n-Oscillator Neural Network based Efficient Cost Function for n-city Traveling Salesman Problem

Shruti Landge
Department of Electrical Engineering
Indian Institute of Technology Indore
Indore, India
shruti.parag.landge@gmail.com

Vivek Saraswat
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai, India
svivek@ee.iitb.ac.in

Srisht Fateh Singh
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai, India
srishtfateh.singh@iitb.ac.in

Udayan Ganguly
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai, India
udayan@ee.iitb.ac.in

***Abstract*—Neural Networks have long been a mainstream technique to solve optimization problems. A classic example is the Travelling Salesman Problem (TSP) which NP-hard. Using a Hopfield-Tank representation, an n-city problem is mapped to a cost function of n^2 interacting neural units. Stochastic gradient descent helps achieve the global minima. Due to the nature of the TSP problem, the cost function has to penalize invalid sub-routes (non-Hamiltonian cycles) and minimize the travel cost simultaneously. In addition, there is a starting point and travel direction associated '$2n$' degeneracy. Previously, a cellular neuronal approach was proposed where the neural units were replaced with oscillators. The phase relations determined the output solution. Multiphase clusters of these oscillators solved the degeneracy issue. This paper proposes an n-oscillator cost function for an n-city TSP. Since a group of single-frequency oscillator phases are naturally ordered and circular in a system, the proposed method exploits the true potential of oscillator nodes. The sub-routes and degeneracy are eliminated by design in addition to massively increasing the scaling potential (n vs. n^2). It was also found that the proposed n-mapping can converge to the optimum tour much faster (about 100 times for a 5-city problem) than for n^2 mapping. Our approach projects hardware efficiency in terms of area footprint, computation time and energy. With coupled single device-based compact nanoscale oscillator systems becoming increasingly viable in hardware, efficient cost function mappings of hard problems using oscillator phases, as shown here, is critical to solving large graphical optimization problems.**

Keywords—TSP, oscillators, optimization, neural networks

I. INTRODUCTION

Algorithmically challenging problems are the ones that do not have a time-efficient method of arriving at the solution on a Boolean logic based sequential microprocessor. The solution space for such problems usually scales exponentially or factorially. In many cases, the exact solution can only be found by a brute force scan of this extensive solution space. Many optimization problems belong to this category. The Max-Cut, Vertex Coloring and Travelling Salesman Problem (TSP) are all examples of NP-complete problems [1], [2]. They are mathematical and graphical abstractions of real-life business and personal level decisions about scheduling, routing and component allocation [3].

In particular, the objective of an n-city TSP is to find the minimum Euclidian distance cycle (closed loop) which visits each city exactly once. There are $n!$ tours among which the optimal tour length is to be found. In addition, there is a '$2n$' degeneracy in the tour lengths related to the starting point ambiguity (in a circular tour – any point can be the starting point for minimum cost) and the direction of tour traversal (in a circular tour – forward and backward travel will not change cost). Accounting the redundancy reduces the effective search space to be $(n-1)!/2$. There have been numerous attempts to solve the TSP. Exact algorithms often relied on the solving the integer linear programming formulation which were among the first to be proposed for the TSP [4]. Most state-of-the-art digital computer solvers today are based on branch and cut algorithms [1], [4], [5]. There are many heuristics like the Lin-Kernighan heuristic, the Polynomial Time Approximation Scheme (PTAS) or the Elastic Nets method which offer a trade-off between the "goodness" of the polynomial time complexity and the quality (how close to optimal) of the solution guaranteed [3], [5]–[8]. Nature inspired optimization algorithms like ant colony optimization and genetic algorithms have also been explored extensively [9], [10].

Neural networks offer an alternative platform to the TSP's solution. Artificial neural networks are powerful for parallel processing of numerous inputs. Hopfield networks are well known for attaining local minimum and are thus utilized for pattern recognition starting from corrupted images [11], [12]. For achieving the global optimum, as is the requirement for the TSP, stochastic Hopfield networks or Boltzmann

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

Machines have become very popular [1], [2], [13]. They perform stochastic gradient descent in simple fully connected networks of neurons (first component in Fig. 1). Gradually diminishing stochasticity or Simulated Annealing is a critical success enhancer in these networks [14]–[16]. It has been a popular belief that the upcoming Neuromorphic hardware utilizing memristor based neurons and synapses can allow Boltzmann Machines to solve optimization problems at "never before seen" speeds [2], [17].

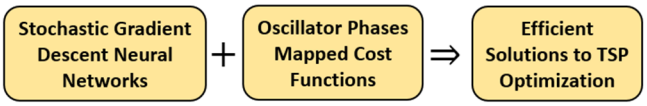

Fig. 1. Oscillator phase based stochastic gradient descent neural networks are key for efficient solutions to TSP optimization

A critical step in solving the TSP, before one can configure a circuit for the same, is mapping the problem constraints to a cost function of the neural units. Oscillations have been a recurring theme in these cost functions (second component in Fig. 1). In 1990, Wells [18] had demonstrated the development of traveling waves in the activation neuron ensembles which visited the degenerate solutions in a periodic manner. This solution used the conventional Hopfield-Tank cost mapping which makes use of n^2-units to encode n-city TSP [12]. The energy function had syntactic constraints which enabled valid tours (Hamiltonian cycles) and distance constraints to optimize tour lengths. Developing this idea, in 2009, Duane [16] replaced the neural activation units with oscillator nodes and encoded the problem in the relative phases of these same-frequency oscillators rather than the voltage levels of the oscillating neuron output. Although the energy function still used n^2 oscillators for an n-city problem, the degenerate solutions, rather than appearing temporally, were all obtained at the same time as multiphase clusters. However, *the requirement of n^2 units is a severe limitation to hardware realization of such networks and is resource expensive especially for large problems.* It is worthwhile to mention that an n neurons digital neural network for an n-city TSP has been reported earlier [19] however it did not utilize oscillator phases for network modelling which, as outlined in this paper, can offer several advantages in capturing the problem constraints efficiently.

