Chaos in the Discretized Analog Hopfield Neural Network and Potential Applications to Optimization

Lipo Wang

School of Computing and Mathematics
Deakin University, 662 Blackburn Road
Clayton, Victoria 3168, Australia
lwang@deakin.edu.au
http://www.cm.deakin.edu.au/~lwang

Kate Smith

Department of Business Systems
Monash University
Clayton, Victoria 3168, Australia
ksmith@bs.monash.edu.au
http://www.bs.monash.edu.au/staff/ks.htm

Abstract - We consider the discretization of the analog Hopfield neural network (DAHNN) using Euler approximation. We suggest an alternative approach to chaotic simulated annealing using the discretizing time-step Δt as the bifurcation parameter, because the DAHNN is chaotic when the time-step Δt is chosen to be sufficiently large and stabilization is guaranteed when the time-step Δt is small enough. It is not necessary to carefully choose other system parameters to assure minimization of Hopfield energy function and network convergence. We argue that this approach should find significant applications in solving combinatorial optimization problems with neural networks.

1. Introduction

It is practically impossible to find *the* optimal solution in many combinatorial optimization problems due to the immense computation involved. It is therefore often desirable to find near-optimal solutions to these problems through some efficient algorithms. Since Hopfield and Tank [10] applied a Hopfield neural network (HNN) [9] to solve the travelling salesman problem, the HNNs [8], [9] have been recognized as powerful tools for optimization (e.g., [25], [28], [22]- [24]).

Hopfield showed [8], [9] that the HNNs have energy functions that strictly decrease whenever a neuron changes its state. Once the cost function in an optimization problem is cast onto the energy function of a HNN, the network is allowed to settle down to a state corresponding to an optimization solution with a low cost. However, a HNN can often settle down at a local minimum with a relatively high energy value and thus is not able to suggest a good optimization solution [35]. Many authors have used various mechanisms to *temporarily* increase the energy, such as *simulated annealing* (e.g. [14], [11]) and hill-climbing (e.g. [22]), thereby allowing the network to search in a larger state space while forcing the network to settle down at the end of certain period of simulation time.

Nozawa [19], [20], [21] showed that by varying the *input bias* for each neuron in the discretized analog HNN (DAHNN) obtained from the Euler approximation, chaos can be generated and used to improve optimization solutions. The searching ability of chaos was also demonstrated by Nara *et al* [18] in memory searching.

Chaotic simulated annealing was recently proposed by Chen and Aihara [5]. They added a *negative self-coupling* to a "chaotic neural network" model proposed earlier by Aihara *et al* [3] based on a simplification of the FitzHugh-Nagumo model [6], [16], [17] and showed that chaos occurs. They then gradually removed this negative self-coupling and used the transient chaos generated by the decreasing negative self-coupling to search the solution space of optimization problems. Compared to other neural network methods, the Chen-Aihara approach significantly increased the probability of finding near-optimal solutions. Nozawa [19], [20], [21] and Chen and Aihara [5] showed that when network parameters satisfy appropriate conditions, the DAHNN becomes very similar to the "chaotic neural network" proposed by Aihara *et al* [3]; however, the "chaotic neural network" of Aihara *et al* [3] has more parameters than the DAHNN.

We will first analyze the dynamics of the DAHNN as a function of the discretization time-step Δt. We will then show why the analysis can be potentially useful for solving optimization problems with neural networks.

0-7803-4859-1/98 $10.00©1998 IEEE

2. Chaos in DAHNN and Optimization

The Chen-Aihara system with transient chaos [5] is a special case of the "chaotic neural network" proposed by Aihara *et al* [3] and is described by, in slightly different notation,

$$u_i(t+1) = k\, u_i(t) + \alpha\, (\textstyle\sum_{j\neq i} T_{ij}V_j + I_i) - z_i(t)\, [\, V_i(t) - I_o \,] \quad (1)$$

where $i = 1, 2, ..., n$, I_i is the external input or input bias to neuron i. The internal state of neuron i, $u_i \in (-\infty, +\infty)$, determines the output of neuron i:

$$V_i = f(u_i) \in [0, 1]\,, \quad (2)$$

or

$$u_i = f^{-1}(V_i) \quad (3)$$

Here f is the neuronal input-output response function and is assumed to be strictly increasing so that f^{-1} exists, such as the sigmoid function,

$$f(x) = \frac{1}{1 + e^{-x/\epsilon}} = \frac{1}{2}\, [\, 1 + \tanh(\lambda x)\,] \quad (4)$$

and $\lambda = 1/(2\epsilon)$ is the gain of the sigmoid function. In eq. 1,

$$z_i(t) \geq 0 \quad (5)$$

is the added *negative self-coupling*,

$$T_{ij} = T_{ji} \quad (6)$$

is the strength of the *symmetric* mutual interactions (or weight matrix) between neurons i and j. In an optimization problem, the weight matrix T_{ij} and the input bias I_i are obtained from the energy function (or cost function) E as follows:

$$\sum_{j\neq i} T_{ij}V_j + I_i = -\frac{\partial E}{\partial V_i} \quad (7)$$

In addition,

$$I_o > 0 \quad (8)$$

is a positive parameter in the Chen-Aihara system.

The analog Hopfield neural network (AHNN) [9] is on the other hand described by

$$\frac{du_i}{dt} = -\frac{u_i}{\tau} + \sum_j T_{ij}V_j + I_i \quad (9)$$

Hopfield [9] showed that the AHNN minimizes the following energy function:

$$E = -\frac{1}{2}\sum_{i,j} T_{ij}V_iV_j - \sum_i I_iV_i + \frac{1}{\tau}\sum_i \int_0^{V_i} f^{-1}(\xi)d\xi \quad (10)$$

and

$$\frac{dE}{dt} = \sum_i \left(\frac{\partial E}{\partial V_i}\right)\left(\frac{dV_i}{dt}\right) = -\frac{du_i}{dt}\cdot f' \cdot \frac{du_i}{dt} < 0 \quad (11)$$

provided that eq. 6 holds.

The discretization of the AHNN (DAHNN) is given by the Euler approximation of eq. 9

$$u_i(t+\Delta t) = u_i(t) + \frac{du_i}{dt}\Delta t$$
$$= u_i(t) - \frac{u_i\Delta t}{\tau} + \Delta t \sum_j T_{ij}V_j + \Delta t I_i$$
$$= (1 - \frac{\Delta t}{\tau})u_i + \Delta t(\sum_j T_{ij}V_j + I_i) \quad (12)$$

By comparing eq. 1 and eq. 12, one sees that Chen-Aihara system with transient chaos [5] is identical to the DAHNN if [19], [20], [21]

$$k = 1 - \frac{\Delta t}{\tau} \quad (13)$$

$$\alpha = \Delta t \quad (14)$$

$$z_i = -\Delta t\, T_{ii} \quad (15)$$

$$\alpha I_i + z_i I_o = I_i \quad (16)$$

Chen and Aihara [5] showed that for large positive z_i values, the system given by eq. 1 is chaotic, whereas for $z_i = 0$ the system becomes stable. Transient chaos can therefore be generated by starting with a large z_i value and letting it decay with time. However, Chen-Aihara system is stable at $z_i = 0$ only when other system parameters take some carefully-selected values.

We observe that in the limit $\Delta t \to 0$, the DAHNN eq. 12 is the same as the AHNN eq. 9. Thus the DAHNN becomes stable and minimizes the Hopfield energy function as $\Delta t \to 0$. In addition, we will show in the rest of this section that chaos exists when the time step Δt is large enough. Therefore it would advantageous to use the following new approach to chaotic simulated annealing: starting from the DAHNN eq. 12 with a large time-step Δt, we gradually reduce the time-step Δt. The network would search a large portion of the state space with chaos and then settles down at a minimum of the right energy function. The advantage of this approach is that there is no need to carefully choose other system parameters so that the Hopfield energy function is minimized and the system eventually settles down to a stable state, thereby indicating a solution to an optimization problem.

Eq. 12 is equivalent to

$$\begin{aligned} u_i(t+\Delta t) &= [(1-\tfrac{\Delta t}{\tau})u_i + \Delta t T_{ii} f(u_i)] \\ &\quad +\Delta t\, [\textstyle\sum_{j\neq i} T_{ij} f(u_j) + I_i] \\ &\equiv G_i(u_i) \end{aligned} \quad (17)$$

If we let

$$t = \tau t' \quad (18)$$

$$\Delta t = \tau \Delta t' \quad (19)$$

$$T_{ij} = \frac{1}{\tau} T'_{ij} \quad (20)$$

$$I_i = \frac{1}{\tau} I'_i \quad (21)$$

Eq. 17 becomes

$$\begin{aligned} u_i(t'+\Delta t') &= [\,(1-\Delta t')\, u_i + \Delta t'\, T'_{ii}\, f(u_i)\,] \\ &\quad + \Delta t'\, [\,\textstyle\sum_{j\neq i} T'_{ij}\, f(u_j) + I'_i\,] \end{aligned} \quad (22)$$

Eq. 22 is identical to eq. 17 with $\tau' = 1$. We thereafter let $\tau = 1$ in eq. 17 without losing generality.

When there are no mutual interactions in the network

$$T_{ij} = 0 \quad \text{, for all} j \neq i \quad (23)$$

eq. 17 can be decoupled:

$$\begin{aligned} u_i(t+\Delta t) &= (1-\Delta t)u_i + \Delta t[T_{ii} f(u_i) + I_i] \\ &\equiv F_i^o(u_i) \end{aligned} \quad (24)$$

Alternatively, we can rewrite the above equation in terms of a new variable

$$s_i = 2u_i - 1 \quad (25)$$

instead of u_i

$$s_i(t+\Delta t) = (\,1 - \Delta t)\, s_i + \Delta t\, T_{ii}\, g(s_i) \quad (26)$$

where we have used

$$I_i = \frac{1}{2}\,(1 - T_{ii}) \quad (27)$$

and

$$\begin{aligned} g(s_i) &= 2f(u_i) - 1 \\ &= \tanh(\lambda u_i) \in [-1, 1] \end{aligned} \quad (28)$$

as opposed to $f(u_i) \in [0, 1]$.

We first analyze the fixed point of eq. 26, i.e.,

$$s_i^* = (\,1 - \Delta t\,)\, s_i^* + \Delta t\, T_{ii}\, g(s_i^*) \equiv G_i^o(s_i^*) \quad (29)$$

Or

$$g(s_i^*) = \frac{s_i^*}{T_{ii}} \quad (30)$$

To assure there exist non-zero fixed points, i.e., there are non-zero intersections between the sigmoid function at the left hand side of eq. 30 and the straight line function at the right hand side of eq. 30, we need

$$g'(0) = \lambda > \frac{1}{T_{ii}} \quad \text{and} \quad T_{ii} > 0 \quad (31)$$

Now let us exam the slope at the non-zero fixed point:

$$G_i^{o'}(s_i^*) = (\,1 - \Delta t\,) + \Delta t\, T_{ii}\, g'(s_i^*) \quad (32)$$

Oscillations and chaos exist if $G_i^{o'}(s_i^*) < -1$, or

$$\Delta t > \frac{2}{1 - T_{ii}\, g'(s_i^*)} \quad (33)$$

Figs.1-5 show the dynamic behaviors of the DAHNN as a function of the discretization time-step Δt, with $\lambda = 2$ and $T_{ii} = 0.7$. The network is chaotic at large time-steps Δt and becomes stable for $\Delta t < 3.78$. The mutual interactions between neurons are expected to increase the complexity of the dynamics. Hence oscillations and chaos exist when the time-step Δt is large enough.

3. Conclusions

In summary, we have suggested an alternative approach to chaotic simulated annealing based on the facts that the DAHNN is chaotic when the discretizing time-step Δt is large enough and, in addition, stabilization and energy reduction are achieved when the time-step Δt is small enough, because the system becomes identical to the HNN at zero time-step limit. Hence by starting with a large time-step in the DAHNN and reducing the time-step to zero, we can achieve chaotic simulated annealing with guaranteed minimization of the Hopfield energy and network stabilization, without the need to select any other system parameters. We have argued that this approach should find significant applications in solving combinatorial optimization problems with neural networks. We are currently applying it to various optimization problems.

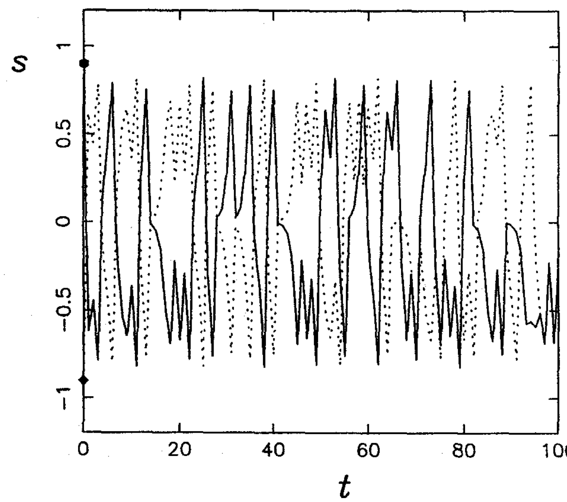

Figure 1: Chaos for large time step $\Delta t = 6.5$ in the discretized analog Hopfield network, according to eq. 26. Initial conditions: $s(0) = 0.9$ (hexagon, solid line) and $s(0) = -0.9$ (diamond, dashed line). The chaotic activities for each trajectory cover *both* positive and negative s.

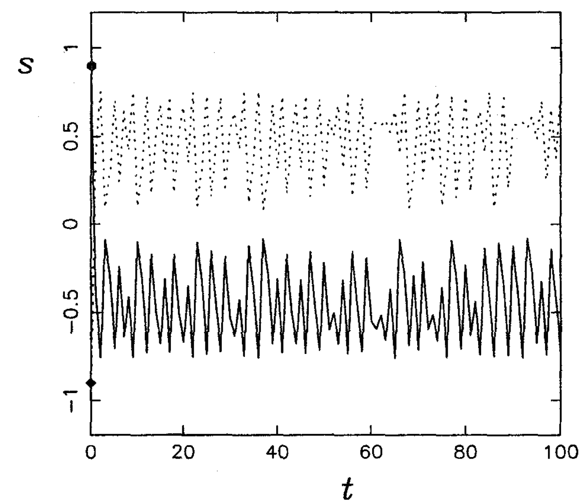

Figure 2: Chaos for large time step $\Delta t = 5.5$ in the discretized analog Hopfield network, according to eq. 26. Initial conditions: $s(0) = 0.9$ (hexagon, solid line) and $s(0) = -0.9$ (diamond, dashed line). The chaotic activities for the trajectory with *positive initial condition* (hexagon, solid line) cover *negative s only*, whereas the chaotic activities for the trajectory with *negative initial condition* cover *positive s only.*

4. Acknowledgement

This work was supported in part by the Australian Research Council and Deakin University.

References

[1] K. Aihara, "Spatio-temporal dynamics and possible computation in chaotic neural networks," *ICNN95*, pp. I-262, 1995.

[2] K. Aihara and G. Matsumoto, "Chaotic oscillations and bifurcations in squid giant axons," in *Chaos*, A.V. Holden, ed., Manchester Univ. Press, Princeton, 1986.

[3] K. Aihara, T. Takabe and M. Toyada, "Chaotic neural networks," *Phys. Letters A*, vol. 144, nos. 6,7, pp. 333-340, March, 1990.

[4] L. Chen and K. Aihara, "Chaotic simulated annealing for combinatorial optimization", In N. Aoki etal (ed.) *Dynamical Systems and Chaos*, vol. 1, pp. 319, Singapore: World Scientific. 1994.

[5] L. Chen and K. Aihara, "Chaotic simulated annealing by a neural network model with transient chaos," *Neural Networks*, vol. 8, no. 6, pp. 915-930, 1995.

[6] R. FitzHugh, in *Biological Engineering*, H.P. Schwan, (ed.) McGraw-Hill, New York, 1969, p.1.

[7] F. Fogelman-Soulié, C. Mejia, Eric Goles, and S. Martinez, "Energy function in neural networks with continuous local functions," *Complex Systems*, vol. 3, pp. 269-293, 1989.

[8] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. USA*, vol. 79, pp. 2554-2558, April, 1982.

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

[10] J.J. Hopfield and D.W. Tank, "Neural computation of decisions in optimisation problems," *Biol. Cybern.*, vol. 52, pp. 141-152, 1985.

[11] L. Ingber, "Simulated annealing: practice versus theory," *Math. Comput. Modelling,* vol. 18, pp. 29-57, 1993.

[12] B. Kamgar-Parsi and B. Kamgar-Parsi, "An efficient model of neural networks for optimization,"

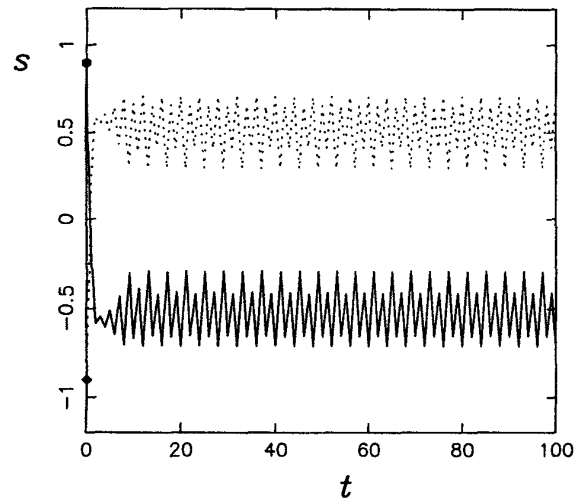

Figure 3: Period-4 oscillations for medium time step $\Delta t = 4.8$ in the discretized analog Hopfield network, according to eq. 26. Initial conditions: $s(0) = 0.9$ (hexagon, solid line) and $s(0) = -0.9$ (diamond, dashed line). The oscillations with *positive initial condition* (hexagon, solid line) cover *negative s only*, whereas the oscillations with *negative initial condition* cover *positive s only.*

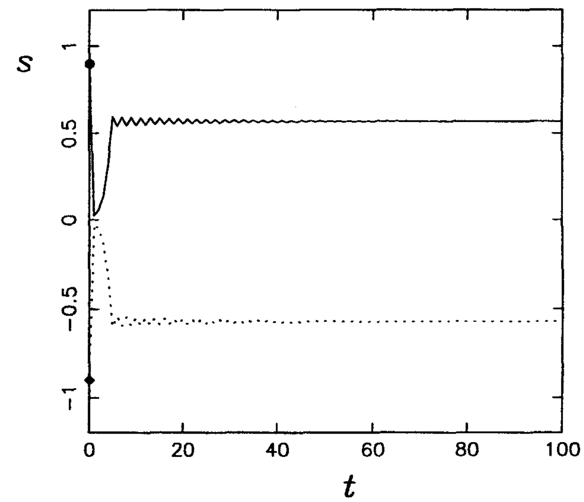

Figure 4: Stable fixed points for small time step $\Delta t = 3.7$ in the discretized analog Hopfield network, according to eq. 26. Initial conditions: $s(0) = 0.9$ (hexagon, solid line) and $s(0) = -0.9$ (diamond, dashed line). The trajectory with *positive initial condition* (hexagon, solid line) cover *positive s only*, whereas the trajectory with *negative initial condition* cover *negative s only.* There are slight overshotting and decaying oscillations at early times.

Proc. IEEE Intl. Conf. Neural Networks, vol. III, pp. 785-790, 1987.

[13] B. Kamgar-Parsi and B. Kamgar-Parsi, "Hopfield model and optimization problems, in Neural Networks for Perception," *Computation, Learning, and Architectures,* vol. 2, Academic Press, 1992, pp. 94

[14] S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671-680, 1983.

[15] P. Koiran, "Dynamics of discrete time, continuous state Hopfield networks," *Neural Computation*, vol. 6, pp. 459-468, 1994.

[16] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proc. IRE*, vol. 50, pp. 2061-2070, 1962.

[17] J. Nagumo and S. Sato, *Kybernetik*, vol. 10, pp. 155, 1972.

[18] S. Nara, P. Davis, and H. Totsuji, "Memory search using complex dynamics in a recurrent neural network model," *Neural Networks*, vol. 6, pp. 963-973, 1993.

[19] H. Nozawa, "A neural network model as a globally coupled map and applications based on chaos," *Chaos*, vol.2, no.3, p. 377-86, 1992.

[20] H. Nozawa, "Solution of the optimization problem using the neural network model as a globally coupled map," *Physica D*, vol.75, no.1-3, p. 179-189, 1994.

[21] H. Nozawa, "Solution of the optimization problem using the neural network model as a globally coupled map," in *Towars the Harnessing of Chaos,* M. Yamaguti (ed), pp. 99-114, 1994.

