An Analysis of High-Capacity Discrete Exponential BAM

Chua-Chin Wang and Hon-Son Don

Abstract— **An exponential bidirectional associative memory (eBAM) using an exponential encoding scheme is discussed. It has higher capacity for pattern pair storage than the conventional BAM's. A new energy function is defined. The associative memory takes advantage of the exponential nonlinearity in the evolution equations such that the signal-noise-ratio (SNR) is significantly increased. The energy of the exponential BAM decreases as the recall process proceeds, ensuring the stability of the system. The increase of SNR consequently enhances the capacity of the BAM. The capacity of the exponential BAM is estimated.**

I. Introduction

KOSKO [7], [8] proposed a two-level nonlinear network, bidirectional associative memory (BAM), which extends a one-directional process to a two-directional process. One beneficial characteristic of the BAM is its ability to recall stored pattern pairs in the presence of noise. However, researchers have discovered a major shortcoming of the BAM, i.e., its limited capacity of stored pattern pairs. In order to overcome this shortcoming, researchers have spent very much efforts to improve the storage capacity of the BAM. Wang *et al.* [13] proposed two alternatives, multiple training and dummy augmentation to enhance BAM's ability to find the global minimum. Wang's approaches enhance the recall probability of pattern pairs in a case-by-case way, but the generalization of these methods is still questionable. Haines and Hecht-Nielsen proposed a nonhomogeneous BAM [5]; Simpson proposed an intraconnected BAM and a high-order autocorrelator [11]; and Tai *et al.* [12] proposed a high-order BAM. All of these works pay a price of increasing the complexity of the network and do not get much increase of the storage of the pattern pairs. In these works, the basic idea to improve the recall probability is to enlarge the attraction radius of a stored pattern pair, or enhance the desired pattern power and reduce the unwanted patterns' influence. Chiueh and Goodman [3] proposed an exponential Hopfield associative memory motivated by the MOS transistor's exponential drain current dependence on the gate voltage in the subthreshold region such that the VLSI implementation of an exponential function is feasible. Based upon the concept of Chiueh's exponential Hopfield associative memory, Jeng *et al.* proposed one kind of exponential BAM [6]. However, the energy function proposed in [6] cannot guarantee that every stored pattern pair will have a local minimum on the energy surface. Moreover, there is no capacity analysis given in [6].

Manuscript received May 7, 1992; revised December 7, 1992. This work was supported by the National Science Council of the Republic of China, Grants NSC82–0408–E-005–084 and NSC83–0408–E-110–015.

C.-C. Wang is with the Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 80424, Republic of China.

H.-S. Don is with the Department of Electrical Engineering, National Chung-Hsing University, Taichung, Taiwan 40227, Republic of China.

IEEE Log Number 9207242.

In this paper, we adopt the exponential form and combine it with the BAM structure in order to enhance the SNR and, consequently, increase the capacity of the BAM. We also propose a new energy function of the BAM system based on the exponential form. The capacity of this exponential BAM (eBAM) is estimated. The simulation result is much more appealing than the previous works.

II. Framework of High-Capacity Exponential BAM

A. Evolution Equations

Suppose we are given N training sample pairs, which are

$$\{(A_1, B_1), (A_2, B_2), \cdots, (A_N, B_N)\} \tag{1}$$

where

$$A_i = (a_{i1}, a_{i2}, \cdots, a_{in}),$$

$$B_i = (b_{i1}, b_{i2}, \cdots, b_{ip}).$$

Let X_i and Y_i be the bipolar mode of the training pattern pairs, A_i and B_i, respectively. That is, $X_i \in \{-1, 1\}^n$ and $Y_i \in \{-1, 1\}^p$. Instead of using Kosko's approach [7], which is

$$(X \rightarrow M \rightarrow Y)$$

$$(X' \leftarrow M^T \leftarrow Y)$$

$$\vdots$$

where $M = \sum_{i=1}^{N} X_i^T Y_i$, we use the following evolution equations in the recall process of the eBAM

$$y_k = \begin{cases} 1, & \text{if } \sum_{i=1}^{N} y_{ik} b^{X_i \cdot X} \geq 0 \\ -1, & \text{if } \sum_{i=1}^{N} y_{ik} b^{X_i \cdot X} < 0 \end{cases} \tag{2}$$

$$x_k = \begin{cases} 1, & \text{if } \sum_{i=1}^{N} x_{ik} b^{Y_i \cdot Y} \geq 0 \\ -1, & \text{if } \sum_{i=1}^{N} x_{ik} b^{Y_i \cdot Y} < 0 \end{cases} \tag{3}$$

where b is a positive number, $b > 1$, "$\cdot$" represents the inner product operator, x_k and x_{ik} are the kth bits of X and X_i, respectively, and y_k and y_{ik} are for Y and Y_i, respectively. The reasons for using an exponential scheme are to enlarge the attraction radius of every stored pattern pair and to augment the desired pattern in the recall reverberation process.

B. Energy Function and Stability

Since every stored pattern pair should produce a local minimum on the energy surface, the energy function is intuitively defined as

$$E(X, Y) = -\sum_{i=1}^{N} b^{X_i \cdot X} - \sum_{i=1}^{N} b^{Y_i \cdot Y}. \tag{4}$$

Assume $E(X', Y)$ is the energy of next state in which Y stays the same as in the previous state. Hence, $\Delta E_x = -\sum_{i=1}^{N} b^{X_i \cdot X'} - (-\sum_{i=1}^{N} b^{X_i \cdot X})$. Assume the ith pair is the target of the recall process. Let d_{xi} be the Hamming distance between X and X_i, d'_{xi} the Hamming distance between the X' and X_i. Hence the ΔE_x can be modified to be

$$\begin{aligned}\Delta E'_x &= -\sum_{i=1}^{N} \log_b (b^{n-2d'_{xi}}) + \sum_{i=1}^{N} \log_b (b^{n-2d_{xi}}) \\ &= -\sum_{i=1}^{N}\sum_{k=1}^{n}(x'_k - x_k)x_{ik}. \end{aligned} \tag{5}$$

Note that log is used, which is a monotonic function. From the recall process shown by (3) and (5), the $\Delta E_x < 0$ is ensured. Because (3) makes $(x'_k - x_k)x_{ik}$ always nonnegative such that $\Delta E'_x \leq 0$, and

$$\begin{aligned}&\Delta E'_x \leq 0 \Rightarrow -\sum_{i=1}^{N} \log_b (b^{n-2d'_{xi}}) \leq -\sum_{i=1}^{N} \log_b (b^{n-2d_{xi}}) \\ &\Rightarrow -\sum_{i=1}^{N} b^{X_i \cdot X'} \leq -\sum_{i=1}^{N} b^{X_i \cdot X} \\ &\Rightarrow \Delta E_x \leq 0.\end{aligned}$$

Obviously, it also holds for the other case: $E(X, Y') \leq E(X, Y)$ if the pair is heading for a stored pair, (X_i, Y_i). Since the $E(X, Y)$ is bounded by $-N(b^n + b^p) \leq E(x, y) \leq -N(b^{-n} + b^{-p})$ for all X and Y, the energy of the exponential BAM will converge to a stable local minimum.

C. Analysis of Capacity of Exponential BAM

We adopt the SNR approach [3] to compute the capacity of the exponential BAM. Equation (2) can be rewritten as

$$y_k = \operatorname{sgn}\left(\sum_{i=1}^{N} y_{ik} b^{X_i \cdot X}\right) = \operatorname{sgn}\left(b^n y_{hk} + \sum_{i \neq h}^{N} y_{ik} b^{X_i \cdot X}\right) \tag{6}$$

where the Y_h is assumed to be the desired pattern, and y_{hk} is its kth bit. The first term in the above equation corresponds to the signal, the second term is the noise. Hence the power of the signal is

$$S = b^{2n}. \tag{7}$$

The second term is actually a sum of $N-1$ independent identically distributed random variables. Therefore, the variance of the second term is $N-1$ times the variance of a single random variable. Let

$$\begin{aligned} v_1 &= y_{1j} b^{X_i \cdot X} \\ v_2 &= y_{2j} b^{X_i \cdot X} \\ &\vdots \\ v_N &= y_{Nj} b^{X_i \cdot X}. \end{aligned}$$

Since all of the v_i's have the same property, we select v_1 as the sample. It is trivial to derive the following probability functions for v_1

$$\Pr(v_1 = b^{n-2-2k}) = \left(\frac{1}{2}\right)^{n-1} C_k^{n-1} \tag{8}$$

$$\Pr(v_1 = -b^{n-2-2k}) = \left(\frac{1}{2}\right)^{n-1} C_k^{n-1} \tag{9}$$

where k is the Hamming distance between X and X_i. The mean of the noise term is obviously zero. Then the variance can be derived as

$$\begin{aligned} E[v_1^2] &= 2\sum_{k=0}^{n-1} b^{2(n-2k-2)} \left(\frac{1}{2}\right)^{n-1} C_k^{n-1} \\ &= 2\sum_{k=0}^{m} b^{2(m-2k-1)} \left(\frac{1}{2}\right)^{m} C_k^{m}, \quad \text{where } m = n-1 \\ &= 2\left(\frac{1}{2}\right)^{m} b^{2(m-1)} \sum_{k=0}^{m} (b^{-4})^k 1^{(m-k)} C_k^m \\ &= 2\left(\frac{1}{2}\right)^{m} b^{-2}(b^2 + b^{-2})^m \\ &= 2\left(\frac{1}{2}\right)^{n-1} \frac{(b^2 + b^{-2})^{n-1}}{b^2}. \end{aligned}$$

Hence we can conclude the following result for the signal-noise-ratio of the exponential BAM, i.e., the capacity of the exponential BAM,

$$\mathrm{SNR}_{\mathrm{eBAM}} = \frac{2^{n-1} b^4}{2(N-1)(1 + b^{-4})^{n-1}}. \tag{10}$$

However, since the BAM is a bidirectional reverberation process, it will be reasonable to use $r = \min(n, p)$ instead of n in the above result. Hence the SNR should be rewritten as

$$\mathrm{SNR}_{\mathrm{eBAM}} = \frac{2^{r-1} b^4}{2(N-1)(1 + b^{-4})^{r-1}}, \tag{11}$$

where $r = \min(n, p)$.

By using the same SNR approach, we can analyze the SNR of Kosko's BAM [8]. The derivation is given in the Appendix.

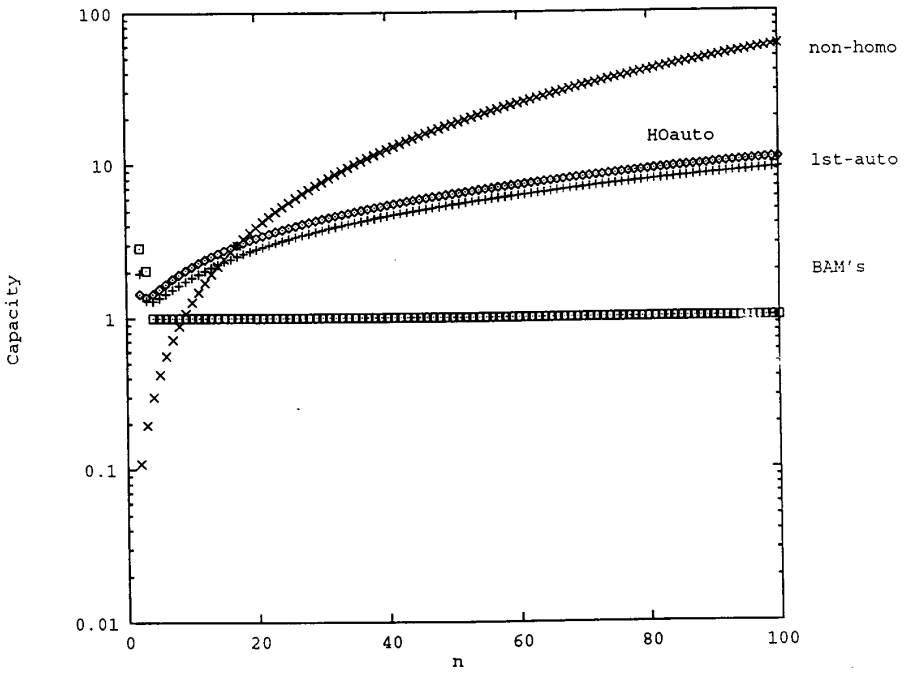

Fig. 1. Capacity of different types of BAM.

$$\begin{aligned}\mathrm{SNR_{BAM}} &= \frac{n^2}{2(N-1)\left(\frac{1}{2}\right)^{n-1}\sum_{k=0}^{n-1}(n-2k-2)^2 C_k^{n-1}} \\ &= \frac{n^2}{2(N-1)\left(\frac{1}{2}\right)^{n-1}2^{n-1}\cdot n} \\ &= \frac{n}{2(N-1)}. \end{aligned} \tag{12}$$

Similarly, the above result will be rewritten as

$$\mathrm{SNR_{BAM}} = \frac{r}{2(N-1)}, \quad \text{where } r = \min(n, p). \tag{13}$$

This result corresponds to the estimation in Kosko's work [7], [8], which states that the upper bound of the BAM's capacity is less than the minimum of the dimensionalities.

Comparing the results shown in (10) and (12), we can tell how significant an increase of capacity the exponential BAM has. If b is sufficiently big and N is large, the denominator of (10) approaches $2N$

$$\mathrm{SNR_{eBAM}} \approx \frac{2^{n-1}b^4}{2N}, \qquad b \gg 1. \tag{14}$$

From the above result, if we want a good recall probability, i.e., a high SNR, we have to pay the price of increasing either the dimensionality n or base b, or decreasing the number N of stored pattern pairs. The result meets what we expect intuitively.

III. Simulation Analysis

Amari and Maginu [1] performed the capacity analysis for the first-order autocorrelator, which has capacity

$$C_{1st\text{-}auto} = \frac{n}{2\log n + \log\log n}. \tag{15}$$

Baldi and Venkatesh [2] did the same analysis for higher-order autocorrelator, which has capacity

$$C_{HOauto} = \frac{n^{g-1}}{2n!\log n}. \tag{16}$$

As for the capacity of the BAM, Haines and Hecht-Nielsen [5] estimated it to be

$$C_{BAM} = \frac{r}{2\log r}, \quad r = \min(n, p). \tag{17}$$

Tai *et al.* [12] proposed a high-order BAM. They did not estimate or prove the possible capacity of the high-order BAM except claiming a better recall probability. However, we still can reasonably expect that the capacity of the high-order BAM will be about that of the high-order autocorrelator, because 1) the BAM is intrinsically a variety of the autocorrelator, so the high-order BAM is certainly a variety of the high-order autocorrelator, and 2) if n is large, the capacity shown in (15) is about the same as that in (17).

Haines and Hecht-Nielsen [5] proposed another variety of BAM, i.e., *nonhomogeneous* BAM. They proved and enlarged the capacity to be

$$C_{non\text{-}homo} = (0.68)\frac{n^2}{(\log_2 n + 4)^2}. \tag{18}$$

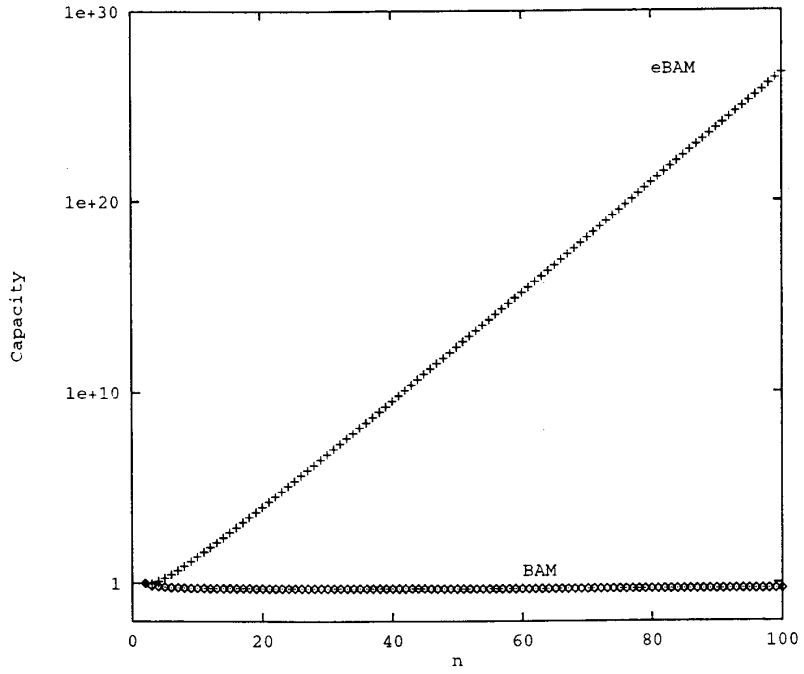

Fig. 2. The capacity of the exponential BAM.

All of the above analyses are plotted in Fig. 1 and Fig. 2. In Fig. 1, the previous works are plotted together to show the comparison, while in Fig. 2, the capacities of the proposed exponential BAM and Kosko's BAM are drawn. Because the numerical values of the exponential BAM is relatively much larger than those of the conventional BAM's, a log scale is used so that the contrast is clear. In Fig. 2, the b is e and $N = n$. From the figures, we can see the great improvement of the proposed exponential BAM in contrast to those previous works.

IV. Conclusion

The exponential BAM provides a significantly higher capacity of storage for pattern pairs. It utilizes an exponential scheme to magnify the SNR. The proposed energy function ensures that every stored pattern pair is located in a local minimum of the energy surface. The monotonic decrease of the energy of the exponential BAM during the recall process ensures its convergence to a local minimum, while maintaining the stability of the system. The capacity of the exponential BAM is estimated, so the size of the exponential BAM can be predetermined by the demand of capacity.

Appendix

According to Kosko's formulation [7]

$$Y = X \cdot M = (X \cdot X_h^T)Y_h + \sum_{i \neq h}^{N} (X \cdot X_i^T)Y_i \qquad (19)$$

$$= nY_h + \sum_{i \neq h} (X \cdot X_i^T)Y_i. \qquad (20)$$

It is natural to assume that the stored pattern pairs are drawn from $\{-1, 1\}^n$ with uniform probability. Hence, the first term corresponds to the signal which has the power n^2. The second term, the noise, has a zero mean and variance as follows

$$E[v_1^2] = 2\sum_{k=0}^{n-1}[n - 2(k+1)]^2 \left(\frac{1}{2}\right)^{n-1} C_k^{n-1}$$

$$= 2\sum_{k=0}^{m}[m - 2k - 1]^2 \left(\frac{1}{2}\right)^{m} C_k^m, \qquad m = n - 1.$$

Assume

$$f(x, y) = \sum_{k=0}^{m} x^k y^{m-k} C_k^m = (x + y)^m$$

$$h(y) = \frac{1}{y} f(y^{-1}, y) = \sum_k y^{m-2k-1} C_k^m.$$

Then, we can get

$$yh'(y) = \sum_k (m - 2k - 1) y^{m-2k-1} C_k^m$$

$$(yh'(y))' = \sum_k (m - 2k - 1)^2 y^{m-2k-2} C_k^m.$$

If $y = 1$, then

$$(h'(1))' = \sum_k (m - 2k - 1)^2 C_k^m.$$

It is trivial to derive $(yh'(y))'$ and substitute in $y = 1$. The result is

$$(h'(1))' = 2(m+1)2^{m-1} = n \cdot 2^{n-1}.$$

Hence the signal-to-noise ratio, i.e., the capacity, of Kosko's BAM is

$$\mathrm{SNR}_{\mathrm{BAM}} = \frac{n^2}{2(N-1)\left(\frac{1}{2}\right)^{n-1}2^{n-1}n} = \frac{n}{2(N-1)}. \quad (21)$$

ACKNOWLEDGMENT

The authors wish to thank Prof. J. Murray of the Department of Electrical Engineering of SUNY at Stony Brook for long and fruitful discussions on the SNR problems.

REFERENCES

[1] S. Amri and K. Maginu, "Statistical neurodynamics of associative memory," *Neural Networks*, vol. 1, pp. 63–74, 1988.
[2] P. Baldi and S. Venkatesh, "Number of stable points for spin glasses and neural networks of higher orders," *Phys. Rev. Lett.*, vol. 58, p. 913, 1987.
[3] T. D. Chiueh and R. M. Goodman, "High-capacity exponential associative memory," in *Proc. IJCNN*, vol. I, 1988, pp. 153–160.
[4] R. Hecht-Nielsen, *Neurocoputing*. Reading, MA: Addison-Wesley, 1990.
[5] K. Haines and R. Hecht-Nielsen, "A BAM with increased information storage capacity," in *Proc. IJCNN*, vol. I, 1988, pp. 181–190.
[6] Y.-J. Jeng, C.-C. Yeh, and T.-D. Chiueh, "Exponential bidirectional associative memories,"*IEEE Electron. Lett.*, vol. 26, no. 11, pp. 717–718, May 1990.
[7] B. Kosko, "Adaptive bidirectional associative memory," *Appl. Opt.*, vol. 26, no. 23, pp. 4947–4960, Dec. 1987.
[8] ——, "Bidirectional associative memory," *IEEE Trans. Syst. Man, Cybern.*, vol. 18, no. 1, pp. 49–60, Jan./Feb. 1988.
[9] R. McEliece, E. Posner, E. Rodemich, and S. Venkatesh, "The capacity of Hopfield model," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 461–482, 1987.
[10] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1986.
[11] P. K. Simpson, "Higher-ordered and intraconnected bidirectional associative memory," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, no. 3, May/June 1990.
[12] H. M. Tai, C. H. Wu, and T. L. Jong, "High-order bidirectional associative memory," *Electron. Lett.*, vol. 25, pp. 1424–1425, 1989.
[13] Y.-F. Wang, J. B. Cruz, Jr., and J. H. Mulligan, Jr., "Two Coding strategies for bidirectional associative memory," *IEEE Trans. Neural Networks*, vol. 1, no. 1, Mar. 1990.