2D Oscillators for Oscillatory Neural Network to solve n-city Travelling Salesman Problem

Anugole Sai Gaurav and Vishwas Bharti
Third year Undergraduates, B.Tech, Electrical Engineering, IIT Bombay
Email: sai1999gaurav@gmail.com, wishwasbharti@gmail.com

Overview
Neural networks have been used to solve multiple optimization techniques, one such class includes constraint satisfaction problems (CSP). The Travelling Salesman problem (TSP) is one of many CSPs that can be mapped onto neural networks. Oscillatory Neural networks (ONN) are a variant of neural networks in which units are dynamical systems that oscillate in absence of inputs from other units, resembling the real neurons. The representation of TSP in such a network seems more natural, wherein the tours are defined in terms of synchronization patterns among the oscillating units. A Hopfield-Tank representation was proposed, which maps n-city problems to a cost function based on n^2 interacting units. The complexity was reduced to n units based ONN with effective cost function. The resultant phase order in both cases is mapped onto a circle. However, the dynamics of the particles in this scenario is restricted and difficult to cross over. This can act as a limiting constraint and a bottleneck while solving for larger values of 'n'. Hence, additional degrees of freedom are introduced and oscillator dynamics is analysed in 2D with the motion on a torus. The cost functions in such cases have been derived and simulated to show improvement in both performance and accuracy over the previous techniques.

Introduction
The objective of an n-city TSP is to find the traversal order for a minimum Euclidean distance path covering all the cities exactly once. The total solution space for such a problem, often referred to as tours, grows exponentially in size with 'n' which makes it difficult to solve on Neumann based architecture systems. Hence, neural networks with its parallel processing ability offer a viable platform to solve such CSPs.
An artificial neural network with connection between its neural units evolves to a state that satisfies such an optimization principle based on its weights defining those connections. Replacing these neural activation units with oscillators based dynamic systems are trending and popular due to their close resemblance with the biological system. These also offer rich possibilities in terms of hardware implementation.
One such proposed method is the Hopfield-Tank based network with an energy cost-function of complexity $O(n^2)$ attempts to solve the TSP, based on the information encoded in the phase of coherent network of oscillators. The solution obtained is one which minimizes the energy function. For this purpose, stochastic gradient descent is used to achieve global minima of the cost function, along with simulated annealing scheme to escape off local minima or tours that are not necessarily shortest.
Corresponding cost function for such a network to be minimized has been suggested which incorporated similar constraints as before. This resulted in substantial decrease in the solution space, improved performance in terms of speed, computation resource, area footprint and accuracy.

Oscillator Cost function for TSP
The oscillator dynamics is mapped to an energy cost function which is minimised to obtain the optimal tours. The stochastic element in the network is further combined with simulated annealing in which random initial phase noise is added to the system that decays with time. Simulated annealing method models the physical process of heating a material and then slowly lowering the temperature to decrease defects, thus minimizing the system energy. For simulation purposes, zero-mean Gaussian noise, $\Theta(\sigma)$ with decaying σ, where σ is reduced by annealing rate α everytime.

$$z(t) \leftarrow z(t)exp(i\theta(\sigma))$$
$$\sigma(t+\Delta t) = \alpha.\sigma(t)$$

The previous methods are discussed briefly in part A and B. Thereafter, the proposed method and dynamics are explained in detail. Multiple variants for 2D Oscillators are derived and discussed too.

A. *n^2 oscillators mapping:*

This representation makes use of an $n \times n$ matrix of coupled oscillators as described in **Fig.1**. A column represents position in the tour, while row indicates the city names. As all oscillators have the same frequency, their relative phase information encodes the solution. A tour is represented by a set of oscillators that have the same/similar phase on the phase circle as shown in **Fig.1**.

