TR - H - 144

Chaotic Potts Spin Model for Combinatorial Optimization Problems

石井　信
Shin Ishii

1995. 4. 21

ATR人間情報通信研究所

〒619-02　京都府相楽郡精華町光台2-2　☎ 0774-95-1011

ATR Human Information Processing Research Laboratories
2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan
Telephone: +81-774-95-1011
Facsimile: +81-774-95-1008

Chaotic Potts Spin model for combinatorial optimization problems

Shin Ishii
ATR Human Information Processing Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 JAPAN

Abstract

In the analog Hopfield network and the mean field theory model of the Boltzmann machine, there occur bifurcations of solutions according to the symmetry of the energy function. This also holds in the Potts neural network. In this report, we investigate the bifurcation processes of the Potts mean field theory equation applied to traveling salesman problems and show some limitations of the annealing procedure. As an alternative approach, we propose a nonequilibrium version of the Potts neural network model, which is called Chaotic Potts Spin (CPS). We show experimental results for a comparison with the mean field annealing and the Potts mean field annealing. We also describe a modified algorithm in which a heuristic method is employed.

1 Introduction

The analog Hopfield network always converges to a local minimum of its Lyapunov function, and when the slope of the sigmoidal output function becomes large, the Lyapunov function is nearly equal to the energy function. By utilizing this feature, the analog Hopfield network can be applied to combinatorial optimization problems defined as a minimization of the quadratic energy function.[2] The analog Hopfield network is equivalent to the mean field theory (MFT) of the Boltzmann machine. The Lyapunov function of the Hopfield network corresponds to the free energy function of the MFT.[4] Wilson and Pawley[10] reported that the Hopfield network is not a good algorithm for solving combinatorial optimization problems when the problem scale is not very small. Neural network approaches need some additional mechanisms for relatively large scale problems. One of them is a gradual slope enlarging of the sigmoidal output function, i.e., a gradual lowering of the system's "temperature," which corresponds to a well-known annealing mechanism. This is the mean field annealing (MFA) algorithm.[4, 9] During the gradual temperature lowering, there occur many bifurcations of the MFT solutions. We have investigated the bifurcation processes in the MFA procedure.[7] Our result implies the MFA has some limitations.

Nozawa[3] showed that the chaotic version of the Hopfield network can find the optimal solution of small scale traveling salesman problems (TSPs) almost 100% of the time. His model is based on an Euler difference equation of the Hopfield network with a self-loop, which is equivalent to the Chaotic Neural Network (CNN) model proposed by Aihara et al.[1] It has been known that an Euler difference equation derived from a differential equation system with a self-loop has a chaotic solution even if the original differential equation system has stable solutions.

In the TSPs, graph bisection problems, N-Queen problems and so on, neural network representations have a common structure, i.e., for some variables, their summation must be 1. Focusing on this property, Peterson and Soderberg[5] employed a Potts spin system in the MFA approach. Their Potts neural network can be applied to fairly large scale problems.[6]

In this report, we briefly discuss the bifurcation processes of the Potts MFT applied to TSPs. The Potts mean field annealing procedure has some limitations in obtaining the optimal solution. As an alternative approach, we propose a chaotic version of the Potts neural network. This is called Chaotic Potts Spin. Some experimental results and a modification are also described.

2 Potts neural networks

Some of the $\mathcal{NP}$-complete optimization problems can be described as a quadratic energy minimization problem for $(M \times N)$-dimensional Potts spin variables $S_{a,n}(=$

0 or 1):

$$E(\mathbf{S}) = \frac{1}{2}\sum_{a,b=1}^{M}\sum_{n,m=1}^{N} W_{a,n;b,m} S_{a,n} S_{b,m} + \sum_{a=1}^{M}\sum_{n=1}^{N} I_{a,n} S_{a,n}, \tag{2.1}$$

where the constraints: $\sum_{n=1}^{N} S_{a,n} = 1 \quad (a = 1, ..., M)$ must be satisfied and $W_{a,n;b,m} = W_{b,m;a,n}$ holds. The values of the parameters $\mathbf{W}$ and $\mathbf{I}$ are determined for each problem. In the MFT of the Potts spin model, analog variables $V_{a,n} \in [0,1]$, each of which represents the probability that the binary variable $S_{a,n}$ takes the value 1, are introduced. These analog variables must satisfy the following constraint:

$$\sum_{n=1}^{N} V_{a,n} = 1 \qquad (a = 1, ..., M). \tag{2.2}$$

In the MFT, the free energy F and the energy E are given as follows:

$$F(\mathbf{V}) = E(\mathbf{V}) + T\left(\sum_{a,n} V_{a,n}\log V_{a,n} + N\log N\right) \tag{2.3a}$$

$$E(\mathbf{V}) = \frac{1}{2}\sum_{a,b,n,m} W_{a,n;b,m} V_{a,n} V_{b,m} + \sum_{a,n} I_{a,n} V_{a,n}, \tag{2.3b}$$

where T corresponds to the temperature in the statistical mechanics. A minimum of the MFT free energy function, which corresponds to an equilibrium in the statistical mechanics of the Potts spin, satisfies the stationary condition of the free energy function:

$$U_{a,n} = \sum_{b,m} W_{a,n;b,m} V_{b,m} + I_{a,n} \tag{2.4a}$$

$$V_{a,n} = H_n(\mathbf{U}_a) = \frac{e^{-U_{a,n}/T}}{\sum_m e^{-U_{a,m}/T}}, \tag{2.4b}$$

which is called Potts MFT equation. The solution of this equation can be obtained by using the continuous-time Potts spin model, which is a differential equation system:

$$\tau\frac{dU_{a,n}}{dt} = -U_{a,n} + \sum_{b,m} W_{a,n;b,m} V_{b,m} + I_{a,n} \tag{2.5a}$$

$$V_{a,n} = H_n(\mathbf{U}_a) = \frac{e^{-U_{a,n}/T}}{\sum_m e^{-U_{a,m}/T}} \tag{2.5b}$$

or the asynchronous Potts MFT equation, which is a difference equation system:

$$U_{a,n}(t) = \sum_{b,m} W_{a,n;b,m} V_{b,m}(t-1) + I_{a,n} \tag{2.6a}$$

$$V_{a,n}(t) = H_n(\mathbf{U}_a(t)) \tag{2.6b}$$

$$V_{b,m}(t) = V_{b,m}(t-1) \quad \text{for} \quad b \neq a. \tag{2.6c}$$

In the continuous-time model (2.5), the free energy (2.3) is the Lyapunov function that always decreases over time, and the MFT equation (2.4) is satisfied at the stable stationary points (minima). If we assume

$$W_{a,n;a,m} = 0 \quad \text{for any} \quad a, n, m, \tag{2.7}$$

this also holds for (2.6).

In the high temperature limit ($T \to \infty$), the free energy (2.3) is dominated by the entropy term, i.e., the second term of the r.h.s. of (2.3a), and there is a unique minimum. In the low temperature limit ($T \to 0$), the free energy function F (2.3a) is nearly equal to the energy function E (2.3b). In addition, the minima of the energy function occur at the hypercube corners ($V_{a,n} \in \{0, 1\}$) if $W_{a,n;a,n} = 0$. Therefore, in the low temperature limit, there are local minima of the free energy function (2.3) that correspond to those of the original energy function (2.1) for binary spins. If the temperature is fixed at a low value, which local minima are found by using (2.5) or (2.6) is dependent on the initial condition.

In order to get a good local minimum of the energy function (2.1), MFT annealing can be used. First, the MFT equation (2.4) is solved for a high temperature, and a unique minimum is obtained. Then slightly lowering the temperature, the MFT equation is solved starting from the higher temperature solution. Continuing this process, one can get a low temperature solution which corresponds to a local minimum of the energy function (2.1). This algorithm[5] is called Potts mean field annealing (PMA).

3 TSP and PMA bifurcations

A Potts spin energy function for an N-city TSP is defined as:

$$E(\mathbf{V}) = \frac{1}{2} \sum_{a,b,n=1}^{N} D_{a,b} V_{a,n} (V_{b,n+1} + V_{b,n-1}) + \frac{\alpha}{2} \sum_{n=1}^{N} (\sum_{a=1}^{N} V_{a,n} - 1)^2 + \frac{\beta}{2} \sum_{a=1}^{N} \sum_{n \neq m}^{N} V_{a,n} V_{a,m}, \tag{3.1}$$

where $V_{a,n}$ represents the probability that the salesman visits city a at the n-th visit, and $D_{a,b}$ denotes the distance between city a and city b.

The Potts MFT equation (2.4) has a unique minimum in the high temperature limit and a lot of minima in the low temperature limit. Therefore, during the course of temperature lowering, there occur bifurcations that generate new minima. These bifurcation processes are dependent on the structural stable symmetries in the problem.

In TPSs, there are two types of symmetries: cyclic symmetry and reverse symmetry. Due to these symmetries, a solution of an N-city TSP has $2N$ equivalent representations. These two types of symmetries affect PMA bifurcation processes. Local bifurcation processes in the PMA for a TSP are classified as follows: saddle-node bifurcations; reverse symmetry breaking bifurcations; cyclic symmetry breaking bifurcations. Let us show a typical bifurcation diagram. Figure 1a shows an example

5-city TSP. Figure 1c shows its bifurcation diagram, where $V_{1,i}$ $(i = 1, ..., 5)$ of every minimum are plotted for each temperature. Figure 1d shows the corresponding free energy diagram. In Figure 1c, we can observe a cyclic symmetry breaking bifurcation at $T \approx 0.32$, where a single minimum bifurcates into 5 minima. Since each bifurcated minimum preserves the reverse symmetry, there are 3 cascades observed in Figure 1c. We can also observe a reverse symmetry breaking bifurcation at $T \approx 0.29$, where each minimum bifurcates into two nonsymmetric minima. At $T \approx 0.27, 0.21$, and 0.19, we can observe saddle-node bifurcations, where 10 nonsymmetric minima simultaneously appear. It is important to note that stable solutions may disappear through a bifurcation.

The PMA procedure is a series of the above-mentioned local bifurcation processes and has the following features.

- **free energy crossing**
  We can observe this phenomenon in Figure 1d. In Figure 1d, the free energy levels of the annealing solution and the new born minima that appear at $T \approx 0.27$ cross each other and the free energy level of the annealing solution becomes higher than that of the new born minima. In fact, the new born minima correspond to the optimal solution shown in Figure 1a. In this case, the PMA fails to obtain the optimal solution, and results in a semi-optimal solution shown in Figure 1b. Moreover, sometimes the PMA deterministically results in a failure that does not correspond to any of the valid Hamilton passes.

- **deterministic/non-deterministic**
  When the annealing solution disappears at some temperature, and there are two or more distinctive minima at this temperature, the PMA may become a non-deterministic procedure.

Accordingly, it can be said that the PMA is, in general, a non-deterministic procedure to obtain a "not-so-bad" solution.

4 Chaotic Potts Spin

4.1 Model description

If we apply the Euler method to the continuous-time Potts spin model (2.5), a difference equation is obtained:

$$U_{a,n}(t) = kU_{a,n}(t-1) + (1-k)\left(\sum_{b,m} W_{a,n;b,m}V_{b,m}(t-1) + I_{a,n}\right) \quad (4.1a)$$

$$V_{a,n}(t) = H_n(\mathbf{U}_a(t)) = \frac{e^{-U_{a,n}(t)/T}}{\sum_m e^{-U_{a,m}(t)/T}}, \quad (4.1b)$$

where $k = 1-\delta t/\tau$ and δt denotes the time interval. The original differential equation system always converges to a local minimum of the free energy (2.3). However, this difference equation often oscillates or produces chaos. If we choose W to be $W_{a,n;a,n} > 0$ and $W_{a,n;a,m} = 0 \quad (m \neq n)$, (4.1) often exhibits chaotic solutions. We thus call the difference equation system (4.1) the Chaotic Potts Spin (CPS). Roughly speaking, the strength of chaos is determined by the self-loop $W_{a,n;a,n} > 0$ and the time interval parameter of the Euler method $0 < 1 - k < 1$. When $W_{a,n;a,n}$ is large and $1-k \approx 1$, the chaos is strong. On the other hand, when $W_{a,n;a,n}$ is small and $1-k$ is small, the system tends to converge. Therefore, the parameter k is a "stabilizer" parameter. The equation (2.6) is a special case of the Euler difference equation (4.1) where $k = 0$, and the system easily becomes chaotic when the self-loop is positive. Peterson and Soderberg[5] mentioned this feature.

4.2 CPS for TSPs

Let us apply the CPS approach to the TSP energy function (3.1). From (3.1), (2.2), (2.3b), and (4.1), the CPS equation for a TSP is defined as a dynamical system:

$$U_{a,n}(t) = kU_{a,n}(t-1) + \sum_b D'_{a,b}(V_{b,(n+1)}(t-1) + V_{b,(n-1)}(t-1)) \quad (4.2a)$$

$$+\alpha(1-k)\sum_b V_{b,n}(t-1) - \beta(1-k)V_{a,n}(t-1)$$

$$V_{a,n}(t) = H_n(\mathbf{U}_a(t)) = \frac{e^{-U_{a,n}(t)/T}}{\sum_m e^{-U_{a,m}(t)/T}}, \quad (4.3b)$$

where $D'_{a,b} = (1-k)D_{a,b}$ for convenience. In this case, $W_{a,n;a,n} = \alpha - \beta$. Therefore, when α is larger than β, the CPS may be chaotic.

Let us show an experimental result. Figure 2a shows the result of the CPS applied to the famous 10-city problem.[2] In a "sweep," all of the variables are updated once and only once. The ordinate denotes the solution obtained at each step (sweep). The indices 1, 2, ..., 5, and 6 denote the optimal solution, the 2nd best, ..., the 5th best, and all the other valid tours, respectively. When no circle is plotted, no valid tour is obtained at that sweep. Figures 2b and 2c show the time-series of the variable V_{11} and the energy function (3.1), respectively. As these figures show, each variable moves chaotically, and the optimal and the semi-optimal solutions are retrieved over time.

In the following, we compare our CPS approach with the PMA, the CNN, and the Ising (binary) spin MFA. A result of the CPS or the CNN is the best tour it finds during the whole sweeps. We prepared five testbeds for evaluation; they are 10-city, 20-city, 30-city, 40-city, and 50-city TSPs. In each problem, the city locations are randomly generated in a unit square. Some normalization is also done. Each testbed consists of 100 sets of city allocations. Table 1 shows the results.

Table 1

	10	20	30	40	50
CPS	100 (3.459)	100 (4.318)	100 (5.252)	100 (5.967)	100 (6.940)
PMA	82 (3.467)	85 (4.325)	96 (5.167)	94 (5.602)	86 (6.185)
CNN	90 (3.578)	97 (4.635)	96 (5.778)	96 (6.775)	92 (7.603)
MFA	74 (3.628)	65 (4.960)	66 (7.517)	96 (8.267)	94 (10.431)

In each column, the upper number and the lower number denote the number of valid tours and the averaged tour length for valid tours, respectively. For example, in the case of the PMA for 20-city TSPs, among the 100 sets of city allocations, valid tours are obtained for 85 sets, and the average tour length over the 85 tours is 4.325. Of course, the ability of each procedure depends on the value of its parameters. As for the CPS, they are shown in Table 2. In general, as the number of sweeps becomes large, the ability of the CPS becomes better.

Table 2

#city	parameter values
10	$\alpha = 2.7, \beta = 0.5, k = 0.7, T = 0.042, \#sweep = 1000$
20	$\alpha = 3.0, \beta = 0.5, k = 0.7, T = 0.035, \#sweep = 1000$
30	$\alpha = 3.3, \beta = 0.6, k = 0.7, T = 0.029, \#sweep = 1500$
40	$\alpha = 3.5, \beta = 0.7, k = 0.7, T = 0.024, \#sweep = 2000$
50	$\alpha = 3.7, \beta = 0.8, k = 0.7, T = 0.020, \#sweep = 2000$

Since the analog Hopfield network corresponds to the MFT without the annealing mechanism, the MFA is a better algorithm than the Hopfield network. Wilson and Pawley[10] reported a quite discouraging result for the Hopfield network in the case $N \geq 10$. Table 1 shows that even with the MFA, the Ising spin model is not so good an algorithm for a fairly large number of cities. In the MFA, the possible solution with $T = 0$ will be on the $N \times N$ dimensional hypercube corners whose number is 2^{N^2}. On the other hand, in the PMA, the possible solution with $T = 0$ will be on the N dimensional hypergrid whose number is N^N. Therefore, the dimension to be searched is smaller in the PMA than in the MFA. We consider that this merit also holds in the CPS. Therefore, as Table 1 shows, the Potts spin models are better than the Ising spin model (MFA).

The PMA deterministically fails to obtain the optimal solution when crossing of the free energy level occur. Even in such a case, the CPS, which is a nonequilibrium dynamical system, has the possibility to find the optimal solution. Actually, for 10-city TSPs, the CPS can obtain the optimal solution at a 98% rate. This merit is also true when the CNN[1] is compared with the Hopfield network. However, as the number of cities increases, even a nonequilibrium CPS cannot find the optimal solution, because of the expanse of the space to be searched. In the CPS, with its nonequilibrium property, even if some part of the solution is good, some other part can be somewhat random. Therefore, with a relatively large number of cities, the CPS becomes inferior to the PMA. On the other hand, although the PMA sometimes fails to obtain any of the valid tours, the CPS can always obtain valid tours for every problem. Accordingly, our CPS approach is fairly good at obtaining the optimal solution for small scale problems and any semi-optimal solutions for relatively large scale problems.

Notice that our CPS approach is significantly dependent on the system's temperature T, because it does not employ the annealing procedure. Although the best temperature value varies according to the city allocation, we fixed it in the above experiment for a practical reason. If we tune the temperature for each problem, the CPS ability can be greatly improved.

4.3 CPS with heuristics

In our CPS approach, a solution tends to have some good parts and some bad/random parts in it; this is due to its nonequilibrium dynamics. Therefore, a local optimization method can improve the obtained solutions. We choose the 2opt algorithm as a local optimization method. The initial state of the modified CPS is a solution of the 2opt. Every time a solution is obtained by the CPS, it is improved by the 2opt, and the result is the best solution among the improved solutions. This modification does not affect the processing time very much. Table 3 shows the experimental results. The first column is the basic CPS result with random initial states. The second column is the modified CPS result. Here, the initial 2opt solution is excluded for evaluation. The third column is the PMA result. The fourth column is the PMA result improved by the 2opt. The averaged tour length of the PMA is only for the valid tours it obtains. If the length of the CPS is averaged over the same allocations as in the PMA, the results are slightly improved. The fifth column is the result of the single 2opt.

Table 3

	CPS	CPS+2opt	PMA	PMA+2opt	2opt
10	3.459	3.458	3.467	3.460	3.480
20	4.318	4.228	4.325	4.244	4.333
30	5.252	4.943	5.176	5.029	5.100
40	5.967	5.440	5.602	5.530	5.652
50	6.940	5.941	6.185	6.072	6.226

Although the CPS dynamics is chaotic, some good partial parts (variables) continue over a long time. In such cases, the CPS can improve only the other bad parts, which is a local search in a domain subspace. This is the merit of the CPS; consequently, the CPS can make an improvement when its initial state is set to be as good as a 2opt solution. The ability of the modified CPS is better than that of the PMA, and it can obtain the optimal solution for every 10-city TSP. On the other hand, the PMA solution cannot be similarly improved by the 2opt.

5 Conclusion

The mean field annealing approaches, including the MFA and the PMA, are very powerful algorithms for combinatorial optimization problems. However, due to their bifurcation properties, they may fail to obtain the optimal solution even for a small scale problem. As an alternative approach, we proposed a nonequilibrium dynamical system whose name is Chaotic Potts Spin. As for very small scale problems, our CPS can solve them almost 100% of the time. As for relatively large scale problems, the CPS is inferior to the PMA but the CPS can obtain any of the semi-optimal solutions for every problem. CPS is inferior because, with its nonequilibrium dynamics, there is some random part left in almost every solution it obtains. To deal with this problem, we propose a CPS combined with local optimization heuristics, namely, 2opt heuristics. With this modification, the CPS becomes better than the PMA.

Our CPS is a fast algorithm. Actually, in the experiment described in Sec. 4.2, the CPS is faster than the PMA, the CNN and the MFA. Furthermore, by modifying the Potts neural network hardware[8], we believe it will not be very difficult to implement the CPS algorithm into hardware. In fact, implementation of the MFA is considered to be more difficult, because our computer simulation shows that the annealing procedure is quite sensitive to its precision, and this is a weak point of a hardware implementation.

References

[1] Aihara, K., T. Takabe, and M. Toyoda (1990). Chaotic Neural Networks. *Physics Letters A*, **144**, 333-340.

[2] Hopfield, J.J., and D.W. Tank (1985). "Neural" computations of decisions in optimization problems. *Biological Cybernetics*, **52**, 141-152.

[3] Nozawa, H. (1992). A neural network model as a globally coupled map and applications based on chaos. *CHAOS*, **2**, 377-386.

[4] Peterson, C., and J.R. Anderson (1988). Neural networks and NP-complete optimization problems: A performance study on the graph bisection problem. *Complex Systems*, **2**, 59-89.

[5] Peterson, C., and B. Soderberg (1989). A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems*, **1**, 3-22.

[6] Peterson, C. (1990). Parallel Distributed Approaches to Combinatorial Optimization: Benchmark Studies on Traveling Salesman Problem. *Neural Computation*, **2**, 261-269.

[7] Sato,M., and S. Ishii. *in preparation.*

[8] Urahama, K., and S-I. Ueno (1993). A gradient system solution to Potts mean field equations and its electronic implementation. *International Journal of Neural Systems*, **4**, 27-34.

[9] Van den Bout, D.E., and T.K. Miller III (1989). Improving the Performance of the Hopfield-Tank Neural Network Through Normalization and Annealing. *Biological Cybernetics*, **62**, 129-139.

[10] Wilson, G.W., and G.S. Pawley (1988). On the Stability of the Travelling Salesman Problem Algorithm of Hopfield and Tank. *Biological Cybernetics*, **58**, 63-70.

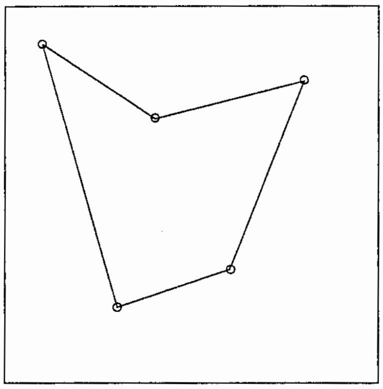

Figure 1a

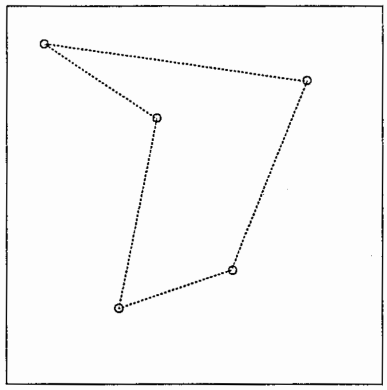

Figure 1b

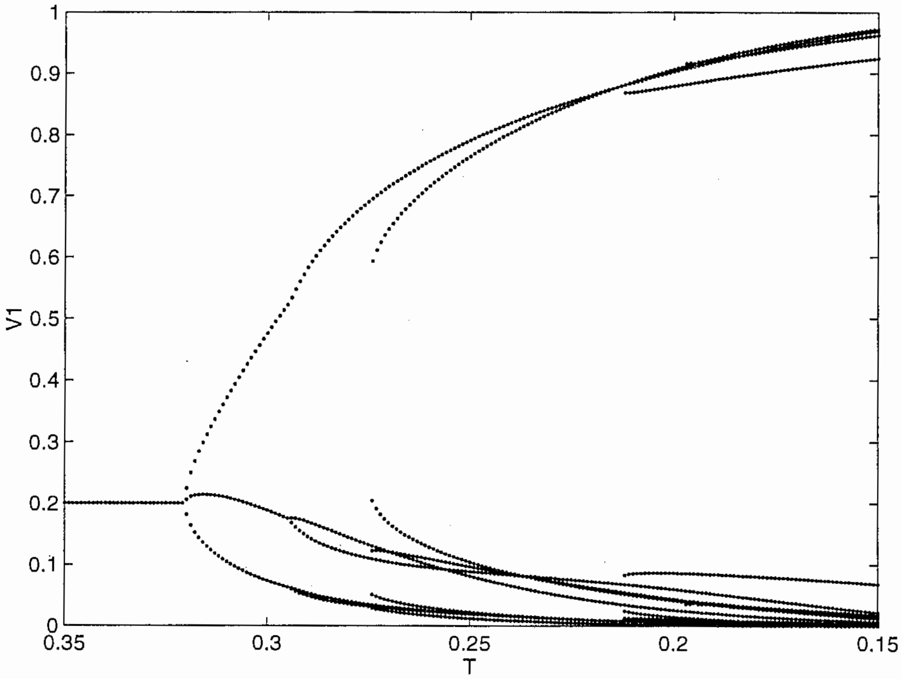

Figure 1c

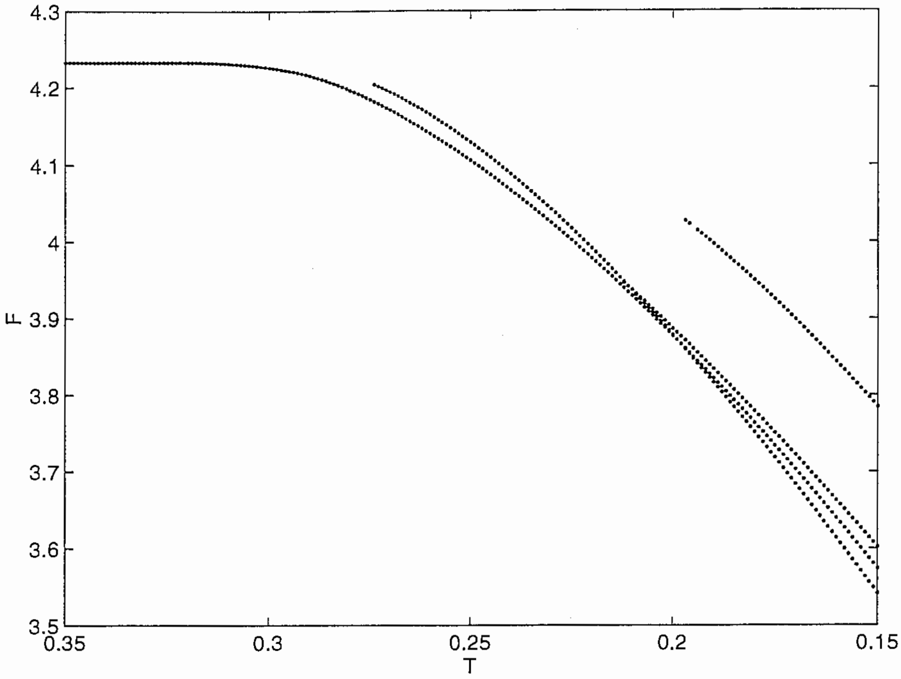

Figure 1d

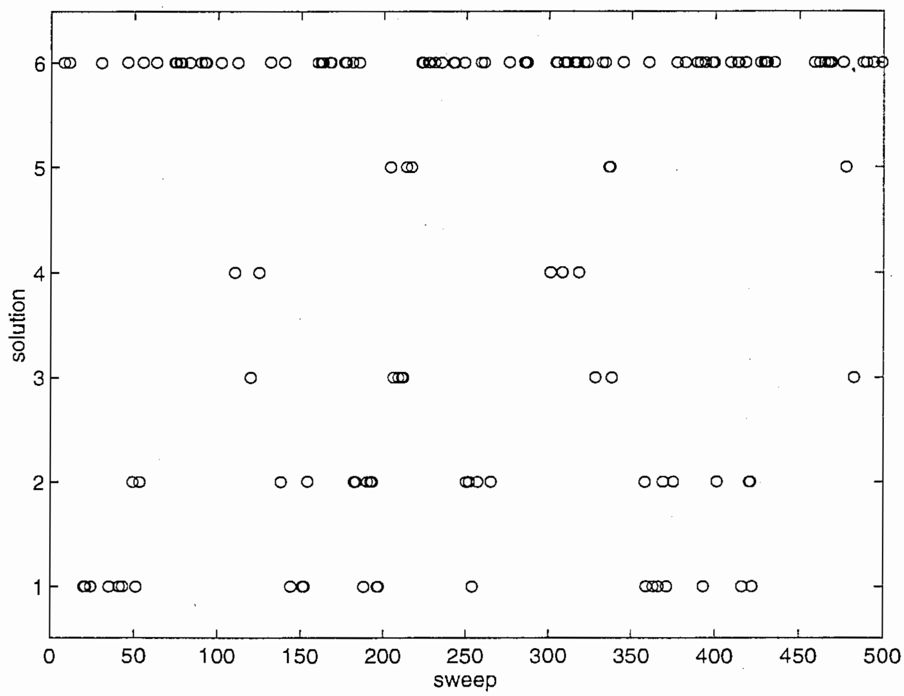

Figure 2a

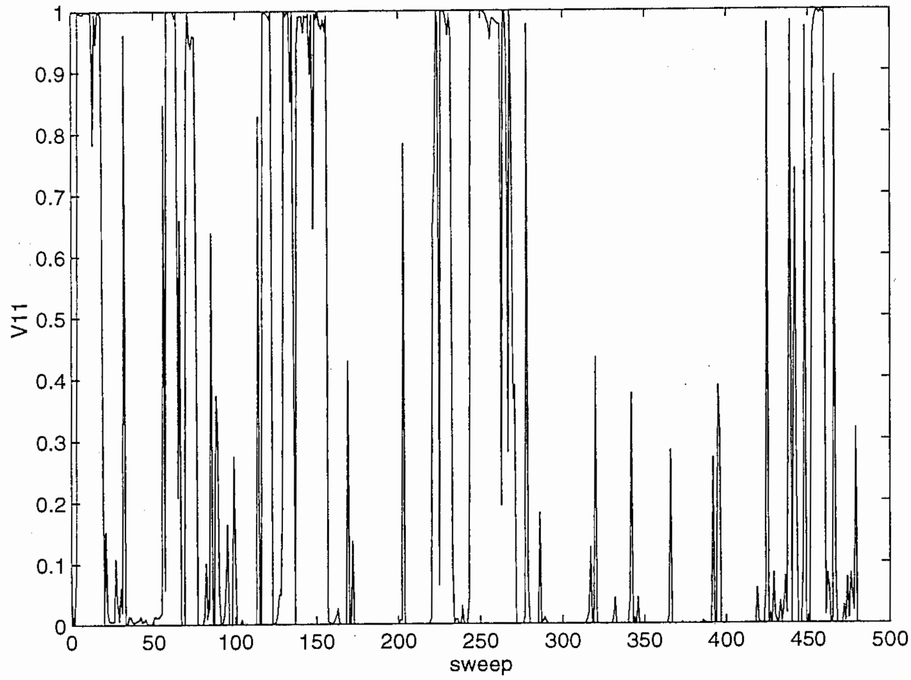

Figure 2b

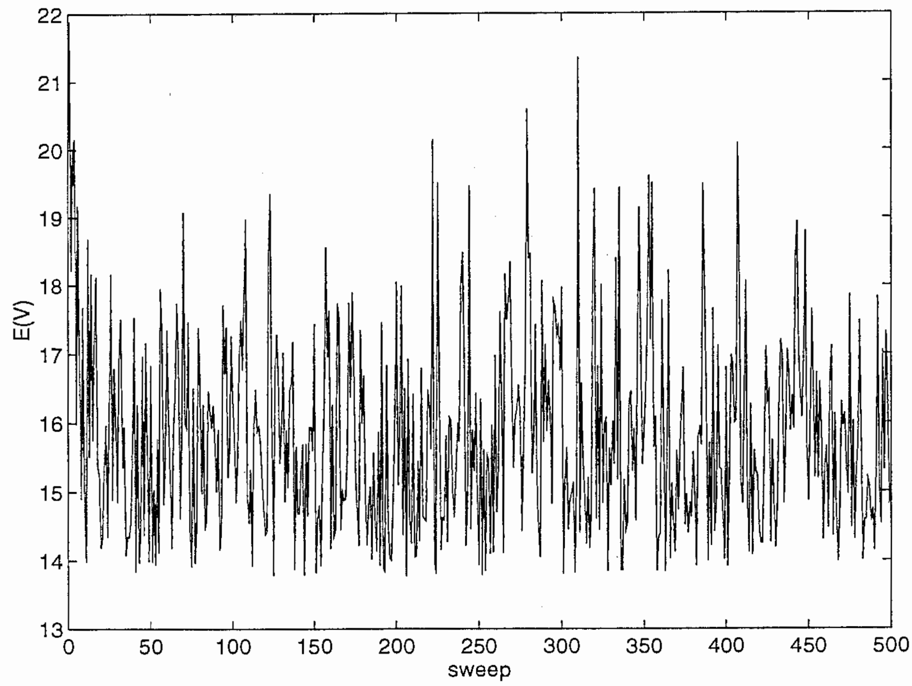

Figure 2c