Coupled networks of compact device level oscillators are becoming increasing feasible in hardware implementation [20]–[22]. Simpler problems with fewer constraints like the vertex coloring and pattern recognition have already been demonstrated on either silicon chips or special memristor arrays [23], [24]. *With the hardware promise of upcoming coupled oscillator circuits and the natural link between the TSP constraints to oscillator phase-based cost functions,* a more effective cost function can have far reaching consequences.

In this paper, we propose an n-oscillator phases mapping where the structure (ordering of oscillator phases) encodes the syntactic constraints directly. This simplified cost function results in a substantial contraction of the solution space. We discuss the two oscillator mappings, n^2 (previous baseline) and n (proposed) and compare their performance. The reduction in oscillator nodes could offer massive improvements in the area footprint, time to converge and energy of computation, which are discussed thereafter.

II. Oscillator Cost Functions for TSP

The general technique to solve the TSP using a network of oscillators involves mapping the problem constraints to an energy function of oscillator variables [16]. The oscillator dynamics should ensure that the energy function minimizes monotonically with time. This criterion describes the analytical nature of the function f which governs the oscillators' time evolution. However, the energy can reduce and get stuck in local minima in energy. Hence, we need a controllable source of phase noise to these oscillators since the solutions are encoded in relative phases of these oscillators. The technique of simulated annealing (SA) is one in which the initial noise source is sufficient to randomize the configuration altogether but reduces over time to eventually arrive at the global optima of the energy function [16]. An example of the TSP is shown in Fig. 2(a). The circuit schematic for the discussed general technique is shown in Fig. 2(b). This technique is discussed explicitly for two energy function mappings in this paper. The n^2 oscillators mapping which has been published earlier [16] and inspired from the well-known Hopfield Tank representation of the problem and, the new proposed n oscillator mapping which exploits the natural phase ordering of oscillators in a system of same-frequency oscillators.

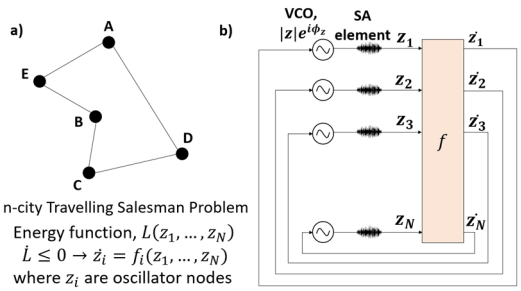

Fig. 2. (a) Travelling Salesman Problem: The problem definition has the city locations as the input. The problem constraints are mapped to an energy function, L of N oscillators (N could be different from n). Monotonic optimization of L over time can be achieved by a specific time dynamics f of the set of oscillators. (b) The network schematic to realize the time dynamics dictated by f. A controllable noise source is required to achieve simulated annealing. Simulated Annealing (SA) element adds controllable stochasticity to voltage-controlled oscillator (VCO) phases

A. n^2 oscillators mapping (Type I)

This representation makes use of an $n \times n$ matrix of coupled oscillators as described in detail in [16] and presented here briefly as the state-of-the-art. A column represents position (e.g. 1-5) in the tour, while row indicates the city names (e.g. A-E) – as shown in Fig. 3(a). As all oscillators have the same frequency, their relative phase information encodes the solution. A tour is represented a set of oscillators that have the same/similar phase on the phase circle as shown in Fig 3(b). For example, 1A, 2E, 3B, 4C, 5D is a tour which has the same phase (color yellow) on the phase circle. A tour is assigned the same color. A valid tour in this representation is one in which

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

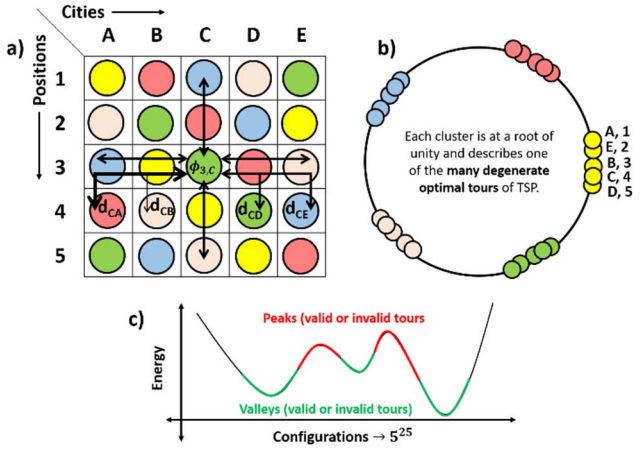

Fig. 3. (a) $n \times n$ matrix of oscillators: Colors represent distinct phases. Black arrows denote the repulsion between phases caused by different terms of the energy function, L. (b) The settling phases clustered at roots of unity corresponding to the valid tours in (a). (c) Qualitative energy profile with peaks and valleys (which can consist of both valid or invalid tours) as a function of the all possible phase configurations which can be 5^{25} states.

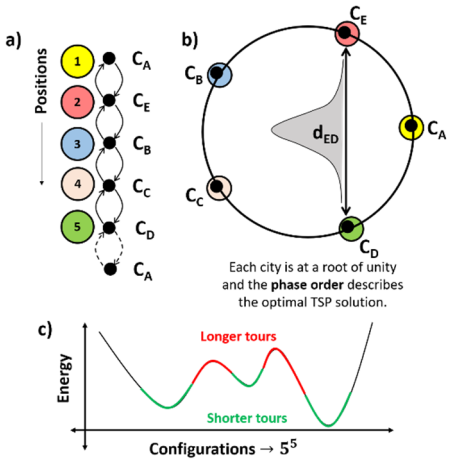

Fig. 4. (a) n oscillator mapping: Colors represent distinct phases and position in a tour. Oscillators mapped to cities can settle at different positions (phases), (b) n oscillators network represented on a unit phase circle with each oscillator having phases at distinct roots of unity. (c) Qualitative energy profile with peaks (longer tours) and valleys (shorter tours) as a function of the all possible phase configurations which can be 5^5 states.

no two oscillators in a row or in a column are in phase which represents the constraint that every city occurs exactly once in a tour. Thus, tours with cities missed or cities repeatedly visited are invalid and need to avoided. For each oscillator, there is only one oscillator in every other row and column which is in the same phase. An example is depicted in Fig. 3(a). Fig. 3(b) yellow cluster denotes AEBCD(A) tour. Other color clusters are the same tours with different starting cities and have the same tour length (degenerate solutions).

In order to solve the TSP, the following energy function, L was used [16] in terms of the oscillator variables, $z_{ij} = |z_{ij}|e^{i\phi_{z_{ij}}}$:

$$L = A\,\Sigma_{ij}\left(|z_{ij}|^2 - 1\right)^2 + B\,\Sigma_{ij}\left|\left(\frac{z_{ij}}{|z_{ij}|}\right)^n - 1\right|^2$$
$$-C\Sigma_i\Sigma_{jj'}\left|\frac{z_{ij}}{|z_{ij}|} - \frac{z_{ij'}}{|z_{ij'}|}\right|^2 - D\Sigma_j\Sigma_{ii'}\left|\frac{z_{ij}}{|z_{ij}|} - \frac{z_{i'j}}{|z_{i'j}|}\right|^2$$
$$+E\Sigma_i\Sigma_j\Sigma_{j'}d_{jj'}Re\left(\frac{z_{ij}.z^*_{i+1,j'}}{|z_{ij}||z_{i+1,j'}|}\right) \qquad (1)$$

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 (vertical/horizontal black arrows in Fig. 3(a)). These terms with parameters C and D are called the syntactic constraints because they minimize the energy for a valid tour. These energy penalties dissuade the system from preferring the invalid tour states. However, the system does visit such invalid tour states as they are not exactly "forbidden" by design. The example of valid degenerate tours in Fig. 3(a) is also plotted on the phase circle in Fig. 3(b) to show the effect that the syntactic constraints try to achieve in the solution. n unique oscillators are clustered at each root of unity which follow the row-column exclusivity. The last term with parameter E represents the distance minimization constraint for cities visited in consecutive order in a tour (angled black arrows in Fig. 3(a)). The following oscillator dynamics ensure that the energy function defined earlier minimizes monotonically with time:

$$\frac{dz_{ij}}{dt} = -2A\left(|z_{ij}|^2 - 1\right)z_{ij} - \frac{n}{2}.\frac{B}{z^*_{ij}}\left[\left(\frac{z_{ij}}{|z_{ij}|}\right)^n - \left(\frac{z^*_{ij}}{|z_{ij}|}\right)^n\right]$$
$$+\frac{C}{2}\Sigma_{j'}\left[\frac{z_{ij}z^*_{ij'}}{z^*_{ij}|z_{ij}||z_{ij'}|} - \frac{z_{ij'}}{|z_{ij}||z_{ij'}|}\right]$$
$$+\frac{D}{2}\Sigma_{i'}\left[\frac{z_{ij}z^*_{i'j}}{z^*_{ij}|z_{ij}||z_{i'j}|} - \frac{z_{i'j}}{|z_{ij}||z_{i'j}|}\right]$$
$$-\frac{E}{4}\Sigma_{j'}d_{jj'}\left[\frac{1}{|z_{ij}|}\left(\frac{z_{i+1,j'}}{|z_{i+1,j'}|} + \frac{z_{i-1,j'}}{|z_{i-1,j'}|}\right)\right.$$
$$\left. - \frac{z_{ij}}{z^*_{ij}|z_{ij}|}\left(\frac{z^*_{i+1,j'}}{|z_{i+1,j'}|} + \frac{z^*_{i-1,j'}}{|z_{i-1,j'}|}\right)\right] \qquad (2)$$

Assuming any oscillator can settle at one of the 5 roots of unity in the example of Fig. 3(a), the possible configurations to choose from are a staggering 5^{25} states as shown in the qualitative energy profile in Fig. 3(c) – which have many local minima in addition to a global minimum. Starting from arbitrary initial phases, the oscillators governed as described above rarely reach the optimal TSP tour or even valid tours as the system may get stuck in local minima due to monotonic energy reduction. A slowly decaying phase noise is added over the deterministic dynamics to allow the occasional escape from the local minima and to eventually arrive at the global minima. If the noise is decayed slow enough, significant probabilities of ending up at the optimal TSP tours can be observed.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

B. n oscillators mapping (Type II)

Instead of using the above described $n \times n$ matrix of oscillators, a mapping to just n 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, $z_i = |z_i|e^{i\phi_{z_i}}$ with $i = 1,2,3 \ldots, n$. Fig. 4(a) shows a typical output solution in this mapping for a 5-city problem. The n-mapping model, hence, uses n oscillators which correspond to the n cities of the problem, and with all oscillators having the same frequency. The oscillators occupy different roots of unity positions and the ordering specifies the tour, AEBCD(A), in this example (Fig. 4(b)).

In order to solve the TSP, the following n-oscillator energy function is proposed in this work:

$$L = A'\Sigma_i(|z_i|^2 - 1)^2 + B'\Sigma_i \left|\left(\frac{z_i}{|z_i|}\right)^n - 1\right|^2 - F'\Sigma_{ij}\left|\frac{z_i}{|z_i|} - \frac{z_j}{|z_j|}\right|^2 + E'\Sigma_{ij} d_{ij} \exp\left[-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2 / k\right] \quad (3)$$

The first term with coefficient A' causes the magnitude of z_i to be unity or makes 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 with coefficient F' prevents more than one oscillator from converging to the same phase value. These syntactic constraints together allow unambiguous reading of the tour. The last term with coefficient E' applies the distance constraint (i.e. it is minimum when the tour length is minimum). This involves a Gaussian repulsion in phases scaled by the intercity distance between each pair of cities. This minimizes the possibility of high intercity distance pairs to carry a small phase difference and run the risk of being consecutive in a tour (Fig. 4(b)). In other words, the Gaussian factor for consecutive cities in a tour (smaller phase separation) is larger and hence smaller $d_{jj'}$ city pairs appearing consecutive are energetically favorable. It is important to note here that if the intercity distances $d_{jj'}$ are too large or small, it can lead to an imbalance between the syntactic and distance constraints, whose contribution is controlled by hyper parameters A', B', F' and E' in the energy function. This limitation, dependent on distance matrix $d_{jj'}$, can be avoided by choosing the cities to be located within a unit square and scaling the distances between the cities to maintain an average separation distance of 0.5 between each pair, as mentioned in [16], so that the same set of hyper parameter values can be used for different problems.

As described earlier, the dynamics of oscillators has to ensure a monotonic decrease in L with time which results in the following update rule:

$$\frac{dz_i}{dt} = -2A'(|z_i|^2 - 1)z_i - \frac{n}{2}.\frac{B'}{z_i^*}\left[\left(\frac{z_i}{|z_i|}\right)^n - \left(\frac{z_i^*}{|z_i|}\right)^n\right] + F'\Sigma_j\left[\frac{z_i z_j^*}{z_i^*|z_i||z_j|} - \frac{z_j}{|z_i||z_j|}\right] - \frac{E'}{2}\Sigma_j d_{ij}\frac{i}{kz_i^*} Im\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)\exp\left[\frac{-\left(Im\left(\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)^{0.5}\right)\right)^2}{k}\right] \quad (4)$$

The n oscillator representation structure produces several advantages. First, the number of neurons (oscillators) are reduced by a factor of n – which makes it significantly more efficient for implementation on any platform. Second, the structure of the network guarantees valid tours as in one complete revolution round the phase circle (Fig. 4(b)), a particular node corresponding to any one city of the problem is encountered only once. Invalid tours of repeated/missed cities are forbidden by structure. The only remaining invalid cases are when multiple cities are assigned the same phase. In comparison, n^2 formulation produces soft penalization for invalid tours by adding terms with parameters C and D. Just by removing these constraints, the search space reduces from 25 oscillators choosing from 5 phases (roots of unity) in n^2 mapping (5^{25} states) to just 5 oscillators choosing from 5 phases (roots of unity) in n mapping (5^5 states) (qualitative energy profile in Fig. 4(c)). Third, same-frequency oscillator phases follow a natural ordering in any system which is also circular in nature. Hence, this structure outputs all degenerate solutions associated with starting city and travelling direction simultaneously. In the n^2 mapping, the degeneracy has to be artificially implemented using separate oscillator clusters at different roots of unity – an additional complexity which is otherwise absent in the TSP problem definition. Fourth, the parameter k in the distance constraint term controls the sharpness of the Gaussian which is absent when we use a simpler $Re\left(\frac{z_i z_j^*}{|z_i||z_j|}\right)$ or $cos(\Delta\theta)$ as in the n^2 mapping, where $\Delta\theta$ is the phase difference between a pair on nodes in the system. This allows to optimize sharpness of distance dependence adding more control. The oscillator evolution would still require addition of time decaying phase noise similar to the procedure of simulated annealing as the previous mapping for escaping the local minima.

III. Results and Discussion

For testing the performance of any mapping, we start with arbitrary initial phases of the oscillators and large phase noise variance ($\sigma^2 \sim \pi^2$). We numerically solve for the oscillator time evolution with a step size of $\Delta t = 0.01$ (arb. units). Every step size we multiply the noise σ by a factor $\alpha < 1$. In each time step iteration, after updating the oscillator variables, $z(t)$ using the rules in eq. (2), (4), a zero mean Gaussian noise $\theta(\sigma)$ with decaying σ is added to the phase as:

$$z(t) \leftarrow z(t)\exp\big(i\theta(\sigma)\big) \quad (5)$$

$$\sigma(t + \Delta t) = \alpha.\sigma(t) \quad (6)$$

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

This α controls how many iterations of time step Δt it takes for the noise σ to become negligible w.r.t the phase values being measured and hence the total annealing duration. In Fig. 5 (a) we show the energy evolution for a 5-city problem using n-mapping. The system is evolved for two different rates of annealing ($\alpha = 0.999$ and 0.9999) and the decay of noise over iterations is significantly different due to this choice. Different annealing rates are important because they offer a trade-off between the total run time and the associated probability of achieving the optimal TSP tour. We perform 100 randomly initialized and independent runs for $\alpha = 0.999, 0.9999$ and 0.99999 which correspond to annealing durations of $10^2, 10^3$ and 10^4 (arb. units). The performance of the proposed n-mapping is tabulated in Table I and in order to estimate the reduction in the occurrence of non-optimal tours we calculate the tour length vs. frequency of occurrence correlation in Table II. We observe that there is a negative correlation which reaches ~ -1 steadily as the annealing duration is increased. Further similar correlation coefficients are observed for n-mapping at much lower annealing durations (20 – 500x improvement) compared to n^2-mapping [16].

TABLE I. PERFORMANCE OF n-MAPPING

Tour Length	No. of occurrences for different annealing rates		
	$\alpha = 0.999$ (Slow)	$\alpha = 0.9999$ (Slower)	$\alpha = 0.99999$ (Slowest)
2.0143	9	13	19
2.0772	9	9	13
2.1356	8	14	16
2.1624	8	11	12
2.2181	7	7	12
2.2543	5	9	11
2.9578	3	4	2
2.9939	2	2	6
3.0496	3	7	2
3.0764	4	2	1
3.1348	7	1	0
3.1977	7	7	1
Invalid tour	28	14	5

TABLE II. TOUR LENGTH/FREQUENCY CORRELATION

	n - mapping			n^2 - mapping	
Total Annealing Duration	10^2	10^3	10^4	5×10^4	2×10^5
Tour Length/Frequency Correlation	-0.67	-0.81	-0.96	-0.62	-0.87

The purpose of any mapping is to (a) Arrive at valid tours (Hamiltonian cycles) and (b) Choose optimal tour length among the valid tours. Hence, in order to quantify the performance, we define the metrics failure probability (FP) and success probability (SP_γ). Failure probability (FP) is the fraction of the total number of initial runs that end up in invalid tours. The term success probability (SP_γ) is the probability of selecting a tour with tour length within $+\gamma\%$ or $\left(1 + \frac{\gamma}{100}\right)$ times the optimum tour length. (e.g. SP_0 is the probability to end at optimum tour):

$$FP = \frac{No.\,of\ runs\ that\ end\ up\ in\ invalid\ tours}{Total\ no.\,of\ randomly\ initialised\ runs} \quad (7)$$

$$SP_\gamma = \frac{No.\,of\ runs\ within\ +\gamma\%\ of\ optimal\ tour\ length}{Total\ no.\,of\ randomly\ initialised\ runs} \quad (8)$$

The parameter γ allows the inclusion of near-optimal tours along with the optimal tours while calculating the success of a particular mapping.

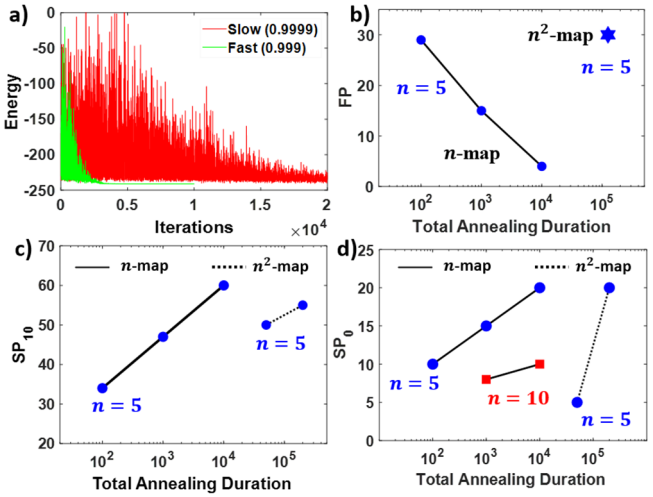

Fig. 5. (a) Energy evolution vs iterations for a 5-city problem using n-mapping for slow and fast anneal rates, (b) Failure Probability (FP) as a function of total annealing duration. Success Probability vs Total annealing duration, (c) SP_{10} and (d) SP_0 for n and n^2 mappings.

We compare the two mappings by calculating FP, SP_0 and SP_{10} for 5-city problems against different annealing times. The total annealing duration is calculated as the Total iterations $\times \Delta t$. For achieving the same noise σ range ($\sim 4 \times 10^{-5}$) at the end of a run (starting from the same initial noise σ), the number of iterations for the slow annealing are much greater and hence correspond to a larger total annealing duration. FP (selecting invalid tours) was observed to reduce with increase in the annealing duration to around 5% at total annealing duration of 10^4 for the proposed n-mapping. This is significantly smaller than the FP of ~30% for the n^2-mapping obtained at even ~10x higher anneal duration (Fig. 5 (b)) indicating that the proposed mapping selects valid tours more often than the previous mapping. Next, we observe that the SP_{10} and SP_0 over 100 runs increase for both mappings on increasing the annealing duration (Fig. 5 (c) and (d)). However, similar success probabilities can be achieved for the n-mapping at much lower annealing durations (100 – 1000x reduction). As noted earlier, the number of possible configurations for n^2-mapping are n^{n^2} while they are n^n for n-mapping. The significant reduction in the explorable solution space can be instrumental in reducing the anneal times required for similar performing oscillator networks that use n-mapping.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

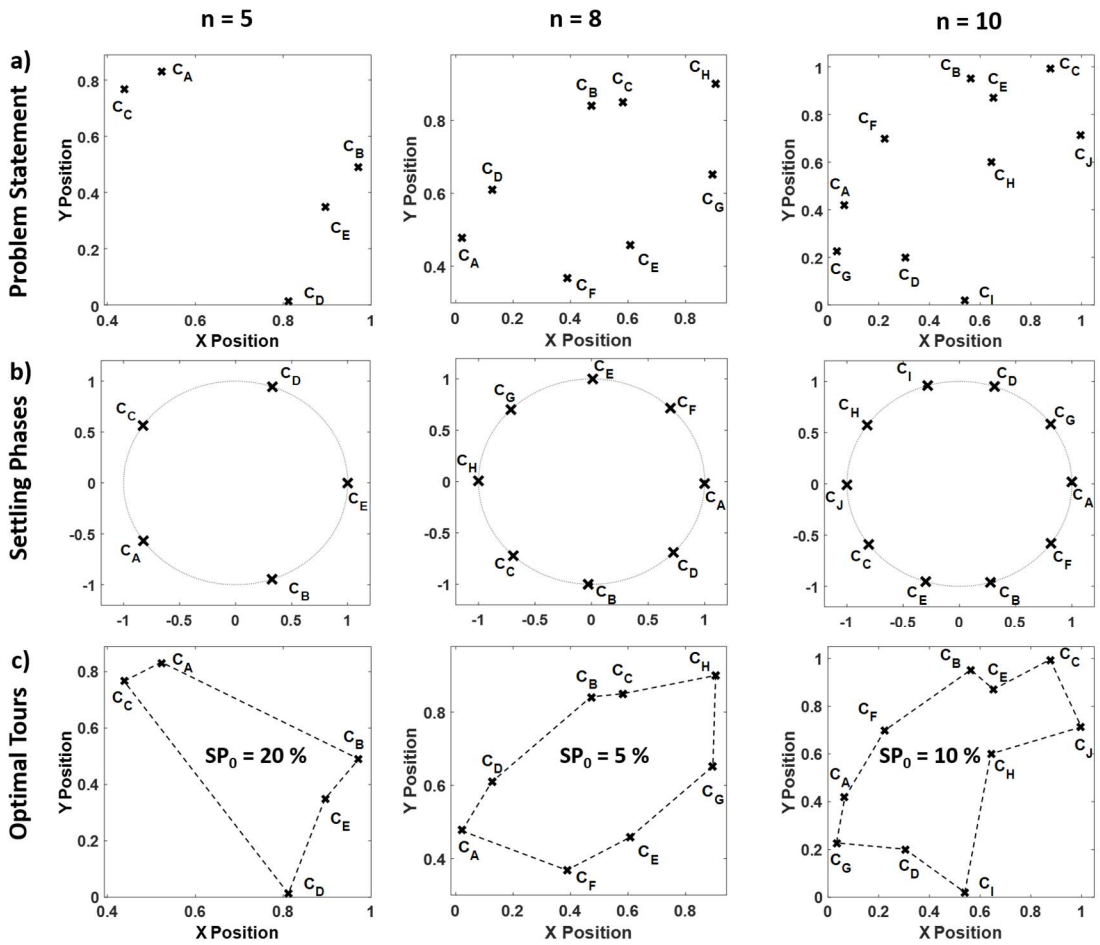

Fig. 6. (a) Problem Statement, (b) Settling phases using the n-mapping and (c) the corresponding optimal tours lengths and their success probabilities SP_0 for the n = 5, 8 and 10 cities TSP problems for $\alpha = 0.99999$.

Next, we test the proposed n-mapping for problems of different number of cities, n = 5, 8, 10 at $\alpha = 0.99999$. In Fig. 6(a), we show the starting problem definitions i.e. the city locations. Followed by the problem statement are the settling phases for the successful cases (Fig. 6(b)). The tours dictated by these settling phases are also included in Fig. 6(c). Despite maintaining an average separation of ~ 0.5 for the distance matrix, the authors recognize that the parameters A', B', F' and E' were originally optimized for n = 5 city problems and they may require recalibration for larger problem sets. Nonetheless, the successful arrival at the optimal tours using the n -mapping demonstrates the applicability to larger problems.

The effect of scaling n on performance by n-oscillator representation is compared to n^2 oscillator. It is worthwhile to note here that success probability SP_0 goes down for larger problems (n = 10 vs. n = 5 in Fig. 5(d)). However, the total annealing durations to obtain significant success probability SP_0 (~ 10 %) even for larger n are significantly smaller for the n-mapping compared to the n^2-mapping. In the baseline work [16], even the implementation for n = 7, 10 grew exponentially high in annealing durations so much so that no optimal tours were reported over the few runs that could be practically performed. This is because the search space for n^2-mapping explodes at n = 7, 10. In fact, for n = 10, the search space is the scale of Googol (i.e. 10^{100}) - one of the largest named numbers. However, n-mapping has a reduced search space (10^{10}) for n = 10 enables a 10x shorter annealing time than even the n = 5 case of n^2-mapping.

IV. Hardware Implications

It has long been recognized that the true power of stochastic Hopfield networks or Boltzmann Machines can only be achieved by dedicated hardware devices and circuit architectures [1], [2], [13]. Analog synaptic arrays are becoming popular for multiply and accumulate functionality [25], [26]. Device level stochasticity in RRAM devices has been harnessed to get rid of the expensive random number generation tasks [2]. Controllable stochasticity has also been reported by reading RRAM states at different bias levels or by using a random telegraph noise generator devices with variable input [17]. Compact device level nano-oscillators and their coupled networks have demonstrated pattern recognition and graphical optimization tasks [20]–[22]. In such a scenario, problem constraints mapping into efficient and resource inexpensive energy functions is a welcome step to realizing application specific hardware optimization chips. The n-mapping has a potential to offer a multitude of benefits (Table III and Fig. 7) if used instead of the n^2-mapping inspired by the Hopfield-Tank network:

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

a) Lower the number of oscillators and hence the area-footprint reduces by at least a factor of n. Since these networks are all to all connected the coupling function will scale the area further down when they are connecting only n elements rather than n^2. As an example, in a Boltzmann Machine with resistive couplings, the crossbar array used to achieve all to all connectivity utilizes n^2 coupling elements. Thus, this leverages oscillator phase in addition to stochasticity to enable a more powerful representation with $1/n$ size reduction.

b) Lower the number of computations being performed per iteration by a factor of n because only n oscillators need to be updated.

c) Lower the time to converge or total anneal time by a factor at least n. This reduction would require more formal test benches, nonetheless, the contraction of the exploration space offers a much higher expectation (n^{n^2} vs. n^n). Even for a 10-city problem the solution space for the n^2 mapping would be a staggering Googol (10^{100}) scale search space.

d) Lower the sampling frequency by a factor of n. Phase detection is a time series task. For a more crowded phase circle, the samples need to be taken much more rapidly in time in order to be able to distinguish them. For a problem of the similar size there are n times fewer oscillator phases in the same 2π range allowing lower sampling rates (or time-keeping clocks) to be sufficient.

e) Lower computations per cycle and time to converge for a given SP_γ along with slower clocking requirement projects that the energy requirement of these circuits can benefit at least by a factor of n^2.

TABLE III. COMPARISON OF MAPPINGS

Comparison Metric	n-mapping	n^2-mapping
Oscillators	n	n^2
Area	1 X	n X
Computations/cycle	$O(n)$	$O(n^2)$
Time to Converge	1 X	n X
Energy Consumption	1 X	n^2 X
Sampling Frequency	1 X	n X
Search Space	n^n	n^{n^2}

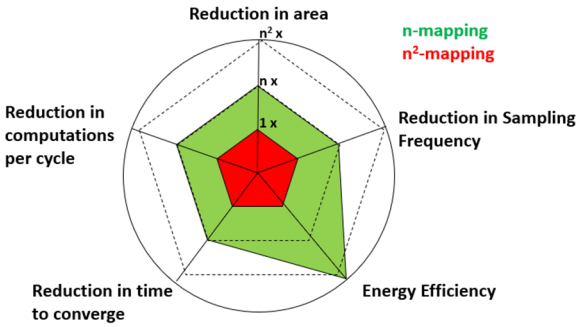

Fig. 7. Spider chart for improvements in various hardware aspects of n-mapping

V. CONCLUSIONS

In this paper, we propose an efficient oscillator phases-based energy function mapping for the n-city TSP. n same-frequency oscillators can efficiently represent the valid tours by encoding the ordering of the cities in the phases. This not only reduces the number of oscillators compared to the state-of-the-art Hopfield Tank network inspired n^2 -oscillator mapping but also reduces the solution search space significantly. The proposed energy function achieves same success probabilities at much lower (~ 100x) annealing durations. With compact device level nanoscale oscillators gaining attention in the neuromorphic hardware space, such efficient mappings can go a long way in application specific circuit level realizations of graphical optimization problems.

ACKNOWLEDGEMENT

The work was supported in part by SRC and NNETRA Project (DST & MeitY, Govt. of India).

REFERENCES

[1] Z. Jonke, S. Habenschuss, and W. Maass, "Solving constraint satisfaction problems with networks of spiking neurons," *Front. Neurosci.*, vol. 10, no. MAR, pp. 1–16, Mar. 2016, DOI: 10.3389/fnins.2016.00118.

[2] D. Khilwani, V. Moghe, S. Lashkare, V. Saraswat, P. Kumbhare, M. Shojaei Baghini, S. Jandhyala, S. Subramoney, and U. Ganguly, "PrxCa1−xMnO3 based stochastic neuron for Boltzmann machine to solve 'maximum cut' problem," *APL Mater.*, vol. 7, no. 9, p. 091112, Sep. 2019, DOI: 10.1063/1.5108694.

[3] J.-Y. Potvin, "State-of-the-Art Survey—The Traveling Salesman Problem: A Neural Network Perspective," *ORSA J. Comput.*, vol. 5, no. 4, pp. 328–348, 2008, DOI: 10.1287/ijoc.5.4.328.

[4] G. Gutin and A. P. Punnen, *The Traveling Salesman Problem and its Variations*, *Paradig. Comb. Optim. Probl. New Approaches 2nd Ed.*, vol. 12, 2004, DOI: 10.1002/9781119005353.ch7.

[5] R. Matai, S. Singh, and M. Lal, "Traveling Salesman Problem: an Overview of Applications, Formulations, and Solution Approaches," *Travel. Salesm. Probl. Theory Appl.*, 2012, DOI: 10.5772/12909.

[6] K. A. Smith, "Neural networks for combinatorial optimization: A review of more than a decade of research," *INFORMS J. Comput.*, vol. 11, no. 1, pp. 15–34, 1999, DOI: 10.1287/ijoc.11.1.15.

[7] J. S. Chen, X. Y. Zhang, and J. J. Chen, "An elastic net method for solving the traveling salesman problem," *Proc. 2007 Int. Conf. Wavelet Anal. Pattern Recognition, ICWAPR '07*, vol. 2, pp. 608–612, 2008, DOI: 10.1109/ICWAPR.2007.4420741.

[8] R. Durbin and D. Willshaw, "An analogue approach to the travelling salesman problem using an elastic net method," *Nature*, vol. 326, no. 6114, pp. 689–691, 1987, DOI: 10.1038/326689a0.

[9] F. Valdez, O. Castillo, and P. Melin, "Ant colony optimization for the design of Modular Neural Networks in pattern recognition," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2016-October, pp. 163–168, 2016, DOI: 10.1109/IJCNN.2016.7727194.

[10] F. Valdez, F. Moreno, and P. Melin, "A Comparison of ACO, GA and SA for Solving the TSP Problem BT - Hybrid Intelligent Systems in Control, Pattern Recognition and Medicine," O. Castillo and P. Melin, Eds. Cham: Springer International Publishing, 2020, pp. 181–189.

[11] J. Mańdziuk, "Solving the Travelling Salesman Problem With a Hopfield-Type Neural Network," *Demonstr. Math.*, vol. 29, no. 1, 2018, DOI: 10.1515/dema-1996-0126.

[12] J. J. Hopfield and D. W. Tank, "'Neural' computation of decisions in optimization problems," *Biol. Cybern.*, vol. 52, no. 3, pp. 141–152, 1985, DOI: 10.1007/BF00339943.

[13] C. C. Liu and C. Chen, "SET based Boltzmann machine and Hopfield neural networks," *Proc. IEEE Conf. Nanotechnol.*, pp. 413–416, 2011, DOI: 10.1109/NANO.2011.6144315.

[14] N. Tezak, T. Van Vaerenbergh, J. S. Pelc, G. J. Mendoza, D. Kielpinski, H. Mabuchi, and R. G. Beausoleil, "Integrated coherent ising machines based on self-phase modulation in microring resonators," *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 1, Jan. 2020, DOI: 10.1109/JSTQE.2019.2929184.

[15] Gang Feng and C. Douligeris, "Using Hopfield networks to solve traveling salesman problems based on stable state analysis technique," no. February, pp. 521–526 vol.6, 2002, DOI: 10.1109/ijcnn.2000.859448.

[16] G. S. Duane, "A 'cellular neuronal' approach to optimization problems," *Chaos*, vol. 19, no. 3, 2009, DOI: 10.1063/1.3184829.

[17] M. Prezioso and D. B. Strukov, "Versatile stochastic dot product circuits based on nonvolatile memories for high performance neurocomputing and neurooptimization," *Nat. Commun.*, no. 2019, pp. 1–10, DOI: 10.1038/s41467-019-13103-7.

[18] D. M. Wells, "Solving degenerate optimization problems using networks of neural oscillators," *Neural Networks*, vol. 5, no. 6, pp. 949–959, Nov. 1992, DOI: 10.1016/S0893-6080(05)80091-7.

[19] A. Varma and Jayadeva, "A novel digital neural network for the travelling salesman problem," *ICONIP 2002 - Proc. 9th Int. Conf. Neural Inf. Process. Comput. Intell. E-Age*, vol. 3, pp. 1320–1324, 2002, DOI: 10.1109/ICONIP.2002.1202835.

[20] S. Li, X. Liu, S. K. Nandi, D. K. Venkatachalam, and R. G. Elliman, "Coupling dynamics of Nb/Nb2O5 relaxation oscillators," *Nanotechnology*, vol. 28, no. 12, 2017, DOI: 10.1088/1361-6528/aa5de0.

[21] S. Lashkare, P. Kumbhare, V. Saraswat, and U. Ganguly, "Transient Joule Heating-Based Oscillator Neuron for Neuromorphic Computing," *IEEE Electron Device Lett.*, vol. 39, no. 9, pp. 1437–1440, 2018, DOI: 10.1109/LED.2018.2854732.

[22] N. Shukla, A. Parihar, E. Freeman, H. Paik, G. Stone, V. Narayanan, H. Wen, Z. Cai, V. Gopalan, R. Engel-Herbert, D. G. Schlom, A. Raychowdhury, and S. Datta, "Synchronized charge oscillations in correlated electron systems," *Sci. Rep.*, vol. 4, p. 4964, May 2014, DOI: 10.1038/srep04964.

[23] A. Parihar, N. Shukla, S. Datta, and A. Raychowdhury, "Computing with dynamical systems in the post-CMOS era," *2016 IEEE Photonics Soc. Summer Top. Meet. Ser. SUM 2016*, vol. 2, pp. 110–111, 2016, DOI: 10.1109/PHOSST.2016.7548777.

[24] T. C. Jackson, A. A. Sharma, J. A. Bain, J. A. Weldon, and L. Pileggi, "Oscillatory neural networks based on TMO nano-oscillators and multi-level RRAM cells," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 5, no. 2, pp. 230–241, 2015, DOI: 10.1109/JETCAS.2015.2433551.

[25] M. N. Bojnordi and E. Ipek, "Memristive Boltzmann machine: A hardware accelerator for combinatorial optimization and deep learning," *2017 5th Berkeley Symp. Energy Effic. Electron. Syst. E3S 2017 - Proc.*, vol. 2018-Janua, pp. 1–3, 2018, DOI: 10.1109/E3S.2017.8246178.

[26] K. Moon, E. Cha, J. Park, S. Gi, M. Chu, K. Baek, B. Lee, S. Oh, and H. Hwang, "High density neuromorphic system with Mo/Pr0.7Ca0.3MnO3 synapse and NbO 2 IMT oscillator neuron," *Tech. Dig. - Int. Electron Devices Meet. IEDM*, vol. 2016-Febru, pp. 17.6.1-17.6.4, 2015, DOI: 10.1109/IEDM.2015.7409721.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE