Empirical Evidence of the Chaotic Behavior of the Hopfield-Tank TSP Model

Stamatis Vassiliadis*
Nikos P. Pitsianis**
Gerald G. Pechanek***

TR 94-1416
March 1994

Department of Computer Science
Cornell University
Ithaca, NY 14853-7501

* Electrical Engineering Dept., T.U. Delft, Delft, The Netherlands.
** Dept. of Computer Science, Cornell University, Ithaca, NY.
*** IBM Microelectronics Division, Research Triangle Park, NC.

Empirical Evidence of the Chaotic Behavior of the Hopfield-Tank TSP Model

Stamatis Vassiliadis * Nikos P. Pitsianis † Gerald G. Pechanek ‡

March 24, 1994

Abstract

Since its introduction, the Hopfield-Tank model for the Traveling Salesman Problem (TSP) has been surrounded by controversial evidence regarding its viability as a model and its capabilities to produce good results to this hard optimization problem. In this paper we investigate the reasons behind the difficulty of obtaining verifiable results and the viability of the model, by investigating the behavior the Hopfield-Tank neural network for the TSP has in circumstances when it is expected to produce identical tours. Our investigations strongly suggest that, when it is expected from the network to converge in a predetermined tour, the neural network converges to *almost* all possible tours when an insignificant perturbation to the initial conditions is applied. The overall consequence of our findings regarding the viability of the Hopfield-Tank model and the cause of the controversy surrounding the Hopfield-Tank model for the TSP can be summarized by the following: The cause of the Hopfield-Tank neural network for the TSP controversy and the difficulties in reproducing results is the chaotic behavior of the model. The finding of useful results for the TSP using the Hopfield-Tank network are purely casual and not to be attributed to the viability of the model. In essence the Hopfield-Tank neural network for the TSP is as viable as chaotic systems can be.

1 Introduction

The Traveling Salesman Problem (TSP) [PS82] is often considered as one of the most, if not the most, representative combinatorial optimization problem. The problem can be described simply by the following: Given a finite set C of cities and a distance function $d_{X,Y} \geq 0$ for each pair of cities $X, Y \in C$, a salesman wants to find a tour visiting all the cities in C and returning to the starting city, that has the minimum length. A tour is an ordering of the set C, or a permutation of the elements of C and its length is defined as the sum of the distances

*Electrical Engineering Dept., T.U. Delft, Delft, The Netherlands.

†Dept. of Computer Science, Cornell University, Ithaca, NY.

‡IBM Microelectronics Division, Research Triangle Park, NC.

of the cities when moving from one to the other and returning to the first city, according to the tour ordering.

The corresponding decision problem to this optimization problem has been shown to be NP-hard and unless NP = P, there is no polynomial time algorithm for finding an optimal tour. Actually when the triangle inequality does not apply, i.e. there exists a triplet of cities X, Y and Z such that $d_{X,Z} > d_{X,Y} + d_{Y,Z}$, there is no ϵ-approximate polynomial time algorithm for the unrestricted TSP [PS82]. That is, for any given ϵ, an algorithm can not be found which determines a tour that is no longer than ϵ% from the optimal tour.

One of the approaches suggested to be promising in providing an acceptable, but not optimal solution, to this classical hard problem is due to Hopfield and Tank [HT85]. The model as proposed by Hopfield and Tank is based in neural networks [HT85] and its validation was performed with a 10 city problem. Based on simulations, Hopfield and Tank have suggested that their neural network can provide results of good quality for the 10 city TSP problem they considered. Since its introduction, a number of investigations have attempted to improve and to validate the Hopfield-Tank neural network model for the TSP. Generally speaking, the research regarding the Hopfield-Tank model for the TSP and Hopfield-Tank related models is characterized by investigating:

- The convergence of the model to a solution within an "acceptable" amount of time.
- The convergence of the model to valid solutions (legitimate tours).
- The quality of the valid solutions.
- The scalability of the model.

Following this investigation paradigm, based on the modification of various parameters in the model, a number of Hopfield-Tank based models have appeared with the intention to improve the performance of the original Hopfield and Tank neural network model, see for example [Biz91, BI89]. Clearly, there is no lack of reports suggesting that simulations of the Hopfield-Tank (HT) and Hopfield-Tank based models have been successful in providing acceptable solutions to the TSP (and other optimization problems) for a number of city configurations [BWLM88, CR89, XT91, Biz91, BI89]. On the other hand, regarding the validation of the Hopfield-Tank model, a number of conflicting findings have also been reported. In a number of occasions, see for example [WP88, KPKP90], it has been suggested —based also on simulations— that the results of the Hopfield and Tank original proposal could not be repeated.

One of the important aspects of the investigations, regarding the reproduction of the results of the Hopfield-Tank model and the improvements of the model, is that what has been found thus far does not provide a consistent view of the behavior of the model. As exemplified by the research reported in [HT85] and [WP88], what has been found by an investigation is partially found, or not found at all, by another investigation. While there is a constant disagreement on the behavior of the proposal, and while it may be becoming more and more evident to the scientific community that the Hopfield-Tank model may in instances produce unpredictable results, what remain as definite open questions is establishing an explanation of the causes of such a behavior and determining whether the model has any possibility of being considered as a viable model.

In this paper, we investigate these open problems and we establish the causes of the unexpected behavior of the model as observed by many. Further we provide an explanation of the reason why results of one investigation are not easily reproducible by other investigations and provide some insights on the viability of the Hopfield-Tank model. We provide the answer to the open problems by showing, using simulations on the original city problem proposed by Hopfield-Tank, that when the network should be producing identical tours, it consistently does not. In particular we show that very small (negligible) variations from the initial set up of the network will produce different tours when in reality it should be converging to identical tours. Our findings strongly suggest that the non-linear recursive Hopfield-Tank equations for the TSP are chaotic, indicating that the cause of the entire controversy is the chaotic behavior of the network and is thus to be expected. Clearly, our investigation suggests, in view of the chaotic nature of the model, that the model can not be considered viable in producing consistent good quality tours. Our investigation also suggests that any other variation of the network equations may have identical behavior and thus it can not be used for the TSP problem, as in general, such models will produce unreliable and unpredictable results.

The presentation of the paper is in the following format: In section 2 we introduce briefly the approach taken by Hopfield-Tank to model the TSP. Additionally we survey the findings of a number of investigations regarding reported improvements and the repeatability of the results for the Hopfield-Tank TSP model. Further, we discuss in some detail the main contributions of the investigation reported in this paper. In section 3 we describe our simulation approach. Consequently we present the results of both simulations and conclude with some remarks regarding the behavior of the Hopfield-Tank network.

2 Background, Open Questions, and Main Results

Briefly, in the Hopfield and Tank Model, a N city TSP is represented by a network of N^2 fully interconnected neurons (X, i) which we can think of as an N by N matrix. Each row of this matrix corresponds to a city X and each column to a particular index i in the ordering. In the Hopfield and Tank Model, the tour ordering is represented as a *permutation* matrix, a matrix that results from the permutation of the rows (or the columns) of the identity matrix. Each row of this matrix corresponds to a city and each column to a particular position in the ordering. For example, consider a five-city TSP, then the tour $DACEB$ (that is, start from city D then visit A, C, E, B and finally return back to D) is represented as the matrix in Table 1. We will say that a Hopfield-Tank network converged to this tour when the network converges to the above permutation matrix. More discussion regarding the specifics of the operation of the model can be found in the simulations section.

Regarding the Hopfield and Tank's method for solving the Traveling Salesman Problem, initially the research has been directed towards two very important issues, namely: How

	1	2	3	4	5
A	0	1	0	0	0
B	0	0	0	0	1
C	0	0	1	0	0
D	1	0	0	0	0
E	0	0	0	1	0.

Table 1: An Example of a Five-city TSP Using the Hopfield and Tank Model

to increase the number of converging valid tours and how to improve the quality of the resulting tours. While some investigations have been successful in their quest [BWLM88, CR89, XT91, Biz91, BI89], others have discovered that the model can be unpredictable and it may not be a viable model for the TSP. We initiate this section, for background purposes, by discussing briefly the controversy for the viability of the Hopfield-Tank neural model for the TSP. Consequently we discuss the open questions left by previous investigations and discuss to some extend our main contributions in answering the open questions.

Previous Experiments: The first investigation reporting questions regarding the Hopfield-Tank model was conducted by Wilson and Pawley. In their investigation [WP88], Wilson and Pawley have found conflicting results with respect to the original Hopfield-Tank experimentation. In particular, Hopfield and Tank have suggested that 16 of their 20 experiments converged to legitimate tours with 50% of the experiments producing the 2 shortest tours. The Wilson and Pawley simulations, using the same 10-city set Hopfield and Tank used[1] suggest that of the 100 experiments they conducted 40 did not converge within an acceptable time (acceptable time was assumed to be 1000 neuron iterations), 15 converged to states corresponding to valid tours and 45 "froze" into invalid configurations. Furthermore, they have also concluded, regarding the length of the valid tours, that the 15 tours obtained by the experimentation are slightly better than randomly selected tours. In conducting additional experimentations with ten different randomly-generated sets of 10 cities, using 50 runs per set, Wilson and Pawley found that 8% of the runs converged to a valid tour state (with proportions for a given set running between 0% and 20%). Furthermore they found that 48% froze to invalid tours and that 44% did not converge within the 1000 timeout iteration limit.

Behzad Kamgar-Parsi and Behrooz Kamgar-Parsi in [KPKP90], by reexamining the Hopfield-Tank TSP model have found –in agreement with the Wilson-Pawley investigation– that the number of times the network succeeds in finding valid solutions is considerably less than that found by Hopfield and Tank. Their findings however disagree with the Wilson and

[1]Note that the coordinates of the 10 cities Wilson and Pawley used in their experimentation has been obtained by photo-expanding a figure of the Hopfield Tank paper [HT85] and by measuring the locations directly.

Pawley findings in regards to the tour lengths. Behzad Kamgar-Parsi and Behrooz Kamgar-Parsi found that when the neural net finds a tour then the solution is "of a remarkably good quality". Additionally, Behzad Kamgar-Parsi and Behrooz Kamgar-Parsi suggest that the success rate of finding valid solutions can be improved. Finally, they suggest that the Hopfield-Tank solution does not scale well with the size of the problem (while the solutions found are still remarkably good, finding valid solutions becomes increasingly difficult as the size of the problem increases).

Van Den Bout and Miller [BI89] report that the setting of the model parameters has an influence on the tour outcome with small value parameters leading to short invalid tours while large parameter values "stiffen" the penalties to a degree that the network tends to converge to any valid tour independent of the tour length. Regarding the model parameters, the experimental work discussed in Hedge et al [HSL88] indicates that good parameter value may exist in difficult to find regions of parameter space. Modifications to avoid the difficulty of finding "good" network parameters have been attempted by [BWLM88], and also by [Szu88] which propose a modification of the Brandt et al [BWLM88] model. However Van Den Bout and Miller [BI89] indicate that such kind of solution improves the generation of valid tours but it produces notably non-optimal solutions as the size of the problem increases. Furthermore, they suggest that incremental changes involving "neural normalization, annealing and annealing with non-uniform temperatures" will improve the quality of the tours found by the Hopfield-Tank network.

Further, regarding the scalability of the model, David S. Johnson in [Joh90, Joh87] suggests that the Hopfield-Tank neural network approach cannot even outperform 2-Opt[2] at 30 cities. Additionally, regarding the validity of the tours produced by the TSP Hopfield-Tank model, Aiyer et al [ANF90] using eigenvalues and a number of approximations, suggest possible relationships among the various parameters in the model. They conclude that "the formulation of the connection matrix sets up a subspace, which contains only the valid solutions of the problem" and suggest that "by carefully selecting the network parameters, it is possible to ensure that the network converges to hypercube corners in this subspace". They test their analytic expressions of their parameters using experimentation and report that the simulation they conducted produces 100% valid solutions.

Regarding the instability and the sensitivity to the initial conditions Kahng [Kah89] has reported symptoms of numerical instability and sensitivity to initial conditions for the model. It is suggested that the final tour is highly dependent on the initial perturbations used for the symmetry breaking of the network. Finally, the investigation reported in W.Lin et al [LDFVP93] suggests that there is a definite influence of the number representation and the precision of the operations on the outcome of the simulations and on the tour convergence.

Open Questions and Main Results: Due to the conflicts regarding the capabilities of the original Hopfield-Tank TSP model, the question whether or not the Hopfield-Tank model [HT85] constitutes a viable approach for the TSP remains still open. Further, while

[2] 2-Opt is a heuristic for improving a TSP tour by exchanging two edges of the tour with two other edges of shorter length.

various investigators, e.g. Kahng [Kah89] and W.Lin et al [LDFVP93], have reported dependability of the final outcome of the neural network on the initial conditions and the computations carried out by the network, they fail to quantify the smallest magnitude of the changes required to produce a deviation of the network, they do not provide a complete conclusive study of such influences, and finally they do not provide an answer for the cause of this network behavior.

In this paper we provide an answer to the open questions by examining closely a related question. The related open question and the essence of our investigation is on determining the behavior the neural network has when it is expected to produce identical tours. Assuming certain initial input values and parameter setting, this open problem can be rephrased for simulation purposes to the following question:

- Does a small (negligible for all computing purposes) variation of the initial conditions have any influence on the convergence of the network and the validity and the length of the tours?

As it will be seen later, the answer to this question is affirmative and it allows us to determine the magnitude of the initial condition changes that produce a deviation from the expected behavior and provide a cause for the behavior of the neural network. In particular, the affirmative answer to this question indicates that the non-linear recursive equations describing the Hopfield-Tank TSP neural model for small (negligible) variations of the input variables cause the production of a different tour than the tour it produced before applying the insignificant initial condition variation. Given that the magnitude of the changes in the initial conditions are insignificant, it is expected that the model "absorbs" the differences and produces identical tours. Because this is not the case, the essence of our experiment is captured by stating that the Hopfield-Tank network has chaotic behavior determined by the non-linear recursive equations and indicated by the "unexpected" results when small input variations are applied in the network with all of the other parameters of the network fixed. Because an instance of the network exhibits chaotic characteristics, we postulate that in general the equations describing the Hopfield-Tank neural network are inherently chaotic.

More in particular, as it will be shown by the actual simulations, we have established the following for the Hopfield-Tank neural model for the TSP:

- The smallest possible perturbation [3] to some values in the initial conditions will alter the tour the network converged before the perturbation.

- Assuming that for a given set of values, before perturbation, the model converges to a legitimate tour, a small insignificant perturbation in some of the values in the set not only does not not produce the identical tour but it may also converge to an illegitimate tour.

[3] The smallest possible perturbation of a value corresponds to the change of the least significant bit of this value's double floating point representation. Because we conducted our experimentation with a SUN 4m/670 MP Sparc computer system, the smallest perturbation of a single value in the set of the initial values corresponds to changing the least significant bit of the mantissa of a double accuracy real.

- Finally, the effect on the initial tour the negligible perturbations on some values in the initial conditions have is that they produce *almost* all possible tour lengths when the running of the model converges to a valid solution.

The overall consequence of our findings regarding the viability of the Hopfield-Tank model for the TSP and the cause of the controversy surrounding the model can summarized by the following:

- The cause of the Hopfield-Tank neural network for the TSP controversy and the difficulties in reproducing results, is the chaotic behavior of the model.

- The finding of useful results for the TSP using the Hopfield-Tank network are purely casual and not to be attributed to the viability of the model. In essence the Hopfield-Tank neural network for the TSP is as viable as chaotic systems can be.

3 The Simulation Strategy and Algorithm

To prove that the Hopfield-Tank neural network model has a chaotic behavior, it is enough to show that there is at least one city problem that forces the model to exhibit such a behavior. We assume the "counter example" city problem for our experimentation to be the same city problem used in the experimentation by Hopfield-Tank in [HT85] having the coordinates determined by Wilson and Pawley in [WP88] and reported in Table 2. Further, given that we are interested in proving via simulation that the neural network corresponding to the Hopfield-Tank city problem [HT85] is chaotic and to show that in general the model is not viable, it is enough to show that an instance of the network for the city problem we consider has these properties. This type of experimentation is sufficient because if there is an instance of the model that exhibits a chaotic behavior, then it can be concluded that in general the model is chaotic and thus also not viable. Note that our goal is not to show that the entire space of the definition of the network (comprising the initial conditions and parameters) is chaotic as such an experimentation is difficult [4] and it possibly does not add to the overall conclusions. Further we are not interested in investigating the possibility of a set of values for which the Hopfield-Tank neural network for the TSP does not exhibit a chaotic behavior. This type of investigation, if it is of any interest, can be viewed as a new research topic and possibly be considered in a later investigation.

In order to conduct our experimentations we consider the instance of the network in which the various parameters of the Energy equation (e.g. the values of the parameters A, B Γ and Δ) of the model [HT85] of the network are invariant and by considering variation on the initial conditions. We assume throughout the experimentation the parameter values to be the same as the ones used by Hopfield-Tank in [HT85]. The specifics of the entire experimentation follows.

The Hopfield-Tank Model : As discussed in the background section, the Hopfield-Tank network for an N-city TSP consists of an N-by-N array of neurons. For every neuron

[4] This experimentation requires the consideration of all parameters considered in their entire space.

0.4000	0.4439
0.2439	0.1463
0.1707	0.2293
0.2293	0.7610
0.5171	0.9414
0.8732	0.6536
0.6878	0.5219
0.8488	0.3609
0.6683	0.2536
0.6195	0.2634

Table 2: Coordinates of the 10 city problem in the unit square from Hopfield and Tank.

(X, i), where X is a variable that takes values from the set of the cities C and i is an index from the range $[1, .., N]$, we consider a current state input voltage $U_{X,i}$ or "mean soma potential" [Hop84] and an output voltage $V_{X,i}$. A sigmoid monotonic function $V_{X,i} = g(U_{X,i})$ describes the output voltage $V_{X,i}$ of neuron (X, i) due to an input voltage $U_{X,i}$. The output of each neuron (X, i) is connected to the input of each other neuron (Y, j) through a resistance of value $T_{(X,i)(Y,j)}$. At time step n, current state voltages $U^{(n)}$ and output voltages $V^{(n)}$ are two N-by-N matrices, (denoted here also as the U and V matrices), while the connectivity matrix T is N^2-by-N^2 matrix.

The equation of motion describing the time evolution of the network, is the differential equation:

$$\frac{\partial U}{\partial t} = -\frac{U}{\tau} + T\mathrm{vec}(V) + I \tag{1}$$

where τ is a global coefficient depending on the neuron capacitance and resistance, vec() is an operator that turns a matrix into a vector by stacking its columns on top of each other starting from the first one [HS81]. I is an N-by-N matrix, $I_{X,i}$ is the offset bias that is externally supplied to the input of neuron (X, i). By integrating this equation, i.e. solving the Initial Value Problem that is defined by Equation 1 and the initial value of matrix U, we can simulate the operation of the HT network on a computer.

Hopfield and Tank in their paper[HT85] propose an energy function E that should describe their network so as to be able to solve the N-city TSP with distances d_{XY}.

$$\begin{aligned} E(V) &= \frac{A}{2}\sum_X \sum_i \sum_{j \neq i} V_{Xi}V_{Xj} + \frac{B}{2}\sum_i \sum_X \sum_{Y \neq X} V_{Xi}V_{Yi} + \frac{\Gamma}{2}\left(\sum_X \sum_i V_{Xi} - N\right)^2 \\ &+ \frac{\Delta}{2}\sum_X \sum_{Y \neq X} \sum_i d_{XY}V_{Xi}(V_{Y,i+1} + V_{Y,i-1}) \end{aligned} \tag{2}$$

where A, B, Γ and Δ are constants chosen *ab initio* in order to "assist" the network converge to a valid and short tour. The minimizers of the energy function E are $2N$ different permutation matrices V that correspond to the shortest TSP tour (assuming that such a tour is unique). That is because the first term of the sum is zero only when every row of V has at most one nonzero value. Similarly, the second term is zero when every column of V has at most one nonzero value. The third term is zero when the sum of the nonzero terms of V is equal to N, the number of cities. Finally, the fourth term is equal to the length of the tour scaled by $\Delta/2$ when V is a permutation matrix. In order to constrain the values of the elements of matrix V in the range $[0, ..., 1]$, Hopfield and Tank propose that the sigmoid monotonic function $V = g(U)$ be the hyperbolic tangent:

$$V_{Xi} = \frac{1}{2}(1 + \tanh(\frac{U_{Xi}}{u_0})) \tag{3}$$

where u_0 is the gain factor of the hyperbolic tangent.

Moreover, by using the energy function $E(V)$ as a guide, Hopfield and Tank derive the value that matrix T should have, in order to make the network operate according to the target energy function defined in equation 2. Thus the general equation of motion in equation 1 is transformed into:

$$\begin{aligned} \frac{dU_{Xi}}{dt} &= -\frac{U_{Xi}}{\tau} - A\sum_{j\neq i} V_{Xj} - B\sum_{Y\neq X} V_{Yi} \\ &- \Gamma\left(\sum_Y\sum_j V_{Yj} - N_+\right) - \Delta\left(\sum_Y d_{XY}(V_{Y,i+1} + V_{Y,i-1}\right) \end{aligned} \tag{4}$$

that is the specific equation of motion for the TSP solving HT network .[5] For a more extensive treatment and derivations the interested reader is refereed to [HT85].

The algorithm we use for the simulation of the model can be briefly described by the following:

The HT Network Simulation : We use the synchronous update scheme.

- Initialize the matrix U such that U_{Xi} gets $1/N_+ + \delta_{Xi}$ where each δ_{Xi} is randomly chosen uniformly in the interval
$$-0.1u_0 \leq \delta_{Xi} \leq 0.1u_0.$$
- Repeatedly iterate
  - Calculate the values of matrix V according to equation 3.
  - Calculate the TSP tour length $\lambda_{d_{XY}}(V)$.

[5]Our equation differs from the one suggested in [HT85, WP88] on the choice of the free variable of the summation of the third term $\Gamma\left(\sum_Y\sum_j V_{Yj} - N_+\right)$. They have used $\sum_X$, instead of $\sum_j$, which appears to be a bad choice for a free variable since the variable X is bound (this term belongs to the calculation of $\frac{dU_{Xi}}{dt}$).

– Using the equations of motion for this network, calculate the quantities $\frac{dU_{Xi}}{dt}$ for all (X, i) from equation 4.

– Update the values of matrix U

$$U_{Xi} = U_{Xi} + \frac{dU_{Xi}}{dt}\delta_t.$$

- Repeat until the length of the TSP tour $\lambda_{d_{XY}}(V)$ remained unchanged for the last 200 iterations.

The various parameters of the simulations and some specifics of the simulation algorithm are discussed bellow.

Initial values: The parameters used for the simulations are the same ones proposed in the original paper [HT85] and they specifically have the following values:

$$A = B = 500,\ \Gamma = 200,\ \Delta = 500,\ N_+ = 15,\ \tau = 1,\ u_0 = 0.02 \tag{5}$$

The value of $\delta_t = 10^{-5}$ was used by Wilson and Pawley in[WP88], Hopfield and Tank were not explicit about it, thus we used this value also in our experiments. Each simulation, unless otherwise stated, begins with all neurons assigned equal "voltages" of value $1/N_+$ plus or minus some noise that does not exceed 10 percent of u_0.

Length function : In order to have a linear measure between the different runs, a length function $\lambda_{d_{XY}}(V)$ is used to calculate the length of a tour for a given activation matrix V and a distance function d_{XY}. When V does not correspond to a valid tour, the value of $\lambda_{d_{XY}}(V)$ is just a real number, and no other significance should be given to it except for differentiating between two illegal tours.

The value of $\lambda_{d_{XY}}(V)$ is calculated by finding the index of the maximum element of every row of the activation matrix V and considering this index as the order in which the cities are visited. The length of this path, including the length of moving from the last city visited to the first one, is the value $\lambda_{d_{XY}}(V)$.

Notice that it is possible two different rows of V to have their maximum element on the same column. In this case the same city will be visited twice and therefore V does not represent a valid TSP tour.

Termination : Instead of terminating the execution of the simulation algorithm after a pre-specified number of updates, we have decided to terminate the simulation after the network does not change the value $\lambda_{d_{XY}}(V)$ for 200 consecutive iterations. In [Hop84], it has been suggested that networks with symmetric connections, i.e. networks with equations of motion with matrix T symmetric, always converge to a stable state. Since the TSP network has symmetric connections when the distance function d_{XY} is symmetric, –as it is our case with the cities having Euclidean distances–, it is therefore suggested that its output converges. For a simulation point of view, the numerical noise might inhibit the detection of convergence. Instead of looking at the values of the matrices U and V, and a matrix norm of them, as Wilson and Pawley in [WP88] did, we are looking at the value of the function $\lambda_{d_{XY}}(V)$. That is, we implicitly project the current V from the interior of the unit hypercube $[0, ..., 1]^{N^2}$ to a near-by vertex of it (that obviously belongs to $\{0,1\}^{N^2}$). Please note that we do this only for the purpose of calculating the length function value and not the update values of U. This action smoothes the numerical noise, so whenever there is no change in the value of $\lambda_{d_{XY}}(V)$ for a long enough period, we can assume we reached the convergence point. The choice to discontinue a simulation after $\lambda_{d_{XY}}(V)$ remained unchanged for 200 iterations is arbitrary and we did not test whether the network will converge to a different state if we had waited a longer time.

We consider this length as the convergence length, and the resulting tour is legitimate only when the final V is a *near-permutation*, that is, whenever element $V_{i,j}$ is the maximum element of row i, to also be the maximum of column j. In this way we know we are really close to a vertex of the N^2-dimensional hypercube that corresponds to a valid tour.

In [HT85] it is required the resulting activation matrix V to be a permutation matrix. For this study we relax this requirement and only want V to be a near-permutation [6] for the simple reason that a near permutation matrix contains enough information to define a tour, even in the case that U fluctuates due to numerical noise, between different values, to still be able to detect convergence, when those fluctuations correspond to equivalent tours.

4 Experiments

To validate our claims regarding the chaotic behavior we performed two diverse sets of experiments. In the first experiment (Experiment # 1) we have attempted to repeat the results of other investigations and to determine data points to be used as a reference for various conclusions for the second experimentation (Experiment # 2). In the second experiment we prove the claim indicating that the smallest variation to initial values allowable by the numerical accuracy of Sun 4m/670 MP Sparc, the computer used in our experiments, can result in almost any imaginable valid and invalid tour. The second experiment clearly suggests that the model produces unpredictable results when it should be "absorbing" the negligible perturbation and behaving the same as before the addition of the perturbation of the initial values occurred. More in detail the following has been been established by the experiments:

Experiment # 1 : As a first experiment, we run 10,000 experiments by starting with an initial state U equal to $1/N_+$ plus-or-minus noise that does not exceed 10% of u_0. In Figure 1, we plot the lengths of the valid tours found. It is interesting to notice that only 2,404 from the 10,000 runs converged to a legal tour despite our relaxation to accept as valid a near permutation. This percentage (24%) is higher than the percentage reported by Wilson et al. in [WP88] that was only 15%, but a lot smaller from the Hopfield and Tank

[6] Whenever element $V_{i,j}$ is the maximum element of row i, then we assume to be also the maximum of column j.

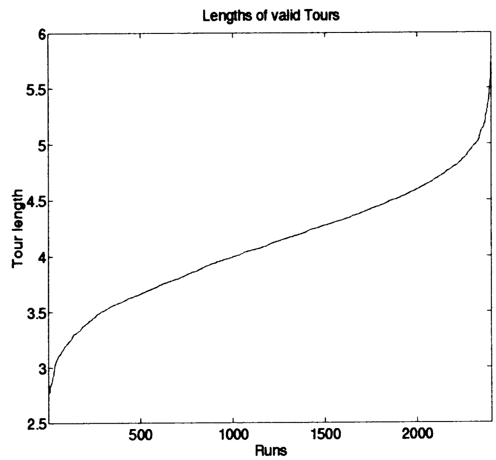


Figure 1: Plot of the lengths of only the valid tours from the 10,000 runs, in ascending order.

paper where they claim that 16 out of 20 runs (80%) converged to legitimate tours. We also observed that the resulting lengths span in the whole range of possible lengths for this specific problem, without giving any indication that the algorithm converges to any near optimum tour. This is evident in the distribution of the lengths of valid tours in figure 2. Clearly the results of other investigations, such as the investigation by Kahng [Kah89], suggesting that the final tour is highly dependent on the initial perturbations used for the symmetry breaking of the network are clearly verified also by this first experiment. One question that can be asked regards the number of runs in an experiment. In essence it can be asked if 10000 runs are enough to characterize the model. To investigate this question we have plotted the statistics of the model for every 50, 100, 500, and 1000 runs. The only basic conclusion that can be reached is that 10000 runs do not have different characteristics than the any of the other runs.

Figures 3 and 4 display some of the statistics of the network simulations for the runs, seen in groups of 1000 runs. The additional plots for the other runs can be found in Appendix A. Almost all the sets of runs have a minimum tour close to the shortest TSP tour found in our experimentation, and a maximum tour of a length around 5.5. Also notice that the mean and the median tour lengths literally coincide at a value just above 4. The overall conclusion is that the averages reported in the Figures, really represent the performance of the network, and they do not come from the accumulation of different simulations. An interesting implication of this finding is that a "small" number of runs is sufficient to characterize the overall behavior of the model.

Experiment # 2 The previous experiment suggests that in general the Hopfield-Tank results are not repeatable. Further, the experiment suggests that the number of runs required to validate the results is not very relevant, as a small number is sufficient to characterize the

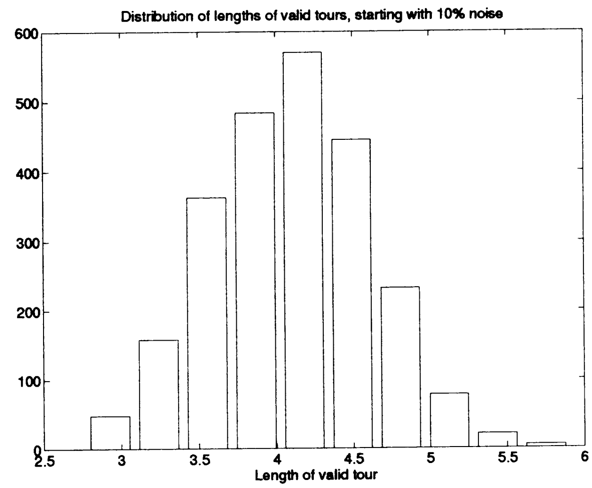


Figure 2: Distribution of the lengths of only the valid tours from the 10,000 runs, where the initial point had 10% noise.

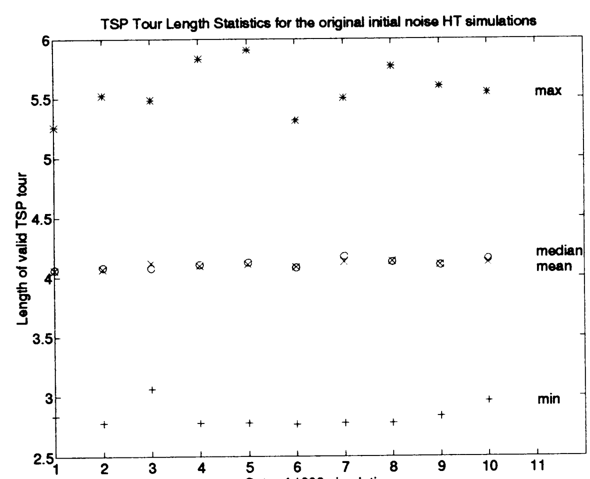


Figure 3: TSP Tour Length Statistics for the HT simulations, where the initial point had 10% noise.

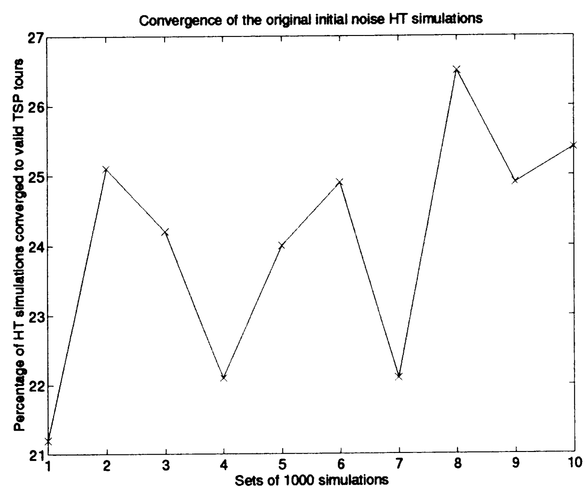


Figure 4: Percentage of HT simulations converged to valid TSP tours, where the initial point had 10% noise.

model. In particular the experiment also indicated that :

- All three investigations (the Hopfield-Tank [HT85], the Wilson/Pawley [WP88], and experiment one) do not agree on the overall behavior of the model.
- Both the Kahng [Kah89] investigation and experiment # 1 , indicated the final tour is highly dependent on the initial perturbations used for the symmetry breaking of the network.

We note here that we conducted additional experiments. In those experiments we reduced the noise level from the initial point of 10% down to 1% or even lower to 10^{-1}%. We run 10000 simulations for each noise level. The overall conclusions for the additional experiments have been the same as in Experiment # 1 and thus we do not report the details of those experiments. The only possible interesting point that may add some additional insight in this neural network paradigm is the fact that we have observed that the number of the runs that converged to valid tours increased and moreover the overall quality of the resulting tours improved with the decreasing of the noise. While the results of the Experiment # 1 and of the additional experiments we run can be viewed as a possible indication for a chaotic behavior, such evidence is not conclusive as it can always be argued that the model does not produce "unexpected" results.

To clarify the previous statement consider the following: Generally speaking, (assuming that the equations describing the model are deterministic, recursive and non linear, and assuming that the equations are modeling the TSP) for the model not to be chaotic the following must occur:

- Assume that on a certain input, the model converges to a certain tour.

- Assume that numerical changes occur in the initial value conditions. Assume that the changes are being numerically insignificant from the initial values.

- Then it is expected that the insignificant changes are ignored by the model and the model should be producing the identical tour with the run that had no perturbation in its input values.

To provide evidence suggesting that the Hopfield-Tank neural network for the TSP has a chaotic behavior, we must show that a small (negligible for all computing purposes) variation of the initial conditions has a significant influence on the convergence of the network, validity and length of the tours. If this is true then it can be concluded that the network has a chaotic behavior, because the experiment indicates that the non linear recursive equations describing the Hopfield-Tank TSP neural model for small (negligible) variations of the input variables, produce "unexpected" results ("unexpected" being the convergence of the model to a state that does not correspond to the tour the network converged with the unperturbed input values).

To show that this is the case, we have to address the question of what constitutes negligible from the computational point of view and then show that the model will produce "unexpected" results. Clearly, assuming a large number representation, the smallest possible "noise" that is allowed by the number representation of the simulation tool can be viewed as a negligible number when used to differentiate two values. In conducting the second experiment we reduced the initial noise to the absolute minimum that our machinery will allow, that is adding or subtracting 1 from the least significant bit (LSB) of the mantissa of the double accuracy (8 byte) floating point representation of the suggested initial value $1/N_+$. This perturbation of the initial condition results in a set of two values $\{\alpha, \beta\}$ that are almost indistinguishable from the initial $1/N_+$ value. We run as many experiments as practicable[7], which in our case it happened to be 21,249 from which 7,185 converged to valid tours (that is 33.81%). This is an improvement from the 24% rate we got initially but it clearly suggests nothing as the model should have converged to the same tour all the runs. Again as we observe in Figure 5, the resulting lengths span the entire range of possible lengths we have found during the experimentation for this model problem. Clearly as the figure suggests the negligible noise of the level of 10^{-14}% is more than enough to generate a perturbation that produces "unexpected" results.

Even though there is no true meaning, (producing better tours when the model should be producing identical tours is hardly of any meaning) it can be observed that the valid tours are of much better quality than the tours observed in the previous experiment! Clearly the distribution of the lengths for the valid tours in figure 6 shows a clear preference to shorter tours. Also similarly to the previous experiment, by breaking the simulations into chunks of 1000 runs we see that the behavior of the network is the same as reported by the total statistics.

[7] As the previous experimentation suggests and as the figures to follow will also suggest, the number of runs does not influence the overall outcome of the conclusions.

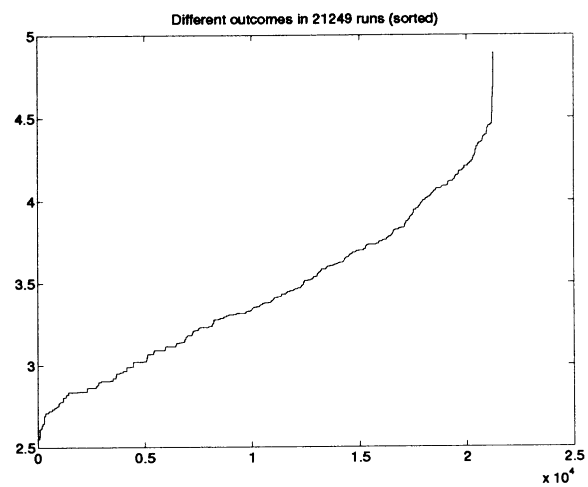


Figure 5: Different outcomes (lengths) resulting from a starting point of only $10^{-14}\%$ noise.

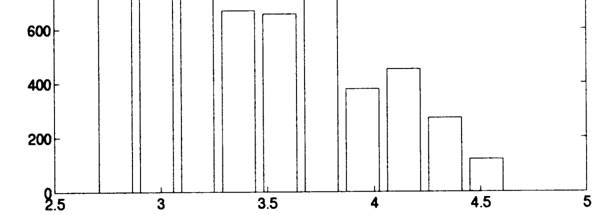


Figure 6: Distribution of the lengths of only the valid tours where the initial point have $10^{-14}\%$ noise.

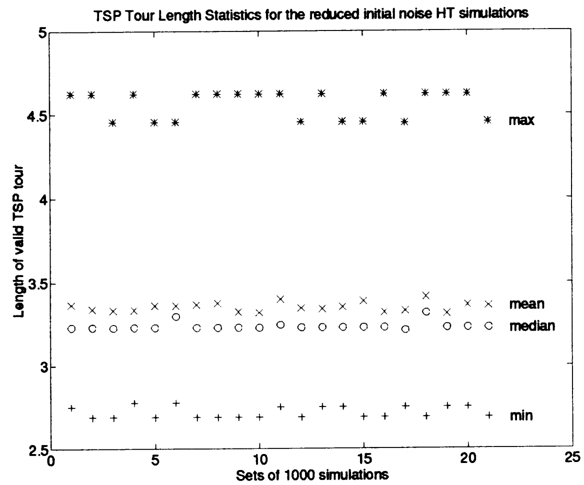


Figure 7: TSP Tour Length Statistics for the HT simulations, where the initial point had minimum noise.

Figures 7 and 8 display those statistics, additional information can be found in Appendix A which includes statistics for every 50, 100, and 500 runs, as for the experiment # 1.

All the sets of 1000 runs contain very good tours, the maximum tours are of length approximately 4.5, compared with the maximums in Experiment #1 that are larger than 5.5; and more impressively, the mean and median tour lengths are at about 3.25, compared with the corresponding ones of Experiment #1 that are above 4.0. Also the median is always smaller than the average. While we do not have any explanation for this outcome, the bottom line for the experimentation is that the negligible perturbation will produce similar results as other larger perturbations. Further, while the network should be converging to a single tour it is observed that the model converges to almost all possible combinations. This behavior with the setting of the experiments allows us to suggest that for the city problem we consider the neural model has a chaotic behavior.

Additional experimentation : In order to appreciate the effects the chaotic behavior has on the outcome of a run we did two additional experiments. The main theme of those experiments is the consideration of some "particular" conditions. In the first experiment we considered three different initial U matrices and observe closely the run. We denoted those matrices as U_1, U_2 and U_3. Their values are shown in Tables 4, 5, and 6 in Appendix B. They consist of two different values $\alpha = (3fb1111111111112)_{16}$ and $\beta = (3fb1111111111110)_{16}$. Note that the α and β representations are in hexadecimal. The numbers for the initial conditions are either α or β and they are obtained by adding and subtracting 1 to the least significant bit of the mantissa of $1/N_+$ with the entire $1/N_+$ 64-bit number being $(3fb1111111111111)_{16}$ in hex.

In following closely the trace (see Figure 9) of the three runs we observe that there are multiple common "convergence" points however at the end two out of three runs converge

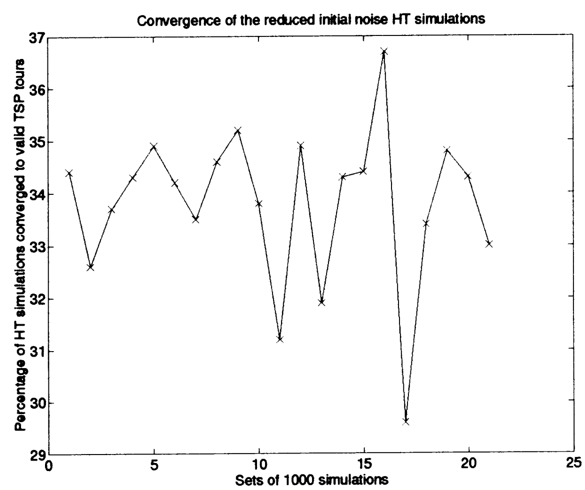


Figure 8: Percentage of HT simulations converged to valid TSP tours, where the initial point had minimum noise.

at the same length value. This scenario occurs even though the difference between the three initial states is, in any numerical respect, insignificant. Actually, the difference between all 21,249 initial states is insignificant. Despite that, the Hopfield-Tank algorithm terminated in almost all possible different tours, valid and invalid; in the same way as when the initial states were coming from the larger initial region proposed by them.

In experiment # 2 we have shown that a small deviation of some initial conditions generate unpredictable results exhibiting a chaotic behavior. In the experiment we have assumed that the initial conditions have been perturbed at random. An interesting question to investigate is the following:

- How many values should be perturbed for the model to exhibit a chaotic behavior?

To answer this question we operated as follows: First we chosed the initial conditions that resulted in a simulation that converged to a legitimate tour. We checked on the neighborhood of the initial starting point U_1 of the previous experiment. The tour the matrix U_1 converges happens to be the shortest tour our runs converged to. We changed exactly one value of the initial matrix: if it was α it was changed to β and vice-versa. We have performed all 100 possible experiments. That is, from the initial matrix U_1 we generate $10 \times 10 = 100$ initial matrices U that only differ from U_1 at the position i, j. The lengths of the tours that resulted from running the Hopfield-Tank network with these matrices as initial states, are displayed on Table 3. In the table the position (i, j) reports the outcome of the simulation for the initial condition matrix with the change of the value $U_1(i, j)$ to the value $U(i, j)$. Invalid tours are indicated by using italics. The outcome of this experimentation is approximately the same as for the experiments # 1 and # 2. The over all conclusion is that it is enough to insignificantly perturb one of the initial values for the network to exhibit a chaotic behavior.

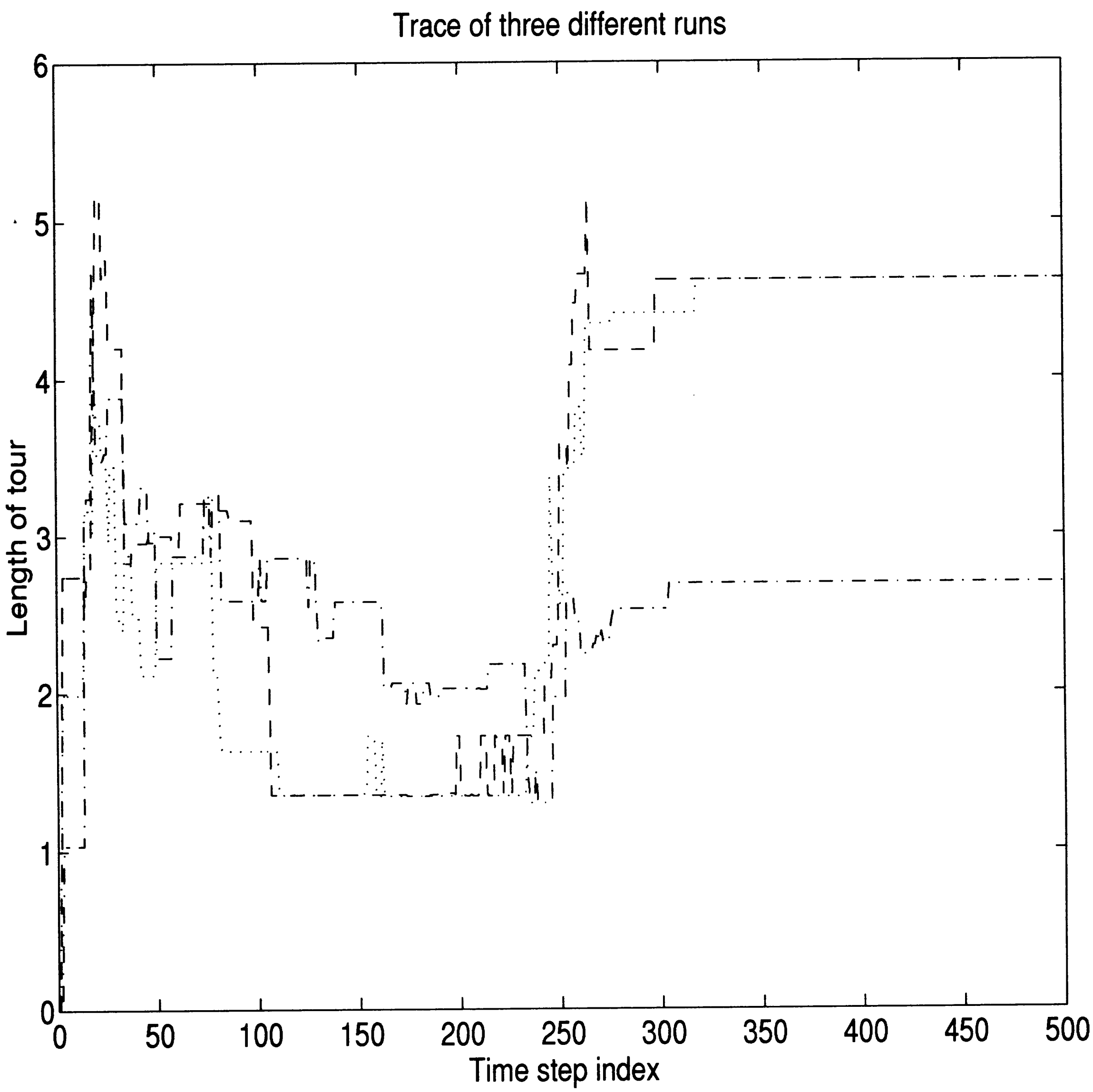


Figure 9: Trace of three different runs.

1	3.902	3.019	*2.803*	4.342	*3.180*	*3.785*	*3.595*	2.836	3.887	*3.360*
2	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691
3	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691
4	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691
5	2.691	2.904	2.691	*2.753*	2.778	3.019	2.752	2.691	2.691	3.019
6	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691	2.691
7	*4.206*	3.427	3.735	3.514	*2.873*	*3.067*	*3.446*	*2.753*	3.735	4.035
8	2.752	2.752	3.019	2.862	3.735	3.584	2.904	2.778	2.862	3.584
9	3.259	3.609	2.839	*2.829*	4.018	*2.884*	2.904	2.904	2.862	3.114
10	3.761	3.735	2.904	2.752	*3.785*	2.778	3.091	2.947	3.833	3.761

Table 3: Tour lengths resulting by changing the (i,j) entry of U_1. Numbers in italic correspond to invalid tours.

In particular we observed that the network converged only 44% of the runs to the expected tour with length 2.691 when it was expected to converge to this tour always. Further there are a number of different tours produced by this insignificant change in the initial conditions. Finally, we have observed that the change of value on the lines 2, 3, 4 and 6 do not change the convergence length. The change in line 5, gave either the same tour as the original starting state or some different tours of good quality. The change in the rest of the lines gave arbitrary and invalid tours. The over all conclusion again is that the model has a chaotic behavior even for the smallest single value change.

5 Conclusions

In this paper, we considered the Hopfield-Tank model for the TSP and investigated its viability as a model and its capabilities to produce good results to the hard optimization problem. Furthermore, we investigated the reasons behind the difficulty of obtaining verifiable results. The basis of the investigations we conducted were instances of the initial values where the behavior of the neural network was expected to produce identical tours. We have assumed that the set of initial values producing identical results need be very close, (close being determined by the tool used in the experimentation), and we simulated the behavior on a 10-city problem. The 10-city problem has been used by the experimentation to show that there is at least one city problem exhibiting chaotic behavior in a certain initial condition settings. In essence our experimentation indicates that the non linear recursive equations describing the Hopfield-Tank TSP neural model for small (negligible) variations of the input variables, cause the production of a different tour than the tour it produced before applying the insignificant initial condition variation. Given that the magnitude of the changes in the initial conditions are insignificant, it is expected that the model produces the same identical tour. Because this is not the case, the essence of our experiment is captured by stating that the Hopfield-Tank network has chaotic behavior determined by the non linear recursive equations and indicated by the "unexpected" results when small input variations are applied in the network with all of the other parameters of the network fixed. Because an instance of the network exhibits chaotic characteristics we postulate that in general the equations describing the Hopfield-Tank neural network are inherently chaotic.

More in particular, after running more than 65000 simulations, we have established the following for the Hopfield-Tank neural model for the TSP:

- The smallest possible perturbation of some initial values will alter the tour corresponding to the initial conditions before the perturbation.

- Assuming that for a given set of values, before perturbation, the model converges to a legitimate tour, a small insignificant perturbation in some of the values may cause the model to converge to an illegitimate tour.

- The effect on the initial tour the negligible perturbations on some values in the initial conditions has, is that it produces *almost* all possible tour lengths when the running of the model converges to a valid solution.

- We have shown that the model behaves as indicated by the previous three items even with the perturbation of a single value (the perturbation also being numerically insignificant for computation purposes).

- Finally, our experiments suggest that a small number of runs is sufficient to characterize the behavior of the Hopfield-Tank network model for the TSP.

The overall consequence of our findings regarding the viability of the Hopfield-Tank model and the cause of the controversy surrounding the Hopfield-Tank model for the TSP can summarized by the following:

- The cause of the Hopfield-Tank neural network for the TSP controversy and the difficulties in reproducing results is the chaotic behavior of the model. Consequently the controversy should have been expected.

- The finding of useful results for the TSP using the Hopfield-Tank network are purely casual and not to be attributed to the viability of the model. In essence the Hopfield-Tank neural network for the TSP is as viable as chaotic systems can be.

References

[ANF90] S. V. B. Aiyer, M. Niranjan, and F. Fallside. A theoretical investigation into the performance of the hopfield model. *IEEE Transactions on Neural Networks*, 1(2):204–215, 1990.

[BI89] D. E. Van Den Bout and T. K. Miller III. Improving the performance of the hopfield-tank neural network through normalization and annealing. *Biological Cybernetics*, 62:129–139, 1989.

[Biz91] A. R. Bizzari. Convergence properties of a modified hopfield-tank model. *Biological Cybernetics*, 64:293–300, 1991.

[BWLM88] R. D. Brandt, Y. Wang, A. J. Laub, and S. K. Mitra. Alternative networks for solving the traveling salesman problem and the list-matching problem. *IEEE International Conference on Neural Networks*, II:333–340, July 1988.

[CR89] R. Cuykendall and R. Reese. Scaling the neural tsp algorithm. *Biological Cybernetics*, 60:365–71, 1989.

[Hop84] J. J. Hopfield. Neurons with graded responce have collective computational properties like those of two state neurons. *Proc. Nat. Acad. Sc. U.S.*, 81:3088–3092, 1984.

[HS81] H. V. Henderson and S. R. Searle. The vec-permutation matrix, the vec operator and kronecker products: A review. *Linear and Multilinear Algebra*, 9:271–88, 1981.

[HSL88] S. U. Hedge, J. L. Sweet, and L. W. Levy. Determination of parameters in a hopfield/tank computational network. *Proc IEEE Int Conf Neural Netw*, II:291–298, 1988.

[HT85] J. J. Hopfield and D. W. Tank. Neural computation of decisions in optimization problems. *Biological Cybernetics*, 52:141–152, 1985.

[Joh87] D. S. Johnson. More approaches to the travelling salesman guide. *Nature*, 330:525, December 10 1987.

[Joh90] D. S. Johnson. Local optimization and the travelling salesman problem. *Proc 17th Colloq on Automata, Languages and Programming*, pages 446–461, 1990.

[Kah89] A. B. Kahng. Traveling salesman heuristics and the embedding dimension in the hopfield model. *Proc IEEE Int Joint Conf Neural Netw*, I:513–520, 1989.

[KPKP90] B. Kamgar-Parsi and B. Kamgar-Parsi. On problem solving with hopfield neural networks. *Biological Cybernetics*, 62:415–423, 1990.

[LDFVP93] W. Lin, J. Delgado-Frias, S. Vassiliadis, and G.G. Pechanek. Machine and bit precision on the hopfield neural network model for the tsp. *Proc IEEE Int Conf Neural Netw, Nagoya, Japan*, pages pp. 1516–1519, 1993.

[PS82] Ch. Papadimitriou and K. Steiglitz. *Combinatorial optimization, Algorithms and complexity.* Prentice Hall, 1982.

[Szu88] Szu. Fast tsp algorithm based on binary neuron output and analog input using the zero-diagonal interconnect matrix and necessary and sufficient constraints of the permutation matrix. *Proc IEEE Int Conf Neural Netw*, II:259–266, 1988.

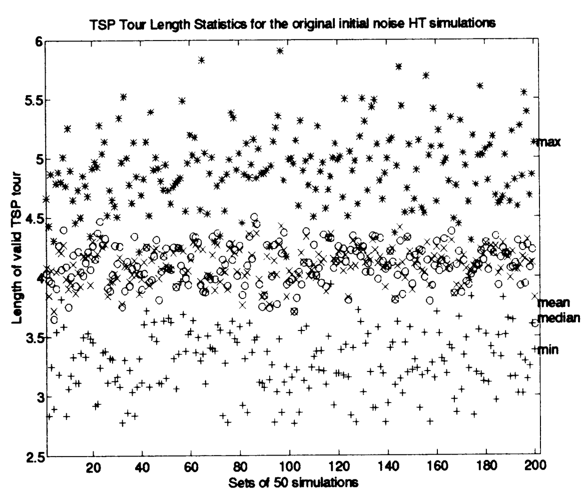
[WP88] G. V. Wilson and G. S. Pawley. On the stability of the travelling salesman problem algorithm of hopfield and tank. *Biological Cybernetics*, 58:63–70, 1988.

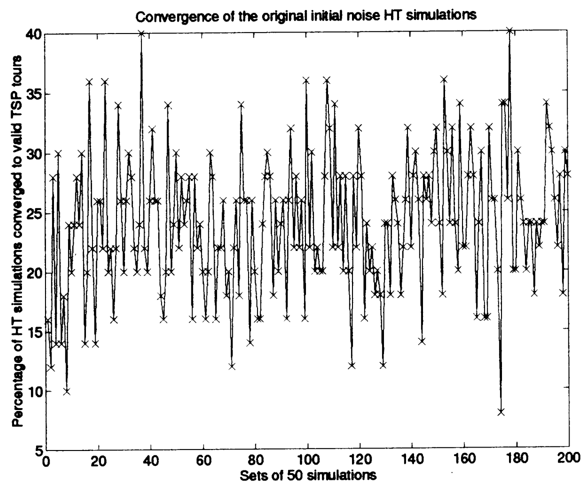
[XT91] X. Xu and W. T. Tsai. Effective neural algorithms for the traveling salesman problem. *Neural Networks*, 4:193–205, 1991.

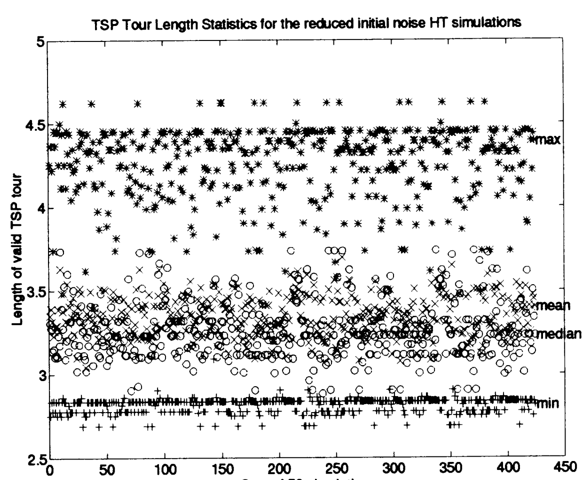
A Appendix

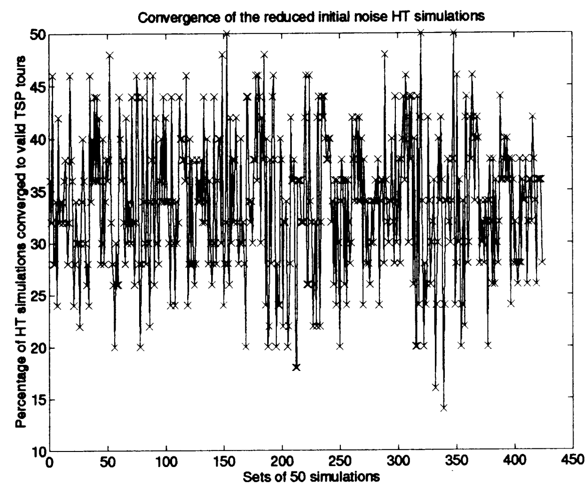
In this Appendix we report additional charts supporting the claim suggesting that: A "small" number of runs is sufficient to characterize the overall behavior of the model. The various figures plot statistics of the TSP model for every 50, 100, and 500 runs. The runs are performed on two different scenarios. The first scenario involves a "noise" level equal to the "noise" level used by the original Hopfield-Tank proposal. The second scenario involves a "noise" level that is negligible for all computing purposes. The charts are "labeled" with either "original noise" or "reduced initial noise" for the first and second scenario respectively. For every scenario we plot two different figures. The first figure regards the lengths of the TSP tour, the second reports the percentages the model converged to. As the figures suggest the model even for a modest number of runs (modest being 50 or more but less than 1000) supports the following conclusions:

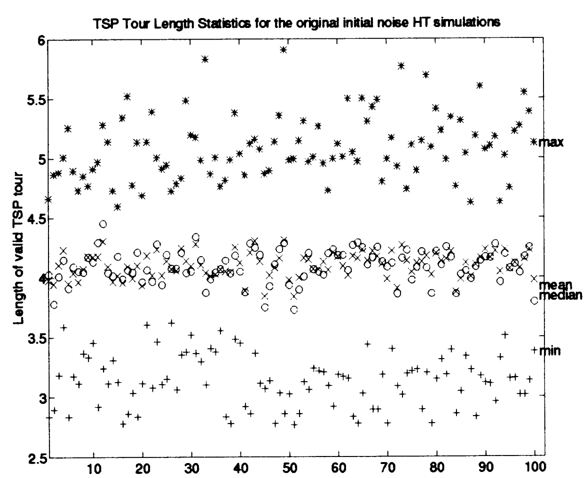
- The model does have a very poor convergence rate (seldom exceeds the 50% convergence rate).
- It does not converge to the same tour when it supposed to.
- It assumes a plurality of tour lengths when it converges to a tour.

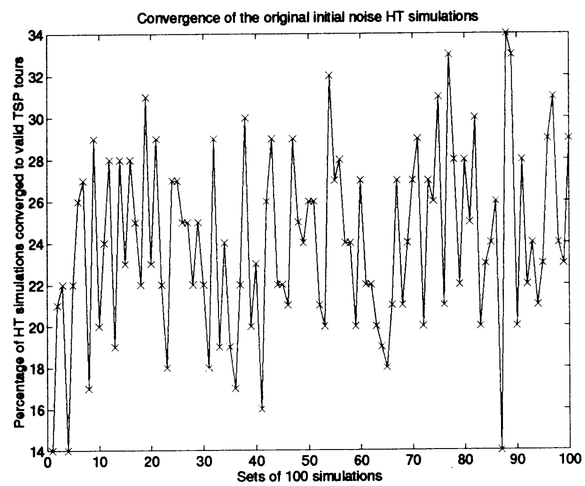


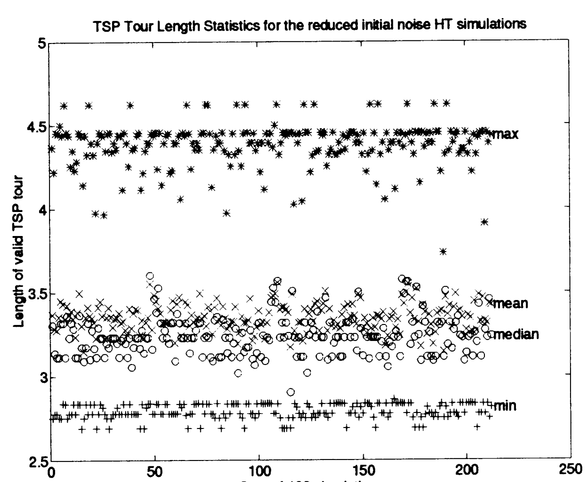

Convergence of the original initial noise HT simulations
Percentage of HT simulations converged to valid TSP tours
Sets of 50 simulations

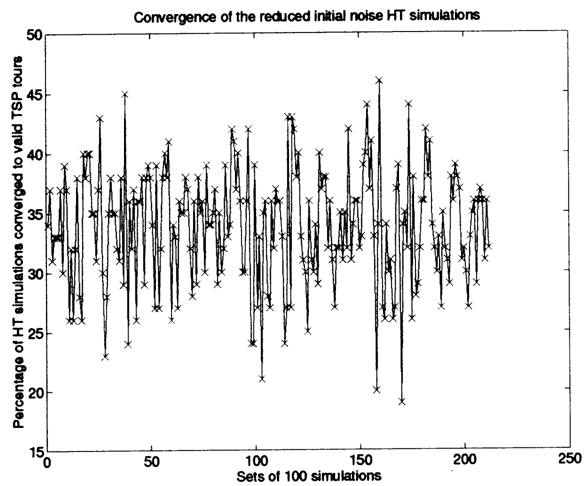

TSP Tour Length Statistics for the reduced initial noise HT simulations
Length of valid TSP tour
max
mean
median
min
Sets of 50 simulations

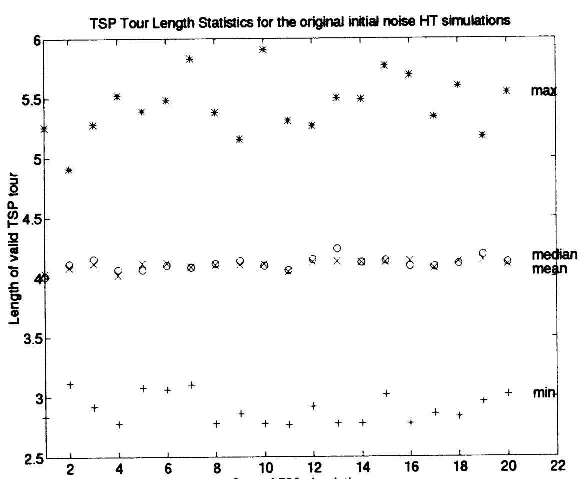

Convergence of the reduced initial noise HT simulations
Percentage of HT simulations converged to valid TSP tours
Sets of 50 simulations

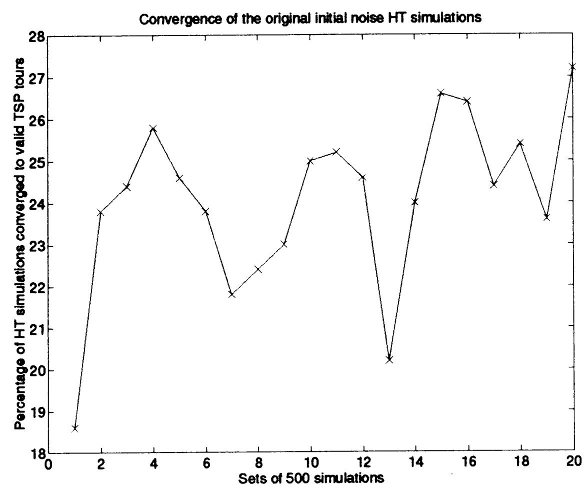

TSP Tour Length Statistics for the original initial noise HT simulations
Length of valid TSP tour
Sets of 100 simulations
max
mean
median
min

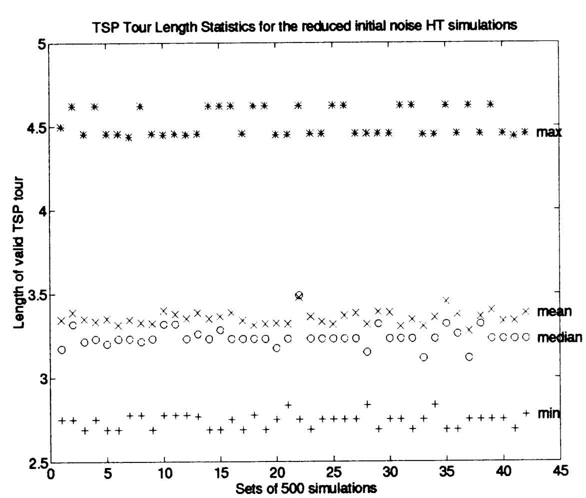

Convergence of the original initial noise HT simulations
Percentage of HT simulations converged to valid TSP tours
Sets of 100 simulations
34
32
30
28
26
24
22
20
18
16
14
0
10
20
30
40
50
60
70
80
90
100

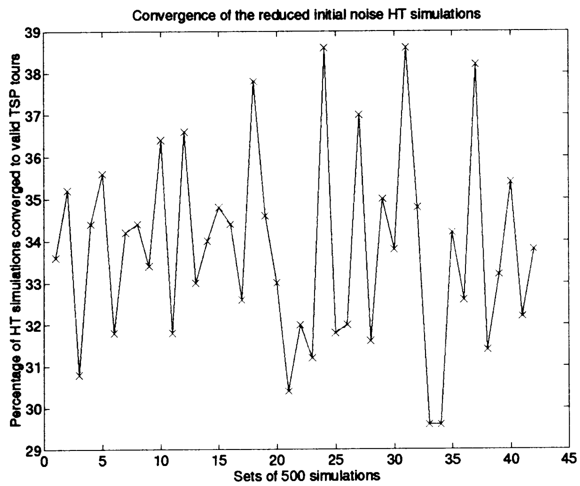

TSP Tour Length Statistics for the reduced initial noise HT simulations
Length of valid TSP tour
Sets of 100 simulations
max
mean
median
min
5
4.5
4
3.5
3
2.5
0
50
100
150
200
250


Convergence of the reduced initial noise HT simulations
Percentage of HT simulations converged to valid TSP tours
Sets of 100 simulations
50
45
40
35
30
25
20
15
0
50
100
150
200
250


TSP Tour Length Statistics for the original initial noise HT simulations
Length of valid TSP tour
Sets of 500 simulations
max
median
mean
min
6
5.5
5
4.5
4
3.5
3
2.5
2
4
6
8
10
12
14
16
18
20
22


Convergence of the original initial noise HT simulations
Percentage of HT simulations converged to valid TSP tours
Sets of 500 simulations


TSP Tour Length Statistics for the reduced initial noise HT simulations
Length of valid TSP tour
Sets of 500 simulations
max
mean
median
min



B Appendix

α	β	β	α	β	β	β	β	α	α
β	α	α	α	β	α	β	β	β	α
β	α	α	β	β	α	α	α	β	β
α	β	α	β	α	β	β	β	α	β
α	α	β	α	α	α	β	α	β	α
α	β	α	α	β	β	α	α	α	β
α	β	β	β	α	β	β	β	α	α
α	α	α	α	β	α	β	α	β	β
β	β	β	β	β	α	β	β	β	β
α	β	α	α	α	α	α	α	β	β

Table 4: Initial matrix U_1

β	β	β	α	β	α	α	β	β	β
β	β	α	α	α	β	β	β	β	α
β	α	α	β	β	β	β	β	β	α
α	α	α	β	β	β	α	α	α	β
β	β	α	β	α	β	β	β	β	β
α	β	β	α	β	α	α	β	β	β
β	β	β	α	α	α	β	β	α	β
α	α	β	α	α	β	α	β	β	α
β	α	α	α	β	β	α	α	β	β
α	β	β	α	α	α	α	α	β	β

Table 5: Initial matrix U_2

β	β	β	β	β	β	α	β	β	β
α	β	β	β	α	α	β	α	β	β
α	β	α	α	β	α	α	α	α	α
β	α	β	α	β	α	β	β	β	β
β	β	β	α	β	β	α	β	β	β
β	β	β	α	α	β	α	β	β	β
β	α	α	β	α	α	α	β	α	β
β	α	α	α	α	α	β	α	β	β
α	β	β	α	β	β	α	β	α	α
α	β	β	β	α	α	α	β	β	α

Table 6: Initial matrix U_3