Capacity Analysis of the Asymptotically Stable Multi-Valued Exponential Bidirectional Associative Memory

Chua-Chin Wang, Shiou-Ming Hwang, and Jyh-Ping Lee

Abstract— **The exponential bidirectional associative memory (eBAM) has been proposed and proved to be a stable and high capacity associative neural network. However, the intrinsic structure and the evolution functions of this network restrict the representation of patterns to be either bipolar or binary vectors. We consider the promising development of multi-valued systems and then design a multi-valued discrete eBAM (MV-eBAM). The multi-valued eBAM has been proved be asymptotically stable under certain constraints. Although MV-eBAM is also verified to possess high capacity by thorough simulations, there are important characteristics to be explored, including the absolute lower bound of the radix, and the approximate capacity. In order to estimate the capacity of the MV-eBAM, a modified evolution equation is also proposed. Hence, an analytic solution is derived. Besides, a radix searching algorithm is presented such that the absolute lower bound of the radix for this MV-eBAM can be found.**

I. Introduction

SINCE Kosko [10], [11], proposed the bidirectional associative memory (BAM), many researchers have invested efforts on exploring the network's properties and limitations. Due to its intrinsic architecture, the capacity of BAM is unfortunately poor [7], [8]. Thus, researchers started to develop different methods to enlarge the capacity of the associative memory, e.g., Wang *et al.* [16], Simpson [14], and Tai *et al.* [15]. Among these works, it is notable that Chiueh and Goodman [4] proposed 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. Chiueh also proposed an exponential correlation associative memory (ECAM) [5] which is an autocorrelator utilizing the mentioned exponential function of VLSI circuits to enlarge the correlation between stored pattern pairs. Based upon the concept of Chiueh's exponential Hopfield associative memory, Jeng *et al.* proposed one kind of exponential BAM [9]. However, the energy function proposed in [9] cannot guarantee that every stored pattern pair will have a local minimum on the energy surface. Moreover, there is no capacity analysis given in [9]. Although we have estimated the impressive capacity of an eBAM [17], the data representation of BAM or eBAM is still limited to be either bipolar vectors or binary vectors. We consider that the expansion of the data range, i.e., from $\{-1,+1\}^n$ to $\{1,2,\ldots,L\}^n, L \gg 1$, is also a feasible method to enlarge the capacity. It also enlarges the data representation. This observation leads to the multi-valued (or called multivalued) exponential bidirectional associative memory (MV-eBAM).

The multi-valued concept has been successfully applied in Hopfield network for A/D conversion [19], [20]. In addition, Chiueh *et al.* proposed the multivalued exponential correlation associative memory (MV-ECAM) [6]. Though many similarity measures were proposed in this work, the proof of convergence of the network was not shown specifically. Besides, critical features about this kind of network are not thoroughly explored. For example, the capacity and the bound of the radix. The multi-valued concept also causes tremendous interest in the digital circuit design, [3], [12], which indirectly shows the hardware implementation of the such a neural network is feasible.

In this paper, we first propose a modified measure and an energy function for MV-ECAM and then prove the stability of the network. The evolution equations and the energy functions of the MV-eBAM, then, will be presented and the asymptotical stability will be proved. Though the high capacity of MV-eBAM is expected, the analytic form of solution is hard to derive. We then propose modified evolution equations of MV-eBAM such that the capacity can be estimated. The radix of the exponential function plays an important role in this network. We are interested in discovering the minimum of the radix which is able to recall every stored pattern pair. This smallest radix is called the absolute lower bound, which will be derived. Finally, when a set of pattern pairs is given, a radix which is small enough recall every pair in this set will be computed basing on a gradient descent method.

II. Multi-Valued Exponential BAM

Before we introduce the evolution functions and energy functions of the multi-valued eBAM (MV-eBAM), we will discuss the convergence of multi-valued exponential correlation associative memory (MV-ECAM). Then, we will discuss the stability and other characteristics of the MV-eBAM.

Manuscript received October 8, 1994; revised August 12, 1995. This work was supported in part by the National Science Council under Grant NSC 83-0404-E-110-014.

C.-C. Wang, S.-M. Hwang, and J.-P. Lee are with the Department of Electrical Engineering, National Sun Yat-Sen University, Taiwan (e-mail: ccwang@ee.nsysu.edu.tw).

Publisher Item Identifier S 1083-4419(96)05363-0.

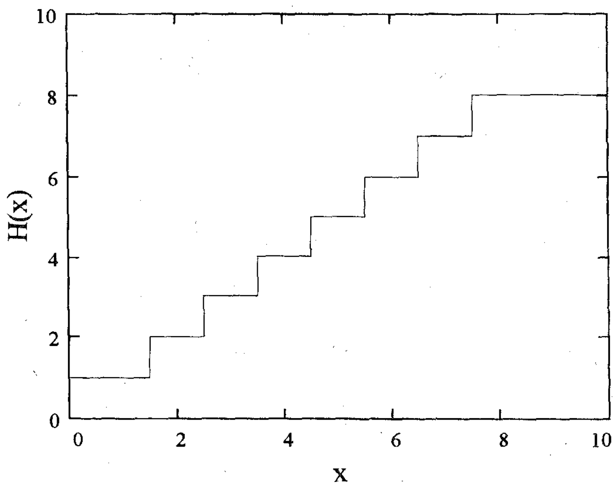


Fig. 1. The staircase function ($L = D = 8$).

A. MV-ECAM

Although Chiueh *et al.* proposed a multivalued exponential recurrent associative memory with several similarity measures [6], they were short of a theoretical proof of the convergence of the network. We proposed a modified similarity measure, which consequently is the measure of the correlation of two pattern vectors, as follows:

$$S = -\|X_1 - X_2\|^2 \tag{1}$$

where $X_1,\ X_2 \in \{1, 2, \ldots, L\}^n$. In the following text, we use the term "digit" to represent the component of the vector.

Thus, S will be largest if $X_1 = X_2$. Based on this norm similarity measure, the evolution function and the energy function of MV-ECAM are, respectively, established below.

Evolution Function:

$$x'_k = H\left(\frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}}\right) \tag{2}$$

where x'_k is the next state of x_k, b is a positive number, $b > 1$, M is the number of patterns in the MV-ECAM, X_i, $i = 1, \ldots, M$, are the stored patterns, X is the initial vector presented to the network, x_k and x_{ik} are the kth digits of X and X_i, respectively, and $H(\cdot)$ is a staircase function shown as the following equation:

$$H(x) = \begin{cases} 1, & x < 1 \\ L, & x > D \\ \lfloor \frac{L}{D} \cdot x + 0.5 \rfloor, & \text{elsewhere} \end{cases} \tag{3}$$

where $l = 1, 2, \ldots, L$, L is the number of finite levels, and D is the finite interval of the staircase function. The graphic representation of the staircase function $H(\cdot)$ is shown in Fig. 1. Note that if $D \to \infty$, and $L \to \infty$, then $H(x) \approx x$, for $x > 0$. The reason why the staircase function is used is that the x in $H(\cdot)$ in (2) is not necessarily a positive integer. Hence, we have to assign this argument to a nearest integer.

Energy Function:

$$E(X) = -\sum_{i=1}^{M} b^{-\|X-X_i\|^2}. \tag{4}$$

This energy function ensures each pattern pair is placed in its own local minimum as long as the radix, b, is large enough.

Stability of MV-ECAM: Based on (4), we can derive the change of energy in every iteration to be negative as shown in the following. Assume x'_k is the next state of x_k, then

$$\begin{aligned}
\triangle_{x_k} E(X) &= \nabla_{x_k} E(X) \cdot \Delta x_k \\
&= \left[2 \sum_{i=1}^{M} (x_k - x_{ik}) b^{-\|X-X_i\|^2} \cdot \ln b\right] \\
&\quad \cdot (x'_k - x_k) \\
&= \ln b \cdot 2 \cdot \left(\sum_{i=1}^{M} b^{-\|X-X_i\|^2}\right) \\
&\quad \cdot \left(x_k - \frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}}\right) \cdot (x'_k - x_k) \\
&= -2 \ln b \cdot \left(\sum_{i=1}^{M} b^{-\|X-X_i\|^2}\right) \\
&\quad \cdot \left(\frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}} - x_k\right) \\
&\quad \cdot (x'_k - x_k).
\end{aligned} \tag{5}$$

According to (2), we have the following inequalities when x'_k is the next state of x_k.

Case 1: If

$$x_k - \frac{1}{2} \leq \frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}} < x_k + \frac{1}{2}$$

then, $x'_k = H(\frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}}) = x_k$ according to (2). Hence, $\triangle_{x_k} E(X) = 0$.

Case 2: If

$$\frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}} \geq x_k + \frac{1}{2}$$

then $x'_k > x_k$. Thus, according to (5), $\triangle_{x_k} E(X) < 0$.

Case 3: If

$$\frac{\sum_{i=1}^{M} x_{ik} b^{-\|X-X_i\|^2}}{\sum_{i=1}^{M} b^{-\|X-X_i\|^2}} < x_k - \frac{1}{2}$$

then, $x'_k < x_k$. Thus, according to (5), $\triangle_{x_k} E(X) < 0$.

Hence, the evolution function ensures the energy of the network will be decreased in the gradient direction. Besides, it is easy to verify that the energy defined by (4) is bounded. Therefore, the MV-ECAM is stable.

B. Framework of MV-eBAM

Suppose we are given M pattern pairs, which are

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_M, Y_M)\} \tag{6}$$

where

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{in})$$
$$Y_i = (y_{i1}, y_{i2}, \ldots, y_{ip})$$

where n is assumed to be smaller than or equal to p without any loss of generality. Hence, the evolution equations of the MV-eBAM are shown as

$$y_k = H\left(\frac{\sum_{i=1}^{M} y_{ik} b^{-\|X_i - X\|^2}}{\sum_{i=1}^{M} b^{-\|X_i - X\|^2}}\right)$$
$$x_k = H\left(\frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y_i - Y\|^2}}{\sum_{i=1}^{M} b^{-\|Y_i - Y\|^2}}\right) \tag{7}$$

where X and Y are input key patterns, b is a positive number, called the radix, $b > 1$, $H(\cdot)$ is the same function as (3), x_k and x_{ik} are the kth digits of X and the X_i, respectively, y_k and y_{ik} are for Y and the Y_i, respectively.

The reasons for using an exponential scheme in (7) are to enlarge the attraction radius of every stored pattern pair and to augment the desired pattern in the recall reverberation process. In the evolution equations, (7), if the given input pattern is close to the desired pattern, the weighting coefficient, $b^{-\|X_i - X\|^2}$, will be close to the maximum, 1, while if the input pattern is far from the desired one, it will approach 0. As for the purpose of the denominator, it makes the y_k and x_k to be the centroids of all of the y_{ik}'s and x_{ik}'s, respectively.

1) Asymptotical Stability: The MV-eBAM is one kind of BAM, bidirectional associative memory. Therefore, we can explore its stability by studying its two phases of evolution, i.e., $X \to Y$ and $Y \to X$.

$X \to Y$ Phase: We define an energy function similar to that of MV-ECAM in (4)

$$E_1(X, Y) = \sum_{i=1}^{M} \|X - X_i\|^2 \, b^{-\|Y - Y_i\|^2}. \tag{8}$$

At the first glance, the above proposed energy function might not show that every pattern pair is encoded in a local minimum on the energy plane. However, it indeed can place the pattern pairs in local minima. The reason is the second term is an exponential term. When the radix b is sufficiently large, the pattern pair (X_i, Y_i) will reside in a local minimum of the E_1 surface.

Thus, we can compute the $\nabla_{x_k} E_1(X, Y)$

$$\begin{aligned}\nabla_{x_k} E_1(X, Y) &= 2\sum_{i=1}^{M} (x_k - x_{ik}) b^{-\|Y - Y_i\|^2} \\ &= 2\left(\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}\right) \\ &\quad \cdot \left(x_k - \frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y - Y_i\|^2}}{\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}}\right).\end{aligned}$$

The change of E_1 due to a digit change, therefore, can be derived to be

$$\begin{aligned}\triangle_{x_k} E_1(X, Y) &= \nabla_{x_k} E_1(X, Y) \cdot \triangle x_k \\ &= 2\left(\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}\right) \\ &\quad \times \left(x_k - \frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y - Y_i\|^2}}{\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}}\right) \\ &\quad \cdot (x_k' - x_k) \\ &= -2\left(\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}\right) \\ &\quad \cdot \left(\frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y - Y_i\|^2}}{\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}} - x_k\right) \\ &\quad \cdot (x_k' - x_k).\end{aligned} \tag{9}$$

According to (7), we have the following inequalities when x_k' is the next state of x_k.

Case 1: If

$$x_k - \frac{1}{2} \le \frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y - Y_i\|^2}}{\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}} < x_k + \frac{1}{2}$$

then, $x_k' = x_k$ according to (7). Thus, $\triangle_{x_k} E_1(X, Y) = 0$.

Case 2: If

$$\frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y - Y_i\|^2}}{\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}} \ge x_k + \frac{1}{2}$$

then, $x_k' > x_k$. Thus, according to (9), $\triangle_{x_k} E_1(X, Y) < 0$.

Case 3: If

$$\frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y - Y_i\|^2}}{\sum_{i=1}^{M} b^{-\|Y - Y_i\|^2}} < x_k - \frac{1}{2}$$

then, $x_k' < x_k$. Thus, according to (9), $\triangle_{x_k} E_1(X, Y) < 0$.

