A Global Optimization Algorithm Based on
Circuit Partitioning Technique

徳島大学工学部電気電子工学科

牛田明夫 (Akio Ushida)

Abstract

Global optimization problems are very important in designing of VLSI circuits and solving combinatorial problems by neural networks. In this paper, we show an efficient algorithm for finding global minimum points of the Hopfield's type energy functions based on a partitioning technique. At the first step, we partition the energy function into the small systems, and find all the local minimum points to each partitioned small functions, where some of them are chosen as the candidates of the global minimum points. At the second step, we estimate the energy of the original function at the points obtained by the combinations of these earlier candidates from each partitioned functions.

1. Introduction

In this paper we show an efficient algorithm for finding all of the global minimum points of Hopfield's type energy functions based on a partitioning technique. In our algorithm, we first partition a given energy function into P small size energy functions. For each energy function, we calculate the approximate energy at near the local minimum points, and some of the lower energies are chosen as candidates of the global minimum points of original energy function [1]. The exact global minimum points are calculated by the Newton-Raphson method starting from the approximate minimum points. Theorem 3 in section 3 is the main result of this paper.

For special types of the Hopfield's networks [2-3] and the cellular neural networks [4-5], all of the solutions lie in a restricted domain such as $0 < x_i < 1$ and/or $-1 \leq x_i \leq 1$ on each axis because of the special type of sigmoid functions, and their local minimum points are located at the corner points of the domains such as 1, 0 or -1.

In this paper, we restrict our discussion on the optimization problem of the Hopfield's type energy functions, because the ideas can be easily applied to the cellular neural networks.

2. Properties of the Hopfield networks

Consider the well known Hopfield network [2-3] having N sigmoids. The circuit equation is described by the following form:

$$C_i \frac{du_i}{dt} = \sum_{j=1}^{N} w_{ij} x_j - \frac{u_i}{R_i} + I_i \tag{1.1}$$

$$x_i = \frac{1}{2}(1 + \tanh \frac{u_i}{a}), \quad 0 < x_i < 1 \tag{1.2}$$

$$u_i = f(x_i) = \frac{a}{2}\log\frac{x_i}{1-x_i} \tag{1.3}$$

$$i = 1,2,...,N$$

where w_{ij} shows the weight of synapse coupling, and has the following properties of $w_{ij}=w_{ji}$ and $w_{ii}=0$. Thus, the coefficient matrix W of (1.1) is a *symmetric matrix* whose diagonal elements are zeros. We assume through this paper that any principal minor of W larger than 2 is full rank. Then, we have following properties [7].

Property 1 : Assume A is a nonsingular symmetric matrix whose diagonal element are all zeros. Then, the eigenvalues contain both positive and negative reals.

Property 2 : Given two real symmetric matrices, A and B, with A positive definite, there exists a nonsingular matrix T such that

$$T^T A T = I \tag{2}$$

where I is a unit matrix.

$$T^T B T = diag[\mu_i] \tag{3}$$

The stiffness of the sigmoid function (1) depends on the value "a", where the stiffness means that the characteristic curve changes in a small domain only as shown in Fig.1. The sigmoid function becomes very stiff; i.e., the slope is very step around "a". For this kind of systems, we can prove that the local minimum points are located around the corner points of [0,1] on each axis.
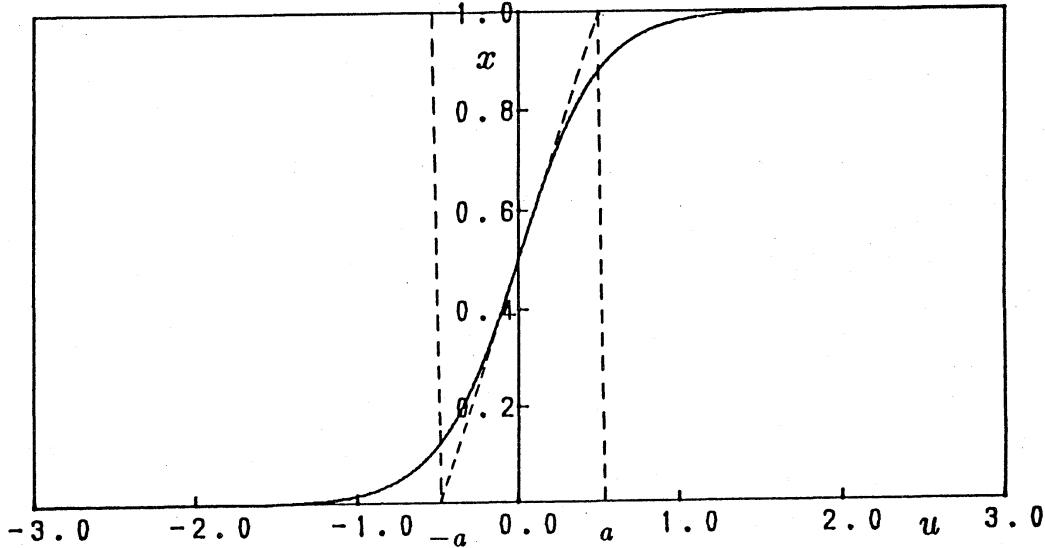

Fig.1 Sigmoid function $a=0.5$.

Let us describe the energy function of Hopfield circuit by

$$\phi(x) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}x_i x_j + \sum_{i=1}^{N}\frac{1}{R_i}\int_0^{x_i} f(x_i)dx_i - \sum_{i=1}^{N} x_i I_i \tag{4}$$

where the second term is given by

$$\int_0^{x_i} f(x_i)dx_i = \frac{a}{2}[x_i\log x_i + (1-x_i)\log(1-x_i)] \tag{5}$$

The stationary points of the energy function $\phi(x)$ are corresponding to the steady-state solutions of (1), and we can prove the following interesting theorem about local minimum points [8-9].

Theorem 1 : The stationary points of (1) will be local minimum points if they satisfy the following relations:

$$\max| \mu_i | < \max|\frac{\partial u_i}{\partial x_i}|/R_i, \quad i=1,2,...,N \tag{6}$$

where μ_i is the eigenvalue of the coefficient matrix W.

Proof: In order to estimate the stability of a stationary point x_0, consider the variational equation of (1) by substituting $x=x_0+\Delta x$: namely

$$C\frac{\partial u}{\partial x}\frac{d\Delta x}{dt} = W\Delta x - \frac{1}{R}\frac{\partial u}{\partial x}\Delta x \tag{7}$$

where we have assumed $C=diag[C_i]$, $R=diag[R_i]$, and

$$\frac{\partial u}{\partial x} = diag[\frac{\partial u_i}{\partial x_i}], \quad \lambda_i\equiv\frac{\partial u_i}{\partial x_i}>0 \tag{8}$$

Now, we introduce a scaler matrix such that

$$S_1 = diag[\frac{\sqrt{R_i}}{\sqrt{\partial u_i/\partial x_i}}] \tag{9}$$

Put $\Delta x=S_1 z$. Then, we have from (7)

$$CR\frac{dz}{dt} = S_1^T WS_1 z - z \tag{10}$$

where $S_1^T WS_1$ is also symmetric matrix, so that there is an orthogonal transformation matrix S_2 such as $z=S_2 y$, which transforms (10) into the following form:

$$S_2^T CRS_2\frac{dy}{dt} = S_2^T S_1^T WS_1 S_2 y - y$$

$$= (diag[\alpha_i]-1)y \tag{11}$$

Using property 2 with $T=Ssub2$, $A=CR$ and $B=S_1^T WS_1$, this is true. Hence, the steady-state solution $\hat{x}$ will be stable if it satisfies

$$\max|\alpha_i| = \| S_2^T S_1^T WS_1 S_2\| \leq \| W\|\times\max\{\frac{R_i}{\partial u_i/\partial x_i}\} < 1 \tag{12}$$

where the norms denote *spectrum norm*. Thus, we proved the theorem. **Q.E.D.**

Note that the derivative $\partial u_i/\partial x_i$ of the sigmoid function around the corner points of [0,1] is always very large. Therefore, we conclude that the stationary points located on around the corners are all local minimum points.

Next, we will prove that, under well defined conditions, the components x_i at the local minimum points are located at corner points of [0,1] only, and no local minimum

exists outside the corner points.

Theorem 2 : Let any sub-matrix of W larger than 2×2 be full rank. Assume that the solution of (1) contains two kinds of components. Namely, the first K-components $(N-K \geq 2)$, $x_1, x_2, \cdots, x_K$ are lying around corner points of the domain [1,0]. Hence, we can assume that the derivatives

$$\beta_i = \partial u_i / \partial x_i, \quad i=1,2,...,K \tag{13}$$

are sufficiently large.

Another (N-K)-components $x_{N-K}, x_{N-K+1}, \cdots, x_N$ are apart from the corner points. Hence, we can assume that the derivatives

$$\varepsilon_i = \partial u_i / \partial x_i, \quad i=K+1,...,N \tag{14}$$

are sufficiently small.

Then, the solution point will never be the local minimum point.

Proof. The first K-variational equations can be written in the following form:

$$\beta_i C_i \frac{d\Delta x_i}{dt} = \sum_{j=1}^{N} w_{ij}\Delta x_j - \frac{\beta_i}{R_i}\Delta x_i \tag{15}$$

For sufficiently large β_i, we can neglect first term of (15) because $1/\beta_i$ is small enough. Thus, we have

$$\frac{d\Delta x_i}{dt} = -\frac{1}{R_i C_i}\Delta x_i \tag{16}$$

The solution Δx_i becomes to zero as "t" infinitive.

On the other hand, the second (N-K)-variational equations can be written in the following form:

$$\varepsilon_i C_i \frac{d\Delta x_i}{dt} = \sum_{j=1}^{K} w_{ij}\Delta x_j + \sum_{j=K+1}^{N} w_{ij}\Delta x_j - \frac{\varepsilon_i}{R_i}\Delta x_i \tag{17}$$

For a sufficiently small ε_i, the third term in the right hand side is negligible compared with other two terms, so that the essential term is written by

$$\frac{d\Delta x_i}{dt} = \frac{1}{\varepsilon_i C_i}\sum_{j=1}^{K} w_{ij}\Delta x_j + \frac{1}{\varepsilon_i C_i}\sum_{j=K+1}^{N} w_{ij}\Delta x_j \tag{18}$$

The system of equations (16) is stable, so that it has zero solutions $(\Delta x_i=0, i=1,2,...,K)$ at steady-state. Substituting the result into (19), we have

$$\frac{d\Delta x_i}{dt} = \frac{1}{\varepsilon_i C_i}\sum_{j=K+1}^{N} w_{ij}\Delta x_j \tag{19}$$

where the coefficient matrix w_{ij} is also symmetric and the diagonal elements are all zeros. Thus, from property 1, the eigenvalues contain both positive and negative reals, so that the system of equations (19) is always unstable.

Thus, we conclude that if more than 2 components of x_i are apart from the corner points of [0,1], then the solution will be unstable and cannot be the local minimum point.**Q.E.D.**

Remark: When R_i, i=1,2,...,N are sufficiently large, diagonal elements will become so small that the solution (u_i=0,i=1,2,...,N) will be unstable from property 1. Furthermore, from these two theorems, we conclude that when the sigmoid function is very stiff, only those solutions whose components are located around corner points of [0,1] could be local or global minimum point.

In our examples, we substitute x_i=0.999, or 0.001 for i=1,2,...,N into (4) for getting the approximate local minimum points. The exact global minimum point can be obtained by solving the algebraic equation

$$\sum_{j=1}^{N} w_{ij}x_j - \frac{a}{2R_i}\log\frac{x_i}{1-x_i} + I_i = 0 \tag{20}$$

with the Newton-Raphson method starting from the approximate lowest energy points.

3. A computational technique of the energy function based on a partitioning technique.

For large scale Hopfield networks, it is time-consuming to calculate the energy function (4). In this section, we show an efficient algorithm based on a partitioning technique. To understand the ideas of our partitioning technique, consider a simple example of partitioning the energy function into 2 functions. Suppose the energy function (4) is written by

$$\phi(x) = -\frac{1}{2}[x_1^T, x_2^T]\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + h_1(x_1)/R_1 + h_2(x_2)/R_2 - x_1^T I_1 - x_2^T I_2 \tag{21}$$

where

$$h_i(x_i)/R_i = \sum_{k=1}^{N_1} f(x_k)/R_k, \quad i=1,2$$

Partition it into two energy functions :

$$\phi_1(x_1,x_2) = -\frac{1}{2}x_1^T W_{11}x_1 - x_2^T W_{21}x_1 + f_1(x_1)/R_1 - x_1^T I_1 \tag{22.1}$$

$$\phi_2(x_1,x_2) = -\frac{1}{2}x_2^T W_{22}x_2 - x_1^T W_{12}x_2 + f_2(x_2)/R_2 - x_2^T I_2 \tag{22.2}$$

where $x_1 \in R^{N_1}$, $x_2 \in R^{N_2}$, $\phi_1(.,.):R^N \to R^{N_1}$, $\phi_2(.,.):R^N \to R^{N_2}$ Note that two energy functions of (22) are coupled with $x_2^T W_{21}x_1$ and $x_1^T W_{12}x_2$ in each other. In this case, the minimal points are stationary points of the following partitioned circuit equations :

$$C_1\frac{dx_1}{dt} = W_{11}x_1 + W_{12}x_2 - \frac{\partial f_1}{R_1 \partial x_1} + I_1 \tag{23.1}$$

$$C_2\frac{dx_2}{dt} = W_{22}x_2 + W_{21}x_1 - \frac{\partial f_2}{R_2 \partial x_2} + I_2 \tag{23.2}$$

which are corresponding to the subnetworks of Hopfield's circuit coupled with $W_{12}x_2$ and $W_{21}x_1$ in each other. Therefore, when the couplings are weak enough, the solutions obtained from each subnetwork by setting the couplings being zeros will give the good approximate solutions. Thus, consider the following energy functions:

$$\phi_1(x_1,0) = -\frac{1}{2}x_1^T W_{11}x_1 + f_1(x_1) - x_1^T I_1 \tag{24.1}$$

$$\phi_2(0,x_2) = -\frac{1}{2}x_2^T W_{22}x_2 + f_2(x_2) - x_2^T I_2 \tag{24.2}$$

Now, we have obtained the energy of (24), and list them in the order as follows:

For local minima of ϕ_1 :

$$\phi_{11} \geq \phi_{12} \geq \cdots\phi_{1H}\cdots \geq \phi_{1,K_1} \quad;K_1 = 2^{N_1} \tag{25.1}$$

For local minima of ϕ_2 :

$$\phi_{21} \geq \phi_{22} \geq \cdots\phi_{2H}\cdots \geq \phi_{2,K_2} \quad;K_2 = 2^{N_2} \tag{25.2}$$

Observe that if the couplings of subnetworks are weak enough, only some lower energy points in (25) will be chosen as candidates (P_1,P_2) of the approximate minimum points.

Now, we have a very important theorem about circuit partitioning technique.

Theorem 3: Let the maximum energies be ϕ_{1H} and ϕ_{2H} satisfying the following conditions, respectively:

$$\| W_{12}\| \geq \phi_{1H} - \phi_{1L}, \quad \| W_{21}\| \geq \phi_{2H} - \phi_{2L} \tag{26}$$

where $\phi_{1L}=\phi_{1,K_1}$ and $\phi_{2L}=\phi_{2,K_2}$ are the lowest energies of $\phi_1(x_1,0)$ and $\phi_2(0,x_2)$, and

$$\| W_{12}\| = \sum_{i=1}^{N_1} \sum_{j=N_1+1}^{N} | W_{ij}|, \quad \| W_{21}\| = \sum_{i=N_1+1}^{N} \sum_{j=1}^{N_1} | W_{ij}| \tag{27}$$

then, the points corresponding to $(\phi_{1,k}, \phi_{2,k}, k=H,...,L)$ in (25) will be chosen as the candidates. The approximate global minimum points are found by the combinations of these points and estimation of energy (4).

Proof: Let us rewrite the partitioned energy function (21) as follows:

$$\phi(x_1,x_2) = \phi_1(x_1,0) + \phi_2(0,x_2) - \frac{1}{2}x_1^T W_{12}x_2 - \frac{1}{2}x_2^T W_{21}x_1 \tag{28}$$

Note that the maximum and minimum values of the third and fourth terms in (28) are attained at $x_{1i}=1$ and $x_{2i}=1$ for all $i=1,...,N$. Put the minimum energy $\phi(x_{1H},x_{2H})$. Then, we have from (28)

$$\phi(x_{1H},x_{2H}) \leq \phi_{1L} + \frac{1}{2}\| W_{12}\| + \phi_{2L} + \frac{1}{2}\| W_{21}\| \tag{29}$$

and

$$\phi(x_{1H},x_{2H}) \geq \phi_1(x_{1H},0) - \frac{1}{2}\| W_{12}\| + \phi_2(0,x_{2H}) - \frac{1}{2}\| W_{21}\| \tag{30}$$

Hence, we have from (29) and (30)

$$\phi_{1L} + \phi_{2L} \geq \phi_1(x_{1H},0) + \phi_2(0,x_{2H}) - \| W_{12} \| - \| W_{21} \| \tag{31}$$

This inequality will be met for the energies satisfying condition (26). **Q.E.D.**

The partitioning technique can also efficiently be applied to weakly coupled circuits as cellular networks [3-4].

4. Illustrative examples

In this section, we show illustrative examples of a Hopfield network for solving a layout problem of printed board [10].

5.1 The first example is the circuit having 12 synapses, and the circuit equation is given by

$$\frac{du_i}{dt} = \sum_{j=1}^{6} w_{ij}x_j - u_i + I_i, \quad i=1,2,...,12 \tag{32}$$

where

$$W = \begin{bmatrix} 0 & -5 & -5 & -5 & -5 & -5 & -5 & -4 & -2 & -2 & -2 & -2 \\ -5 & 0 & -5 & -5 & -5 & -5 & -4 & -5 & -2 & -2 & -2 & -2 \\ -5 & -5 & 0 & -5 & -5 & -5 & -2 & -2 & -5 & -2 & -2 & -2 \\ -5 & -5 & -5 & 0 & -5 & -5 & -2 & -2 & -2 & -5 & -2 & -2 \\ -5 & -5 & -5 & -5 & 0 & -5 & -2 & -2 & -2 & -2 & -5 & -4 \\ -5 & -5 & -5 & -5 & -5 & 0 & -2 & -2 & -2 & -2 & -4 & -5 \\ -5 & -4 & -2 & -2 & -2 & -2 & 0 & -5 & -5 & -5 & -5 & -5 \\ -4 & -5 & -2 & -2 & -2 & -2 & -5 & 0 & -5 & -5 & -5 & -5 \\ -2 & -2 & -5 & -2 & -2 & -2 & -5 & -5 & 0 & -5 & -5 & -5 \\ -2 & -2 & -2 & -5 & -2 & -2 & -5 & -5 & -5 & 0 & -5 & -5 \\ -2 & -2 & -2 & -2 & -5 & -4 & -5 & -5 & -5 & -5 & 0 & -5 \\ -2 & -2 & -2 & -2 & -4 & -5 & -5 & -5 & -5 & -5 & -5 & 0 \end{bmatrix} \tag{33}$$

$$I = [21\ 21\ 20\ 20\ 21\ 21\ 21\ 21\ 20\ 20\ 21\ 21]^T \tag{34}$$

and the sigmoid functions are given by

$$u_i = \frac{a}{2}\log\frac{x_i}{1-x_i}, \quad i=1,2,...,12 \tag{35}$$

Then energy function is described by

$$\phi(x) = -\frac{1}{2}\sum_{i=1}^{12}\sum_{j=1}^{12} w_{ij}x_ix_j + \frac{a}{2}\sum_{i=1}^{12}[x_i\log x_i + (1-x_i)\log(1-x_i)] - \sum_{i=1}^{12} x_iI_i \tag{36}$$

For a=1.5, we found using a homotopy method that there are at least 335 stationary points, and they are traced from an initial guess

$$x_0 = (0.8\ 0.2\ 0.8\ 0.7\ 0.8\ 0.9\ 0.7\ 0.8\ 0.9\ 0.7\ 0.8\ 0.8)$$

and the step size was h=0.03. It takes about 5 minutes by using SUN(SPARC station IPC). We have also tried to get the approximate global minimum points according to section 3 by choosing x_i=0.001 and 0.999. We were rapidly obtained 4 global minimum

points,

and the exact solutions are obtained at the energy -76.13433 by the Newton-Raphson method which are the same from the homotopy method.

5.2 The second example of N=24 whose coefficient matrix is given by

$$W = \begin{bmatrix} W_{11} & 0 & 0 \\ & w_{12} & 0 \\ 0 & w_{12}^T & W_{11} \\ 0 & 0 & \end{bmatrix} \tag{37}$$

$$I = [I_2 \; I_2]^T \tag{38}$$

where the sub-matrices W_{11} is equal to (33), and

$$W_{12} = \begin{bmatrix} 0 & 1 & -2 & -2 & -2 & -2 \\ 1 & 0 & -2 & -2 & -2 & -2 \\ -2 & -2 & 0 & -2 & -2 & -2 \\ -2 & -2 & -2 & 0 & -2 & -2 \\ -2 & -2 & -2 & -2 & 0 & 1 \\ -2 & -2 & -2 & -2 & 1 & 0 \end{bmatrix} \tag{39}$$

Furthermore, the current components I_2 is equal to that of (34). In this example, we take 11276.0 sec by the direct method, and 67.2 sec by the partitioning technique given in section 3, where 4.7 sec has been used for getting local minimum points and 62.5 sec for the combinational calculation of energy function from the candidate points. The ratio is about 180. Thus, the partitioning technique can be quite efficiently applied to the sparse systems.

5.3 The algorithm can be also applied to a problem of the image cording [11].

5. Conclusions

We have proved very important theorems concerning the global minimum points of Hopfield networks containing stiff sigmoid characteristics. They are all located exclusively around corner points of x_i=[0,1]. The exact minimum points are found by the Newton method starting from the approximate minimum points.

For relatively large scale energy functions, the partitioning technique in section 3 can be efficiently applied to get the minimum point when the coupling strength are weak enough, satisfying a well defined condition (Theorem 3). In many practical problems such as circuit designs, the partitioning technique can also be usefully applied because of circuit equations being generally sparse.

optimization techniques.

References

[1] L.O. Chua, "Global optimization: A naive approach," *IEEE Trans. Circuits Syst.*, Vol.CAS-37, July, pp. 966-969, 1990.

[2] J.J. Hopfield and D.W. Tank, " Neural computation of decision in optimization problems," *Bio. Cybern.* vol.52, pp.141-152, 1985.

[3] J.J. Hopfield and D.W. Tank, "Computing with neural circuits: A model," *Science*, vol.233, pp.625-633, 1986.

[4] L.O. Chua and L.Yang, "Cellular neural networks: Theory", *IEEE Trans. Circuits Syst.* , Vol.CAS-35, No.10, pp.1257-1272, 1988.

[5] L.O. Chua and L.Yang, "Cellular neural networks: Applications", *IEEE Trans. Circuits Syst.* , Vol.CAS-35, No.10, pp.1273-1290, 1988. [6] A.Torn and A.Zilinskas, *Global Optimization*, Springer-Verlag, 1989.

[7] R. Bellman *Introduction to Matrix Analysis*, McGraw-Hill, 1970.

[8]J.J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons", *Proc. Natl. Acad. Sci.* Vol.81, pp.3088-3092, 1984.

[9]M.P.Kennedy and L.O. Chua,"Neural networks for nonlinear programming", *IEEE Trans. on Circuits and Systems*, Vol.35, No.5, pp.554-562, 1988.

[10] Private communication to prof. A. Sakamoto at Tokushima university.

[11]L.O. Chua and T. Lin, "A neural network approach to transform image cording",*Int. J. of Circuit theory and applications*, Vol. 16, pp.317-324, 1988.