A

Presented to
the faculty of the School of Engineering and Applied Science
University of Virginia

in partial fulfillment
of the requirements for the degree

by

APPROVAL SHEET

This

is submitted in partial fulfillment of the requirements
for the degree of

Author:

Advisor:

Advisor:

Committee Member:

Committee Member:

Committee Member:

Committee Member:

Committee Member:

Committee Member:

Accepted for the School of Engineering and Applied Science:

Jerry L. West

Jennifer L. West, School of Engineering and Applied Science

To my family and friends

Abstract

While the Von-Neumann architecture-based digital computing framework has been the foundation for modern information processing, there are certain classes of computational problems that are still considered intractable to solve using digital computers. Such problems entail exponentially increasing computing resources with increasing problem size. This has motivated the exploration of alternate computing paradigms. Ising machines have emerged as a promising candidate for solving such computationally challenging problems e.g., the NP-Hard problem, Maximum Cut (Max-Cut) of a graph. However, the effectiveness of this paradigm and its ability to outperform state-of-the-art digital computers will ultimately be decided by the area, scalability, energy efficiency, and performance of the underlying hardware implementation. Current implementations using CMOS, optical phenomena as well as qubits have limitations in terms of these performance parameters.

In this dissertation, the electronic oscillator-based Ising machine (OIM) is investigated due to its promise of CMOS compatibility, compactness, and room-temperature operation. Prior works on coupled oscillator Ising machines have focused on small-scale prototype implementations (4-8 oscillators), while larger designs have been limited to solving only planar and sparse graphs. Therefore, to establish the coupled oscillator platform as a promising candidate for realizing physical Ising machines, its scalability needs to be evaluated- which remains missing and is the primary focus of this dissertation. To accomplish this, a 672 oscillator-based Ising machine IC (65nm GP CMOS technology) with 30,896 programmable coupling elements is developed. Using this platform, a

fundamental trade-off between the solution quality and the time-to-compute is experimentally revealed. Further, we evaluate methods to overcome this tradeoff and propose a hybrid (OIM + heuristic local search) approach. Our projections show that the hybrid approach offers 3-100× improvement in experimentally measured time-to-compute at iso-solution quality, compared to a digital algorithm. Finally, we expand on the dynamical system model to enable solving a broad range of problems requiring multi-valued 'spin' states such as the traveling salesman problem, the graph coloring problem, the Hamiltonian path/cycle problem, etc.

Acknowledgment

First and foremost, I am extremely grateful to my advisor, Dr. Nikhil Shukla, who has been with me throughout my Ph.D. journey, offering critical theoretical and experimental insights. In the years I have spent earning my Ph.D., I believe to have made significant academic, professional, and personal growth. For this, I would like to sincerely express my gratitude to my mentor, Dr. Nikhil Shukla and I truly consider myself lucky for having the opportunity to be mentored by him.

I am thankful to my committee members, Dr. Avik Ghosh, Dr. Zongli Lin, Dr. Mona Zebarjadi, and Dr. Kyusang Lee for their constructive feedback and guidance which has been instrumental in the preparation of my dissertation proposal and defense.

I would like to thank my fellow graduate students, past and present, from the CHRL group, specially Mohammad Khairul Bashar.

Finally, but most importantly, I would like to thank my family and friends. I am eternally grateful to my parents, Debabrota Mallick and Shilpi Mallick, for their constant and unconditional support. I would like to express my gratitude to my sister, Anushua Mallick, for always being encouraging in my endeavors. I want to thank two of my friends, Fazlay Rabbi Masum and Moogdho Mahzab for being a constant source of inspiration and joy.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1. Motivation

The idea of high-performance computing with applications that have transcended various facets of our society has primarily developed around the Von-Neumann architecture [1] which implements the "Turing Machine" model of computation [2] with a clear distinction between processing units and memory. Such a computing framework—the work-horse of modern information technology— serves as a ubiquitous computational model that solves problems by decomposing them into a set of sequential binary operations [3-11]. However, not all computing problems are created equal. Despite the tremendous success of digital computers, it becomes increasingly inefficient in solving certain classes of computational problems. Such problems, commonly known as NP-hard (non-deterministic polynomial times hard) problems, require an exponentially increasing number of operations for increasing sizes of the input problem - a fundamental property of the NP-hard computational classes of problems. Many Combinatorial optimization problems belong to this NP-Hard class of computational complexity [12-16]. It entails the maximization/minimization of an objective function in the combinatorial/discrete domain under a given constraint and finds a wide range of applications from artificial intelligence [17], and communication networks [18], data mining [19] to molecular biology [20], biocomplexity [21], molecular chemistry [22][23], scientific discovery [24]. However, in order to find the optimal solution to these problems, the solution space that needs to be

investigated increases exponentially with the problem size. As a result, these problems are still considered intractable to solve using digital computers as it requires exponentially increasing resources (computation time and memory) [25][26]. This is further aggravated by the Von-Neumann bottleneck [27-32] arising from the limited transfer rate between the processing and memory unit (the processing unit is idle while the memory block is accessed). Therefore, this has naturally motivated the quest to explore beyond-digital computing fabrics [33-43] that can confront the challenges associated with solving such a computationally hard class of problems.

One potential approach is to map such problems to the Ising model [44] [45] originally developed for interacting spin systems. Objective functions of many combinatorial optimization problems (such as the Max-Cut problem, illustrated in chapter 2) can be formulated as the minimization of the Ising Hamiltonian [46]. This has motivated the investigation of physical Ising machine platforms for solving computationally hard problems. In this dissertation, the coupled electronic oscillator-based Ising machine (OIM) is extensively evaluated as it facilitates CMOS compatibility, compactness, and room-temperature operation, compared to other approaches (detailed in the next section). OIMs exploit the fact that the energy function of the oscillator dynamical system can be mapped to the Ising model and consequently, as the system evolves to minimize its energy and reach its ground state (corresponding to optimal solution), it will naturally compute the solution to the problem [47]. However, while the coupled oscillator-based approach is promising, system scalability will be a critical factor in deciding its eventual success; the evaluation of which remains missing in the existing literature.

Therefore, in this dissertation, to achieve this overarching goal of evaluating the system scalability, first the implementation and demonstration of a scaled oscillator Ising machine IC of 672 oscillators and 30896 coupling elements is performed. The hardware platform uniquely facilitates categorical investigation of the challenges and the critical scaling trends not only in terms of implementation but also in terms of computational performance (solution accuracy, compute-time) at scaled nodes. In addition, a hardware-algorithm-based hybrid approach is demonstrated to circumvent these challenges and to facilitate a successful realization of scaled oscillator Ising machines. Moreover, it is to be noted that prior works on oscillator Ising machines have been limited to solving a small subset of optimization problems such as the MaxCut. Solutions to such problems constitute binary spin states (± 1). However, there is a broad range of problems in combinatorial optimization, computing solutions of which require multiple spin states such as traveling salesman, graph coloring, Hamiltonian cycle, etc., and utilization of OIMs to compute the solution to such problems remains missing. Therefore, in the dissertation, the applications of OIM are expanded by presenting its capability to compute problems, with solutions beyond binary spin configurations.

1.2. Prior Studies

There has been an active research effort to realize a physical Ising machine that inherently evolves to minimize its energy and, thus, naturally solves the Ising problem. Examples of such demonstrations include the D-Wave quantum annealer [48–53], optical parametric oscillator-based coherent Ising machines (CIMs) [54–59], and SRAM-based Ising machines that use CMOS annealing [60], as well as the CMOS annealing processors that use processing—in-/near—memory [61–63]. Quantum annealers proposed by D-Wave systems require cryogenic temperature (-273 ^{o}C) to operate and necessitate high cooling power (~25 KW) and a large volume (700 ft^3) for implementing the necessary cooling system [64] [65]. On the other hand, parametric oscillator-based (optical) coherent Ising machines have been proposed where Ising spins are represented as laser pulses traveling on multiple km long optical fiber, which limits the energy and area efficiency [66]. Moreover, digital implementations require stochastic random number generators as an annealing mechanism over thousands of cycles where the spin states need to be repeatedly evaluated and updated [62].

Coupled oscillators have also been explored as an alternate non-Boolean approach to solving computationally hard problems [67–79] and, more importantly, have recently been shown to behave as Ising machines [47,80–83]. Wang *et al.* [47], and Chou *et al.* [81] demonstrated Ising machines using four resistively coupled sinusoidal oscillators operating under the influence of a second-harmonic injection signal, and Dutta et al. [82] showed a similar functionality in four capacitively coupled injection-locked VO_2 oscillators. Moreover, Ahmed *et al.* [83] recently demonstrated a scaled IC of 560 hexagonally connected CMOS-based oscillators to solve the Max-Cut in large planar graphs. In

contrast to other approaches mentioned above, such oscillator Ising machines (OIMs) provide a room-temperature solution that can be realized in a compact integrated circuit using CMOS as well as emerging technologies and attest to the increasing interest in exploring coupled oscillators to solve computationally hard problems.

