# Incorporating Chaos into the Hopfield Neural Network for Combinatorial Optimisation

Terence Kwok, Kate Smith<sup>\*</sup>, Lipo Wang<sup>+</sup>

School of Business Systems, Monash University Clayton, Victoria 3168, Australia. Email: tkwok@fcit -m1.fcit.monash.edu.au \*Email: ksmith@bs.monash.edu.au \*URL: http://www.bs.monash.edu.au/staff/ks.htm

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

# ABSTRACT

Various approaches of incorporating chaos into artificial neural networks have recently been proposed, and used successfully to solve combinatorial optimisation problems. This paper investigates three such approaches: 1) Chen & Aihara's transiently chaotic neural network with chaotic simulated annealing, which has a gradually decaying negative self-coupling term; 2) Wang & Smith's chaotic simulated annealing, which employs a gradually decreasing time-step; 3) Hayakawa et al's method of adding chaotic noise to a Hopfield network. The N-Queen problem is used as an application to test and compare the performance and robustness of the three methods. The traditional simulated annealing is also included for comparison in order to contrast the effectiveness of the various approaches.

**Keywords:** chaos, chaotic neural networks, combinatorial optimisation, Hopfield network, N-Queen, constraint satisfaction problems.

# **1. INTRODUCTION**

Since Hopfield & Tank proposed using artificial neural networks [1]-[3] to solve combinatorial optimisation problems, many extensions and modifications to the original formulation were suggested, e.g. Boltzmann, Cauchy machine, etc. [4]. These models were designed mainly to overcome the problem of being trapped in local minima which are often encountered when the Hopfield & Tank model, being a steepest descent method, is used to solve combinatorial optimisation problems. With the discovery of chaos theory and the subsequent novel applications in engineering [5]-[7], it was also found that chaos plays important roles in information processing in the biological brain [8]. Based on chaotic properties of biological neurons, Aihara et al [9] proposed a chaotic neural network (CNN) which includes some conventional neural network models as special cases. Using this model, Yamada et al [10] solved the Travelling Salesman Problem (TSP) which is a NPhard problem. They found that relative refractoriness in the model improves optimisation by escaping from local minima. By using Euler's method to discretise the Hopfield network having negative self-feedback connections, Nozawa [11] compared it to Aihara et al's CNN and then transformed it to a

globally coupled map (GCM) formulation. With this approach, he solved the TSP more efficiently than other neural networks with probabilistic search, and the chaotic search of the model was able to escape from local minima. Using the same method, Ohta et al [12] found that it significantly outperforms a random-walk approach in solving larger TSP problems, which further showed the inherent ability of the CNN to escape from local minima. To explore the optimisation characteristics of the CNN, Hasegawa et al [13] used a CNN with two internal states, namely refractoriness and mutual interaction, to solve the TSP. By adjusting the decay parameters of the two internal states, they concluded that the network has the highest optimisation capability at the edge of chaos, which exists between the ordered phase and disordered phase. For TSP, N-Queen, and similar problems having a common form of constraints, Ishii and Sato [14],[15] developed the chaotic Potts spin (CPS) network, which can effectively obtain optimal solutions for the TSP with better performance than some CNN models and other spin annealing techniques.

Another approach is to add chaotic noise to the conventional Hopfield network [16],[17]. The purpose is to help the network escaping from local minima more efficiently than mere white noise, and the structure of the chaotic noise was found to be useful in finding the global minimum of the TSP.

In order to introduce convergence properties into CNNs, some developments make use of the transient process of gradually tuning a neural network from a chaotic to a non-chaotic state, which is analogous to lowering the temperature in the traditional simulated annealing algorithm. Kasahara & Nakagawa [18] employed exponentially decaying functions to decrease various parameters in their CNN for convergent purpose. Unfavourable trappings at certain spurious states are avoided in solving the TSP and memory association problems. Chen & Aihara [19] proposed chaotic simulated annealing (CSA) to illustrate the features and effectiveness of a transiently chaotic neural network (TCNN) in solving optimisation problems. Based on a previous CNN model [9], the self-feedback connection weight or refractory strength is gradually decreased to achieve a transiently chaotic annealing process. They found that the chaotic behaviour provides a more efficient search for the global minimum while escaping from local minima. As the annealing parameter gradually decreases, the network becomes non-chaotic through a reverse bifurcation process, and approaches a network similar to the Hopfield

neural network which has convergence properties [1]-[3]. The TSP and a maintenance-scheduling problem were solved and the chaotic annealing process was shown to be effective in avoiding local minima. Conditions for convergence to a fixed point and the existence of chaotic structures, as observed in TCNN and discrete-time recurrent neural networks in general, were subsequently proved [20].

Another chaotic annealing approach was proposed by Wang & Smith [24]. They used Euler's method on the continuous Hopfield neural network model to construct a chaotic annealing scheme. By tuning the time-step, the network can exhibit chaotic behaviour as well as approaching a Hopfield network. A transiently chaotic version of the aforementioned CPS network was also devised by Ishii & Sato [15], which has even better performance than the original CPS.

In section 2, three different ways of integrating chaos into the Hopfield network to solve combinatorial optimisation problems are described in relation to the model, convergence criteria, chaotic properties, and respective implementations to solve the N-Queen problem. A brief introduction of using simulated annealing to solve the NQueen problem is also included. In section 3, computer simulation results of solving the N-Queen problem with various approaches are presented. Comparisons of solution quality, optimisation performance, etc. are discussed in section 4.

# 2. THREE APPROACHES TO INCORPORATE CHAOS INTO THE HOPFIELD NETWORK

In this section, chaotic simulated annealing using a decaying self-feedback term is described, followed by an alternative CSA scheme with a decreasing time-step, and then a Hopfield network with added chaotic noise. A brief account of using simulated annealing to solve the NQueen problem is also included.

#### 2.1 Chaotic Simulated Annealing with Decaying Selfcoupling

TCNN with CSA as formulated by Chen & Aihara [19] involves adding a large negative self-feedback term into a Hopfield-like network, and its magnitude is then gradually decreased with time. As  $t \rightarrow \infty$ , this self-coupling term would tend to zero, thus approaching a Hopfield-like network which converges to a fixed point [20], and minimises the energy function in the process. The formulation is as follows:

$$x_i(t) = \frac{1}{1 + e^{-y_i(t)/e}}$$
(1)

$$y_i(t+1) = ky_i(t) + \left(\sum_{j=1}^N w_{ij} x_j(t) + I_i\right) - z(t) (x_i(t) - I_0)$$
<sup>(2)</sup>

$$z(t+1) = (1 - \mathbf{b})z(t)$$
(3)

where  $i, j = 1, \dots, N.$ , and

 $x_i$  = output of neuron *i* 

- $y_i$  = internal state of neuron *i*
- a = positive scaling input parameter
- k = damping factor of nerve membrane,  $0 \le k \le 1$

$$W_{ij}$$
 = connection weight matrix;  $\sum_{j=1}^{N} w_{ij} x_j + I_i = -\frac{\P E}{\P x_i}$ 

 $= W_{ii}$ , i.e. symmetric.

Ε

$$= -\frac{1}{2}\sum_{i}\sum_{j}x_{i}x_{j}w_{ij} - \sum_{i}I_{i}x_{i}$$

 $z(t) = \text{self-feedback connection weight (refractory strength)} \ge 0$ 

**b** = damping factor of  $z(t) \cdot 0 \le \mathbf{b} \le 1$ 

 $I_i$  = input threshold of neuron *i* 

 $I_0$  = positive parameter

It is the self-feedback parameter, z(t), that contributes to the decaying negative self-coupling term in Eq. (2) [19]. Chaotic simulated annealing is achieved by evaluating Eq. (1) to (3) iteratively, starting with an initial z(0) which is large enough for the network to be chaotic, and then gradually decreasing z(t) according to Eq. (3). The speed of this annealing process is determined by **b**. When the self-coupling term is small enough, the network would become stable and converge to a fixed point, depending on some stability criteria described in the next part.

**Stability Criteria:** For asynchronous updating of Eq. (2), and with a symmetric weight matrix, the network asymptotically converges to a fixed point if one of the followings is satisfied [20]:

1. 
$$\frac{1}{3} \ge k \ge 0, 4(1-k)e > -\min\{a_{W_{ii}}\}, \text{ or }$$

2. 
$$1 \ge k \ge \frac{1}{3}, 8ke > -\min\{aw_{ii}\}, or$$

$$3. k > 1, 8e > -\min\{a_{W_{ii}}\}.$$

Furthermore, a state  $\mathbf{x}$  is asymptotically stable for Eq. (1) and (2) if and only if the state is at least a local minimum of E for  $0 \le x_i \le 1$ , for i = 1..N. Chen & Aihara [20] also found the conditions for determining which of the local minimum of the energy function is an attractor (theorem 3.4, 3.5 of [20]). Most importantly, they also proved the existence of chaos in TCNN (theorem 4.2 of [20]), and the fact that TCNN has only one bounded fixed point when the absolute value of the self-feedback weight,  $W_{ii}$ , is sufficiently large (theorem 4.1 of [20]).

**N-Queen Problem Formulation:** The N-queen problem is the problem of how to place N queens onto an N by N chessboard without attacking each other.  $x_{ij} = 1$  means a queen is placed on row *i*, column *j*. This is a constraint satisfaction problem and the constraints are:

- Each row can only have one queen.
- Each column can only have one queen.
- Each diagonal (there are many) can only have one queen.
- There are exactly N queens on the chessboard.

Related researches on solving the Nqueen problem using neural networks are, for example, using Cauchy machines [21], Gaussian machines [22], and Hopfield network [23].

In order to satisfy the N-queen constraints mentioned above, a cost function is constructed such that its value is increased if a constraint is violated: (A, B, C, D are positive parameters)

$$f = \frac{A}{2} \sum_{i} \sum_{k} d_{ik} \sum_{j} \sum_{l} (1 - d_{jl}) x_{ij} x_{kl}$$

$$+ \frac{B}{2} \sum_{j} \sum_{l} d_{jl} \sum_{i} \sum_{k} (1 - d_{ik}) x_{ij} x_{kl}$$

$$+ \frac{C}{2} \left( \sum_{i} \sum_{j} \sum_{k} \sum_{l} x_{ij} x_{kl} - 2N \sum_{i} \sum_{j} x_{ij} + N^{2} \right)$$

$$+ \frac{D}{2} \left( \sum_{i} \sum_{j} \sum_{k} \sum_{l} x_{ij} x_{kl} \left( d_{i+j,k+l} + d_{i-j,k-l} \right) \right) (1 - d_{ik})$$
(5)

neglecting the term  $CN^2/2$ , this becomes the function to be minimized using the CSA, which has the same form of energy function as the Hopfield network:

$$E = -\frac{1}{2} \sum_{i} \sum_{j} \sum_{k} \sum_{l} x_{ij} x_{kl} w_{ijkl} - \sum_{i} \sum_{j} I_{ij} x_{ij}$$
(6)

by equating (5) and (6), it follows that

 $I_{ij} =$ 

$$w_{ijkl} = -A\boldsymbol{d}_{ik} \left(1 - \boldsymbol{d}_{jl}\right) - B\boldsymbol{d}_{jl} \left(1 - \boldsymbol{d}_{ik}\right) - C$$
  
-  $D\left(1 - \boldsymbol{d}_{ik}\right) \left(\boldsymbol{d}_{i+j,k+l} + \boldsymbol{d}_{i-j,k-l}\right)$  (7)

and,

with,

$$E = \frac{A}{2} \sum_{i} \sum_{j} \sum_{k} \sum_{l} \mathbf{d}_{ik} (1 - \mathbf{d}_{jl}) x_{ij} x_{kl} + \frac{B}{2} \sum_{j} \sum_{l} \sum_{i} \sum_{k} \sum_{k} \mathbf{d}_{jl} (1 - \mathbf{d}_{ik}) x_{ij} x_{kl} + \frac{C}{2} \left( \sum_{i} \sum_{j} \sum_{k} \sum_{l} x_{ij} x_{kl} - 2N \sum_{i} \sum_{j} x_{ij} \right) + \frac{D}{2} \left( \sum_{i} \sum_{j} \sum_{k} \sum_{l} x_{ij} x_{kl} (\mathbf{d}_{i+j,k+l} + \mathbf{d}_{i-j,k-l}) \right) (1 - \mathbf{d}_{ik})$$
(9)

with N  $\,\times$  N neurons, the following equations are evaluated iteratively,

$$x_{ij}(t) = \frac{1}{1 + e^{-y_{ij}(t)/e}}$$
(10)

$$y_{ij}(t+1) = k y_{ij}(t) + \boldsymbol{a} \left( \sum_{k,l=1}^{N} w_{ijkl} x_{kl}(t) + I_{ij} \right) - z(t) \left( x_{ij}(t) - I_0 \right)$$
(11)

$$z(t+1) = (1 - \mathbf{b})z(t)$$
(12)

For good optimisation ability, the main parameters needed to be adjusted are a, b and z(0).

#### 2.2 Chaotic Simulated Annealing with Decreasing Timestep

Noticing that the Chen-Aihara system converges to a Hopfieldlike network [19] rather than an exact Hopfield network, Wang & Smith [24] devised an alternative approach for CSA. Considering the Euler approximation of the continuous Hopfield network,

$$y_i(t + \Delta t) = \left(1 - \frac{\Delta t}{t}\right) y(t) + \Delta t \left(\sum_j w_{ij} x_j(t) + I_i\right)$$
(13)

In their approach, CSA is achieved by using a large initial time-step  $\Delta t(0)$ , and then gradually decrease it as the network iterates, for example, by using the exponential decaying rule,

$$\Delta t(t+1) = (1-\mathbf{b})\Delta t(t) \tag{14}$$

where  $0 < \mathbf{b} < 1$ . This causes the network to go through a reverse bifurcation process as it starts with a chaotic state and ends with a stable convergent state. As  $\Delta t \rightarrow 0$ , the network becomes the continuous Hopfield network, which is convergent and minimises the energy function *E*. Overall, there are fewer parameters in this approach when compared to Chen & Aihara's system.

**Stability Criteria:** [24] For an asynchronous updating of Eq. (13), together with a symmetric w, t > 0, and an increasing function with a maximum slope of  $\mathbf{m}_{\text{max}}$  for the activation function, the network stabilises if the following conditions are all satisfied:

1. 
$$\Delta t \leq \mathbf{t}$$
,  
2.  $\Delta t < \frac{2\mathbf{t}}{1 - \mathbf{m}_{\max} w_{ii}}$ , and  
3.  $w_{ii} \leq 0$ . (15)

The existence of chaos is guaranteed by the violation of condition 2 above [24].

**N-Queen Implementation:** Since the same energy function E is minimised, the formulation with the same weight matrix  $W_{ijkl}$  and  $I_{ij}$  as in section 2.1 are used. The main parameters to be adjusted are  $\Delta t(0)$ , t and b.

### 2.3 Hopfield Network with Chaotic Noise

Hayakawa et al [16] investigated adding chaotic noise to the discretised continuous Hopfield network to solve the TSP. Because of the chaotic noise, this technique has hill-climbing features to escape from local minima. The better optimisation ability of a chaotic noise source over a non-chaotic or random one has been observed by Asai et al [17] and Hayakawa et al [16]. Chaotic time series with different initial values is associated with each neuron, which can be generated from the logistic map [15]

$$x(n+1) = a x(n)(1-x(n)).$$
(16)

A normalised series can be obtained as follows,

$$\boldsymbol{h}_{i}(t) = \frac{\boldsymbol{x}_{i}(t) - \langle \boldsymbol{x} \rangle}{\boldsymbol{s}_{x}}$$
(17)

where  $S_x$  is the standard deviation of the series x over time. A time series of normalised chaotic noise,  $h_i(t)$ , is then added to the network as follows,

$$x_i(t) = f\left(y_i(t) + A\boldsymbol{h}_i(t)\right) \tag{18}$$

where f is the sigmoidal activation function in Eq. (1), and A is a positive parameter.  $y_i$  is the internal state of the Hopfield network, and is updated according to Eq. (13) with a default t value of 1.

The use of other mappings and strange attractors as chaotic noise were investigated by Zhou et al [26] and Asai et al [17].

**N-Queen Implementation:** Since the Hopfield network with chaotic noise does not converge, a sufficiently large number of iterations should be pre-determined for the network to seek the global minima. To determine the solution feasibility at each iteration, a discretised state variable is introduced,

$$x_{ij}^{d} = \begin{cases} 1 & \text{if } x_{ij} \geq \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ij} \end{cases}$$
(19)

A corresponding discretised energy  $E^{d}$  is then calculated according to Eq. (9) to determine the feasibility of the solution. Parameters to be adjusted are the logistic map parameter a in Eq. (16), and the amplification factor A in Eq. (18).

#### 2.4 Simulated Annealing and N-Queen Implementation

Simulated annealing is a well-known algorithm developed by Kirkpartick et al [27]. It minimises the objective function by a stochastic annealing scheme with guaranteed convergence [27],[28]. A decreasing temperature parameter T controls the cooling schedule which allows the system to jump out of local minima with decreasing likelihood as time goes on.

To solve the N-Queen problem, an initial N  $\times$  N identity matrix is used to ensure that only 1 queen is placed in each row and column, and that there are exactly N queens. Any new state is obtained by randomly choosing 2 rows and swap. The cost to be minimised is constructed such that its value increases if any queen can attack another diagonally, which can just be the last term in Eq. (9). Note that we have discrete states updating for this formul ation while the neural network approaches mentioned above iterate with continuous states.

The following parameters should be carefully adjusted to improve optimisation performance: Markov chain length L, initial temperature  $T_0$ , and cooling rate r.

