LETTER — **Communicated by Richard Golden**

An Optimization Approach to Design of Generalized BSB Neural Associative Memories

Jooyoung Park
Department of Control and Instrumentation Engineering, Korea University, Chochiwon, Chungnam, 339-800, Korea

Yonmook Park
Department of Information Engineering, Graduate School, Korea University, Chochiwon, Chungnam, 339-800, Korea

This article is concerned with the synthesis of the optimally performing GBSB (generalized brain-state-in-a-box) neural associative memory given a set of desired binary patterns to be stored as asymptotically stable equilibrium points. Based on some known qualitative properties and newly observed fundamental properties of the GBSB model, the synthesis problem is formulated as a constrained optimization problem. Next, we convert this problem into a quasi-convex optimization problem called GEVP (generalized eigenvalue problem). This conversion is particularly useful in practice, because GEVPs can be efficiently solved by recently developed interior point methods. Design examples are given to illustrate the proposed approach and to compare with existing synthesis methods.

1 Introduction

Recently, the problem of realizing associative memories via recurrent neural networks has received much attention (Hassoun, 1993; Michel & Farrell, 1990). In general, the key properties required for these neural associative memories are the following (Lillo, Miller, Hui, & Zak, 1994):

- The desired binary patterns, which will be called the prototype patterns in this article, are stored as asymptotically stable equilibrium points of the system.
- When a vector sufficiently close to a stored prototype pattern is applied as an initial condition, then the system trajectory converges to the prototype pattern.
- The number of asymptotically stable equilibrium points that do not correspond to the prototype patterns (i.e., spurious states) is minimal.
- The system is globally stable; every trajectory of the system converges to some equilibrium point.

Among the various kinds of promising neural models suitable for achieving these properties are the so-called brain-state-in-a-box (BSB) neural networks. This model was first proposed by Anderson, Silverstein, Ritz, and Jones (1977) and has attracted substantial interest among researchers. Cohen and Grossberg (1983) proved a theorem on the global stability of the continuous-time, continuous-state BSB dynamical systems with real symmetric weight matrices. Golden (1986) proved that all trajectories of the BSB dynamical systems with real symmetric weight matrices approach the set of equilibrium points under a certain condition, which will be used as the global stability condition in this article. Marcus and Westervelt (1989) reported results on global stability for a large class of BSB model type systems. Perfetti (1995) analyzed qualitative properties of the BSB model and formulated the design of the BSB-based associative memories as a constrained optimization in the form of a linear programming with an additional nonlinear constraint. Also, he proposed an ad hoc iterative algorithm to solve the constrained optimization and illustrated the algorithm with some design examples. Park, Cho, and Park (1999) recast Perfetti's formulation into a semidefinite programming problem (SDP) by converting its nonlinear constraints into a set of linear matrix inequalities (LMIs). We are concerned here with developing a synthesis procedure for associative memories based on an advanced form of the BSB model, which is often referred to as GBSB (generalized BSB). Our synthesis procedure will be given in the form of an LMI-based optimization problem called a generalized eigenvalue problem (GEVP). This article may be viewed as extending the design method of Park et al. (1998) for the GBSB memories. Both articles use a readily available software (MATLAB LMI Control Toolbox) rather than implementing ad hoc algorithms for the memory synthesis. The most significant difference between them is that the synthesis procedure we present is based on the GEVP, while the method of Park et al. (1999) uses the SDP. A design example in section 4 will demonstrate the superiority of the proposed GEVP-based method over the SDP method.

The GBSB model was proposed and studied by Hui and Zak (1992) and Golden (1993). Lillo, Miller, Hui, and Zak (1994) analyzed the dynamics of the GBSB model and presented a novel synthesis procedure for the GBSB-based associative memories. Their procedure uses a decomposition of the weight matrix, which results in an asymmetric interconnection structure, an asymptotic stability of the prototype patterns, and a small number of spurious states. Later, Zak, Lillo, and Hui (1996) incorporated the learning and forgetting capabilities into the synthesis method of Lillo et al. (1994). Also, Chan and Zak (1997) proposed a "designer" neural network to solve the problem of realizing associative memories via the GBSB neural networks.

In this article, attention is focused on how to find the parameters of the optimally performing GBSB given a set of the prototype patterns. After introducing fundamental properties of the GBSB model, we formulate the synthesis of the GBSB that can store the prototype patterns with optimal performance as a constrained optimization problem. Next, we convert the synthesis problem into an optimization problem called a generalized eigenvalue problem (GEVP). This conversion is particularly useful in practice, because GEVPs can be efficiently solved by recently developed interior point methods (Boyd, El-Ghaoui, Feron, & Balakrishnan, 1994; Vandenberghe & Balakrishnan, 1997). For specific algorithms belonging to these interior point methods, one can refer to Boyd and El-Ghaoui (1993) and Nesterov and Nemirovskii (1994), for example. An implementation of the Nesterov and Nemirovskii algorithm is also provided in the LMI Control Toolbox (Gahinet, Nemirovskii, Laub, & Chilali, 1995). Since GEVPs can be efficiently solved by interior point methods, converting the synthesis problem into a GEVP is equivalent to finding a solution to the original problem. As an optimum searcher for the GEVP appearing in the proposed approach, we use the function "gevp" of the LMI Control Toolbox.

Throughout this article, we use the following definitions and notation, in which R^n denotes the set of real n-vectors. A symmetric matrix $\mathbf{A} \in R^{n\times n}$ is positive definite if $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for any $\mathbf{x} \neq \mathbf{0}$, and $\mathbf{A} > \mathbf{0}$ denotes this. Also, $\mathbf{A} > \mathbf{B}$ denotes that $\mathbf{A} - \mathbf{B}$ is positive definite. H^n denotes the hypercube $[-1, +1]^n$. By a binary vector, we mean a bipolar binary vector whose element is either $+1$ or -1, and the set of all these binary vectors in H^n is denoted by B^n. The usual Hamming distance between two vectors $\mathbf{x}^*$ and $\mathbf{x}$ in B^n is denoted by $H(\mathbf{x}^*, \mathbf{x})$. By an l-neighborhood of a vertex $\gamma \in B^n$, we mean the set $\{\gamma^* \in B^n \mid H(\gamma^*, \gamma) \leq l\}$.

In section 2, we present some known qualitative properties and newly observed fundamental properties of the GBSB model. Also, based on the properties, we formulate the synthesis of the GBSB with optimal performance as a constrained optimization problem. Section 3 describes how this synthesis problem can be converted into a GEVP. In section 4, we present numerical examples to show the superiority of the proposed method over existing synthesis methods. Finally, in section 5 contains concluding remarks.

2 Background Results

The dynamics of the GBSB model is described by the following state equation:

$$\mathbf{x}(k+1) = g(\mathbf{x}(k) + \alpha(\mathbf{W}\mathbf{x}(k) + \mathbf{b})), \tag{2.1}$$

where $\mathbf{x}(k) \in R^n$ is the state vector at time k, $\alpha > 0$ is the step size, $\mathbf{W} \in R^{n\times n}$ is the weight matrix, $\mathbf{b} \in R^n$ is the bias vector, and g: $R^n \to R^n$ is a linear saturating function whose i^{th} component is a function defined as follows:

$$g_i([x_1 \cdots x_i \cdots x_n]^T) = \begin{cases} 1 & \text{if} \quad x_i \geq 1, \\ x_i & \text{if} \quad -1 < x_i < 1, \\ -1 & \text{if} \quad x_i \leq -1. \end{cases}$$

This GBSB model is a generalized version of the BSB network proposed by Anderson et al. (1977), and it differs from the original network in the presence of the bias vector **b**.

In the discussion on the stability of the GBSB model, we use the following definitions:

- A point $\mathbf{x_e} \in R^n$ is an equilibrium point of system 2.1 if $\mathbf{x}(0) = \mathbf{x_e}$ implies $\mathbf{x}(k) = \mathbf{x_e}, \forall k > 0$.
- An equilibrium point $\mathbf{x_e}$ of system 2.1 is stable if for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$\|\mathbf{x}(0) - \mathbf{x_e}\| < \delta \text{ implies } \|\mathbf{x}(k) - \mathbf{x_e}\| < \epsilon, \forall k > 0.$$

- An equilibrium point $\mathbf{x_e}$ of system 2.1 is asymptotically stable if it is stable and there exists $\delta > 0$ such that

$$\mathbf{x}(k) \to \mathbf{x_e} \text{ as } k \to \infty \text{ if } \|\mathbf{x}(0) - \mathbf{x_e}\| < \delta.$$

- System 2.1 is globally stable if every trajectory of the system converges to some equilibrium point.

The criteria on the stability of the GBSB model are now well established (Lillo et al., 1994; Golden, 1993; Perfetti, 1995):

- A vertex γ of the hypercube H^n is an equilibrium point of system 2.1 if and only if

$$\gamma_i \left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) \geq 0, \forall i \in \{1, \ldots, n\}. \tag{2.2}$$

- A vertex γ of the hypercube H^n is an asymptotically stable equilibrium point of system 2.1 if

$$\gamma_i \left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) > 0, \forall i \in \{1, \ldots, n\}. \tag{2.3}$$

- If condition 2.3 holds for a vertex $\gamma \in B^n$ and $w_{ii} = 0, \forall i \in \{1, \ldots, n\}$, then only on the vertices $\gamma^* \in B^n$ such that $H(\gamma, \gamma^*) > 1$ can other asymptotically stable equilibrium points of system 2.1 exist.[1]

[1] In fact, this was proved only for the BSB model in theorem 1 and corollary 2 of Perfetti (1995). However, one can easily prove that this still holds true for the GBSB model by following the arguments of Perfetti (1995).

- System 2.1 is globally stable if the weight matrix $\mathbf{W}$ is symmetric and

$$\lambda_{\min} > -1, \tag{2.4}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $\mathbf{I} + \alpha\mathbf{W}$.

Designing GBSB neural associative memories based on the above criteria only is not good enough in general. The guidelines for addressing the attraction domains of the prototype patterns and the number of spurious states should be added. We present our main theorem, which will be used to provide an additional guideline for performance optimization:

Theorem. *Let $\gamma \in B^n$ be an asymptotically stable equilibrium point of system 2.1 and let l be an integer in $\{1, \ldots, n\}$. If $\mathbf{W}$ and $\mathbf{b}$ satisfy*

$$w_{ii} = 0, \forall i \in \{1, \ldots, n\} \tag{2.5}$$

$$\text{and} \quad \gamma_i \left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) > 2(l-1) \max_j |w_{ij}|, \; i = 1, \ldots, n, \tag{2.6}$$

then any binary vector $\gamma^* \in B^n$ such that $1 \le H(\gamma^*, \gamma) \le l$ has the following properties:

- *γ^* is not an equilibrium point.*
- *If $\mathbf{x}(0) = \gamma^*$ and $\gamma_i^* \neq \gamma_i$, then at the next time step, x_i moves toward γ_i (i.e., $x_i(1)$ is closer to γ_i than $x_i(0)$ is).*

Proof. Let $\gamma^* \in B^n$ be any binary vector satisfying $1 \le H(\gamma^*, \gamma) \le l$, and let $\gamma_i^* \neq \gamma_i$ (i.e., $\gamma_i^* = -\gamma_i$). Since $\delta \stackrel{\triangle}{=} \gamma^* - \gamma$ satisfies

$$\begin{aligned} \left| \sum_{j=1}^{n} w_{ij}\delta_j \right| &= |w_{i1}\delta_1 + \cdots + 0 \times \delta_i + \cdots + w_{in}\delta_n| \\ &\le 2(l-1) \max_j |w_{ij}|, \end{aligned}$$

we have

$$\begin{aligned} \gamma_i^* \left(\sum_{j=1}^{n} w_{ij}\gamma_j^* + b_i \right) &= -\gamma_i \left(\left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) + \left(\sum_{j=1}^{n} w_{ij}\delta_j \right) \right) \\ &\le -\gamma_i \left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) + \left| \gamma_i \left(\sum_{j=1}^{n} w_{ij}\delta_j \right) \right| \end{aligned}$$

$$= -\gamma_i \left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) + \left| \sum_{j=1}^{n} w_{ij}\delta_j \right|$$

$$\leq -\gamma_i \left(\sum_{j=1}^{n} w_{ij}\gamma_j + b_i \right) + 2(l-1)\max_j |w_{ij}|.$$

This, together with inequalities 2.6, implies that

$$\gamma_i^* \left(\sum_{i=1}^{n} w_{ij}\gamma_j^* + b_i \right) < 0. \tag{2.7}$$

Thus, condition 2.2 does not hold for $\gamma^* \in B^n$. Therefore, γ^* cannot be an equilibrium point of system 2.1. Now, let system 2.1 start from $x(0) = \gamma^*$. From inequality 2.7, we see that γ_i^* and $(\sum_{j=1}^{n} w_{ij}\gamma_j^* + b_i)$ have different polarities. Since

$$x_i(1) = g\left(x_i(0) + \alpha \left(\sum_{j=1}^{n} w_{ij}x_j(0) + b_i \right) \right)$$

$$= g\left(\gamma_i^* + \alpha \left(\sum_{j=1}^{n} w_{ij}\gamma_j^* + b_i \right) \right),$$

we have

$$|x_i(1) - \gamma_i| < |x_i(0) - \gamma_i| = 2.$$

From our theorem, we see that $\mathbf{W}$ and $\mathbf{b}$ satisfying conditions 2.5 and 2.6 ensure that the vertices in the l-neighborhood of the prototype pattern γ cannot become spurious states and are very likely to be in the attraction domain of γ. Thus, a possible strategy for decreasing the number of spurious states and increasing the attraction domains of the prototype patterns is to maximize the l in the constraints 2.6 for each prototype pattern $\gamma \in B^n$. This observation and other background results allow us the following formulation for the optimal design: Given a set of the prototype patterns $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}$, the parameters of the optimally performing GBSB can be found by solving:

$$\begin{array}{ll} \max & \delta\ (>0) \\ \text{s.t.} & x_i^{(k)} \left(\sum_{j=1}^{n} w_{ij}x_j^{(k)} + b_i \right) > 2\delta \max_j |w_{ij}|, \\ & \quad i = 1, \ldots, n,\ k = 1, \ldots, m, \\ & w_{ii} = 0, i = 1, \ldots, n, \\ & \mathbf{W} = \mathbf{W}^T, \\ & \lambda_{\min} > -1, \end{array} \tag{2.8}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $\mathbf{I} + \alpha\mathbf{W}$. This optimization problem has two groups of nonlinear constraints, which prevent us from applying readily available optimization techniques. In the next section, we show that this problem can be converted into an optimization problem that can be efficiently solved.

3 A GEVP Approach to Design of GBSBs

In this section, we convert problem 2.8 into an optimization problem called GEVP. The general form of a GEVP is given as follows (Boyd et al., 1994):

$$\begin{array}{ll} \min & \lambda \\ \text{s.t.} & \lambda\mathbf{B}(\mathbf{z}) - \mathbf{A}(\mathbf{z}) > 0, \\ & \mathbf{B}(\mathbf{z}) > 0, \\ & \mathbf{C}(\mathbf{z}) > 0, \end{array}$$

where $\mathbf{A}(\mathbf{z})$, $\mathbf{B}(\mathbf{z})$, and $\mathbf{C}(\mathbf{z})$ are symmetric matrices that are affine functions of the variable $\mathbf{z}$. This GEVP is to minimize the maximum generalized eigenvalue of the pair $(\mathbf{A}(\mathbf{z}), \mathbf{B}(\mathbf{z}))$ subject to the linear matrix inequality (LMI) constraints $\mathbf{B}(\mathbf{z}) > 0$ and $\mathbf{C}(\mathbf{z}) > 0$. It is a quasi-convex optimization problem and can be efficiently solved by recently developed interior point methods (Boyd et al., 1994; Vandenberghe & Balakrishnan, 1997). Before proceeding further, note that multiple LMIs $\mathbf{C}^{(1)}(\mathbf{z}) > \mathbf{0}, \ldots, \mathbf{C}^{(p)}(\mathbf{z}) > \mathbf{0}$ can be expressed as the single LMI $\mathbf{diag}(\mathbf{C}^{(1)}(\mathbf{z}), \ldots, \mathbf{C}^{(p)}(\mathbf{z})) > \mathbf{0}$.

In order to convert problem 2.8 into a GEVP, we first introduce additional variables q_i, $i = 1, \ldots, n$ satisfying

$$x_i^{(k)} \left(\sum_{j=1}^{n} w_{ij} x_j^{(k)} + b_i \right) > 2\delta q_i > 2\delta \max_j |w_{ij}|, \quad i = 1, \ldots, n,\ k = 1, \ldots, m. \tag{3.1}$$

Note that inequalities 3.1 are only a restatement of the first constraint set of problem 2.8 with the deliberately chosen terms $2\delta q_i$ placed between $x_i^{(k)}(\sum_{j=1}^{n} w_{ij} x_j^{(k)} + b_i)$ and $2\delta \max_j |w_{ij}|$. Also, note that inequalities 3.1 can be divided into LMIs

$$\begin{array}{l} q_i - w_{ij} > 0,\ i = 1, \ldots, n,\ j = 1, \ldots, n, \\ w_{ij} + q_i > 0,\ i = 1, \ldots, n,\ j = 1, \ldots, n, \end{array}$$

and a set of inequalities,

$$-2\delta q_i + x_i^{(k)} \left(\sum_{j=1}^{n} w_{ij} x_j^{(k)} + b_i \right) > 0,\ i = 1, \ldots, n,\ k = 1, \ldots, m.$$

Next, we consider the eigenvalue condition, inequality 2.4. Since $\mathbf{I} + \alpha\mathbf{W}$ is real symmetric, its eigenvalues are real, and corresponding eigenvectors can be chosen to be real orthonormal (Strang, 1988). Thus, its spectral decomposition can be written as

$$\mathbf{I} + \alpha\mathbf{W} = \mathbf{U}\Lambda\mathbf{U}^T,$$

where the eigenvalues of $\mathbf{I} + \alpha\mathbf{W}$ appear on the diagonal of Λ, and $\mathbf{U}$, whose columns are the real orthonormal eigenvectors of $\mathbf{I} + \alpha\mathbf{W}$, satisfies $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$ (Strang, 1988). Note that $\lambda_{\min} > -1$ is equivalent to

$$\Lambda > -\mathbf{I}. \tag{3.2}$$

By pre- and postmultiplying matrix inequality 3.2 by $\mathbf{U}$ and $\mathbf{U}^T$, respectively, we obtain an equivalent condition $\mathbf{I} + \alpha\mathbf{W} > -\mathbf{I}$, which is again equivalent to the following LMI:

$$2\mathbf{I} + \alpha\mathbf{W} > \mathbf{0}.$$

As a result of the conversion processes, the problem of finding the GBSB that can store the prototype patterns $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}$ with optimal performance can now be reformulated as the following GEVP:

$$\begin{array}{ll} \min & (-\delta) \\ \text{s.t.} & (-\delta)(2q_i) + x_i^{(k)} \left(\sum_{j=1}^{n} w_{ij} x_j^{(k)} + b_i \right) > 0, \\ & \quad i = 1, \cdots, n, k = 1, \cdots, m, \\ & \left\{ \begin{array}{l} q_i - w_{ij} > 0, \ i = 1, \ldots, n, \ j = 1, \ldots, n, \\ w_{ij} + q_i > 0, \ i = 1, \ldots, n, \ j = 1, \ldots, n, \\ 2\mathbf{I} + \alpha\mathbf{W} > \mathbf{0}, \\ w_{ii} = 0, \ i = 1, \ldots, n, \\ \mathbf{W} = \mathbf{W}^T. \end{array} \right\} \end{array} \tag{3.3}$$

Note that the constraints inside "$\{\cdot\}$" are LMIs in q_i, $i = 1, \ldots, n$, and the independent scalar entries of $\mathbf{W}$ satisfying $w_{ii} = 0$, $i = 1, \ldots, n$ and $\mathbf{W} = \mathbf{W}^T$.

As a further requirement that should be considered in designing GBSB neural associative memories, the magnitudes of the weights $|w_{ij}|$ should have a reasonable upper bound so that the resulting networks could be implemented without difficulty. To reflect this practical concern, the following LMIs are added in the list of constraints of problem 3.3:

$$L < q_i < U, \ i = 1, \ldots, n. \tag{3.4}$$

Note that the resulting optimization problem is still in the form of GEVP.

4 Design Examples

To demonstrate the applicability of the GEVP approach proposed in this article and to compare with existing synthesis methods, we consider three design examples. In these examples, we consider the GBSB model or the BSB model with $n = 10$ neurons. The performance of each designed memory system is to be evaluated by simulations, in which every possible binary vector is applied as an initial condition for the system, and the system is allowed to evolve from the initial condition to a final state. Analyzing the simulation results, one can obtain average recall probabilities and the convergence rate to correct patterns, which are defined as follows:

- Given the pair of a prototype pattern $\mathbf{x}^{(k)}$ and an integer $l \in \{0, 1, \ldots, n\}$, the probability that a binary initial condition vector at the Hamming distance of l away from $\mathbf{x}^{(k)}$ is attracted to $\mathbf{x}^{(k)}$ is denoted by $P(\mathbf{x}^{(k)}, l)$. By the average recall probability $Prob(l)$, we mean the average of the $P(\mathbf{x}^{(k)}, l)$ over all the prototype patterns $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}$ (i.e., $Prob(l) \triangleq \{\sum_{k=1}^{m} P(\mathbf{x}^{(k)}, l)\}/m$).
- By the convergence rate to correct patterns, we mean the percentage that binary initial condition vectors are attracted to the prototype patterns nearest in the sense of Hamming distance.

In the first example, which is taken from Lillo et al. (1994), we wish to store the following six prototype patterns:

$$
\begin{array}{lcl}
\mathbf{x}^{(1)} & = & [\ -1\ \ +1\ \ +1\ \ +1\ \ +1\ \ +1\ \ -1\ \ -1\ \ -1\ \ -1\ ]^T \\
\mathbf{x}^{(2)} & = & [\ +1\ \ +1\ \ -1\ \ -1\ \ -1\ \ +1\ \ -1\ \ -1\ \ +1\ \ -1\ ]^T \\
\mathbf{x}^{(3)} & = & [\ -1\ \ +1\ \ +1\ \ +1\ \ -1\ \ -1\ \ +1\ \ -1\ \ -1\ \ -1\ ]^T \\
\mathbf{x}^{(4)} & = & [\ -1\ \ +1\ \ -1\ \ -1\ \ -1\ \ -1\ \ +1\ \ -1\ \ +1\ \ +1\ ]^T \\
\mathbf{x}^{(5)} & = & [\ +1\ \ -1\ \ -1\ \ +1\ \ +1\ \ -1\ \ +1\ \ +1\ \ +1\ \ -1\ ]^T \\
\mathbf{x}^{(6)} & = & [\ +1\ \ +1\ \ -1\ \ +1\ \ -1\ \ +1\ \ +1\ \ +1\ \ -1\ \ -1\ ]^T
\end{array}
$$

The step size of the GBSB model was set to be $\alpha = 0.3$ as in Lillo et al. (1994). Solving the corresponding GEVP with the bounds in inequalities 3.4 set to $L = 0.1$ and $U = 0.2$, we obtained a GBSB memory for the first example. According to the simulation results, the GBSB designed by the proposed method is with no spurious state, while the GBSB of Lillo et al. (1994) has two spurious states in B^n. The convergence rates to correct patterns and the average recall probabilities of these GBSBs are shown in Table 1 and Figure 1, respectively.

In the second example, which is taken from Chan and Zak (1997), we want to store the following five prototype patterns in the GBSB model:

$$
\begin{array}{lcl}
\mathbf{x}^{(1)} & = & [\ -1\ \ +1\ \ -1\ \ +1\ \ +1\ \ +1\ \ -1\ \ +1\ \ +1\ \ +1\ ]^T \\
\mathbf{x}^{(2)} & = & [\ +1\ \ +1\ \ -1\ \ -1\ \ +1\ \ -1\ \ +1\ \ -1\ \ +1\ \ +1\ ]^T
\end{array}
$$

Table 1: Convergence Rates to Correct Patterns of the GBSBs Designed for the First Example.

	Convergence Rate to Correct Patterns (%)
Lillo et al. (1994)	13.6
Proposed method	92.7

$$\begin{aligned}\mathbf{x}^{(3)} &= [\ -1\ \ +1\ \ +1\ \ +1\ \ -1\ \ -1\ \ +1\ \ -1\ \ +1\ \ -1\ ]^T\\ \mathbf{x}^{(4)} &= [\ +1\ \ +1\ \ -1\ \ +1\ \ -1\ \ +1\ \ -1\ \ +1\ \ +1\ \ +1\ ]^T\\ \mathbf{x}^{(5)} &= [\ +1\ \ -1\ \ -1\ \ -1\ \ +1\ \ +1\ \ +1\ \ -1\ \ -1\ \ -1\ ]^T\end{aligned}$$

As in Chan and Zak (1997), the step size of the GBSB model was set to be $\alpha = 0.3$. Solving the corresponding GEVP with $L = 0.1$ and $U = 0.2$, we obtained a GBSB memory for the second example. From the simulations, we found that the GBSB designed by the proposed method and the GBSB of Chan and Zak (1997) are both with no spurious state. The convergence rates to correct patterns and the average recall probabilities of these GBSBs are shown in Table 2 and Figure 2, respectively.

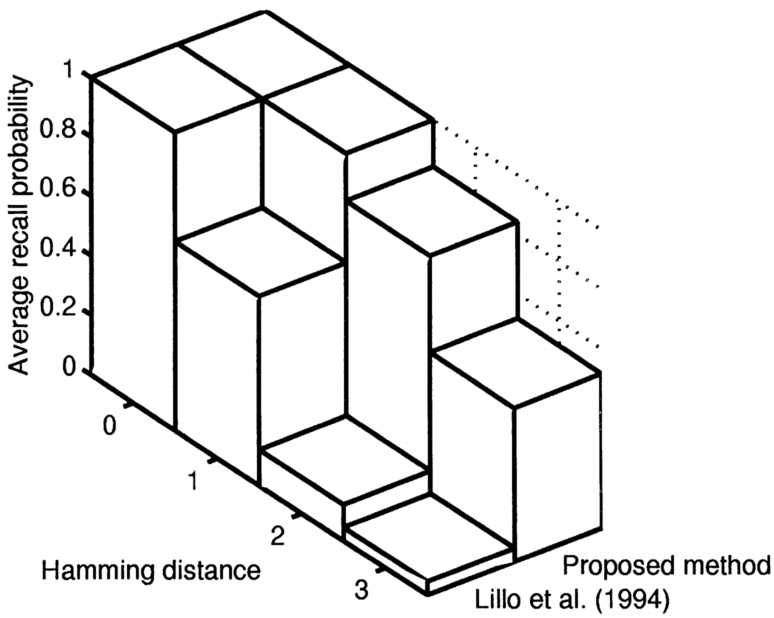

Figure 1: Average recall probabilities of the GBSBs designed for the first example.

Table 2: Convergence Rates to Correct Patterns of the GBSBs Designed for the Second Example.

	Convergence Rate to Correct Patterns (%)
Chan and Zak (1997)	80.1
Proposed method	85.3

In the third and final example, which was considered in Perfetti (1995) and Park et al. (1999), the prototype patterns of the second example are again considered. The design objective of this example is to store the prototype patterns in the BSB model instead of the GBSB model. For a fair comparison, the bias vector and the step size were fixed at $\mathbf{b} = \mathbf{0}$ and $\alpha = 1$, respectively, as in Perfetti (1995) and Park et al. (1999). Solving the corresponding GEVP with $L = 0.1$ and $U = 0.2$, we obtained a BSB memory. From the simulations for the BSB designed by the proposed method, the BSB of Perfetti (1995), and BSB I of Park et al. (1999), we observed that each BSB is with five spurious states, which consist of the negatives of the prototype patterns. The

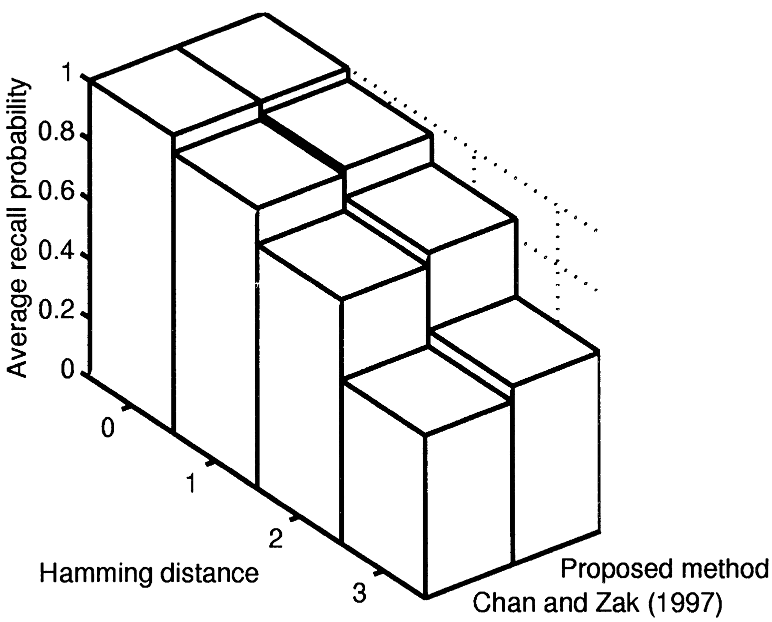

Figure 2: Average Recall probabilities of the GBSBs designed for the second example.

Table 3: Convergence Rates to Correct Patterns of the BSBs Designed for the Third Example.

	Convergence Rate to Correct Patterns (%)
Perfetti (1995)	46.7
BSB I of Park et al. (1998)	46.4
Proposed method	48.9

convergence rates to correct patterns and the average recall probabilities of these BSBs are shown in Table 3 and Figure 3, respectively.

The performance comparisons in the above design examples show that the memories designed by the proposed method generally have higher average recall probabilities, higher convergence rates to correct patterns, and fewer or a comparable number of spurious states than the memories designed by other recently developed synthesis methods.

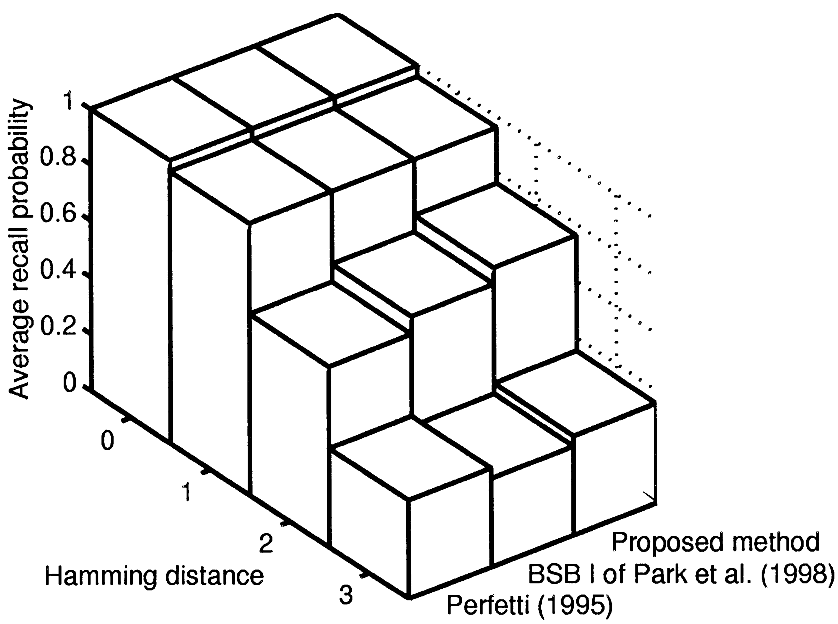

Figure 3: Average recall probabilities of the BSBs designed for the third example.

5 Conclusion

We have addressed the problem of finding the parameters of the GBSB neural networks that can store given prototype patterns with optimal performance. Based on known qualitative properties and newly derived fundamental properties of the GBSB model, the problem was formulated as a nonlinear constrained optimization problem. Then this problem was converted into a quasi-convex optimization problem called a GEVP. This conversion is particularly useful in practice, because GEVPs can be efficiently solved by recently developed interior point methods. Three design examples were presented to illustrate the proposed method and to compare with existing methods. The simulation results showed that the memories obtained by the proposed method have better performance than the memories designed by existing methods.

References

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review, 84*, 413–451.

Boyd, S., & El-Ghaoui, L. (1993). Method of centers for minimizing generalized eigenvalues. *Linear Algebra and Applications, 188*, 63–111.

Boyd, S., El-Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in systems and control theory*. Philadelphia: SIAM.

Chan, H. Y., & Zak, S. H. (1997). On neural networks that design neural associative memories. *IEEE Transactions on Neural Networks, 8*, 360–372.

Cohen, M. A., & Grossberg, S. (1983). Absolute stability of global pattern formulation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man and Cybernetics, 13*, 815–826.

Gahinet, P., Nemirovskii, A., Laub, A. J., & Chilali, M. (1995). *LMI control toolbox*. Natick, MA: Mathworks.

Golden, R. M. (1986). The brain-state-in-a-box neural model is a gradient descent algorithm. *Journal of Mathematical Psychology, 30*, 73–80.

Golden, R. M. (1993). Stability and optimization analyses of the generalized brain-state-in-a-box neural network model. *Journal of Mathematical Psychology, 37*, 282–298.

Hassoun, M. H. (Ed.). (1993). *Associative neural memories: Theory and implementation*. New York: Oxford University Press.

Hui, S., & Zak, S. H. (1992). Dynamical analysis of the brain-state-in-a-box (BSB) neural models. *IEEE Transactions on Neural Networks, 3*, 86–100.

Lillo, W. E., Miller, D. C., Hui, S., & Zak, S. H. (1994). Synthesis of brain-state-in-a-box (BSB) based associative memories. *IEEE Transactions on Neural Networks, 5*, 730–737.

Marcus, C. M., & Westervelt, R. M. (1989). Dynamics of iterated-map neural networks. *Physical Review A, 40*, 501–504.

Michel, A. N., & Farrell, J. A. (1990). Associative memories via artificial neural networks. *IEEE Control Systems Magazine, 10*, 6–17.
Nesterov, Y., & Nemirovskii, A. (1994). *Interior point polynomial methods in convex programming*. Philadelphia: SIAM.
Park, J., Cho, H., & Park, D. (1999). On the design of BSB neural associative memories using semidefinite programming. *Neural Computation, 11*, 1985–1994.
Perfetti, R. (1995). A synthesis procedure for brain-state-in-a-box neural networks. *IEEE Transactions on Neural Networks, 6*, 1071–1080.
Strang, G. (1988). *Linear algebra and its applications* (3rd ed.). Orlando, FL: HBJ Publishers.
Vandenberghe, L., & Balakrishnan, V. (1997). Algorithms and software for LMI problems in control. *IEEE Control Systems Magazine, 17*, 89–95.
Zak, S. H., Lillo, W. E., and Hui, S. (1996). Learning and forgetting in generalized brain-state-in-a-box (GBSB) neural associative memories. *Neural Networks, 9*, 845–854.

Received February 4, 1999; accepted April 20, 1999.