However, even though OIMs provide a promising pathway to solving computationally expensive combinatorial optimization problems, several challenges need to be addressed in order to develop this platform as a key candidate for Ising hardware, which remains missing in the existing literature. Prior works [80-83] have focused on demonstrating the functionality of OIM for extremely sparse and planar graphs. Such graphs are not NP-Hard and can be easily solved using digital algorithms in polynomial time [84]. Therefore, one of the challenges that need to be addressed is the implementation of a fully reconfigurable OIM for non-trivial system size and evaluating the impact of graph connectivity (both dense and sparse non-planar graphs) on the solution accuracy of OIMs. Moreover, the assessment of the scalability of OIM is paramount for its successful realization as a technology which remains missing in the prior works. As we scale up the system size, two key challenges evolve that are required to be tackled. Firstly, the number of coupling elements that are required increases quadratically with system size, which can considerably limit the area and reconfigurability that can be achieved. Secondly, the complexity of the solution space that needs to be evaluated by OIM at scaled nodes increases significantly, and subsequently, there can be a significant degradation in solution quality. As a result, the evaluation of these fundamental trade-offs in scaled OIMs need to be addressed for its eventual success. Furthermore, most of the works in the literature involving OIMs have focused on a small subset of combinatorial optimization

problems, such as the Max-Cut, the solution of which entails only binary spin configurations. Consequently, a key challenge is to demonstrate the applicability of OIM to solve a vast subset of problems whose solutions entail multi-valued spin configurations such as the Max-K-Cut, Travelling Salesman, Graph Coloring, Hamiltonian cycle/path, Graph Partitioning, etc., which remains unexplored in the prior works. Therefore, in this dissertation, we extensively evaluate each of these challenges and their potential solutions, which are detailed in the following chapters, as illustrated in the next section (dissertation organization).
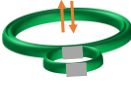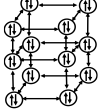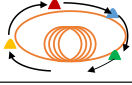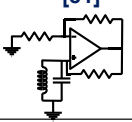
	D-Wave [66]	CMOS Annealing [62]	Coupled Oscillator-based Ising Machines			
			Coherent Ising Machine [66][85]	LC Oscillators [81]	IMT Oscillators [82]	CMOS Ring Oscillators [83]
Basic Computing Unit (Variable)	Qubit (Quantum Annealing potentially reduces complexity)	SRAM (Charge)	Optoelectronic Parametric Oscillators (Coherent Light)	LC Tank Oscillators (Charge)	IMT Relaxation Oscillators (Charge)	Ring Oscillators
Coupling	Quantum Tunneling	SRAM cells store coefficients	Parametric Conversion	Resistive (Discrete)	Capacitive (Discrete)	Latch
Coupling Scheme	Nearest Neighbor	2-layer 2D lattice (limited graph configurations)	All-to-All (Reconfigurable)	Discrete Components	Discrete Components	-Hexagonal planar coupling -Only planar graphs can be solved
Power	25 KW (w/ Cooling)	Low Power 57mW (total power)	3.58 W	11 mW/oscillator [86]	Low Power 20 µW/ oscillator	Low Power 41 µW/ oscillator
Operating Temperature	0.015 K	300K	300K	300K	300K	300K
Scalability Bottleneck	-Cryogenic operation -Large power required for cooling	-Limited programmability -Limited no. of graph can be processed	~Km long optical fiber needed	-Bulky oscillators due to inductor requirement -Discrete coupling components	-Discrete coupling components	-Planar coupling scheme limits programmability

Fig. 1.1. Summary of different approaches for physical implementation of the Ising model.

1.3. Dissertation Organization

The dissertation is organized as follows: Chapter 1, as was already presented, discusses the motivation behind this research and the prior studies in this research area.