[22] K. Smith and M. Palaniswami, "Static and dynamic channel assignment using neural networks," *IEEE J. Selected Areas Commun.*, vol. 15, no. 2, pp.238-249, 1997.

[23] K. Smith, M. Palaniswami, and M. Krishnamoorthy, "Traditional heuristic versus Hopfield neural network approaches to a car sequencing problem," *European J. Op. Research*, vol. 93, pp. 300-316, 1996.

[24] K. Smith, M. Palaniswami, and M. Krishnamoorthy, "A hybrid neural approach to combinatorial

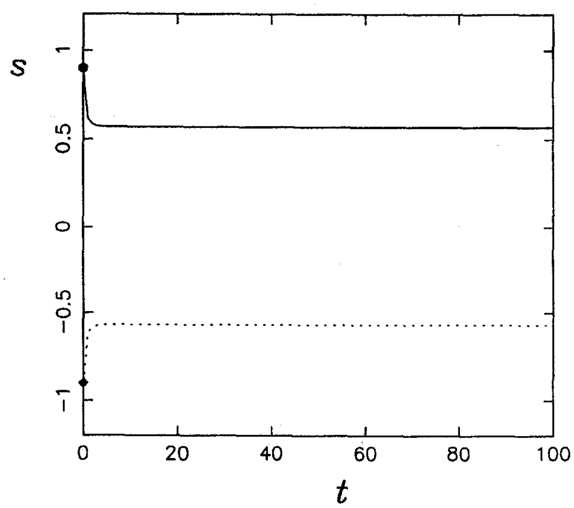

Figure 5: Stable fixed points for small time step $\Delta t = 1.2$ in the discretized analog Hopfield network, according to eq. 26. Initial conditions: $s(0) = 0.9$ (hexagon, solid line) and $s(0) = -0.9$ (diamond, dashed line). The trajectory with *positive initial condition* (hexagon, solid line) cover *positive s only*, whereas the trajectory with *negative initial condition* cover *negative s only*. There are no overshotting or decaying oscillations at early times. The fixed points are the same as those in Fig.4 and are independent of Δt.

optimization," *Computers Ops. Res.*, vol. 23, no. 6, pp. 597-610, 1996.

[25] D. W. Tank and J. J. Hopfield, "Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circ. Sys.*, vol. 33, pp. , May, 1986.

[26] L. Wang, "Oscillatory and chaotic dynamics in neural networks under varying operating conditions," *IEEE Trans. Neural Networks*, vol. 7, no. 6, pp. 1382-1388, November, 1996.

[27] L. Wang, "Suppressing chaos with hysteresis in a higher-order neural network," *IEEE Trans. on Circuit and Systems-II: Analog and Digital Signal Processing*, vol. 43, no. 12, pp. 845-846, December, 1996.

[28] L. Wang, "Discrete-time convergence theory and updating rules for neural networks with energy functions," *IEEE Trans. on Neural Networks*, vol. 8, no. 2, pp. 445-447, March, 1997.

[29] L. Wang, "Noise injection into inputs in sparsely-connected Hopfield and winner-take-all neural networks," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 27, no.5, October, 1997, accepted.

[30] Lipo Wang, "On the dynamics of discrete-time, continuous-state Hopfield neural networks," *IEEE Trans. on Circuit and Systems-II: Analog and Digital Signal Processing*, accepted.

[31] L. Wang and D.L. Alkon, Eds. *Artificial Neural Networks: Oscillations, Chaos, and Sequence Processing*, Los Alamitos, California: IEEE Computer Press, 1993.

[32] L. Wang, E. E. Pichler, and J. Ross, "Oscillations and chaos in neural networks: an exactly solvable model," *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 9467 - 9471, December, 1990.

[33] L. Wang and J. Ross, "Synchronous neural networks of nonlinear threshold elements with hysteresis," *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 988-992, February, 1990.

[34] L. Wang and J. Ross, "Chaos, multiplicity, crisis, and synchronicity in higher order neural networks," *Phys. Rev. A*, vol. 44, no. 4, pp. R2259 - 2262, August, 1991.

[35] G. V. Wilson and G. S. Pawley, "On the stability of the traveling salesman problem algorithm of Hopfield and Tank," *Biol. Cybern.*, vol. 58, pp. 63, 1988.