The $X \to Y$ phase of the network is proved to be asymptotically stable.

$Y \to X$ Phase: By the similar derivation as shown in $X \to Y$ phase, we also can prove that $Y \to X$ is asymptotically stable. The only difference is the definition of the energy function in this phase. The energy function of this phase is similar to (8)

$$E_2(X, Y) = \sum_{i=1}^{M} \|Y - Y_i\|^2 \, b^{-\|X - X_i\|^2}. \tag{10}$$

Since the procedure of the derivation is very much the same as that of the $X \to Y$ phase. There is no need to repeat the lengthy discussion.

Note that the energy functions defined in (8) and (10) are both bounded. In short, the $X \to Y$ phase always drags down the $E_1(X, Y)$, while the $Y \to X$ phase always reduce the $E_2(X, Y)$. The evolution will be terminated when both $E_1(X, Y)$ and $E_2(X, Y)$ reach their respective local minima at which the pattern pairs are stored.

2) Absolute Lower Bound of the Radix: The definition of the absolute lower bound of the radix can be stated as the smallest radix which is able to recall every unique stored pattern pair. In other words, we are interested in discovering what is the minimal radix that is good enough to recall every pattern pair as long as these stored pairs are one-to-one associated. Therefore, we have to consider the worst case in order to derive this minimal radix, which is called the absolute lower bound.

Assume all of the stored pattern pairs are unique, and the given input pattern is the same as either X vector or Y vector of one of the pairs. Thus, according to (7), we can take one of the evolution equations as an illustration

$$\begin{aligned} y_k &= H\left(\frac{\sum_{i=1}^{M} y_{ik} b^{-\|X_i - X\|^2}}{\sum_{i=1}^{M} b^{-\|X_i - X\|^2}}\right) \\ &= H\left(\frac{y_{hk} \cdot b^{-\|X_h - X\|^2} + \sum_{i \neq h}^{M} y_{ik} b^{-\|X_i - X\|^2}}{b^{-\|X_h - X\|^2} + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}}\right) \\ &= H\left(\frac{y_{hk} + \sum_{i \neq h}^{M} y_{hk} \cdot b^{-\|X_i - X\|^2}}{1 + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}} \right. \\ &\qquad \left. + \frac{\sum_{i \neq h}^{M} (y_{ik} - y_{hk}) \cdot b^{-\|X_i - X\|^2}}{1 + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}}\right) \\ &= H\left(\frac{y_{hk}\left(1 + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}\right)}{1 + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}} \right. \\ &\qquad \left. + \frac{\sum_{i \neq h}^{M} (y_{ik} - y_{hk}) \cdot b^{-\|X_i - X\|^2}}{1 + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}}\right) \\ &= H\left(y_{hk} + \frac{\sum_{i \neq h}^{M} (y_{ik} - y_{hk}) \cdot b^{-\|X_i - X\|^2}}{1 + \sum_{i \neq h}^{M} b^{-\|X_i - X\|^2}}\right) \end{aligned} \tag{11}$$

where the first term in the $H(\cdot)$, i.e., y_{hk}, is the signal, and the second term is deemed as the noise. In order not to make the y_k jump to either of y_{hk}'s adjacent levels, the sufficient condition based on (3) and Fig. 1 is

$$-\frac{1}{2} < \frac{\sum_{i \neq h} (y_{ik} - y_{hk}) \cdot b^{-\|X_i - X\|^2}}{1 + \sum_{i \neq h} b^{-\|X_i - X\|^2}} < \frac{1}{2}. \tag{12}$$

That is, the noise must be bounded. Therefore, the discussion can be divided into two parts. The right part and the left part of the inequality shown in (12) can be simplified, respectively, to the following inequalities:

$$\sum_{i \neq h} [2(y_{ik} - y_{hk}) - 1] \cdot b^{-\|X_i - X\|^2} < 1$$
$$\sum_{i \neq h} [2(y_{hk} - y_{ik}) - 1] \cdot b^{-\|X_i - X\|^2} < 1.$$

In summary, the above two inequalities are rewritten as the following equation:

$$\sum_{i \neq h} [2|y_{hk} - y_{ik}| - 1] \cdot b^{-\|X_i - X\|^2} < 1. \tag{13}$$

We consider the distribution under the worst condition, $|y_{hk} - y_{ik}| = L - 1$. Then, (13) is again to be restated as

$$\sum_{i \neq h} [2L - 3] \cdot b^{-\|X_i - X\|^2} < 1. \tag{14}$$

The worst case for the pairs distribution happens when those X_i, $i \neq h$, are located as close to X_h as possible. This will produce the largest noise to the signal, y_{hk}. For instance, if $n = p = 2$, the worst condition to X_h is shown in Fig. 2. Assume t is the largest number of different digits between any Y_i, $i \neq h$ and Y_h in the worst case of pattern pairs distribution. m_r is the largest number of patterns satisfying $\|X - X_i\|^2 = r$, where r is square of the distance between X and X_i. Then, the worst case of the pattern pairs distribution must be

$$\sum_{r=1}^{t-1} m_r \leq M - 1 < \sum_{r=1}^{t} m_r. \tag{15}$$

Therefore, according to (15), (14) can be rewritten to be the following equation:

$$(2L - 3)\left[\sum_{r=1}^{t-1} m_r \cdot b^{-r} + \left(M - 1 - \sum_{r=1}^{t-1} m_r\right) \cdot b^{-t}\right] < 1. \tag{16}$$

Unfortunately, the solution of this result is not an analytical form. However, if n and L are given, the minimal b can still be derived numerically. Some numerical analysis will be required in order to solve this lower bound.

3) Capacity Analysis: Following the same scenario as the previous subsection, we consider the worst distribution of the pattern pairs in order to assess the capacity of MV-eBAM. Then according to (14), the capacity can be derived when the equality holds. Therefore,

$$M < \left(\frac{1}{2L - 3} - \sum_{r=1}^{t-1} m_r \cdot b^{-r}\right) \cdot b^t + \sum_{r=1}^{t-1} m_r + 1. \tag{17}$$

This result also shows that it cannot be expressed as a closed form solution. Hence, some numerical analysis is also required in order to solve the above inequality.

4) Radix Searching Algorithm: We are also interested in a problem: given a limited set of pattern pairs in which every pair is uniquely associated. What is the smallest radix able to recall all of the pairs in the set? Obviously this radix must be no larger than the lower bound of the radix which is able to recall every unique pair in the worst case of pattern distribution.

As we mentioned in the discussion for the asymptotical stability, the sufficient condition for a stored pattern pair to be recalled correctly is it has to be encoded in respective local minima of E_1 and E_2. This observation leads to the following criterion for the stability of the MV-eBAM.

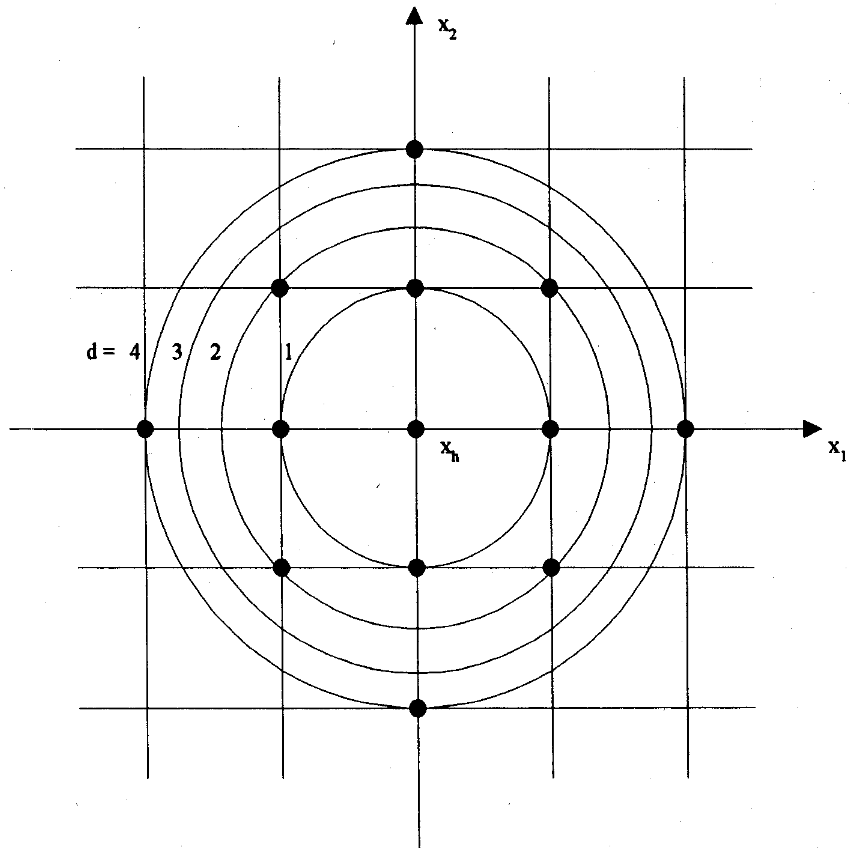


Fig. 2. The worst distribution of MV-eBAM (dimension = 2).

For a pattern pair (X_h, Y_h), the sufficient conditions to be recalled correctly are

$$-\frac{1}{2} \leq x_{hk} - \frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y_i - Y_h\|^2}}{\sum_{i=1}^{M} b^{-\|Y_i - Y_h\|^2}} < \frac{1}{2}, \forall k = 1, \ldots, n$$
$$-\frac{1}{2} \leq y_{hk} - \frac{\sum_{i=1}^{M} y_{ik} b^{-\|X_i - X_h\|^2}}{\sum_{i=1}^{M} b^{-\|X_i - X_h\|^2}} < \frac{1}{2}, \forall k = 1, \ldots, p. \quad (18)$$

Hence, a binary searching method can be adopted to iteratively compute the smallest radix which is still large enough to recall all of the given pattern pairs. A cost function is defined as follows:

$$J(b) = \sum_{h=1}^{M} \left(\sum_{k=1}^{n} \alpha_{hk} + \sum_{k=1}^{p} \beta_{hk} \right) \quad (19)$$

where

$$\alpha_{hk} = \left| x_{hk} - x'_{hk} \right| = \left| x_{hk} - H\left(\frac{\sum_{i=1}^{M} x_{ik} b^{-\|Y_i - Y_h\|^2}}{\sum_{i=1}^{M} b^{-\|Y_i - Y_h\|^2}} \right) \right|$$
$$\beta_{hk} = \left| y_{hk} - y'_{hk} \right| = \left| y_{hk} - H\left(\frac{\sum_{i=1}^{M} y_{ik} b^{-\|X_i - X_h\|^2}}{\sum_{i=1}^{M} b^{-\|X_i - X_h\|^2}} \right) \right|.$$

The cost function possesses several nice features to be a good measure of the radix searching.

1) If (18) is satisfied, then $J(b) = 0$. That is, the good radix, b, is found.
2) If either of the inequalities shown in (18) is not satisfied, then $J(b) > 0$.
3) If b increases, $J(b)$ decreases.

Considering the computation complexity, we will use the binary search method to find the minimal radix instead of using gradient descent method. The radix searching algorithm is summarized below.

Step 1. Set $b_l = 1.0$ and $b_u = 2.0$, where b_l indicate the lower bound of the searching interval, and b_u means the upper bound. Substitute b_u into (19) with the variable b. If $J(b) > 0$, set $b_l = b_u$, $b_u = 2 \cdot b_u$, and repeat the step until $J(b_u) = 0$.

Step 2. Suppose err denotes a predetermined tolerable error. Thus, if $b_u - b_l \geq$ err, $b_m = \frac{1}{2}(b_u + b_l)$. If $J(b_m) \neq 0$, $b_l = b_m$; else, bu = bm.

Step 3. Repeat Step 2 until the difference between b_u and b_l is smaller than a predetermined error. Finally, b_u is the radix to be used.

C. Modified MV-eBAM

Although the MV-eBAM shown in Section II-B possesses amazing recall ability, the capacity still remains unclear because the analytic solution of (17) is hard to derived. Hence, we proposed a modified MV-eBAM with different evolution

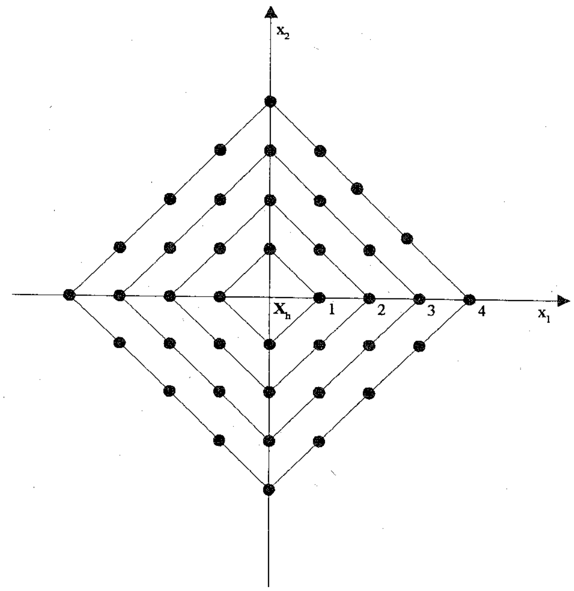


Fig. 3. The worst distribution of modified MV-eBAM (without the limit of L).