In order to solve the TSP, the following energy function, L is used in terms of the oscillator variables, $z_{ij} = \left|z_{ij}\right| e^{i\varphi_{z_{ij}}}$:

$$L = A\Sigma_{ij}(\left|z_{ij}\right|^2 - 1)^2 + B\Sigma_{ij}\left|\left(\frac{z_{ij}}{|z_{ij}|}\right)^n - 1\right|^2 - C\Sigma_i\Sigma_{ij'}\left|\frac{z_{ij}}{|z_{ij}|} - \frac{z_{ij'}}{|z_{ij'}|}\right|^2 - D\Sigma_j\Sigma_{ii'}\left|\frac{z_{ij}}{|z_{ij}|} - \frac{z_{i'j}}{|z_{i'j}|}\right|^2$$

$$+ E\Sigma_i\Sigma_j\Sigma_{j'}d_{jj'}Re\left(\frac{z_{ij}.z^*_{i+1,j}}{|z_{ij}||z_{i+1,j'}|}\right)$$

The term with parameter A tries to bring oscillators to the unit circle. The term with parameter B tries to cluster oscillators at the locations of the roots of unity in an effort to evenly distribute clusters on the phase circle akin to roots of unity. The terms with parameters C and D try to repel oscillator phases that lie within a row or column. These terms with parameters C and D are called the syntactic constraints because they minimize the energy for a valid tour. The last term with parameter E represents the distance minimization constraint for cities visited in consecutive order in a tour.

B. *n oscillator mapping:*

Instead of using the above described n̈ × n̈ matrix of oscillators, a mapping to just n̈-coupled oscillators is more suitable for representing tours in a graph. The relative phase ordering of these cities denotes a valid tour. Again, the oscillator variables are complex numbers, $z_i = \left|z_i\right| e^{i\varphi_{z_i}}$ with ï = 1,2,3 … n. In order to solve the TSP, the following n̈-oscillator energy function L is proposed in this work:

$$L = A'\Sigma_i(\left|z_i\right|^2 - 1)^2 + B'\Sigma_i\left|\left(\frac{z_i}{|z_i|}\right)^n - 1\right|^2 - F'\Sigma_{ij}\left|\frac{z_i}{|z_i|} - \frac{z_j}{|z_j|}\right|^2 + E'\Sigma_{ij}d_{ij}exp\left[-\left(Im\left(\left(\frac{z_iz_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2/k\right]$$

The first term with coefficient A' causes the magnitude of z_i to be unity or make the phases settle on the unit circle. The second term with coefficient B' forces the oscillators to converge at phases equal to any one of the n roots of unity and the third term F' prevents more than one oscillator from converging to the same phase value. The last term with coefficient E' applies the distance constraint (i.e. it is minimum when the tour length is minimum).

C. *2D Oscillator: Variant I*

The phase information in the previous method has been encoded in terms of a single parameter in the previous method and hence, was restricted to a circle. With the introduction of an additional degree of freedom, each parameter spanning an angle of 2π, the resultant space is now restricted to a torus surface as shown in Fig(3). In all cases the dimension of torus is fixed with the main radius, R and cross-sectional radius, r as shown in Fig(4). The corresponding equation of torus in cartesian coordinates is:

$$\left(R - \sqrt{x^2 + y^2}\right)^2 + z^2 - r^2 = 0$$

With the parameters Θ_1 and Θ_2, the coordinates can also be represented in parametric form as (**Fig 5**):

$$x = (R + rcos\theta_2)\,cos\theta_1$$
$$y = (R + rcos\theta_2)sin\theta_1$$
$$z = rsin\theta_1$$

In the first variant, which is similar to the case of 1D oscillator, motion dynamics as maintained the same as before for Θ_2 with unity magnitude constraint on Θ_1. The cost function for this case is:

$$L = A\left(\Sigma_i\left[\left|z_{\theta_1}\right|^2 - 1\right] + \Sigma_i\left[\left|z_{\theta_2}\right|^2 - 1\right]\right) + B\Sigma_i\left[\left(\frac{z_{\theta_2}}{\left|z_{\theta_2}\right|}\right)^n - 1\right] - C\Sigma_{ij}\left[\frac{z^i_{\theta_2}}{\left|z^i_{\theta_2}\right|} - \frac{z^j_{\theta_2}}{\left|z^j_{\theta_2}\right|}\right]^2$$
$$+ D\Sigma_{ij} d_{ij} exp\left[-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2 / k\right]$$

where A coefficient term causes the phase to each parameter to lie on unit circle, B coefficient term forces Θ_2 to converge to one of the phases of unity, C coefficient term is a syntactic constraint to prevent values converging to a phase value and the last term applies distance constraint, scaling a Gaussian repulsion based on phase difference. The following dynamics ensure that the energy function minimizes monotonically with time:

$$\frac{dz_i}{dt} = -2A\left(\left|z^i_{\theta_1}\right|^2 - 1\right) z^i_{\theta_1}\hat{\theta_1} + \left[-2A\left(\left|z^i_{\theta_2}\right|^2 - 1\right) z^i_{\theta_2} - \frac{n}{2}\frac{B}{\left(z^i_{\theta_2}\right)^*}\left\{\left(\frac{z^i_{\theta_2}}{\left|z^i_{\theta_2}\right|}\right)^n - \left(\frac{\left(z^i_{\theta_2}\right)^*}{\left|z^i_{\theta_2}\right|}\right)^n\right\}\right]\hat{\theta_2}$$

$$+\left[C\Sigma_j\left\{\frac{z^i_{\theta_2}\left(z^j_{\theta_2}\right)^*}{\left(z^i_{\theta_2}\right)^*\left|z^i_{\theta_2}\right|\left|z^j_{\theta_2}\right|} - \frac{z^j_{\theta_2}}{\left|z^i_{\theta_2}\right|\left|z^j_{\theta_2}\right|}\right\} - \frac{D}{2}\Sigma_j d_{ij}\frac{i}{k\left(z^i_{\theta_2}\right)^*} Im\left(\frac{z^i_{\theta_2}\left(z^j_{\theta_2}\right)^*}{\left|z^i_{\theta_2}\right|\left|z^j_{\theta_2}\right|}\right) exp\left\{-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2 / k\right\}\right]\hat{\theta_2}$$

Since there are two parameters, the gradient is a two dimensional vector. The final solution is expected to lie along the unity root phases w.r.t. Θ_2 at any angle of cross-section (Θ_1).

D. 2D Oscillator: Variant II

In this variant, the magnitude constraint of both the parameters has been relaxed while others remain the same. This is replaced by the constraint (L_1) that the points should lie on the torus surface. Hence, the resultant cost function is:

$$L = A\Sigma_i\left(\left(R - \sqrt{x^2 + y^2}\right)^2 + z^2 - r^2\right) + B\Sigma_i\left[\left(\frac{z_{\theta_2}}{\left|z_{\theta_2}\right|}\right)^n - 1\right] - C\Sigma_{ij}\left[\frac{z^i_{\theta_2}}{\left|z^i_{\theta_2}\right|} - \frac{z^j_{\theta_2}}{\left|z^j_{\theta_2}\right|}\right]^2$$
$$+ D\Sigma_{ij} d_{ij} exp\left[-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2 / k\right]$$

Since, the surface constraint (L_1) is in cartesian form, while rest are complex functions, some processing is done to convert the derivative to complex numbers. The coordinates are first converted to their respective parametric form. The derivative is then taken along the orthogonal vector space (Θ_1, Θ_2). The gradient in this space can be written as:

$$\nabla L = \frac{1}{a}\frac{\partial L}{\partial \theta_1}\hat{\theta_1} + \frac{1}{R+acos\theta_1}\frac{\partial L}{\partial \theta_2}\hat{\theta_2}$$

These are then converted to complex derivative based on chain rule:

$$\frac{\partial L_1}{\partial z} = \frac{\partial L_1}{\partial \theta}\frac{\partial \theta}{\partial z}\text{ , where } z = \cos\theta + i\sin\theta$$

E. 2D Oscillator: Variant III

The syntactic constraints in the above two variants are weaker in higher dimensions of parametric space as the distance relies on phase difference with respect to single parameter. Hence, it is replaced with Euclidean distance between points along the toroidal surface. With the parameters known, the distance between any two points (i, j) along the surface is as shown in **Fig(6(a))**. With the cross-section of torus cut and open, the resultant shape is a truncated cylinder with the line on its surface as shown in **Fig(6(b))**. With the cylinder open to form a sheet, the resultant line can be seen as the diagonal of the trapezoid as shown in **Fig(7)**. The formulae for the sides and diagonal of the trapezoid is as shown below:

$$d_1 = 2\pi\left(R + rcos\theta_1^i\right)\left(\theta_2^j - \theta_2^i\right)$$

$$d_2 = 2\pi\left(R + rcos\theta_1^j\right)\left(\theta_2^j - \theta_2^i\right)$$

$$a = 2\pi r\left(\theta_1^j - \theta_2^i\right)$$

$$diagonal\ length = dist(i,j) = \sqrt{a^2 + d_1 d_2}$$

The syntactic constraint term is now replaced with this distance term. The resultant cost function is as shown below, with complex derivative for gradient calculated using chain rule as done in the previous variant:

$$L = A\left(\Sigma_i\left[\left|z_{\theta_1}\right|^2 - 1\right] + \Sigma_i\left[\left|z_{\theta_2}\right|^2 - 1\right]\right) + B\Sigma_i\left[\left(\frac{z_{\theta_2}}{\left|z_{\theta_2}\right|}\right)^n - 1\right] - C\Sigma_{ij}dist(i,j)^2$$
$$+ D\Sigma_{ij}d_{ij}exp\left[-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2/k\right]$$

F. 2D Oscillator: Hybrid

This is a hybrid of variant II and III with distance and syntactic constraint from the original replaced with the newly derived equation set. The resultant cost function is as shown below:

$$L = A\Sigma_i\left(\left(R - \sqrt{x^2 + y^2}\right)^2 + z^2 - r^2\right) + B\Sigma_i\left[\left(\frac{z_{\theta_2}}{\left|z_{\theta_2}\right|}\right)^n - 1\right] - C\Sigma_{ij}dist(i,j)^2$$
$$+ D\Sigma_{ij}d_{ij}exp\left[-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2/k\right]$$

Results

For the purpose of comparison between different techniques, the computation time is estimated based on the number of iterations taken to reach the optimal solution. The annealing rates are also varied to visualise the trend with changing rates. Each technique is estimated for 5, 8 and 10 cities. The success probability as a accuracy metric for each case and each rate is measured over 200 iterations :

$$SP = \frac{No.\ of\ iterations\ with\ optimal\ tour\ length}{Total\ no.\ of\ randomly\ initialised\ runs\ (200)}$$

Five techniques are simulated for comparison: 1D Oscillator of order n and the four variants of 2D Oscillator. For different numbers of cities(n), the success probability is calculated for 4 different annealing rates (1 - α values are mentioned in the below tables).

1D Oscillator: Table 1

n vs (1 - α)	**5×10^{-4}**	**1×10^{-3}**	**5×10^{-3}**	**1×10^{-2}**
5	0.34	0.34	0.30	0.24
8	0.12	0	0	0
10	0.06	0.05	0	0

For 2D Oscillators, the annealing rates mentioned are tuned for parameter Θ_2. Now significant change in accuracy or performance is observed upon varying this for Θ_1 as it converges faster within fewer iterations owing to single magnitude/surface constraint alone in all cases.

2D Oscillator Variant I: Table 2

n vs (1 - α)	**5×10^{-4}**	**1×10^{-3}**	**5×10^{-3}**	**1×10^{-2}**
5	0.34	0.28	0.28	0.30
8	0.08	0.06	0	0.02
10	0.06	0.05	0.02	0

2D Oscillator Variant II: Table 3

n vs (1 - α)	**5×10^{-4}**	**1×10^{-3}**	**5×10^{-3}**	**1×10^{-2}**
5	0.32	0.30	0.28	0.28
8	0.08	0.04	0.04	0.04
10	0.15	0.04	0.04	0

2D Oscillator Variant III: Table 4

n vs (1 - α)	**5×10^{-4}**	**1×10^{-3}**	**5×10^{-3}**	**1×10^{-2}**
5	0.32	0.42	0.32	0.4
8	0.08	0.12	0.10	0.04
10	0.02	0.02	0.03	0

2D Oscillator Hybrid: Table 5

n vs (1 - α)	**5 x 10^{-4}**	**1 x 10^{-3}**	**5 x 10^{-3}**	**1 x 10^{-2}**
5	0.30	0.32	0.30	0.36
8	0.14	0.12	0.13	0.08
10	0.05	0.04	0.02	0.02

The number of iterations taken to converge to the optimal solution for each technique has been calculated and given to the nearest estimate. The time taken for each iteration is almost the same for all cases. These are almost independent of the number of cities and hence compared w.r.t. the annealing rates alone. (Duplicacy of similar numbers for different techniques has been avoided as shown below).

Table 6

(1 - α)	**1D Oscillator, Variant I & II**	**Variant III**	**Hybrid**
5 x 10^{-4}	16000	7000	11000
1 x 10^{-3}	8000	4000	5500
5 x 10^{-3}	1600	800	1000
1 x 10^{-2}	800	400	500

Significance

The success probability only showed marginal improvement with the new dynamics. The results were also consistent with an increasing number of cities over all variants. The accuracy decreases with the number of cities and slower annealing rate tend to give more optimal tour solutions most of the time.

Maintaining the same accuracy, significant improvement in performance of the system is seen in this 2-D Oscillators dynamics. The computation time compared to 1-D Oscillator almost reduced drastically by 40-50% for the later variants as seen in Table(6).

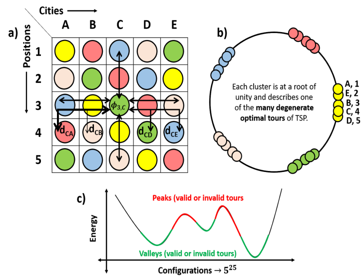
Hence, this new technique resembles 1D n-Oscillator mapping in terms of lesser number of oscillator units, reduced hardware area footprint, reduced search space and improved accuracy of optimal tours compared to n^2 mapping. The convergence time is additionally reduced and thereby lower energy consumption in the device is also expected.

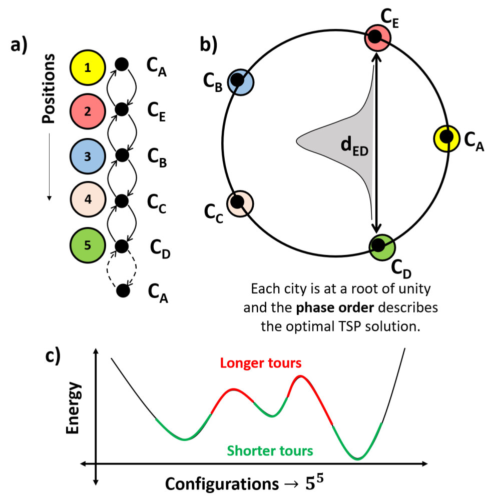
Conclusion

In this paper, an efficient mechanism of 2D n-Oscillator mapping has been proposed to find optimal tours in TSP. The success probability is achieved at low annealing durations also as compared to n^2 Oscillator mapping[1] and reduced computation time and energy consumption w.r.t 1D Oscillator[2]. With usage of an oscillator based neural network in the Neuromorphic hardware space modelling real neurons, such techniques incorporated in compact devices can prove to be very much efficient in solving various optimization problems.

[1] Gregory S. Duane, A ‘Cellular Neuronal’ Approach to Optimization Problems

[2] Langde et al, Oscillator Network based Efficient Cost function for TSP





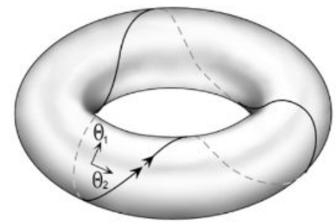
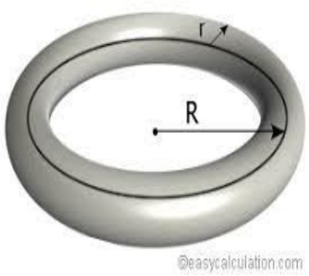


Fig. 1 n^2 oscillator figure[1]

Fig. 2 n-oscillator figure[2]

Fig. 3 2D Oscillator phase parameters on a torus

Parametrization



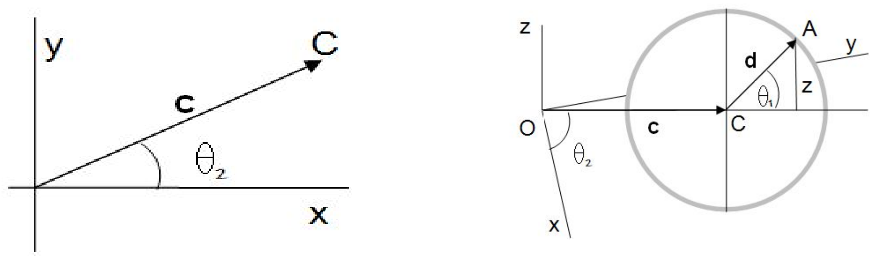
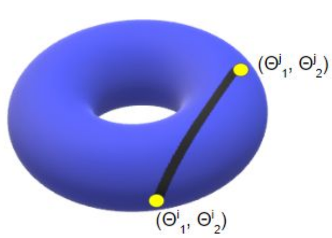
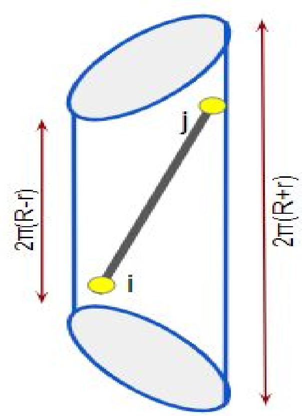


Fig. 4 Torus: main and cross-sectional radii

Fig. 5(a), (b)Derivation of parametric equation for a torus





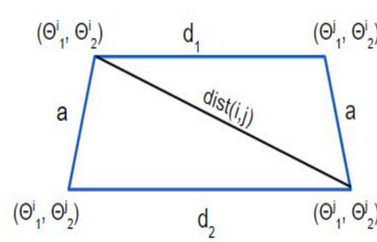
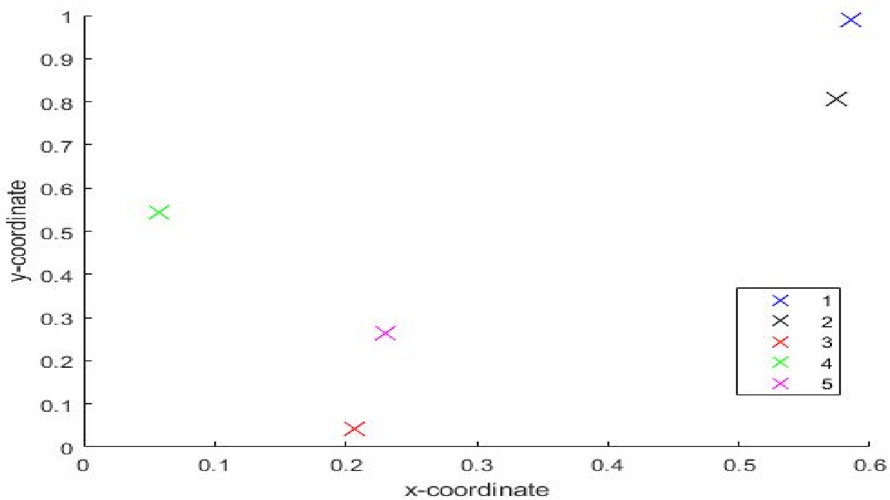


Fig. 6 Line on the toroidal surface between points ‘i’ and ‘j’

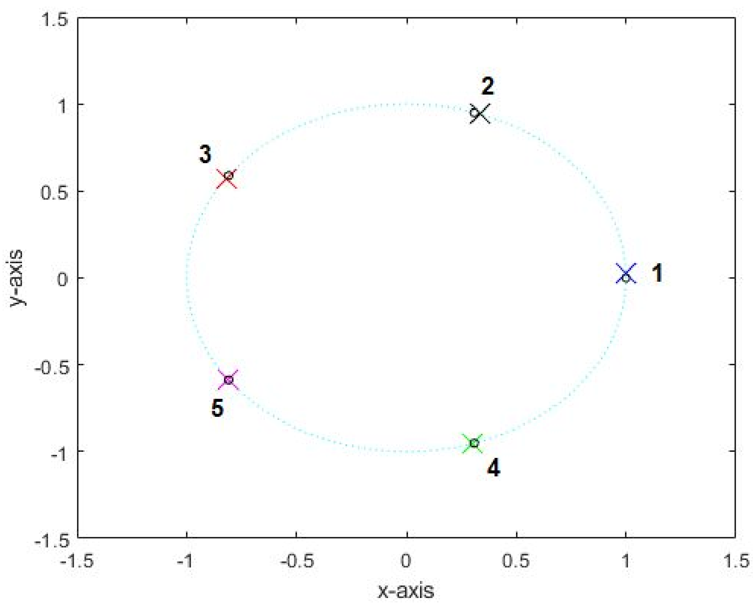
Fig. 7 Truncated cylinder obtained upon cutting-open the torus with the line between ‘i’ and ‘j’

Fig. 8 Distance between ‘i’ and ‘j’ as a diagonal of trapezoid obtained upon the surface of the cylinder

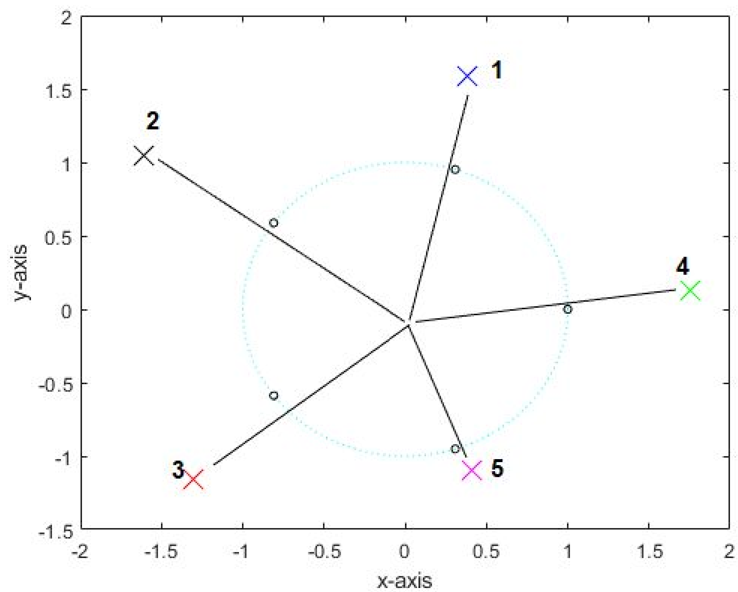
Simulation Plots:



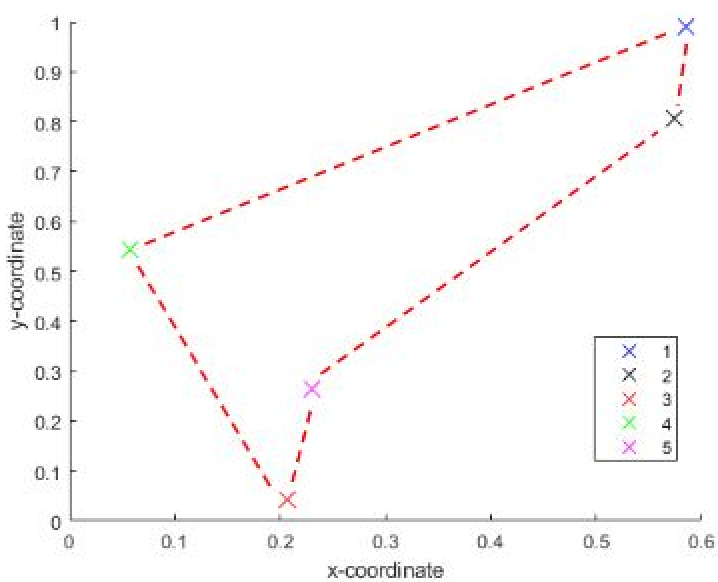
Solution at roots of unity for those cases with unit magnitude constraint (1D oscillator, Variant I & III)



Without unit magnitude constraint (variant II & Hybrid):

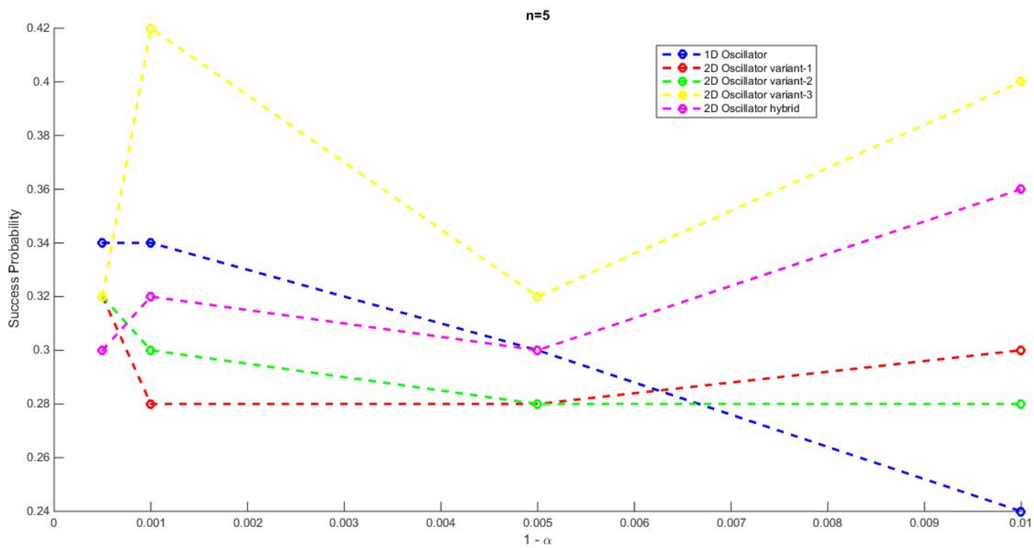


Optimal Tour:

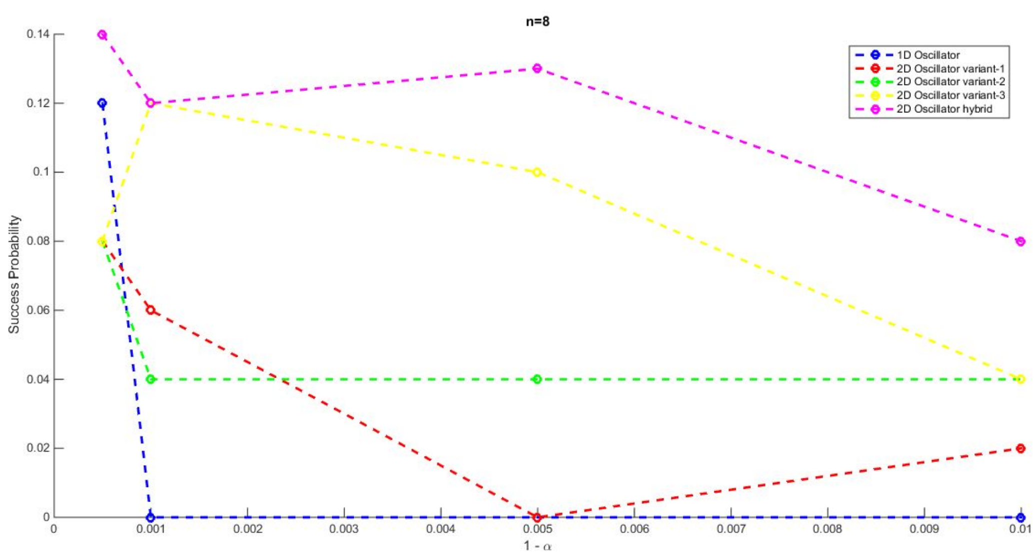


Comparison between different approaches:

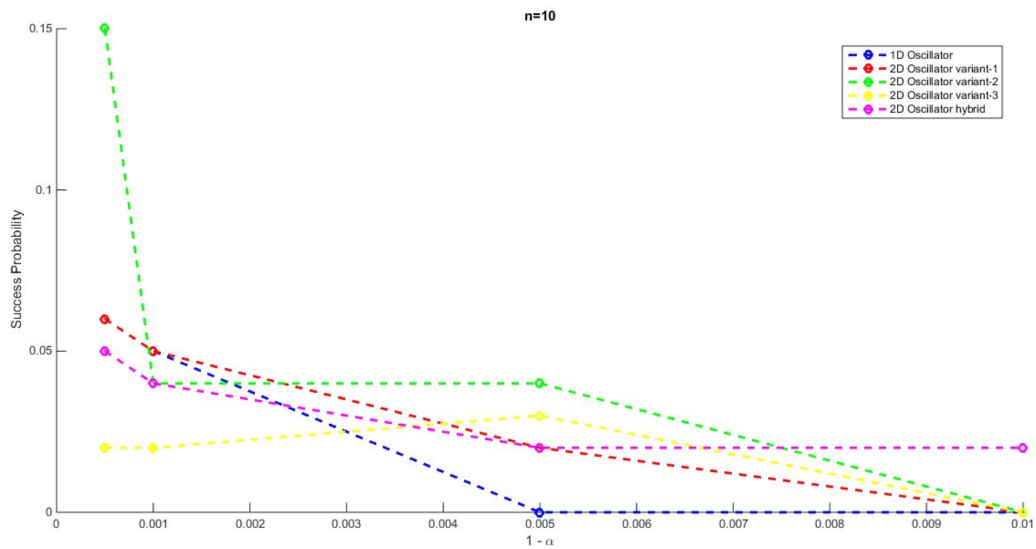
For n=5



For n = 8



For n = 10



Comparison between number of iterations: