A Larger Family of Objective Functions to which Hopfield Neural Networks can give Globally Optimal Solutions

Yoshinori Uesaka
Department of Information Sciences, Faculty of Science and Technology,
Science University of Tokyo
2641 Yamazaki, Noda-City, Chiba, 278-8510, JAPAN
Phone : +81-471-24-1501, Fax : +81-471-23-9764
email : uesaka@is.noda.sut.ac.jp

ABSTRACT : A family of objective functions is discussed for which the conjecture, stating that globally optimal solution (not a local one) of Hopfield neural networks may be obtained by starting from an initial point sufficiently close to the origin, truly holds. To the conjecture, Nishi (1998) has been giving an interesting family of objective functions through the eigenvector analysis. The present paper extends his result to the larger family which suggests that the conjecture really holds for a wide range of objective functions.

KEYWORDS : Combinatorial optimization, Global optimum, Hopfield neural networks, Dynamical system

1. INTRODUCTION

Since Hopfield and Tank (1985) pointed out the possibility such that their neural networks (i.e., Hopfield machine) might be able to solve intractable combinatorial problems such as the traveling salesman problem, a considerable amount of investigations has been done so far (for examples, Wilson and Pawley, 1988; Aiyer, Niranjan and Fallside, 1990; Kamgar-Parsi, 1990; Abe, 1993; Matsuda, 1994). Several years ago Uesaka (1991) gave a mathematical formulation to solve combinatorial problems by means of Hopfield neural netwoks and showed that almost all problems may be reduced to the minimization of a quadratic form of objective function with a dynamical system derived from Hopfield neural networks.

When we solve the problem by simulation of the dynamical system, we encounter the problem of selecting a "good" initial state approaching to the global optimum point. Based on numerical simulation of the dynamical system, Uesaka (1991) obtained the conjecture stating that globally optimal solution (not a local one) may be obtained by starting from an initial point sufficiently close to the origin. Though this conjecture has been left a long time without mathematical analysis, recently Nishi (1998) has began to analyze this conjecture from a mathematical point of view and found an interesting family of objective functions for which the conjecture holds.

In the present paper it is tried to extend the Nishi's result to a larger family of objective functions which includes the Nishi's family as a special case. The obtained family suggests that the conjecture really holds for a considerably wide range of objective functions.

2. COMBINATORIAL OPTIMIZATION PROBLEMS

Let X be a Cartesian product of n $\{-1,+1\}$'s, that is, $X := \{[x_1 \ \cdots \ x_n]^T \,|\, x_1,\ldots,x_n = \pm 1\}$, where T denotes the transpose of a vector and/or matrix. Consider a real-valued function (objective function) on X such that

(1) $F(x_1,\ldots,x_n) = -\frac{1}{2}\sum_{i,j=1}^{n} a_{ij} x_i x_j - \sum_{i=1}^{n} b_i x_i$,

where a_{ij} and b_i are constants satisfying that for any $i, j = 1,\ldots,n$ $a_{ii} = 0$ and $a_{ij} = a_{ji}$. First we discuss the following problem :

Problem A *Let S be a subset of X. Find the minimum value $F_{\min}$ and minimum point $x_{\min}$ of F with a constraint of " $x \in S$ ", that is, find*

(2) $F_{\min} := \min\{F(x) \mid x \in S\}$, $x_{\min} := \arg\min\{F(x) \mid x \in S\}$.

It is well known that most combinatorial optimization problems such as the traveling salesman problem (TSP) may be formulated as Problem 1 (Hopfield and Tank, 1985). In almost all cases the constraint "$x \in S$" might be, however, removed away :

Theorem 1 (Uesaka, 1991) *Assume that there exists a real-valued quadratic function, induced from S, on X such that $G(x) = 0$ if $x \in S$ and $G(x) > 0$ otherwise. Let H be a real-valued function, made of F and G, on X :*

(3) $H(x) = F(x) + cG(x)$,

where c is a positive constant. Then, there exists a c such that if $x_{\min}$ is a minimum point of H on X, it is also a minimum point of F on S.

Though this theorem requires the existence of G, for typical optimization problems, e.g., TSP, G is well known to be easily found (Uesaka and Ozeki, 1990). In view of Theorem 1 we discuss hereafter the minimum search problem without constraint :

Problem B *Find the minimum value $F_{\min}$ and minimum point $x_{\min}$ of F, that is, find*

(4) $F_{\min} := \min\{F(x) \mid x \in X\}$, $x_{\min} := \arg\min\{F(x) \mid x \in X\}$.

It has been proved that we can drop the linear term in (1) without loss of generality :

Theorem 2 (Uesaka, 1991) *Using the matrix and vector made up of coefficients in (1) as follows :*

(5) $A := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, $b := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$,

we define the square matrix of order $n+1$ as

(6) $B := \begin{bmatrix} 0 & b^T \\ b & A \end{bmatrix}$

and consider the following quadratic form :

(7) $G(y) = -\frac{1}{2} y^T By$, $y = [x_0 \quad x_1 \quad \cdots \quad x_n]^T$.

If $[x_0^* \quad x_1^* \quad \cdots \quad x_n^*]^T$ is a minimum point of G on $\{-1,+1\} \times X$, then a minimum point of F on X is given by the n dimensional vector $x_0^*[x_1^* \quad \cdots \quad x_n^*]^T$.

In view of those results above we consider the minimum search problem of a quadratic form on X as combinatorial problems without loss of generality :

Problem C *Let F be a real-valued quadratic function on X :*

(8) $F(x) = -\frac{1}{2} x^T Ax$,

where $A = [a_{ij}]$ is a matrix of order n, satisfying that for any $i, j = 1, \ldots, n$ $a_{ii} = 0$ and $a_{ij} = a_{ji}$. Then, find the minimum value $F_{\min}$ and minimum point $x_{\min}$ of F, that is,

(9) $F_{\min} := \min\{F(x) \mid x \in S\}$, $x_{\min} := \arg\min\{F(x) \mid x \in S\}$.

3. DYNAMICAL SYSTEM FOR OPTIMIZATION PROBLEMS

In order to solve Problem C, we prepare a real-valued function E which is obtained by extension of F to the n dimensional Euclidean space and consider the following dynamical system :

(10) $E(x) = -\frac{1}{2} x^T Ax$, $x \in \Re^n$,

(11) $\frac{dx_i}{dt} = -(1 - x_i^2)\frac{\partial E}{\partial x_i} = (1 - x_i^2)\sum_{i,j=1}^{n} a_{ij} x_j$, $i = 1, \ldots, n$,

or in the matrix form

(12) $\frac{dx}{dt} = -R(x)\,\mathrm{grad}E = R(x)Ax$, $R(x) = \mathrm{diag}[1 - x_1^2 \quad \cdots \quad 1 - x_n^2]$.

This dynamical system is closely related to the Hopfield neural network (Hopfield and Tank, 1985), and the function E is called its *energy function*. Note that the n dimensional vector x, of which components are outputs of *neural elements*, is regarded as a *state* of the dynamical system. As easily seen in (11) or (12), the state moves with time in the

n dimensional cube $C := [-1,+1]^n$, drawing a *trajectory*, i.e., a *solution* of the differential equation (12). If, after starting from an initial sate, the state asymptotically approaches to the vertex of C which is a minimum point of F, then the optimization Problem C is able to be solved by means of simulation of the dynamical system above. In fact the following results have been obtained :

Theorem 3 (Uesaka, 1991) *For the objective function F, a vertex $p = [p_1 \ \cdots \ p_n]^T$ of the cube C is said to be minimal if the inequalities hold for $i = 1,\ldots,n$:*

(13) $F(p_1,\ldots,p_i,\ldots,p_n) < F(p_1,\ldots,-p_i,\ldots,p_n)$.

If at least one of $<$'s is replaced with $\leq$, p is said to be minimal in wide sense. If p is neither minimal nor minimal in wide sense, it is said to be nonminimal. Then, the dynamical system (12) has the following properties :

1° *Any minimal point of F is asymptotically stable,*

2° *Any minimal point in wide sense of F may be stable or unstable,*

3° *Any nonminimal point of F is unstable,*

4° *Any inner point of C except the origin is unstable when the matrix A in Problem C is nonsingular.*

In the case of A being singular, the right-hand side of (12) becomes zero at some inner points of C. If we choose such points as initial points, then we can get no minimal point of the objective function F. However the probability to choose such an initial point is zero except the trivial case of A.

4. MAIN RESULTS

Theorem 3 guarantees that minimal points of F can be obtained as asymptotically stable ones of the dynamical system if the state starts from an appropriate inner point of C. What we want to get is however globally minimal (minimum) points of F but not locally minimal ones. Since the globally minimal points lie, of course, among their locally minimal ones, it is possible to obtain them if "good" initial states are selected. Thus it becomes very important to choose an appropriate initial point in order to get globally minimal points of F.

On this *problem of selecting good initial points,* Uesaka (1991) has been giving the following conjecture based on simulation results :

Conjecture (Uesaka, 1991) *Let $S(d)$ be an n dimensional hypercube of size $d \in (0,1]$ and let $P(d)$ denote the probability that an initial point randomly selected from the cube $S(d)$ falls in the basin of globally minimal points of F, that is, the probability for a minimum search with randomly selected initial points to achieve success. Then $P(d)$ converges to 1 as d becomes to 0 .*

This conjecture is very important from not only theoretical but also practical point of view. So it is an interesting task to characterize a family of objective functions for which the conjecture holds. In fact Nishi (1998) has given the following attractive family :

Theorem 4 (Nishi, 1998) *Let $d_1,\ldots,d_n$ be eigenvalues of A, assuming without loss of generality that*

(14) $d_1 \geq \cdots \geq d_n$.

Then, the conjecture holds for the case that the eigenvector corresponding to the largest eigenvalue d_1 is composed only of $\pm\left(1/\sqrt{n}\right)$.

In the present paper we extend Theorem 4, that is, we will give the larger family, than one specified by Theorem 4, of objective functions for which the conjecture holds.

Since A is a real symmetric matrix, it may be diagonalized as

(15) $$A = UDU^T, \quad D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$$

where U is the orthogonal matrix and D is the diagonal matrix composed of eigenvalues of A, assumed as (14) without loss of generality. Let u_j denote the eigenvector corresponding to the eigenvalue d_j, being also the j th column vector of U, that is,

(16) $u_j := [u_{1j} \ \cdots \ u_{nj}]^T$, $j = 1,\ldots,n$, where $U = [u_{ij}]$.

Note that $u_1,\ldots,u_n$ is an orthonormal basis, and hence

(17) $U^T U = UU^T = I$,

where I is the identity matrix. Furthermore, we have

(18) $d_1 + \cdots + d_n = 0$

since the trace of A is zero.

Our main result is now stated as follows :

Theorem 5 *If the matrix A of an objective function F satisfies that*

(19) $d_1 = \cdots = d_s > d_{s+1} \geq \cdots \geq d_n$,

and there exist real numbers $a_1, \ldots, a_s$ such that

(20) $a_1 u_1 + \cdots + a_s u_s \in X$,

that is, the vector $a_1 u_1 + \cdots + a_s u_s$ is composed only of ± 1, then the conjecture holds for the objective function F specified by A as in (8).

The proof for Theorem 5 will be given in the next section after preparing a few lemmata.

As easily seen by taking $s = 1$, Theorem 4 is included in Theorem 5 as a special case. In order to get examples satisfying the premise of Theorem 5, we note the following :

Lemma 1 *If the matrix A satisfies that*

(21) $d_1 = \cdots = d_s > d_{s+1} \geq \cdots \geq d_n$,

and there exist real numbers $a_1, \ldots, a_s$ such that

(22) $a_1 u_1 + \cdots + a_s u_s \in X$,

then

(23) $\forall i = 1, \ldots, n, \ \sum_{j=s+1}^{n} u_{ij}^2 \left(1 - \frac{d_j}{d_1}\right) = 1,$

and

(24) $\forall j = s+1, \ldots, n, \ u_j^T p = \sum_{i=1}^{n} u_{ij} p_i = 0$,

where $p := a_1 u_1 + \cdots + a_s u_s$.

