2009 International Symposium on Nonlinear Theory and its Applications
NOLTA'09, Sapporo, Japan, October 18-21, 2009

The Quartic Form Energy Function for General Combinatorial Optimization Problems

Takahiro Sota†, Yoshihiro Hayakawa‡ and Koji Nakajima†

†Laboratory for Brainware/Laboratory for Nanoelectronics and Spintronics
Research Institute of Electrical Communication, Tohoku University
2-1-1 Katahira, Aoba-ku, Sendai-shi, 980-8577, Japan
‡Department of Electronic Engineering, Sendai National College of Technology
4-16-1 AyashiChuoh, Aoba-ku, Sendai-shi, 989-3128, Japan
Email: sota@nakajima.riec.tohoku.ac.jp, hayakawa@cc.sendai-ct.ac.jp, hello@riec.tohoku.ac.jp

Abstract—The Inverse function Delayed model (ID model) is a neuron model that has negative resistance dynamics. The negative resistance can destabilize local minimum states, which are undesirable network responses, so that the ID network can remove these states. Actually, we have demonstrated that the ID network can perfectly remove all local minima with N-Queen problems or 4-Color problems, where stationary states are only correct answer. Moreover, we have also applied the same method to the case of Traveling Salesman Problems (TSP) by expanding the energy function to the quartic form that means higher order synaptic connections. However, we need general energy functions to solve the other combinatorial optimization problems.

In this paper, we redefine the quartic form energy function to be able to apply not only TSPs but also Quadratic Assign Problems (QAP). After that we show that the ID network has only global minima, which are located on the vertices of the output space, as the stationary states, and the parameter region is discussed.

1. Introduction

Hopfield et al. proposed a neural network that has dynamics of moving along the gradient of the quadratic form energy function. This network can find the solution of combinatorial optimization problems (COPs) by assigning the optimal solution to the global minima of this energy function. However, the network state is often trapped in local minima. This is the local minimum problem.

As an improvement way of this problem, the method of using the Inverse function Delayed model (ID model) [1] has been proposed. One of important properties of ID model is a negative resistance, and the negative resistance region is controllable. Hence we can destabilize the local minimum states of the energy function by setting the appropriate negative resistance region. Actually, N-Queen problems or 4-Color problems are solved with 100% success rate by using the ID model[2]. However, in Traveling Salesman Problems (TSP) or Quadratic Assignment Problems (QAP), this method cannot always destabilize only the local minimum states. Hence, we have used the quartic form energy function [3] to the ID model[4] for the purpose of using the same method to apply with the TSP. In this case, only the global minimum states can reaches vertices of the hypercube of output space as a stationary state. Unfortunately, we cannot use this quartic form energy function to other problems.

In this paper, we aim to expand this quartic form energy function for the ID network to deal with not only TSPs but also other COPs. Moreover, by estimating the equilibrium points of the local minima, we aim to obtain the minimum parameter region analytically to be able to destabilize them perfectly. Finally, we confirm analytical result by dealing with the TSP and the QAP numerically.

2. Higher-order Connection ID model

2.1. Basic Equations

A Higher order Connection ID model (HC-ID model) is a ID model with higher order synaptic connections. To apply the quartic form energy function, we consider the HC-ID network with 3rd-order synaptic connections, described as follows[4]:

$$\tau_u \frac{du_i}{dt} = \sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} w_{ijkl}x_j x_k x_l + \sum_{j=1}^{N}\sum_{k=1}^{N} w_{ijk}x_j x_k + \sum_{j=1}^{N} w_{ij}x_j + h_i, \quad (1)$$

$$\tau_x \frac{dx_i}{dt} = u_i - g(x_i), \quad (2)$$

where N is the number of neurons, and u_i, x_i and h_i are the internal state, the output and the bias of neuron i, respectively. $w_{ijk\cdots}$ is the synaptic weight from neurons $j, k, \cdots$ to neuron i. τ_u and $\tau_x (\ll \tau_u)$ is the time constant of the internal state and the output, respectively. Moreover from Eqs.(1) and (2),

$$\tau_x \frac{d^2 x_i}{dt^2} + \eta(x_i)\frac{dx_i}{dt} = F_i \quad (3)$$

is derived, where

$$\eta(x_i) = \left.\frac{dg(x)}{dx}\right|_{x=x_i} + \frac{\tau_x}{\tau_u}, \quad (4)$$

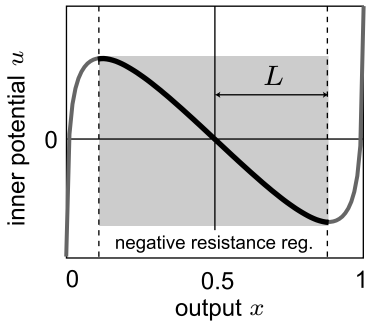

Figure 1: The g-function and negative resistance region.

$$F_i = -\frac{1}{\tau_u}\Big(\sum_j\sum_k\sum_l w_{ijkl}x_jx_kx_l + \sum_j\sum_k w_{ijk}x_jx_k + \sum_j w_{ij}x_j + h_i\Big). \tag{5}$$

Eq.(3) shows that the HC-ID model is updated by dynamics of particles, as well as the ID model. At this time, $\eta(x_i)$ means the friction, and F_i is force derived from the potential.

2.2. Negative Resistance Region and Energy Function

$g(x)$ in Eq.(4) is[2]

$$g(x) = \frac{1}{\beta}\ln\left(\frac{x}{1-x}\right) - \alpha\left(x - \frac{1}{2}\right), \tag{6}$$

where α and β is a control parameter of the range and the gain of the negative resistance, respectively. Because $g(x)$ has an N-shape as shown in Fig. 1, $\eta(x_i)$ has a negative value. This region is called the negative resistance region, and the length of this region L is defined as

$$L = \sqrt{\frac{1}{4} - \frac{1}{\beta(\alpha - \tau_x/\tau_u)}}. \tag{7}$$

Next, when the synaptic weights are symmetric ($w_{ijkl\cdots} = w_{jikl\cdots} = w_{jilk\cdots} = \cdots$), the energy function of HC-ID model can be defined.

$$E_{\text{HC-ID}} = -\frac{1}{4\tau_u}\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N w_{ijkl}x_ix_jx_kx_l - \frac{1}{3\tau_u}\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N w_{ijk}x_ix_jx_k - \frac{1}{2\tau_u}\sum_{i=1}^N\sum_{j=1}^N w_{ij}x_ix_j - \frac{1}{\tau_u}\sum_{i=1}^N h_ix_i + \frac{\tau_x}{2}\sum_{i=1}^N\left(\frac{dx_i}{dt}\right)^2 \tag{8}$$

From Eqs. (1) and (2), its time derivative is

$$\frac{dE_{\text{HC-ID}}}{dt} = -\sum_i \eta(x_i)\left(\frac{dx_i}{dt}\right)^2. \tag{9}$$

Hence if the network state is in the negative resistance region, $E_{\text{HC-ID}}$ increases with time like the normal connection ID model. Moreover, from Eqs.(5) and (8), the following equation is derived:

$$F_i = -\frac{1}{\tau_u}\cdot\frac{\partial E_{\text{HC-ID}}}{\partial x_i}. \tag{10}$$

Therefore if $E_{\text{HC-ID}}/\partial x_i < 0$, the output x_i receives the force derived from the potential toward 1, and if $E_{\text{HC-ID}}/\partial x_i > 0$, x_i receives the force toward 0.

3. Design of Quartic Form Energy Function for COP

3.1. Quartic Form Energy Function for COP

We improve the quartic form energy function to apply it for general COPs:

$$E_{4\text{TH}} = \frac{A}{2}\sum_{i=1}^n\left(\sum_{x=1}^n x_{xi} - 1\right)^2 + \frac{A}{2}\sum_{x=1}^n\left(\sum_{i=1}^n x_{xi} - 1\right)^2 + \frac{B}{2}\sum_{x=1}^n\sum_{i=1}^n\sum_{y=1}^n\sum_{j=1}^n b_{xi,yj}x_{xi}x_{yj}(1 - x_{xi}x_{yj}) + \frac{C}{2}\left(\sum_{x=1}^n\sum_{i=1}^n\sum_{y=1}^n\sum_{j=1}^n b_{xi,yj}x_{xi}x_{yj}\right)^2, \tag{11}$$

where A,B and C are positive coefficients, n is the problem size of COP ($n^2 = N$), and $b_{xi,yj} \in \boldsymbol{b}^{[N\times N]}$ is a cost value when the neuron (x, i) and (y, j) fire. The first and second terms of Eq.(11) have minimum value if only one neuron fires in each row and column, and the third term is minimized when output values have 0 or 1. These terms mean constrained condition, and it will be 0 when the condition is satisfied. Under this condition, the fourth term expresses a squared value of cost shown by the network state. Therefore if we can denote a COP by N-dimension vector of neuron and detect the cost matrix $\boldsymbol{b}$, we can apply the quartic form energy function to this COP, like the case of the quadratic form function in the reference [5].

3.2. Coefficients Condition to Stabilize Only Global Minimum States

From Eq.(11), the gradient of $E_{4\text{TH}}$ is calculated as

$$\frac{\partial E_{4\text{TH}}}{\partial x_{ah}} = A\Big(\sum_x x_{xh} - 1\Big) + A\Big(\sum_i x_{ai} - 1\Big) + 2B\sum_y\sum_j b_{ah,yj}x_{yj}\Big(\frac{1}{2} - x_{ah}x_{yj}\Big) + 2C\sum_z\sum_k\sum_w\sum_l b_{zk,wl}x_{zk}x_{wl} \times \sum_y\sum_j b_{ah,yj}x_{yj}. \tag{12}$$

First, to activate only one neuron in each row and column, coefficient A has to satisfy the condition:

$$A > \max\left[B \cdot Mb_{\max}\,,\; M/2 \cdot b_{\max}\left\{B + 2C\left\{M(N-1) - 1\right\}b_{\max}\right\}\right], \quad (13)$$

where M is maximum number of connection between a neuron and others without the same row and column neurons. Also, $b_{\max}$ is maximum term of cost matrix $\boldsymbol{b}$.

Next, we assume that the negative resistance region is enough large to cover intermediate value of the output. Under this assumption, all output values are 0 or 1, so that $x_i^2 = x_i$ holds. Hence, Eq.(12) is described as

$$\begin{aligned}\frac{\partial E_{4\mathrm{TH}}}{\partial x_{ah}} &= \sum_y \sum_j b_{ah,yj} x_{yj} \times \\ &\quad \left\{2B\left(\frac{1}{2} - x_{ah}\right) + 2C \sum_z \sum_k \sum_w \sum_l b_{zk,wl} x_{zk} x_{wl}\right\} \\ &= \sum_y b_{ah,yJ(y)} \left\{2B\left(\frac{1}{2} - x_{ah}\right) + 2C \sum_z \sum_w b_{zJ(z),wJ(w)}\right\} \\ &= \begin{cases} c_{ah}(\boldsymbol{x}) \cdot 4C\left\{B/4C + c_{\mathrm{sol}}(\boldsymbol{x})\right\} & (x_{ah} = 0) \\ c_{ah}(\boldsymbol{x}) \cdot 4C\left\{-B/4C + c_{\mathrm{sol}}(\boldsymbol{x})\right\} & (x_{ah} = 1) \end{cases}. \quad (14)\end{aligned}$$

where $J(y)$ means row index of firing neuron in column y, and

$$c_{ah}(\boldsymbol{x}) = \sum_y b_{ah,yJ(y)}, \quad (15)$$

$$c_{\mathrm{sol}}(\boldsymbol{x}) = \frac{1}{2} \sum_z \sum_w b_{zJ(z),wJ(w)}. \quad (16)$$

From Eq.(15), $c_{\mathrm{sol}}(\boldsymbol{x})$ means the cost value of the network state $\boldsymbol{x}$. At this time, from Eq.(14), $\partial E_{4\mathrm{TH}}/\partial x_{ah} > 0$ is always satisfied when $x_{ah} \sim 0$. Hence x_{ah} receives force toward 0 and this state is stable. On the other hand, when $x_{ah} \sim 1$, the sign of $\partial E_{4\mathrm{TH}}/\partial x_{ah}$ depends on magnitude relation between $B/4C$ and $c_{\mathrm{sol}}(\boldsymbol{x})$. If $B/4C > c_{\mathrm{sol}}(\boldsymbol{x})$, the output x_{ah} receives force toward 1, and this state is stable. However, if $B/4C < c_{\mathrm{sol}}(\boldsymbol{x})$, this state becomes unstable. Therefore, if the following equation is satisfied, the network can stabilize only global minimum states:

$$c_{\mathrm{sol}}(\boldsymbol{x}_0) < B/4C < c_{\mathrm{sol}}(\boldsymbol{x}_1), \quad (17)$$

where $c_{\mathrm{sol}}(\boldsymbol{x}_0)$ and $c_{\mathrm{sol}}(\boldsymbol{x}_1)$ are the cost value of optimal solution and the 2nd optimal solution, respectively.

3.3. Estimate the Equilibrium Points of Local Minimum States

If we assume that the negative resistance region covers the output space except the vertices, the outputs can be only 0 or 1. Hence the local minimum states always destabilized. However, because the parameter β in Eq. (6) has finite value, the regions where the network state can be stable exist apart from the vertices. Hence we estimate the furthest equilibrium point of local minimum states from the vertices to cover all of the local minimum states with the negative resistance region.

To estimate that point, we define ξ_{xi} as the distance from 0 or 1 to output value of the neuron (x, i). Moreover, we define firing neuron $(a, J(a))$ as the furthest neuron from 1, and $\xi_{\max} = \xi_{aJ(a)}$ as the distance from 1 to output value of the neuron $(a, J(a))$. Under these conditions, $\partial E_{4\mathrm{TH}}/\partial x_{aJ(a)}$ is calculated as

$$\begin{aligned}\frac{\partial E_{4\mathrm{TH}}}{\partial x_{aJ(a)}} &\approx A\left\{-2\xi_{\max} + \left(\sum_{x \neq a} \xi_{xJ(a)} + \sum_{i \neq J(a)} \xi_{ai}\right)\right\} \\ &+ B \sum_{y \neq a}\left\{b_{aJ(a),yJ(y)}(-1 + 3\xi_{yJ(y)} + 2\xi_{\max}) + \sum_{j \neq J(y)} b_{aJ(a),yj}\xi_{yj}\right\} \\ &+ 2C \sum_{y \neq a} b_{aJ(a),yJ(y)} \\ &\times \left[\sum_{z \neq a}\left\{2b_{aJ(a),zJ(z)}(1 - \xi_{\max} - \xi_{zJ(z)}) + 2\sum_{k \neq J(z)} b_{aJ(a),zk}\xi_{zk}\right\}\right. \\ &\quad \left. + \sum_{z \neq a}\sum_{w \neq a}\left\{b_{zJ(z),wJ(w)}(1 - 2\xi_{zJ(z)}) + 2\sum_{l \neq J(w)} b_{zJ(z),wl}\xi_{wl}\right\}\right] \\ &+ 2C \sum_{y \neq a}\left[-b_{aJ(a),yJ(y)}\xi_{yJ(y)} + \sum_{j \neq J(y)} b_{aJ(a),yj}\xi_{yj}\right] \\ &\quad \times \sum_{z \neq a}\left[2b_{aJ(a),zJ(z)} + \sum_{w \neq a} b_{zJ(z),wJ(w)}\right]. \quad (18)\end{aligned}$$

In this calculation, we consider that $b_{xi,xi} = 0$ and $b_{xi,yj} = b_{yj,xi}$. Also, we consider a linear approximation about ξ because ξ is negligible small. In Eq.(18), when the ξ of positive terms are 0 and the ξ of negative terms are $\xi_{\max}$, the right-side of Eq. (18) has minimum value. Thus, following inequality is satisfied at least:

$$\begin{aligned}\frac{\partial E_{4\mathrm{TH}}}{\partial x_{aJ(a)}} &> -2A\xi_{\max} - B\sum_{y \neq a} b_{aJ(a),yJ(y)}(1 - 2\xi_{\max}) \\ &\quad + 2C\sum_{y \neq a} b_{aJ(a),yJ(y)}\left[(1 - 3\xi_{\max})\left\{\sum_z \sum_w b_{zJ(z),wJ(w)}\right\}\right] \\ &= c_{aJ(a)}(\boldsymbol{x})\{-B + 4Cc_{\mathrm{sol}}(\boldsymbol{x})\} \\ &\quad - \xi_{\max}[2A + c_{aJ(a)}(\boldsymbol{x})\{-2B + 12Cc_{\mathrm{sol}}(\boldsymbol{x})\}]. \quad (19)\end{aligned}$$

Because the network state is the local minimum states now($\boldsymbol{x} = \boldsymbol{x}_1$), output x_{ah} has to be received the force toward 0. Hence, from Eq.(10), $\partial E_{4\mathrm{TH}}/\partial x_{aJ(a)}$ must have positive value. From this constraint, $\xi_{\max}$ has to hold as:

$$\xi_{\max} < \frac{c_{aJ(a)}(\boldsymbol{x}_1)\{-B + 4Cc_{\mathrm{sol}}(\boldsymbol{x}_1)\}}{2A + c_{aJ(a)}(\boldsymbol{x}_1)\{-2B + 12Cc_{\mathrm{sol}}(\boldsymbol{x}_1)\}}. \quad (20)$$

If the network state is a local minimum state, at least the distance of firing neuron from 1 to the output value is larger than maximum value of $\xi_{\max}$. Therefore, the length of the negative resistance region has to be larger than $(1 - \max[\xi_{\max}])$ to destabilize the local minimum states. Moreover, from Eq.(20), we can estimate $(\max[\xi_{\max}])$ value independently of the problem size.

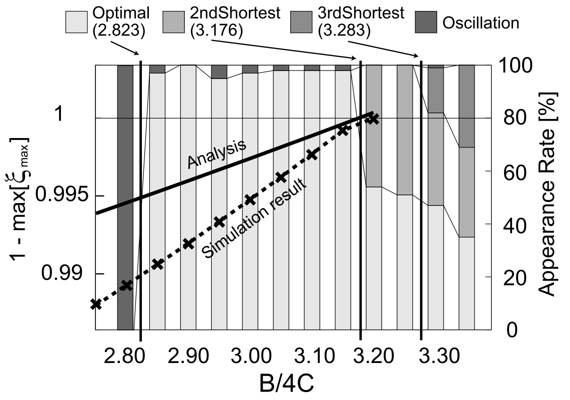

Figure 2: The result of the 4-TSP ($\beta = 200, C = 0.1, A = 8.05$).

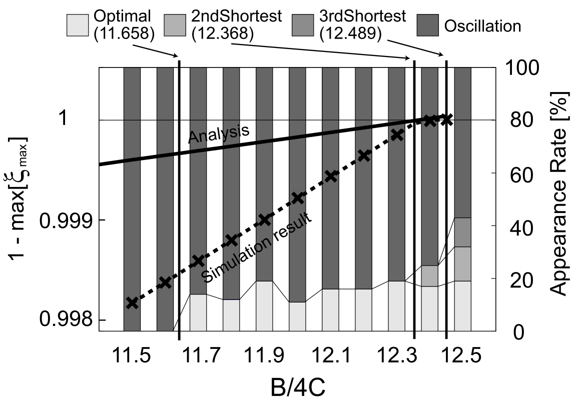

Figure 3: The result of the 4-QAP ($\beta = 2000, C = 0.1, A = 386.38$).

4. Simulation Results

In this section, we apply the proposed energy function to a 4city-TSP and a 4size-QAP, and confirm the analytical result of previous sections by using computer simulations. In all simulations, we tested by 100 trials with random initial values.

4.1. Equilibrium Points of Local Minimum States

First, we investigated about the furthest equilibrium point of local minimum states from the vertices. To investigate the equilibrium point, we set no negative resistance region by setting $\alpha = 0$. Moreover, we fixed the coefficients A and C, and changed B value. Figs. 2 and 3 show the simulation results of the 4-TSP and the 4-QAP, respectively, and the horizontal axis shows the value of $B/4C$. Moreover, the solid line and the dot line mean the calculated value from Eq.(20) and the simulated value of $(1 - \max[\xi_{\max}])$, respectively.

From the lines of figures, about maximum value of $\xi_{\max}$, the simulated values are always larger than analytical values. Therefore, we can confirm that there are no equilibrium points of the local minimum states in the region decided by the analytical result.

4.2. Coefficients Dependency of Network States

Next, we investigated the relationship between the appearance rate of the network states and the coefficients value of $B/4C$. In this test, the length of the negative resistance region was set in accordance with the analytical result of $\xi_{\max}$, and the histogram of figures show the appearance rate of the network states.

In both cases of the TSP and the QAP, the states whose cost is larger than $B/4C$ have never appeared as stationary state. Therefore we can obtain only optimal solutions when the network state reaches stationary state if the condition Eq.(17) is satisfied.

5. Conclusion

In this paper, we proposed the improved quartic form energy function to deal with the general COP. In the case of using this improved energy function, we can show that only the global minimum states are stationary states in the vertices of the hypercube of output space if Eq.(17) is satisfied. Moreover, we estimated the equilibrium points of the local minimum states, and found regions where the equilibrium points of the local minimum states did not exist. Therefore if we cover all of the output space except these regions with the negative resistance region, we can destabilize the local minimum states perfectly. Finally, we confirmed the analytical result by dealing with the TSP and the QAP.

References

[1] K. Nakajima and Y. Hayakawa, "Characteristics of inverse function delayed model for neural computation", *Proc. NOLTA'02*, pp. 861–864, 2002.

[2] A. Sato, Y. Hayakawa and K. Nakajima, "The parameter dependence of the inverse function delayed model on the success rate of combinatorial optimization problems", *IEICE Trans. (JAPANESE EDITION)*, vol. J89-A, no. 11, pp. 960–972, 2006.

[3] S. Matsuda, ""optimal" neural representation of higher order for traveling salesman problems", *Electronics and Communications in Japan, Part2*, vol. 85, no. 9, pp. 32–42, 2002.

[4] T.Sota, Y.Hayakawa and K.Nakajima, "ID model with higher-order connections for the traveling salesman problem", *Proc. NOLTA'08*, pp. 548–551, 2008.

[5] G.Feng and C.Douligeris, "On the convergence and parameter relation of discrete-time continuous-state hopfield networks with self-interaction neruons", *IEICE Trans.*, vol. E84-A, no. 12, pp. 3162–3173, 2001.