#### **3. RESULTS**

A 10-Queen problem is solved by computer simulation of each of the neural networks described above, followed by simulated annealing. For each method, 10 sets of different randomised initial states with values centered around the unit hyper-cube are used. A feasibility measure of 0/10 means no global minimum is found with the 10 sets of initial states, and 10/10 means global minimum is obtained in every case. For networks with convergent properties, the average number of iterations required to converge to a stable state (within a tolerance of  $5 \times 10^{-5}$ ) is used to measure performance. In this study, we only consider asynchronous cyclic updating for the neural network models.

Unless otherwise stated, these values as in Eq. (9) are used: A = B = C = D = 1.

Focus of the simulations is on the optimisation performance (feasibility measure) and efficiency (average number of iterations) with various values of the parameters.

#### 3.1 CSA with Decaying Self-coupling

Chen & Aihara's CSA scheme mentioned in section 2.1 is employed. Note that  $k = 1 - \mathbf{a}$  is used [11]. Parameters that require adjustments are  $\mathbf{a}$ ,  $\mathbf{b}$ , and z(0). Table 1 & 2 shows the feasibility and performance variation when different parameter values are used. Fig. 1 & 2 shows a typical time evolution of the neuron states and the energy function respectively.

TABLE 1			
Feasibility variation with	а	and	z(0)

 $(\mathbf{b}=0.001, \mathbf{I}_0=0.65; \mathbf{e}=0.004)$ 

Feasibility			z(0)	
		0.0799	0.0800	0.0810
а	0.010	10/10	10/10	10/10
	0.015	7/10	4/10	7/10

 TABLE 2

 Feasibility and performance variation with **b**

$(a = 0.01, z(0) = 0.08, I_0 =$	$0.65; \ \theta = 0.004$	1)
---------------------------------	--------------------------	----

		Feasibility	Average no. of Iterations
b	0.01	5/10	789
	0.001	10/10	~ 4000



**Fig. 1** Neuron states against the number of iterations, t. The top three are chosen to show neurons having a final state of zero, and the rest are 10 states giving a converged value of 1



Fig. 2. The corresponding energy function to Fig. 1

#### 3.2 CSA with Decreasing Time-step

To implement Wang & Smith's decreasing time-step CSA described in section 2.2 with an exponential decay rule as Eq. (14), the initial time-step  $\Delta t(0)$  and the annealing factor **b** should be adjusted carefully. According to the stability criteria (15) in section 2.2, **t** should be close to but less than  $\Delta t(0)$  to start the network with chaotic dynamics.

Table 3 & 4 shows the feasibility and performance variation with different initial time-steps and annealing factors.

Fig. 3 & 4 respectively shows the time evolution of the energy function and a typical neuron state corresponding to an optimal solution.

TABLE 3 Feasibility variation with initial time -step and  $\boldsymbol{b}$ 

(t = 0.5 and e = 0.01)

Feasibility		oility		b	
			0.001	0.005	0.01
	$\Delta t(0)$	1.0	10/10	1/10	4/10
		1.2	8/10	1/10	2/10

TABLE 4 Performance variation with initial time-step and  $\boldsymbol{b}$ 

(t = 0.5 and e = 0.01)

Averag	ge no.	b		
of Itera	ations	0.001	0.005	0.01
$\Delta t(0)$	1.0	965	222	130
	1.2	1129	240	134



Fig. 3. Energy plotted against t



Fig. 4. Neuron state  $x_{2,1}$  plotted against t

## 3.3 Hopfield Network with Chaotic Noise

To add chaotic noise to the Hopfield network to solve the 10-Queen problem, the normalised logistic map in Eq. (16)-(18) in section 2.3 is employed. Since the network is no longer convergent, the feasibility measure (Table 5) is now a count of runs which encounter the optimal solution at least once from iteration 1000 to 2000. Iterations less than 1000 correspond to the transient period, and 2000 is chosen to be the maximum number of iterations. The average frequency of arriving at optimal solutions is also counted after the transient period (Table 6).

The discretised energy function is shown in Fig. 5, and a typical neuron state is shown in Fig. 6.

The logistic map parameter a is chosen to be 3.81 and 3.93 since they belong to the chaotic region of the logistic map and were found to have good optimisation ability on the TSP by Hayakawa et al [16]. Amplification parameter A determines the relative magnitude of the noise.

TABLE 5 Feasibility variation with a and A $(e = 0.1, \Delta t = 0.1)$ 

Feasibility			Α	
		0.05	0.5	5.0
a	3.81	0/10	3/10	0/10
и	3.93	0/10	7/10	0/10

TABLE 6 Optimisation ability variation with a and A $(e = 0.1, \Delta t = 0.1)$ 

No. Iteration	of ns with	A		
Glo Min	bal ima	0.5	0.7	1.0
a	3.81	981	24	1
u	3.93	960	13	1



Fig. 5. Discretised energy versus t



**Fig. 6.** Neuron state  $X_{7,1}$  plotted against t

## **3.4 Simulated Annealing**

Since simulated annealing is very different from neural network approaches, it is included in the present study only to give us a rough idea of how well the neural networks with chaos can be compared to a widely used and robust heuristics in terms of overall performance. For this reason, a set of well tuned parameters are chosen for this method to maximise its ability.

TABLE 7 Results of simulated annealing

Initial temperature	3
Cooling rate	0.9
Markov chain length	200
Average number of Markov chains	11
Feasibility	10/10

## 4. DISCUSSION

In this section, we illustrate the properties and optimisation ability of chaos in neural networks by analysing the results obtained by solving the 10-Queen problem.

With Chen & Aihara's method of CSA with decaying selfcoupling, we obtain a typical time evolution picture of neuron states in Fig. 1, which corresponds to an evolution towards an optimal solution. One noticeable feature happens at around t =750, where the neuron states transit into a seemingly chaotic region from a metastable state with values close to 0. The unstable wandering of the states ceases when they are sufficiently close to 1, which later converge to a stable and optimal solution. The state transition from the inherent instability to order has the effect of chaotic wandering in search of the global minima. This effect has also been observed by Hasegawa et al [13] in solving the TSP. When the selfcoupling term is eliminated by setting z(t) = 0 for all t, no optimal solution is found. This reinforces the role of disorderorder transition in seeking the global minima. Although a more convincing evidence of the existence of chaos could be obtained by calculating the Lyapunov exponent, the above characteristic transition from disorder to order, together with the proof of the existence of chaos in this CSA scheme by Chen & Aihara [20], strongly suggest a chaotic role in the process. As expected for any method capable of escaping from local minima, uphill moves of the energy can be observed in Fig. 2, which occurs in the same unstable region described above. From Eq. (9), a global minimum would have E = -50, as in Fig. 2.

From Table 1, a feasibility of 10/10 is obtained for some combinations of the parameters, which is rather sensitive to the value of  $\mathbf{a}$ . In Table 2, we find that better feasibility is obtained with a smaller  $\mathbf{b}$ , which corresponds to a slower annealing schedule. This can be interpreted as a prolonged chaotic search due to a lengthened region of instability. However, the tradeoff is the increased number of iterations required for convergence. This compromise is analogous to the determination of an efficient cooling schedule in simulated annealing.

Note than with  $\boldsymbol{a} = 0.015$  and  $\boldsymbol{b} \approx 0.001$ , Chen & Aihara [19] obtained a much higher feasibility for the 10-city TSP, compared to around 40% found in Table 1 for our solution to the 10-Queen problem. This may suggest an inadequate robustness of the method. However, more initial value sets should be used in our investigation for a more convincing conclusion.

Fig. 3 and 4 are the results of obtaining an optimal solution using Wang & Smith's CSA with decreasing time-step. Fig. 4 shows how a neuron state follows its path contributing to an optimal solution. Initially it oscillates between 0 and 1, but starts to visits intermediate values when t is between 400 and 600. The two branches finally merge together towards 1, which corresponds to a global minimum. The network is proved to be chaotic before the merging of the two branches by using the criterion mentioned in section 2.2. The result suggests a search of the global minimum through chaotic dynamics. The corresponding wandering of the energy among local minima can be observed in Fig. 3.

In terms of feasibility (Table 3), very good results are obtained with b = 0.001. But it has a sensitive dependence on b. Increasing b (a faster annealing) does not necessarily decrease feasibility as shown in Table 3. This is clearly different from the case of Chen & Aihara's method in which a slower annealing is preferable for good feasibility. This illustrates the fact that decaying the time-step (an inherent quantity to the network) is fundamentally different from decaying the selfcoupling term (an externally introduced quantity). On the other hand, this method is more efficient probably by the same reason. Table 4 shows that an optimal solution with 10/10 feasibility can be obtained with ~1000 iterations. In general, this method requires fewer parameters to be adjusted, but the difficulty arises when choosing an effective annealing rate. Future work could investigate whether the same optimal annealing rate applies to other combinatorial problems of various sizes.

Since the Hopfield network with chaotic noise, i.e. the third approach in this study, is not convergent, no single stable state is attained by the neuron states. A typical neuron state iteration corresponding to an optimal solution is illustrated in Fig. 6. Although noisy, the state tends to have a denser distribution around the value of 1, except at the beginning where 0 is often visited. The final state of this neuron is 1 according to Eq. (19). The energy diagram is shown in Fig. 5. Because of the noisy nature of the network, a discretised energy is used, as explained in section 2.3. In Fig. 5, a transient region exists for t < 1000 where the energy attains non-optimal values. When t > 1000, the energy dramatically drops to a global minimum and stays around that value with occasional jumps to a local minimum. The ability of a network with chaotic noise to have a persistent attraction towards an optimal state is also reported by Hayakawa et al [16]. They also found that such a behaviour is absent if white noise is used instead. This feature may be described as a quasi-convergent behaviour, although the network is never converged.

Two values of the logistic parameter a found to have high optimisation ability by Hayakawa et al [16] are used, which yield a maximum of 7/10 feasibility as shown in Table 5. This is comparatively low against the other methods discussed so far. Also in Table 5 is the existence of an optimal value (or range of values) for the amplification factor A for better feasibility. This can be explained by the fact that when A is too large, it becomes too noisy for the network to be attracted to an optimal point; when A is too small, the dynamics of the network would be dominated by the steepest-descent mechanism of the Hopfield network, which is unlikely to give optimal solutions. The effect of increasing the noise amplitude can be found in Table 6, where the increasing A causes the global minimum to be visited less frequently.

In terms of efficiency, this method is comparable to Wang & Smith's CSA. This method requires fewer parameters to be adjusted, and in fact good values of a can be expected to perform as well on other problems. However, the feasibility is lowest among other networks in this study. Improvements may be resulted if an annealing scheme is employed to this model, for example, by decaying A gradually. Also, the question remains as to whether different chaotic noise sources affect the optimisation ability of the network.

For the case of simulated annealing, good feasibility is obtained with the well-chosen parameters shown in Table 7. There are three parameters to be chosen for each optimisation problem, which is comparable to Chen & Aihara's method. It is inappropriate to compare the number of iterations of this method to other methods in this investigation because it is a method of entirely different nature. For practical purpose, the actual run time on a computer for this method is comparable to all others discussed above. A bigger and more complicated problem would be expected to make a difference. Future works to investigate the use of chaos in neural networks for combinatorial optimisation should include:

- How the range of optimal parameters varies with problem size;
- The effect of synchronous updating on the chaotic dynamics;
- The effect of using discrete activation functions;
- Generalisation to other optimisation problems.

## 5. CONCLUSIONS

In this paper, we have investigated three methods of incorporating chaos into the Hopfield network for combinatorial optimisation. Two such methods, namely Chen & Aihara's CSA with a decaying self-coupling term and Wang & Smith's CSA with decaying time-steps, make use of chaotic annealing schemes analogous to the traditional simulated annealing. Both have features like chaotic wandering and transitions from instability to a stable state, which are found to have a novel ability to improve the optimisation performance of the network. Existence of chaos and convergence to a stable solution in both methods are well established by the respective proofs. Wang & Smith's method is more efficient, requires fewer parameters, but solution optimality depends sensitively on the annealing rate, while Chen & Aihara's method allows better solution with a slower annealing rate. The Hopfield network with chaotic noise has the least number of adjustable parameters and is relatively efficient, but lacks convergence properties and is the least capable of reaching for an optimal solution. All models in this study are comparable to, in some cases better than, the traditional simulated annealing in terms of optimisation efficiency, solution quality and ease of choosing parameters.

# 6. REFERENCES

- Hopfield, J. J., (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc. Nat. Acad. Sci., Vol. 79, pp. 2554-2558.
- [2] Hopfield, J. J., (1984). Neurons with Graded Response have Collective Computational Properties like those of Two-state Neurons. *Proc. Nat. Acad. Sci.*, Vol. 81, pp. 3088-3092.
- [3] Hopfield, J. J., Tank, D. W. (1985). "Neural" Computation of Decisions in Optimization Problems. *Biological Cybernetics*, 52, pp. 141-152.
- [4] Looi, C. K., (1992). Neural Network Methods in Combinatorial Optimization. *Computers Ops Res.*, Vol. 19, No. 3/4, pp. 191-208.
- [5] Chen, G., (1998). Chaos: Control and Anti-Control. *IEEE Circuits and Systems*, Vol. 9, Iss. 1, pp. 4-5.
- [6] Ditto, W., Munakata, T., (1995). Principles and Applications of Chaotic Systems. *Communications of the* ACM, Vol. 38, Iss. 11, pp. 97-102.
- [7] Aihara, K., Katayama, R. (1995). Chaos Engineering in Japan. *Communications of the ACM*, Vol. 38, Iss. 11, pp. 103-107.
- [8] Freeman, W. J. (1992). Tutorial on Neurobiology: from Single Neurons to Brain Chaos. *International Journal of Bifurcation and Chaos*, Vol. 2, Iss. 3, pp. 451-482.
- [9] Aihara, K., Takabe, T., Toyoda, M. (1990). Chaotic Neural Networks. *Physics Letter A*, Vol. 144, No. 6, 7, Mar., pp.333-340.
- [10] Yamada, T., Aihara, K., Kotani, M. (1993). Chaotic Neural Networks and The Travelling Salesman Problem. *Proceedings of 1993 International Joint Conference on Neural Networks*, pp. 1549-1552.

- [11] Nozawa, H. (1994). Solution of the Optimization Problem using the Neural Network Model as a Globally Coupled Map. *Towards the Harnessing of Chaos*, pp. 99-114.
- [12] Ohta, M., Ogihara,A., Takamatsu,S., Fukunaga,K., (1995). A Study on the Mechanism of the Minimum Searching by the Chaotic Neural Network, *IEEE International Conference on Neural Networks Proceedings*, Vol. 3, pp. 1517-1520.
- [13] Hasega wa, M., Ikeguchi, T., Matozaki, T. (1995). Solving Combinatorial Optimization Problems by Nonlinear Neural Dynamics. *IEEE International Conference on Neural Networks Proceedings, Perth*, Vol. 6, pp. 3140-3145.
- [14] Ishii, S. (1995). Chaotic Potts Spin. IEEE International Conference on Neural Networks Proceedings, Vol. 4, pp. 1578-1583.
- [15] Ishii, S., Sato, M. A. (1997). Chaotic Potts Spin Model for Combinatorial Optimization Problems. *Neural Networks*, Vol. 10, No. 5, pp. 941-963.
- [16] Hayakawa, Y., Marumoto, A., Sawada, Y. (1995). Effects of the chaotic noise on the performance of a neural network model for optimization problems. *Physical Review E*, Vol. 51, No. 4, Apr., pp. 2693-2696.
- [17] Asai H., Onodera, K., Kamio, T., Ninomiya, H. (1995). A Study of Hopfield Neural Networks with External Noises. *IEEE International Conference on Neural Networks Proceedings, Perth*, Vol. 4, pp. 1584-1589.
- [18] Kasahara, T., Nakagawa, M. (1995). Parameter-Controlled Chaos Neural Networks. *Electronics and Communications in Japan*, Part 3, Vol. 78, No. 7, pp. 1-10.
- [19] Chen, L., Aihara, K. (1995). Chaotic Simulated Annealing by a Neural Network Model with Transient Chaos. *Neural Networks*, Vol. 8, No. 6, pp. 915-930.
- [20] Chen, L., Aihara, K. (1997). Chaos and Asymptotical Stability in Discrete-time Neural Networks. *Physica D*, Vol. 104, pp. 286-325.
- [21] Takefuji, Y., Szu, H. (1989). Design of Parallel Distributed Cauchy Machines. Proceedings of the International Joint Conference on Neural Networks. I-529 – I-532.
- [22] Akiyama, Y., Yamashita, A., Kajiura, M., Aiso, H. (1989). Combinatorial Optimization with Gaussian Machines. *IJCNN*, Washington, Vol. 1, pp. 533-540.
- [23] Tagliarini, G. A., Page, E. W. (1987). Solving Constraint Satisfaction Problems with Neural Networks. *IEEE International Conference on Neural Networks Proceedings, San Diego*, Jun., III-741 - III-747.
- [24] Wang, L., Smith, K. (1998). On Chaotic Simulated Annealing. *IEEE Transactions on Neural Networks*, Vol. 9, No. 4, July.
- [25] Wang, L., Smith, K. (1998). Chaos in the Discretized Analog Hopfield Neural Network and Potential Applications to Optimization. *International Conference* on Neural Networks Proceedings, pp. 1679-1684.
- [26] Zhou, D., Yasuda, K., Yokoyama. R. (1997). A Method to Combine Chaos and Neural Network Based on the Fixed Point Theory. *T.IEE Japan*, Vol. 117-C, Iss. 5, pp. 599-608.
- [27] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P. (1983). Optimisation by Simulated Annealing. *Science*, Vol. 220, pp. 671-680.
- [28] Aarts, E. H. L., Korst, J. (1989). Simulated Annealing and Boltzmann Machines, John Wiley & Sons., Essex.