Proof In view of (15), any component of A is calculated as

(25) $a_{ij} = \sum_{k,h=1}^{n} u_{ik} d_k d_{kh} u_{jv} = \sum_{k=1}^{n} u_{ik} u_{jk} d_k$,

where d_{kh} is Kronecker's delta. Hence, with $a_{ii} = 0$ and $u_j^T u_j = 1$, we have that

(26) $0 = \sum_{j=1}^{n} u_{ij}^2 d_j = d_1 \sum_{j=1}^{s} u_{ij}^2 + \sum_{j=s+1}^{n} u_{ij}^2 d_j = d_1 - \sum_{j=s+1}^{n} u_{ij}^2 \left(d_1 - d_j\right)$

which gives (23). Noting that $u_1, \ldots, u_n$ is an orthonormal basis, we have for $j = s+1, \ldots, n$

(27) $u_j^T p = a_1 u_j^T u_1 + \cdots + a_s u_j^T u_s = a_1 d_{j1} + \cdots + a_s d_{js} = 0$. Q.E.D.

Consider the case of $n = 3$. Suppose that

(28) $d_1 = d_2 > d_3$.

Applying this to (23) and noting that $d_3 = -2d_1$ because of $d_1 + d_2 + d_3 = 0$, we have $u_{i,3} = \pm\left(1/\sqrt{3}\right)$ for $i = 1, 2, 3$. On the other hand, from (24), $u_{13} p_1 + u_{23} p_2 + u_{33} p_3 = 0$, where $p = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}^T = a_1 u_1 + a_2 u_2$, which implies that

(29) $\pm p_1 \pm p_2 \pm p_3 = 0$.

Since p_i takes only ± 1, (29) is not satisfied by any p, which means that $s = 1$ is a only possible case except the trivial A.

However we can have a matrix A satisfying the premise of Theorem 5 for the case of $n \geq 4$. For example, let

(30) $d_1 = d_2 = 3, \quad d_3 = -1, \quad d_4 = -5$.

Then, from (23), we have that

(31) $\forall i = 1, \ldots, 4, \ \frac{4}{3} u_{i,3}^2 + \frac{8}{3} u_{i,4}^2 = 1$.

Hence, for example, we may take u_3 and u_4 as

(32) $u_3 = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}^T, \ u_4 = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}^T$,

which satisfy (24) for $p = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$. Taking two vectors u_1', u_2', say $u_1' = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T$ and $u_2' = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^T$, for which u_1', u_2', u_3, u_4 are linearly independent, and orthonormalizing them by Gram-Schmitt method, we obtain an orthonormal matrix :

$$(33)\quad U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

In view of (30) and (33) we finally get a matrix A satisfying the premise of Theorem 5 :

$$(34)\quad A = UDU^T = \begin{bmatrix} 0 & -1 & 1 & 3 \\ -1 & 0 & 3 & 1 \\ 1 & 3 & 0 & -1 \\ 3 & 1 & -1 & 0 \end{bmatrix}.$$

5. PROOF FOR THEOREM 5

In this section we will give a complete proof for Theorem 5. In order to do so we first prepare a few lemmata. From (8), (15) and (17), it is easily confirmed that

Lemma 2 *For any vector x in C, there are real numbers $a_1,\ldots,a_n$ such that*

(35) $x = a_1u_1+\cdots+a_nu_n$ and $F(x) = -\frac{1}{2}(a_1^2d_1+\cdots+a_n^2d_n)$

Lemma 3 *Suppose that $d_1=\cdots=d_s>d_{s+1}\geq\cdots\geq d_n$ for some s between 1 and n. Let p be in X. If there exist real numbers $a_1,\ldots,a_s$ such that*

(36) $p = a_1u_1+\cdots+a_su_s$,

then the objective function F takes a minimum value at p.

Proof Let q be any vector in X. Since the eigenvectors $u_1,\ldots,u_n$ of A are an orthonormal basis of the n dimensional Euclidean space, there exist real numbers $b_1,\ldots,b_n$ such that $q=b_1u_1+\cdots+b_nu_n$. Hence, by calculating the norms of p and q, we have that

(37) $n = \|p\|^2 = p^Tp = a_1^2+\cdots+a_s^2$, $n = \|q\|^2 = q^Tq = b_1^2+\cdots+b_s^2$.

Thus, in view of Lemma 2 and $d_1=\cdots=d_s>d_{s+1}\geq\cdots\geq d_n$,

$$(38)\quad F(p)-F(q) = -\frac{1}{2}(a_1^2d_1+\cdots+a_s^2d_s)+\frac{1}{2}(b_1^2d_1+\cdots+b_n^2d_n) \leq -\frac{1}{2}(a_1^2+\cdots+a_s^2)d_1+\frac{1}{2}(b_1^2+\cdots+b_n^2)d_1$$
$$= -\frac{1}{2}nd_1+\frac{1}{2}nd_1 = 0.$$ Q.E.D.

Lemma 4 *The dynamical system (12) may be rewritten as*

$$(39)\quad \frac{dx}{dt} = \mathrm{diag}\left[1-x_1^2 \quad \cdots \quad 1-x_n^2\right]\left[d_1(u_1^Tx)u_1+\cdots+d_n(u_n^Tx)u_n\right]$$

For any p in X, we define a function on C such that

$$(40)\quad \forall x\in C,\ Q(x) = \frac{1}{2}\|x-p\|^2 = \frac{1}{2}(x-p)^T(x-p).$$

When x lies on the solution of the dynamical system (12),

$$(41)\quad \frac{dQ}{dt} = \sum_{i,j=1}^{n} u_{ij}(x_i-p_i)(1-x_i^2)d_ju_j^Tx.$$

Proof Applying (15) to (12), we have (39) straightforward. Inserting (12) and (15) into the relation $dQ/dt=(x-p)^T(dx/dt)$, its right-hand side is calculated as in (41). Q.E.D.

Lemma 5 *Suppose that $d_1=\cdots=d_s>d_{s+1}\geq\cdots\geq d_n$ for some s between 1 and n, and that there exist real numbers $a_1,\ldots,a_s$ such that $p=a_1u_1+\cdots+a_su_s$. Note that p is a minimum point of the objective function F in view of Lemma 3. Hence, by Theorem 3, p is also an asymptotically stable.*

Then, the straight line G connecting the origin and p, that is, the set $G := \{x \mid \exists b \in [0,1)\ x = bp\}$ is contained in the basin of p.

Proof Suppose that for some time $t \ge 0$ the solution $x(t)$ lies on the line G, that is, there is a real number $b \in [0,1)$ such that $x(t) = bp$. Then, in view of Lemma 4, we have that

(42) $$\frac{dx}{dt} = \mathrm{diag}\left[1 - b^2 p_1^2 \quad \cdots \quad 1 - b^2 p_n^2\right] b \left\{d_1\left(u_1^T p\right)u_1 + \cdots + d_n\left(u_n^T p\right)u_n\right\}$$

(because of $p_j^2 = 1$ and p being orthogonal to u_j for $j = s+1,\ldots,n$)

$$= \left(1 - b^2\right) b \left\{d_1\left(u_1^T p\right)u_1 + \cdots + d_s\left(u_s^T p\right)u_s\right\}$$

(because of $d_1 = \cdots = d_s$ and $p = a_1 u_1 + \cdots + a_s u_s = u_1^T p u_1 + \cdots + u_s^T p u_s$)

$$= \left(1 - b^2\right) b d_1 p.$$

Noting that $x(t + \Delta t) = x(t) + \frac{dx}{dt}\Delta t$, we can therefore conclude that for any $t' \ge t$ $x(t')$ also lies on the line G. On the other hand, (41) may be evaluated as

(43) $$\frac{dQ}{dt} = \sum_{i,j=1}^{n} u_{ij}\left(p_i - bp_i\right)\left(1 - b^2 p_i^2\right) d_j u_j^T (bp)$$

(because of $p_j^2 = 1$)

$$= b(1-b)\left(1 - b^2\right) p^T \sum_{j=1}^{n} u_j^T p u_j d_j$$

(because of p being orthogonal to u_j for $j = s+1,\ldots,n$ and $d_1 = \cdots = d_s$)

$$= b(1-b)\left(1 - b^2\right) d_1 p^T \sum_{j=1}^{s} u_j^T p u_j$$

(because of $p = a_1 u_1 + \cdots + a_s u_s = u_1^T p u_1 + \cdots + u_s^T p u_s$)

$$= b(1-b)\left(1 - b^2\right) d_1 p^T p = b(1-b)\left(1 - b^2\right) d_1 n < 0.$$

Note that, by Theorem 3, there is no stable point on the line G. Hence (42) and (43) guarantees that $x(t)$ approaches to p as $t \to \infty$. Q.E.D.

Lemma 6 *Suppose that $d_1 = \cdots = d_s > d_{s+1} \ge \cdots \ge d_n$ for some s between 1 and n, and that there exist real numbers $a_1,\ldots,a_s$ such that $p = a_1 u_1 + \cdots + a_s u_s \in X$. If the initial point $x(0)$ is sufficiently close to the origin, then the solution $x(t)$ of the dynamical system (12) passes a point sufficiently close to the line G within a sufficiently small time. More precisely, for any $\varepsilon > 0$ there exist $\delta > 0$, $t > 0$ and $b \in [0,1)$ such that $\|x(0)\| < \delta$ implies $\|x(t) - bp\| < \varepsilon$.*

Proof First we take respectively t and δ sufficiently small so that the diagonal matrix $R(x)$ of (12) may be considered to be the identity. Then, (12) can be easily solved as

(44) $$x(t) = Ue^{Dt}U^T x(0) = \left(e^{d_1 t} u_1 u_1^T + \cdots + e^{d_n t} u_n u_n^T\right) x(0).$$

Now we evaluate $\|x(t) - bp\|$ as follows :

(45) $$\|x(t) - bp\| \le \left\|x(t) - \left(e^{d_1 t} u_1 u_1^T + \cdots + e^{d_s t} u_s u_s^T\right) x(0)\right\| + \left\|\left(e^{d_1 t} u_1 u_1^T + \cdots + e^{d_s t} u_s u_s^T\right) x(0) - bp\right\|.$$

In view of (44) the first term in the right-hand side of (45) is evaluated as

(46) $$\left\|x(t) - \left(e^{d_1 t} u_1 u_1^T + \cdots + e^{d_s t} u_s u_s^T\right) x(0)\right\| = \left\|\left(e^{d_{s+1} t} u_{s+1} u_{s+1}^T + \cdots + e^{d_n t} u_n u_n^T\right) x(0)\right\|$$

(using the triangle inequality and Chaucy's inequality)

$$\le e^{d_{s+1} t}\left\|u_{s+1}^T x(0)\right\| + \cdots + e^{d_n t}\left\|u_n^T x(0)\right\| \le \left(e^{d_{s+1} t} + \cdots + e^{d_n t}\right)\|x(0)\|$$

(noting that $d_1 > d_{s+1} \ge \cdots \ge d_n$ and $\|x(0)\| < \delta$)

$$\le (n - s)e^{d_1 t}\delta.$$

Since $u_1 u_1^T + \cdots + u_s u_s^T$ is a projection onto the subspace spanned by $u_1,\ldots,u_s$, and p is in the subspace, p is equal to $\left(u_1 u_1^T + \cdots + u_s u_s^T\right)p$. Hence the second term in the right-hand side of (45) is evaluated with $d_1 = \cdots = d_s$ as

(47) $\left\|\left(e^{d_1t}u_1u_1^T+\cdots+e^{d_st}u_su_s^T\right)\left(x(0)-bp\right)\right\| = \left\|\left(u_1u_1^T+\cdots+u_su_s^T\right)\left(e^{d_1t}x(0)-bp\right)\right\|$

$\le \left\|e^{d_1t}x(0)-bp\right\| \le e^{d_1t}\left\|x(0)\right\|+b\left\|p\right\| \le e^{d_1t}d+b\sqrt{n}$.

Combining (46) and (47) with (45), we have the evaluation such that

(48) $\left\|x(t)-bp\right\| \le (n-s+1)e^{d_1t}d+b\sqrt{n}$.

Here we take t sufficiently small and, after that, select $d>0$ and $b\in(0,1)$ for a given $e>0$ so that the right-hand side of (48) becomes less than e . Q.E.D.

Proof for Theorem 5 For (20) let $p=a_1u_1+\cdots+a_su_s$. Then, by Lemma 3, p is a minimum point of the objective function F. Hence, in view of Theorem 3, p is also asymptotically stable, meaning that there is a neighborhood V of p contained in its basin. Note that the right-hand side of the dynamical system (12) is continuously differentiable. Therefore, as well known in the standard theory of differential equations (e.g., Hirsch and Smale, 1974), the neighborhood V is mapped to a neighborhood W of bp ($0<b<1$) contained in the basin of p by the flow map of the system (12) because bp is in the basin in view of Lemma 5. So any solution passing a point in the neighborhood W approaches to the globally optimal point p .

On the other hand, Lemma 6 guarantees that, if the initial point is sufficiently close to the origin, then the solution of the dynamical system (12) passes a point in W, and hence it finally approaches to p. The amount of sufficient closeness of the initial state to the origin depends on the matrix A, that is, the objective function F. Thus, the probability for a minimum search with a randomly selected initial point to achieve success converges to one as the size d of the hypercube $S(d)$ becomes to zero. Q.E.D.

6. CONCLUDING REMARKS

Though it is very difficult to characterize the family for which the conjecture holds, it may be possible to get much more wider family. The first step to do so might be to analyze the case where the first s eigenvalues $d_1,\ldots,d_s$ are nearly equal each other and a global minimum point lies in the subspace spanned by the eigenvectors $u_1,\ldots,u_s$. Another open problem is to characterize the family of objective functions for which the conjecture holds for a case of small n .

REFERENCES

Abe, S., 1993, "Global convergence and suppression of spurious states of the Hopfield neural networks", IEEE Trans. Circuits Syst. I, Vol.40, pp.246-257.

Aiyer, S.V.B., Niranjan, M. and Fallside, F., 1990, "A theoretical investigation into the performance of the Hopfield model", IEEE Trans. Neural Networks, Vol.1, pp.204-215.

Hirsch, M.W. and Smale, S., 1974, "Differential equations, dynamical systems, and linear algebra", Academic Press, New York.

Hopfield, J.J. and Tank, D.W., 1985, ""Neural" computation of decisions in optimization problems", Biol. Cybern., 52, pp.141-152.

Kamgar-Parsi, B., 1990, "On the problem solving with Hopfield neural networks", Biol. Cybern., Vol.62, pp.415-423.

Matsuda, S., 1994, "Theoretical characterizations of possibilities and impossibilities of Hopfield neural networks in solving combinatorial optimization problems", ICNN'94, pp.4563-4566.

Nishi, T., 1998, "On the Uesaka's conjecture as to the optimization by means of neural networks for combinatorial problems", IEICE Trans. Fundamentals, Vol.E81-A, pp.1811-1817.

Uesaka, Y., 1991, "Mathematical aspects of neuro-dynamics for combinatorial optimization", IEICE Trans. Fundamentals, Vol.E-74, No.6, pp.1368-1372.

Uesaka, Y. and Ozeki, K, 1990, "Algorithms for pattern recognition and learning machines", Bunichi-Sougou Shuppan, Tokyo (in Japanese).

Wilson, G.W. and Pawley, G.S., 1988, "On the stability of the traveling salesman problem algorithm of Hopfield and Tank", Biol. Cybern., Vol.58, pp.63-70.