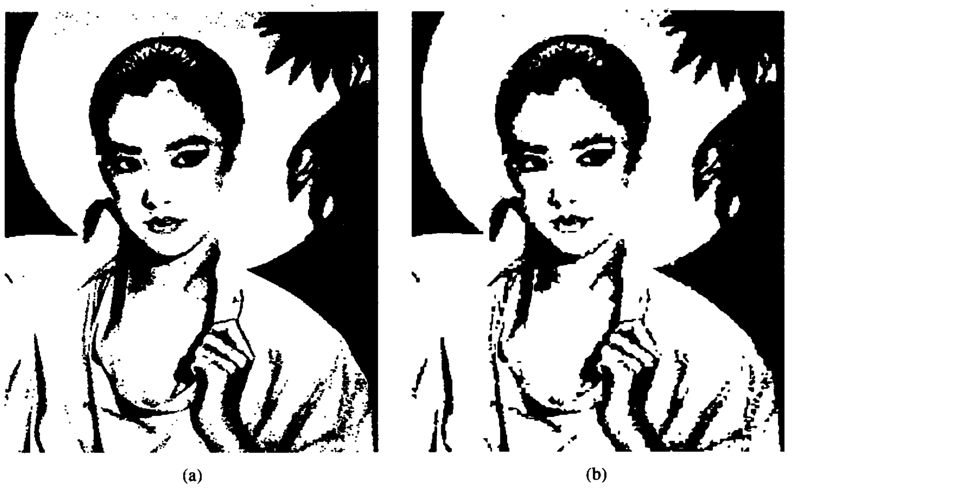

Fig. 7. Comparison of (a) the original girl image and (b) the reconstructed girl image after vector quantization by the ECAM chip.

Appendix I

Lemma 1: Let $f(t)$ be continuous and monotone nondecreasing over $[-N, N]$; then the RCAM with the following evolution equation,

$$x' = \operatorname{sgn}\left\{\sum_{k=1}^{M} f(\langle u^{(k)}, x\rangle) \cdot u^{(k)}\right\}$$

is stable in both synchronous and asynchronous (sequential) update modes.

Proof: At first, define

$$g(x) \equiv \int^{x} f(t)\,dt.$$

By the mean-value theorem, for any x and y, $y \neq x$, there exists some z lying between x and y so that

$$g(y) - g(x) = g'(z) \cdot (y - x) = f(x) \cdot (y - x).$$

By the assumption that $f(t)$ is monotone nondecreasing, we have the following inequality:

$$g(y) - g(x) \geq f(x) \cdot (y - x) \qquad \forall x, y.$$

Now let the Liapunov ("energy") function of the RCAM be defined as

$$E(x) \equiv -\sum_{k=1}^{M} g(\langle u^{(k)}, x\rangle).$$

Suppose all neurons now update themselves according to the above evolution equation at the same time; the difference in the Liapunov function between the current state and the next state is then given by

$$\begin{aligned}
\Delta E &= E(x') - E(x) \\
&= -\sum_{k=1}^{M} [g(\langle u^{(k)}, x'\rangle) - g(\langle u^{(k)}, x\rangle)] \\
&\leq -\sum_{k=1}^{M} f(\langle u^{(k)}, x\rangle) \cdot [\langle u^{(k)}, x'\rangle - \langle u^{(k)}, x\rangle] \\
&= -\sum_{k=1}^{M} f(\langle u^{(k)}, x\rangle) \cdot \langle u^{(k)}, x' - x\rangle \\
&= -\sum_{k=1}^{M} f(\langle u^{(k)}, x\rangle) \cdot \sum_{i=1}^{N} [u_i^{(k)} \cdot (x_i' - x_i)] \\
&= -\sum_{i=1}^{N} \left[\left\{\sum_{k=1}^{M} f(\langle u^{(k)}, x\rangle)\, u_i^{(k)}\right\} \cdot (x_i' - x_i)\right] \\
&\leq 0.
\end{aligned}$$

The last inequality comes about because x_i' and $\Sigma_{k=1}^{M} f(\langle u^{(k)}, x\rangle)\, u_i^{(k)}$ are of the same sign. Also, x_i' and $(x_i' - x_i)$ are of the same sign if $x_i' \neq x_i$. Moreover, if $\Delta E = 0$, then for $i = 1, 2, \cdots, N$ either $x_i' = x_i$ or $x_i' \neq x_i$ and $\Sigma_{k=1}^{M} f(\langle u^{(k)}, x\rangle)\, u_i^{(k)} = 0$. Note that in the latter case $x_i' = +1$ and $x_i = -1$. Hence, after a finite number of -1 or $+1$ changes, no more state changes are possible if $\Delta E = 0$. Also, note that E is bounded from below; therefore, the RCAM eventually becomes stable at a fixed point. By the same token, if only one neuron updates itself in every iteration, the RCAM also converges to a fixed point. ■

Appendix II

Theorem 3: Suppose an ECAM is loaded with

$$M(N) = \begin{cases} [a^4/(4T)]\, 2^{N(1-\mathcal{H}(\rho'))} + 1 \\ \qquad \text{if } \rho' \geq (1 + a^2)^{-1} \\ [a^4/(4T)]\, 2^N[(1 + a^{-2})a^{2\rho'}]^{-N} + 1 \\ \qquad \text{if } \rho' < (1 + a^2)^{-1} \end{cases}$$

N-bit memory patterns, where $\rho' = \rho + (1/N)$, $0 \leq \rho < 1/2$, and $\mathcal{H}(\rho')$ is the binary information entropy of ρ'. If the current-state pattern x is ρN bits away from the nearest memory pattern, then as $N \rightarrow \infty$, the bit-error probability (P_e) is less than $(4\pi T)^{-1/2}e^{-T}$.

Proof: For a given ρ, $0 \leq \rho < 1/2$, suppose the ECAM is initialized with an N-bit bipolar $(+1$ or $-1)$ state pattern $\boldsymbol{x}$ that is $r(r = \rho N)$ bits away from the nearest memory pattern, say $\boldsymbol{u}^{(l)}$; in other words,

$$\boldsymbol{x} = \boldsymbol{u}^{(l)} + \boldsymbol{e}$$

where $\boldsymbol{e}$ has r nonzero $(+2$ or $-2)$ components.

Assume, without loss of generality, that $u_i^{(l)} = -1$; the ith component of $\boldsymbol{x}'$ is then given by

$$x_i' = \operatorname{sgn}\left\{-a^{N(1-2\rho)} + \sum_{k=1, k \neq l}^{M} a^{\langle u^{(k)}, x\rangle} u_i^{(k)}\right\}.$$

Now define

$$\omega_k \equiv a^{\langle u^{(k)}, x\rangle} u_i^{(k)}, \qquad k = 1, 2, \cdots, M$$

and let

$$\omega \equiv \sum_{k=1, k\neq l}^{M} \omega_k$$

$$v \equiv \sum_{k=1}^{M} \omega_k = -a^{N(1-2\rho)} + \omega.$$

Since $\boldsymbol{u}^{(l)}$ is the nearest memory pattern to x, all other $(M - 1)$ memory patterns must be at least $r + 1$ bits away from $\boldsymbol{x}$. Now define

$$r' \equiv r + 1 \quad \text{and} \quad \rho' \equiv r'/N = \rho + (1/N).$$

Furthermore, the bit-error probability is larger for the case where $x_i \neq u_i^{(l)}$ $(e_i = +2)$ than when $x_i = u_i^{(l)}$ $(e_i = 0)$; hence only the former case will be studied. The probability distribution function of the random variable ω_1 when $e_i = +2$ (i.e., $x_i = +1$) can be formulated as

$$\text{Prob}\,[\omega_1 = a^{N-2j}] = (1/K)\binom{N-1}{j}, \qquad j = r', r'+1, \cdots, N-1$$

$$\text{Prob}\,[\omega_1 = -a^{N-2j-2}] = (1/K)\binom{N-1}{j}, \qquad j = r'-1, r', \cdots, N-1.$$

The first formula applies to the case where $u_i^{(1)} = +1$ and $\boldsymbol{u}^{(1)}$ and $\boldsymbol{x}$ differs at j positions, while the second applies to the case where $u_i^{(1)} = -1$ and $\boldsymbol{u}^{(1)}$ and $\boldsymbol{x}$ differ at $j + 1$ positions. The constant K is a normalizing factor and

$$\begin{aligned}K &= \sum_{j=r'}^{N-1}\binom{N-1}{j} + \sum_{j=r'-1}^{N-1}\binom{N-1}{j}\\ &= \sum_{j=r'}^{N-1}\left[\binom{N-1}{j} + \binom{N-1}{j-1}\right] + \binom{N-1}{N-1}\\ &= \sum_{j=r'}^{N-1}\binom{N}{j} + \binom{N}{N} = \sum_{j=r'}^{N}\binom{N}{j}.\end{aligned}$$

Note that $\rho < 1/2$; hence $r' = \rho N + 1 \leq \lceil N/2 \rceil$ and

$$K \geq \sum_{j=\lceil N/2\rceil}^{N}\binom{N}{j} \geq 2^{N-1}. \tag{8}$$

Next let us bound the expectation of ω_1 from above:

$$\begin{aligned}E[\omega_1] &= (1/K)\left\{\sum_{j=r'}^{N-1}\binom{N-1}{j}a^{N-2j} - \sum_{j=r'-1}^{N-1}\binom{N-1}{j}a^{N-2j-2}\right\}\\ &< 2^{-(N-1)}a^N\left\{\sum_{j=r'}^{N-1}\binom{N-1}{j}a^{-2j}\right\}\\ &< 2^{-(N-1)}a^N\left\{\sum_{j=r'}^{N}\binom{N}{j}a^{-2j}\right\}.\end{aligned}$$

In order to express the upper bound analytically, the Chernoff method is applied. Multiplying each term in the summation by a number greater than or equal to 1 $(e^{t(j-r')}, t \geq 0)$ and summing from $j = 0$ instead of from $j = r'$ gives

$$\begin{aligned}E[\omega_1] &< 2^{-(N-1)}a^N\left\{\sum_{j=0}^{N}\binom{N}{j}a^{-2j}e^{t(j-r')}\right\}\\ &< 2^{-(N-1)}a^Ne^{-tr'}\left\{\sum_{j=0}^{N}\binom{N}{j}(a^{-2}e^t)^j\right\}\\ &= 2^{-(N-1)}a^Ne^{-tr'}(1 + a^{-2}e^t)^N, \qquad \text{where } t \geq 0.\end{aligned}$$

Similarly, the expectation of ω_1^2 can be bounded from above:

$$\begin{aligned}E[\omega_1^2] &= (1/K)\left\{\sum_{j=r'}^{N-1}\binom{N-1}{j}a^{2N-4j} + \sum_{j=r'-1}^{N-1}\binom{N-1}{j}a^{2N-4j-4}\right\}\\ &= (a^{2N}/K)\left\{\sum_{j=r'}^{N-1}\binom{N-1}{j}a^{-4j} + \sum_{j=r'}^{N}\binom{N-1}{j-1}a^{-4j}\right\}\\ &= (a^{2N}/K)\left\{\sum_{j=r'}^{N}\binom{N}{j}a^{-4j}\right\}\\ &< 2^{-(N-1)}a^{2N}\left\{\sum_{j=0}^{N}\binom{N}{j}a^{-4j}e^{t(j-r')}\right\}\\ &= 2^{-(N-1)}a^{2N}e^{-tr'}(1 + a^{-4}e^t)^N, \qquad \text{where } t \geq 0.\end{aligned}$$

Accordingly, the variance of ω_1 is

$$\begin{aligned}\text{Var}\,[\omega_1] &= E[\omega_1^2] - E[\omega_1]^2\\ &\leq E[\omega_1^2]\\ &< 2^{-(N-1)}a^{2N}e^{-tr'}(1 + a^{-4}e^t), \qquad \text{where } t \geq 0.\end{aligned}$$

Since ω is the sum of $(M - 1)$ i.i.d. random variables, the expectation and the variance of ω are both $(M - 1)$ times those of ω_1; namely,

$$\begin{aligned}E[\omega] &= (M-1)E[\omega_1]\\ &< (M-1)2^{-(N-1)}a^Ne^{-tr'}(1 + a^{-2}e^t)^N, \qquad (9)\\ &\quad \text{where } t \geq 0\end{aligned}$$

$$\begin{aligned}\mathrm{Var}\,[\omega] &= (M-1)\,\mathrm{Var}\,[\omega_1]\\ &< (M-1)2^{-(N-1)}a^{2N}e^{-tr'}(1+a^{-4}e^{t})^{N}, \qquad (10)\\ &\quad \text{where } t \geq 0.\end{aligned}$$

To estimate the bit-error probability, we need to deal with two cases separately. The first is for $\rho' \geq (1+a^2)^{-1}$. Since (9) and (10) are valid for all nonnegative t, we can find an optimal t so that the right-hand sides of both inequalities are minimized. In (9), let

$$\begin{aligned}e^{t} &= (a^2\rho')/(1-\rho')\\ &\geq [a^2/(1+a^2)][1-(1+a^2)^{-1}]^{-1} = 1.\end{aligned}$$

Then

$$\begin{aligned}E[\omega] &< (M-1)2^{-(N-1)}a^{N}[(1-\rho')/(a^2\rho')]^{\rho'N}(1-\rho')^{-N}\\ &= (M-1)2^{-(N-1)}a^{N(1-2\rho')}(\rho')^{-\rho'N}(1-\rho')^{-(1-\rho')N}\\ &= 2(M-1)a^{N(1-2\rho')}2^{N(\mathcal{H}(\rho')-1)}\end{aligned}$$

where $\mathcal{H}(x) = -x\log_2 x - (1-x)\log_2(1-x)$, the binary information entropy of ρ'. Assume that T is large and let

$$M(N) = [a^4/(4T)]\,2^{N(1-\mathcal{H}(\rho'))} + 1. \qquad (11)$$

It then follows that

$$E[\omega] < [a^2/(2T)]\,a^{N(1-2\rho)} \ll a^{N(1-2\rho)}. \qquad (12)$$

Similarly, the variance of ω can be bounded from above by substituting

$$e^{t} = (a^4\rho')/(1-\rho') \geq a^2 > 1$$

in (10). Hence,

$$\begin{aligned}\mathrm{Var}\,[\omega] &< (M-1)2^{-(N-1)}a^{2N(1-2\rho')}(\rho')^{-\rho'N}\\ &\quad\cdot(1-\rho')^{-(1-\rho')N}\\ &= 2(M-1)a^{2N(1-2\rho')}2^{N(\mathcal{H}(\rho')-1)}.\end{aligned}$$

Substituting (11) in the above equation yields

$$\mathrm{Var}\,[\omega] < a^{2N(1-2\rho)}/(2T). \qquad (13)$$

The second case is for $\rho' < (1+a^2)^{-1}$. Substituting $e^{t} = 1$ in (9) gives

$$\begin{aligned}E[\omega] &< (M-1)2^{-(N-1)}a^{N}(1+a^{-2})^{N}\\ &= (M-1)2^{-(N-1)}a^{N(1-2\rho')}[(1+a^{-2})a^{2\rho'}]^{N}.\end{aligned}$$

Now suppose that T is large and

$$M(N) = [a^4/(4\mathrm{T})]\,2^{N}[(1+a^{-2})a^{2\rho'}]^{-N} + 1. \qquad (14)$$

Then

$$E[\omega] < [a^2/(2T)]\,a^{N(1-2\rho)} \ll a^{N(1-2\rho)}. \qquad (15)$$

Next, an upper bound of Var $[\omega]$ can be found by setting $e^{t} = a^2$ in (10):

$$\begin{aligned}\mathrm{Var}\,[\omega] &< (M-1)2^{-(N-1)}a^{2N(1-\rho')}(1+a^{-2})^{N}\\ &= (M-1)2^{-(N-1)}a^{2N(1-2\rho')}[(1+a^{-2})a^{2\rho'}]^{N}.\end{aligned}$$

Combining (14) and the above equation leads to

$$\mathrm{Var}\,[\omega] < a^{2N(1-2\rho)}/(2T). \qquad (16)$$

We have shown in both cases that $E[\omega]$ is significantly smaller than $a^{N(1-2\rho)}$ when T is large and thus can be ignored when compared with $a^{N(1-2\rho)}$. Also, Var $[\omega]$ is found to be bounded from above by the same quantity in both cases. We now estimate the bit-error probability (P_e) of the ECAM, namely, the probability that $v > 0$. Since the random variable ω is the sum of $(M-1)$ i.i.d. random variables, as $N, M \to \infty$, ω can be approximated by a normal distribution (the central limit theorem [14]). Therefore,

$$\begin{aligned}\mathrm{Prob}\,[v > 0] &= \mathrm{Prob}\,[\omega > \mathrm{a}^{N(1-2\rho)}]\\ &\simeq \mathrm{Prob}\,[\omega - E[\omega] > a^{N(1-2\rho)}]\\ &= Q(a^{N(1-2\rho)}/\sigma_\omega) < Q(\sqrt{2T}) \qquad (17)\end{aligned}$$

where σ_ω is the standard deviation of ω and

$$Q(t) = \frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-x^2/2}\,dx.$$

Note that since T is fixed, we do not have to worry about the large-deviation problem in applying the central limit theorem. If T is large, we can use the asymptotic formula for $Q(\cdot)$:

$$Q(t) \simeq \frac{1}{\sqrt{2\pi}}t^{-1}e^{-t^2/2}.$$

By the above formula and (17), one has

$$\begin{aligned}P_e &= \mathrm{Prob}\,[v > 0]\\ &< (4\pi T)^{-1/2}e^{-T}. \qquad (18)\end{aligned}$$

■

References

[1] J. J. Hopfield, "Neural network and physical systems with emergent collective computational abilities," *Proc. Nat. Acad. Sci. U.S.*, vol. 79, pp. 2554-2558, 1982.

[2] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci. U.S.*, vol. 81, pp. 3088-3092, 1984.

[3] R. J. McEliece, E. C. Posner, E. R. Rodemich, and S. S. Venkatesh, "The capacity of the Hopfield associative memory," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 461-482, 1987.

[4] D. Psaltis and C. H. Park, "Nonlinear discriminant functions and associative memories," in *Neural Networks for Computing*, J. S. Denker, Ed. New York, NY: American Institute of Physics, 1986, pp. 370-375.

[5] A. Dembo and O. Zeitouni, "High density associative memories," in *Neural Information Processing Systems*, D. Z. Anderson, Ed. New York, NY: American Institute of Physics, 1988, pp. 211-218.

[6] A. Dembo and O. Zeitouni, "General potential surfaces and neural networks," *Phys. Rev. A*, vol. 37, no. 6, pp. 2134-2143, 1988.

[7] M. R. Sayeh and J. Y. Han,"Pattern recognition using a neural network," in *Proc. SPIE Cambridge Symp. Opt. and Optoelec. Eng.* (Cambridge, MA), Nov. 1987.

[8] T. D. Chiueh and R. M. Goodman, "High-capacity exponential associative memory," in *Proc. IEEE Int. Conf. Neural Networks* (San Deigo, CA), vol. 1, 1988, pp. 153-160.

[9] T. D. Chiueh, "Pattern classification and associative recall by neural networks," Ph.D. dissertation, Department of Electrical Engineering, California Institute of Technology, 1989.

[10] P. A. Chou, "The capacity of the Kanerva associative memory is exponential," in *Neural Information Processing Systems*, D. Z. Anderson, Ed. New York, NY: American Institute of Physics, 1988, pp. 184-191.

[11] T. Kohonen, "Correlation matrix memories," *IEEE Trans. Comput.*, vol. C-21, pp. 353-359, 1972.

[12] J. A. Anderson, "A simple neural network generating an interactive memory," *Math. Biosci.*, vol. 14, pp. 197-220, 1972.

[13] J. Bruck and J. W. Goodman, "A generalized convergence theo-

rem for neural networks and its applications in combinatorial optimization," in *Proc. IEEE Int. Conf. Neural Networks* (San Diego, CA), vol. III, 1987, pp. 649-656.

[14] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd ed. New York, NY: Wiley, 1971.

[15] L. A. Glasser and D. W. Dopperpuhl, *The Design and Analysis of VLSI Circuits.* Reading, MA: Addison-Wesley, 1985.

[16] T. D. Chiueh and R. M. Goodman, "VLSI implementation of a high-capacity neural network associative memory," in *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, 1990, pp. 793-800.

[17] C. A. Mead, *Analog VLSI and Neural Systems.* Reading, MA: Addison-Wesley, 1989.

[18] R. M. Gray, "Vector quantization," *IEEE ASSP Magazine*, vol. 1, pp. 4-29, 1984.

*

Tzi-Dar Chiueh (S'86-M'89) received the B.S.E.E. degree from the National Taiwan University in 1983 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology in 1986 and 1989, respectively.

Since 1989, he has been with the Department of Electrical Engineering, National Taiwan University, where he is presently an Associate Professor. His research interests include neural networks, analog VLSI design, and VLSI signal processing.

Dr. Chiueh is a member of the International Neural Network Society.

*

Rodney M. Goodman (M'85) was born in London, England, on February 22, 1947. He received the B.Sc. degree in electrical engineering from Leeds University, Yorkshire, U.K., in 1968, and the Ph.D. in electronics from the University of Kent at Canterbury, U.K., in 1975.

In 1985 he joined the faculty of the Department of Electrical Engineering at the California Institute of Technology as Associate Professor. His research interests are in error control coding, cryptography, neural networks, and expert systems—from both a theoretical and a VLSI implementation viewpoint. He has consulted for a wide variety of government and commercial organizations, and is a founder of two advanced technology research and development companies in the U.K. He is currently a consultant for the Jet Propulsion Laboratory and for Pacific Bell.

Dr. Goodman is a Chartered Electrical Engineer of the I.E.E. in the U.K.