equations as follows:

$$y_k = H\left(\frac{\sum_{i=1}^{M} y_{ik} b^{-\sum_{j=1}^{n}|x_{ij}-x_j|}}{\sum_{i=1}^{M} b^{-\sum_{j=1}^{n}|x_{ij}-x_j|}}\right)$$
$$x_k = H\left(\frac{\sum_{i=1}^{M} x_{ik} b^{-\sum_{j=1}^{p}|y_{ij}-y_j|}}{\sum_{i=1}^{M} b^{-\sum_{j=1}^{p}|y_{ij}-y_j|}}\right). \tag{20}$$

The major difference between the evolution equations of MV-eBAM, as shown in (7), and those of modified MV-eBAM, as shown in (20) is the measure of the distance of the retrieval pattern and the stored patterns. That is, the MV-eBAM employs a measure of spatial distance of two vectors, i.e., $\|X_i - X\|^2$, while the modified MV-eBAM uses the measure of Manhattan distance of the vectors, i.e., $\sum_{j=1}^{n}|x_{ij} - x_j|$. The pattern pair distribution in the worst case of the former is like concentric circles as shown in Fig. 2. In contrast, the latter will arrange the pattern pairs in a family of rhombuses, as shown in Fig. 3. Comparing these two structures, we can easily find out that the number of patterns of the modified MV-eBAM in the worst case distribution of pattern pairs is more than that of MV-eBAM when given the same dimension and number of levels. Thus, the capacity of the modified MV-eBAM can provide us the bound of the MV-eBAM in an analytic form solution. Besides, since the exponent part of the modified MV-eBAM is smaller than that of the MV-eBAM, the modified MV-eBAM unavoidably needs a bigger radix.

1) Capacity of Modified MV-eBAM: The key reason that the analytic solution of the capacity of MV-eBAM cannot be found is the m_r terms in (15) and (16) cannot be expressed in a general form when the exponent of the evolution equations is $\|X_i - X\|^2$. For the sake of clarity, we denote the total amount of the difference of the digits of the retrieval pattern and the stored patterns with d for the modified MV-eBAM. Using the similar derivation shown in subsections B-2 and B-3 of Section II, (12)–(17), we can find a m_d for the modified MV-eBAM which is similar to the m_r for the MV-eBAM. Since m_d is not only a function of d, the difference of number of digits of the retrieval pattern and the stored patterns, but also a function of of n, the dimension of the pattern vector. Hence we define $m(n, d)$ as the largest number of patterns satisfying $\sum_{j=1}^{n}|x_{ij} - x_j| = d$, where x_{ij} and x_j is the jth component of X_i and X, respectively, and $X \in \{1, 2, \ldots, L\}^n$. In order to precisely compute the capacity of the modified MV-eBAM, again we have to consider the worst case of pattern distribution, as shown in Fig. 3. Thus, the equation of the worst case distribution is the same as (15)

$$\sum_{d=1}^{s-1} m(n, d) \le M - 1 < \sum_{d=1}^{s} m(n, d) \tag{21}$$

where M is the capacity, the number of pattern pairs to be stored, and s is assumed to be the largest number of different digits between the retrieval pattern and the stored patterns for the modified MV-eBAM.

Obviously, the $m(n, d)$ is also affected by the number of levels, i.e., L. Hence, we have to consider two situations: L is sufficiently large, which won't be a factor to calculate $m(n, d)$; and L is not sufficiently large.

Case I – Large L: Assume L is large enough so that it won't restrict the number of the pattern vectors in the worst case distribution. We have to define some terms in order to solve the $m(n, d)$.

$SN(0)$: the number of solutions of the equation, $d_1 + d_2 + \cdots + d_n = d$, where d_i's are nonnegative integers which are the distance of x_j to the corresponding x_{ij}, for $j = 1, \ldots, n$.

$SZ(0)$: the number of solutions of the equation, $|d_1| + |d_2| + \cdots + |d_n| = d$, where d_i's are integers. In fact, the $SZ(0)$ is the same as $m(n, d)$.

$SN(k)$: the number of solutions of the equation, $d_1 + d_2 + \cdots + d_n = d$, where d_i's are nonnegative integers, and at least k of the d_i's are 0's.

$SZ(k)$: the number of solutions of the equation, $|d_1| + |d_2| + \cdots + |d_n| = d$, where d_i's are integers, and at least k of the d_i's are 0's.

According to the definition of $SN(0)$, the solution is the same as randomly allocate d 1's in n positions. In other words, it is also the same as to insert $(n - 1)$ dividers into d 1's. That is,

$$SN(0) = C_{n-1}^{d+n-1} = C_d^{d+n-1}. \tag{22}$$

As for the $SN(k)$, it is defined that there are at least k 0's in the x_i's of the solution. Hence,

$$SN(k) = C_k^n \cdot C_d^{d+(n-k)-1}. \tag{23}$$

Considering the relation between $SZ(0)$ and $SN(0)$, we find that the x_i can be either positive or negative. Thus, $SZ(0)$ should be $2^n \cdot SN(0)$. However, since $+0 = -0 = 0$, we have to delete the these solutions which are counted twice. In other words, $SZ(0)$ is the same as $2^n \cdot SN(0)$ subtracts those solutions containing at least one 0

$$SZ(0) = 2^n \cdot SN(0) - SZ(1). \tag{24}$$

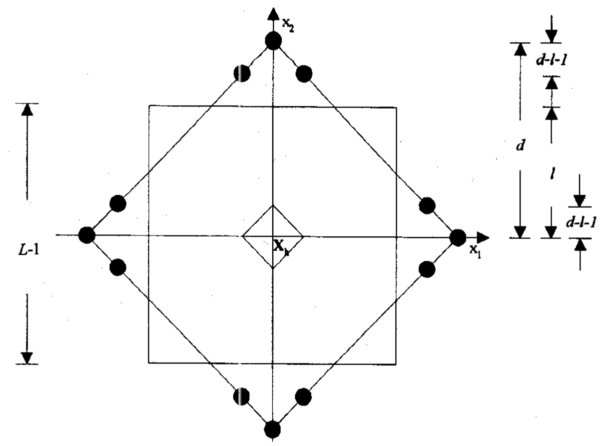


Fig. 4. The worst distribution of modified MV-eBAM (L is odd, $L = 2l + 1$).

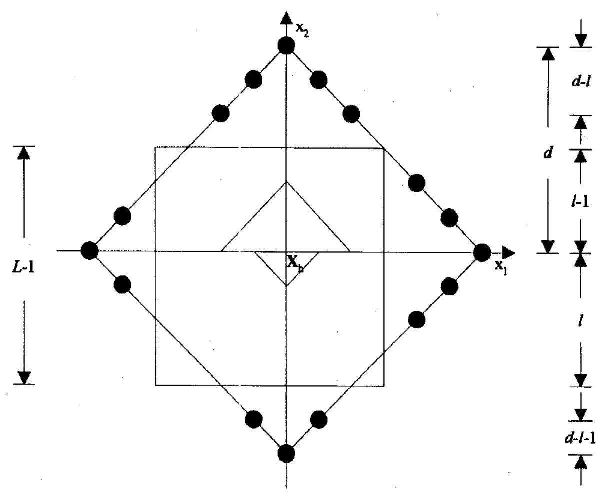


Fig. 5. The worst distribution of modified MV-eBAM (L is even, $L = 2l$).

By the same observation

$$SZ(k) = 2^{n-k} \cdot SN(k) - SZ(k+1). \tag{25}$$

Assume X_h is the pattern located at the center of Fig. 3, i.e., the pattern supposed to be recalled. Then $SZ(n)$ is 0, since X_h is assumed to be unique, and this solution has at most $n-1$ 0's. Hence, we can conclude the following equations:

$$\begin{aligned} SZ(0) &= 2^n \cdot SN(0) - 2^{n-1} \cdot SN(1) \\ &\quad + 2^{n-2} \cdot SN(2) \ldots 2^1 \cdot SN(n-1) \\ &= \sum_{k=0}^{n-1} (-1)^k \cdot 2^{n-k} \cdot C_k^n \cdot C_d^{n-k+d-1} \\ &= m(n,d). \end{aligned}$$

Another way to compute $m(n,d)$ and solve it recursively is shown as follows:

$$m(n,d) = m(n,d)|_{x_n<0} + m(n,d)\,|_{x_n=0} + m(n,d)|_{x_n>0}$$

b	recalled	failed	prob.
1.5	295	205	59.0%
2.0	474	26	94.8%
2.5	491	9	98.2%
3.0	492	8	98.4%
3.5	496	4	99.2%
4.0	496	4	99.2%
4.5	496	4	99.2%
5.0	496	4	99.2%

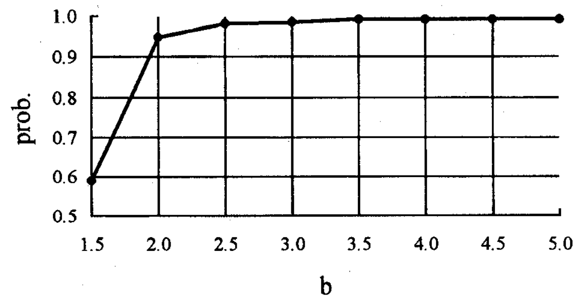


Fig. 6. MV-eBAM simulation $M = 1000$, $n = L = 8$, 500 tested patterns.

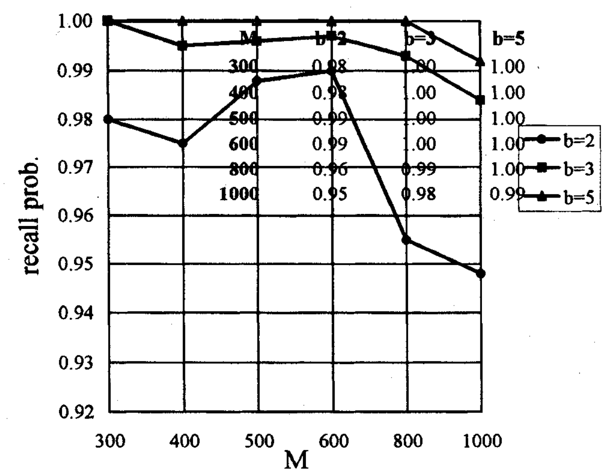


Fig. 7. Summary of MV-eBAM simulations ($n = L = 8$).

where

$$\begin{aligned} m(n,d)|_{x_n<0} &= \sum_{j=-d}^{-1} m(n,d)|_{x_n=j} \\ &= \sum_{j=-d}^{-1} m(n-1, d-|j|) \\ &= \sum_{k=0}^{d-1} m(n-1,k) \\ &= \sum_{j=d}^{1} m(n,d)|_{x_n=j} \\ &= m(n,d)|_{x_n>0} \end{aligned}$$

and

$$m(n,d)|_{x_n=0} = m(n-1,d).$$

TABLE I
$m(n, d)$: $n = 1, \ldots, 8$, $d = 1, \ldots, 8$

n	d=1	d=2	d=3	d=4	d=5	d=6	d=7	d=8
1	2	2	2	2	2	2	2	2
2	4	8	12	16	20	24	28	32
3	6	18	38	66	102	146	198	258
4	8	32	88	192	360	608	952	1408
5	10	50	170	450	1002	1970	3530	5890
6	12	72	292	912	2364	5336	10836	20256
7	14	98	462	1666	4942	12642	28814	59906
8	16	128	688	2816	9424	27008	68464	157184

TABLE II
$m(n, d, l)$: $n = 1, \ldots, 8$, $d = 1, \ldots, 8$, $L = 8$

n	d=1	d=2	d=3	d=4	d=5	d=6	d=7	d=8
1	2	2	2	1	0	0	0	0
2	4	8	12	14	12	8	4	1
3	6	18	38	63	84	92	84	63
4	8	32	88	188	328	480	600	646
5	10	50	170	445	952	1720	2680	3650
6	12	72	292	906	2292	4904	9084	14799
7	14	98	462	1659	4844	11956	25580	48265
8	16	128	688	2808	9296	25984	62960	134684

iteration	bm	J(bm)
1	3.00	0.00
2	2.50	0.00
3	2.25	2.00
4	2.38	0.00
5	2.31	2.00
6	2.34	2.00
7	2.36	0.00
8	2.35	0.00
final b	2.35	0.00

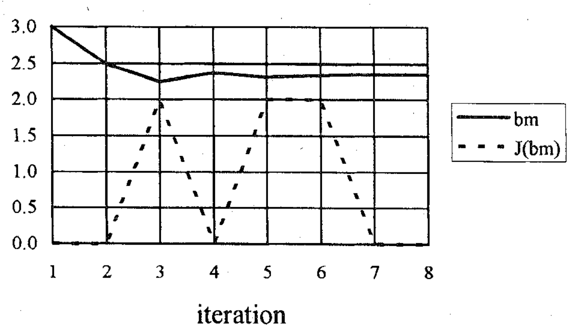


Fig. 8. MV-eBAM radix searching ($M = 400$, $n = L = 8$, tolerance $= 0.01$).

b	p1	p2
1.5	0.0%	59.4%
2.0	74.0%	94.8%
2.5	97.7%	98.3%
3.0	98.3%	98.6%
3.5	99.6%	99.6%
4.0	99.6%	99.6%
4.5	99.6%	99.6%
5.0	99.6%	99.6%

p1: recall probability of modified MV-eBAM
p2: recall probability of MV-eBAM

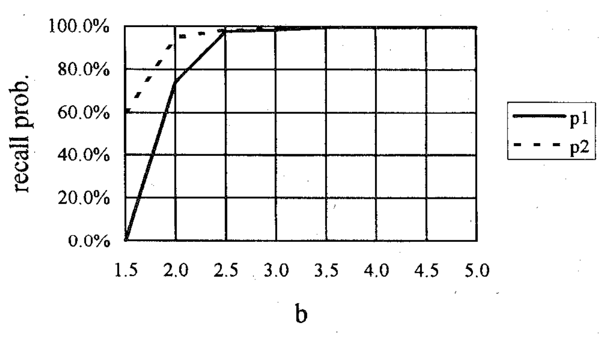


Fig. 9. Modified MV-eBAM simulation $M = 1000$, $n = L = 8$, fully tested.

Hence, we can conclude that the following equation holds:

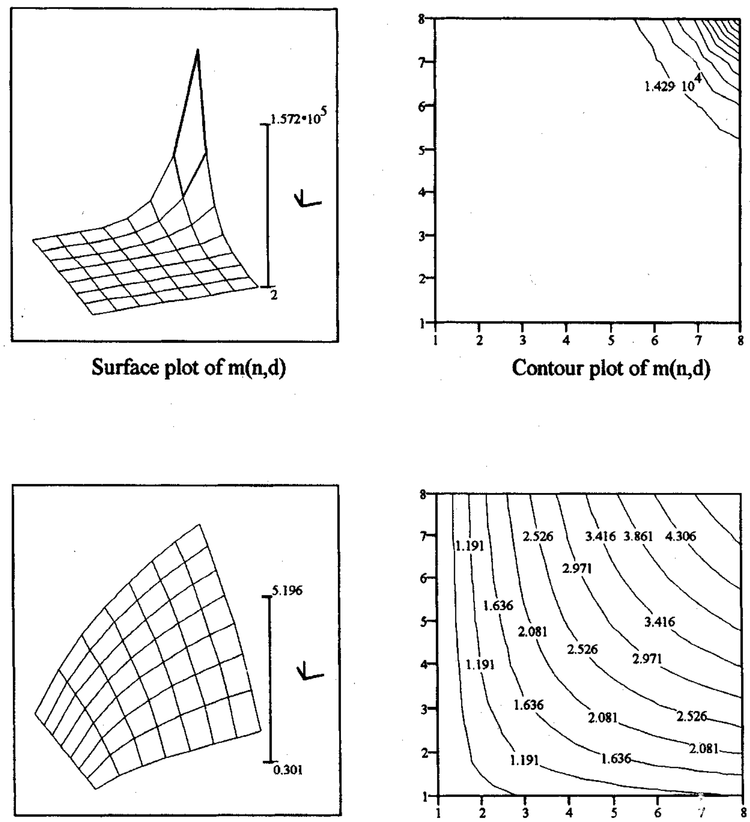
$$SZ(0) = m(n, d) = 2 \cdot \left\{ \sum_{i=0}^{d-1} m(n-1, i) \right\} + m(n-1, d). \tag{26}$$

Case II – Small L: In this case, the L undoubtedly will reduce the number of solutions, i.e., the number of patterns in the worst case distribution, given a fixed d. Thus, the function $m(n, d)$ in the previous case should be correctly rewritten as $m(n, d, l)$, where $L = 2l + 1$ or $L = 2l$ depending on whether L is even or odd.

$L = 2l + 1$: Take the $n = 2$ shown in Fig. 4 as an example. $m(n, d, l)$ is the equal to $m(n, d)$ subtracting those solution

TABLE III
$m(n,d,l)$: $n = 1,\ldots,8$, $d = 1,\ldots,8$, $L = 9$

n	d=1	d=2	d=3	d=4	d=5	d=6	d=7	d=8
1	2	2	2	2	0	0	0	0
2	4	8	12	16	16	12	8	4
3	6	18	38	66	96	116	120	108
4	8	32	88	192	352	552	752	904
5	10	50	170	450	992	1880	3120	4600
6	12	72	292	912	2352	5204	10104	17484
7	14	98	462	1666	4928	12460	27624	54628
8	16	128	688	2816	9408	26768	66656	147984



Surface plot of m(n,d)

Contour plot of m(n,d)

Surface plot of log(m(n,d))

Contour plot of log(m(n,d))

Fig. 10. Plots of $m(n,d)$and $\log(m(n,d))$.

points located outside the square area with the side length $L-1$. In Fig. 4, the number of the solution points outside of the square in the positive x_2 direction is $m(n,d-l-1)|_{x_2>0} + m(n,d-l-1)|_{x_2=0}$. And

$$m(n,d-l-1)|_{x_n>0} + m(n,d-l-1)|_{x_n=0} = \frac{1}{2}[m(n,d-l-1) + m(n-1,d-l-1)].$$

The boundary condition of the above equation is shown as follows:

$$\begin{aligned} m(n,d) &= 1, \quad \text{if } d=0 \\ m(n,d) &= 0, \quad \text{if } d<0 \\ m(n,d) &= 0, \quad \text{if } n=0, d\neq 0. \end{aligned}$$

Hence, since the dimension of X vector is assumed to be n and there are two directions along each dimension, the $m(n,d,l)$ can be concluded to be

$$\begin{aligned} m(n,d,l) = m(n,d) - n \\ \cdot [m(n,d-l-1) + m(n-1,d-l-1)]. \end{aligned} \quad (27)$$

$L = 2l$: Referring to Fig. 5, the solution points located outside of the box area are $m(n,d-l)|_{x_n>0} + m(n,d-l)|_{x_n=0}$ and $m(n,d-l-1)|_{x_n>0} + m(n,d-l-1)|_{x_n=0}$. Hence,

$$\begin{aligned} m(n,d,l) = m(n,d) - \frac{n}{2} \\ \cdot \{m(n,d-l) + m(n-1,d-l) \\ + m(n,d-l-1) + m(n-1,d-l-1)\}. \end{aligned} \quad (28)$$

III. Simulation Analysis

Example 1 – Asymptotical Stability of MV-eBAM: In order to verify the high capacity and the stability of the ML-eBAM, we have conducted simulations with $n = p = 8$, $L = D = 8$, $M = 1000$. Note that M is the number of stored pattern pairs. In our simulations, the pattern pairs are randomly generated, which are all unique pairs. The number of patterns to be tested is 500. The detailed results are shown in Fig. 6. In the Fig. 6, we have shown the effect of b to the recall probability given fixed M, n, and L. If the b is increased, the probability of successful recall is also increased.

In order to comprehend the function of the radix to the recall probability, we also conducted a series of simulations in which M is different, as shown in Fig. 7. Obviously, the larger radix will give better recall probability.

Example 2 – Radix Searching Algorithm: Fig. 8 shows the result of the radix searching algorithm. As shown in Fig. 8, bm will converge to a fixed value, and the cost function $J(b_m)$ will finally converges to 0. Then, we use this found radix to repeat the recall simulation as those in Example 1. The recall probability is 100%. This simulation shows the feasibility of our searching algorithm of the radix.

Example 3 – Capacity of Modified MV-eBAM: We first conduct the recall simulation of the modified MV-eBAM which is also compared with that of the MV-eBAM, as shown in Fig. 9. As we explained in Section II, the modified MV-eBAM will have a worse recall probability when the b is the same, because the exponent part of the modified MV-eBAM is smaller than that of the MV-eBAM.

In order to verify the capacity analysis described in of Section II-B3, we use programs to compute $m(n, d)$ and $m(n, d, l)$ when given the number of levels. The results are shown in Tables I–III. The important thing we have to point out is that the entries of the tables are not derived by (26)–(28). In contrast, these numerical entries are computed by programs when the the distribution of the patterns is the worst case. They are matched with the prediction of (26)–(28). Fig. 10 shows the contour of the magnitude of $m(n, d)$ in 3D natural scale and log scale.

After the $m(n, d)$ and $m(n, d, l)$ are derived, then the capacity of the modified MV-eBAM can be computed by (21).

IV. Conclusion

In addition to the proof of the stability of MV-ECAM and the asymptotical stability of MV-eBAM, we have derived the absolute lower bound of the radix and the capacity of the MV-eBAM. The derivation of the lower bound of the radix provide us the information how large the radix should be such that every unique pair can be recalled in the worst case of pattern distribution. As for the radix searching algorithm, this method provide the information how large the radix should be when a certain set of pattern pairs is given. The simulation results indicate a convincing performance regarding the encoding and retrieving of pattern pairs. Since the analytic form of the capacity of the MV-eBAM is hard to derive, a modified MV-eBAM is proposed to estimate the capacity of these multi-valued neural networks. Though the modified MV-eBAM needs a larger radix to recall pattern pairs, it shows the relationship among L, n and the capacity.

References

[1] S. Amari 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] K. W. Current, "Current-mode CMOS multiple-valued logic circuits," *IEEE J. Solid-State Circuits*, vol. 29, no. 2, Feb. 1994.

[4] T. D. Chiueh and R. M. Goodman, "High-capacity exponential associative memory," in *Proc. IJCNN*, 1988, vol. I, pp. 153–160.

[5] ——, "Recurrent correlation associative memories," *IEEE Trans. Neural Networks*, vol. 2, no. 2, pp. 275–284, 1991.

[6] T.-D. Chiueh and H.-K. Tsai, "Multivalued associative memories based on recurrent networks," *IEEE Trans. Neural Networks*, vol. 4, no. 2, pp. 364–366, Mar. 1993.

[7] R. Hecht-Nielsen, *Neurocomputing*. Reading, MA: Addison-Wesley, 1990.

[8] K. Haines and R. Hecht-Nielsen, "A BAM with increased information storage capacity," in *Proc. IJCNN*, 1988, vol. I, pp. 181–190.

[9] Y.-J. Jeng, C.-C. Yeh, and T.-D. Chiueh, "Exponential bidirectional associative memories," *IEE Electron. Lett.*, vol. 26, no. 11, pp. 717–718, May 1990.

[10] B. Kosko, "Adaptive bidirectional associative memory," *Appl. Opt.*, vol. 26, no. 23, pp. 4947–4960, Dec. 1987.

[11] ——, "Bidirectional associative memory," *IEEE Trans. Syst., Man, Cybern.*, vol. 18, no. 1, pp. 49–60, Jan./Feb. 1988.

[12] S. Kawahito, M. Ishida, T. Nakamura, M. Kameyama, and T. Higuchi, "High-speed area-efficienct multiplier design using multiple-valued current-mode circuits," *IEEE Trans. Comput.*, vol. 43, no. 1, Jan. 1994.

[13] 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.

[14] P. K. Simpson, "Higher-ordered and intraconnected bidirectional associative memory," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, no. 3, May/June 1990.

[15] H. M. Tai, C. H. Wu, and T. L. Jong, "High-order bidirectional associative memory," *Electron. Lett.*, vol. 25, pp. 1424–1425, 1989.

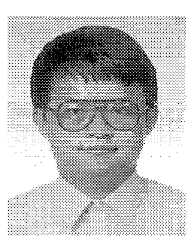
[16] Y.-F. Wang, J. B. Cruz Jr., and J. H. Mulligan, Jr., "Two coding strategies for bidirectional associative memory," *IEEE Trans. Neural Networks*, vol. 1, no. 1, Mar. 1990.

[17] C.-C. Wang and H.-S. Don, "An analysis of high-capacity discrete exponential BAM," *IEEE Trans. Neural Networks*, vol. 6, no. 2, pp. 492–496, Mar. 1995.

[18] T. Wang, X. Zhuang, and X. Xing, "Weighted learning of bidirectional associative memories by global minimization," *IEEE Trans. Neural Networks*, vol. 3, no. 6, pp. 1010–1018, Nov. 1992.

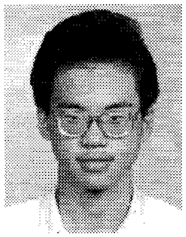
[19] J. Yuh and R. W. Newcomb, "Circuits for multi-level nonlinearities," presented at *Inter. Joint Conf. Neural Networks*, Baltimore, MA, June 1992.

[20] ——, "A multi-level neural network for A/D conversion," *IEEE Trans. Neural Networks*, vol. 4, no. 3, pp. 470–483, May 1993.



Chua-Chin Wang was born in Taiwan in 1962. He received the B.S. degree in electrical engineering from National Taiwan University in 1984, and the M.S. and Ph.D. degrees in electrical engineering from State University of New York, Stony Brook, in 1988 and 1992, respectively.

Currently, he is an Associate Professor in the Department of Electrical Engineering, National Sun Yat-Sen University, Taiwan. His research interests include neural networks and implementations, low-power logic and circuit design, and VLSI design.



Shiou-Ming Hwang was born in Kaohsiung, Taiwan, in 1971. He received the B.S. degree in eectronic engineering from Tamkang University in 1993, and the M.S. degree in electrical engineering from National Sun Yat-Sen University in 1995. His research areas include neural networks and VLSI architectures.

Jyh-Ping Lee was born in Pingtong, Taiwan, in 1971. He received the B.S. and M.S. degrees in electrical engineering from National Sun Yat-Sen University, Taiwan, in 1992 and 1994, respectively. His research interests include neural networks, pattern recognition, expert system, and VLSI design.

He is currently serving as an officer in the army. After he is discharged from the military service, he will pursue a Ph.D. degree in the above-mentioned areas.