Chapter 2 first discusses the Ising model and the equivalence between the energy function of the coupled oscillator network and the Ising Hamiltonian. Subsequently, an integrated circuit (IC) consisting of 30 coupled oscillators with fully programmable capacitive coupling ($^{30}P_2$=870) among them (*first demonstration of fully reconfigurable electronic oscillator Ising machine IC*) is demonstrated, which allows the processing of graphs (up to 30 nodes) with arbitrary connectivity. The fully reconfigurable feature of the OIM implementation enables extensive evaluation of the impact of graph connectivity (a varied range of sparse and dense graphs were randomly generated) on the solution quality and the computational properties of OIM, to solve the NP-Hard Max-Cut problems in non-planar graphs.

Chapter 3 presents the investigation of the scalability of OIMs and addresses the key challenges with scaling up the system both with regard to implementation and computational performance (solution time & accuracy). To enable scalable OIM implementation (672 oscillators and 30896 coupling elements) we adopt a tile-based coupling architecture that facilitates a maximum degree of 111 i.e. an oscillator can be connected to 111 other oscillators in the network. Moreover, this platform uniquely enables us to characterize non-planar graphs of varying connectivity (NP-Hard) at scaled nodes and experimentally reveal the accuracy vs. time-to-compute trade-off of OIMs, as

we increase the system size. Subsequently, a hybrid approach is proposed and experimentally demonstrated to optimize the trade-off in OIMs.

While the previous chapters primarily focus on the oscillator Ising machines to solve the MaxCut problem (solution constituting only binary spin configurations), in chapter 4, oscillator Ising machine work is extended from computing Max-Cut to solving a large subset of problems in combinatorial optimization such as Max-K-Cut, Travelling Salesman, Graph Coloring, Hamiltonian cycle/path, etc. (solution constituting multi-valued spin configurations) by developing novel synchronized oscillator based computing models. Therefore, synchronized oscillator system dynamics are expanded and showed that it can be engineered to solve different optimization problems by appropriately designing the coupling function and the external injection to each oscillator.

 Chapter 5 presents future avenues of research topics that can be a potential extension of this dissertation work. Chapter 6 concludes by summarizing the key outcomes and contributions of the dissertation.

Chapter 2

Coupled Oscillator Based Ising Machines (OIM)

In this chapter, the direct mapping of the Max-Cut problem to the Ising Hamiltonian: $H = -\sum_{i,j}^{N} J_{ij} s_i s_j$ is shown, where each spin s corresponds to a node of the graph and can take binary values $s \in \{\pm 1\}$, N is the total number of nodes in the problem, and J_{ij} is the interaction coefficient between nodes i and j. Computing the Max-Cut solution then corresponds to minimizing H [47]. In addition, a coupled relaxation oscillator IC to solve the Max-Cut problem in non-planar (NP-Hard) graphs is demonstrated. The platform incorporates: (a) 30 programmable CMOS Schmitt-trigger-based relaxation oscillators that operate under the influence of a second-harmonic injection signal ($f_{inj} \cong 2*f_R$ where f_R is the resonant frequency); (b) Reconfigurable and symmetric capacitive coupling among the oscillators i.e. any oscillator can be coupled to any and all other oscillators which allows us to process a graph (up to 30 nodes) with arbitrary connectivity. Moreover, the reconfigurability incorporated in the design provides a unique opportunity to characterize and evaluate the dynamics and computational properties of the system over a range of graph sizes and connectivity.

2.1. The Ising Model

The Ising model is named after German physicist Ernest Ising who in 1925 developed a mathematical model for explaining domain formation in ferromagnets [44]. It comprises a

group of discrete variables spins (s_i), each taking a binary value ±1, such that an associated "energy function", known as the Ising Hamiltonian, is minimized:

$$H = -\sum_{i,j}^{N} J_{ij} s_i s_j - \sum_{i}^{N} h_i s_i$$

where N represents the number of nodes in a particular problem set, J_{ij} represents the weight values interconnecting the nodes, $s_i = [s_i \ldots s_n]$ represents the solution space where s_i can take the value of either +1 (spin ↑) or −1 (spin ↓), and h_i is the local field parameter. For the problem of interest here we drop the h_i term and under this simplification, the Ising Hamiltonian becomes $H = -\sum_{i,j}^{N} J_{ij} s_i s_j$. The Ising Model has become particularly interesting as many hard optimization problems can be mapped to it. Recently, in 2014, A. Lucas et al. have shown that all of Karp's 21 NP-complete problems can be mapped to the Ising model by assigning appropriate values to the coefficients [46]. As a result, developing a physical system that can minimize Ising Hamiltonian (Ising Machines) has become an active area of research for solving such computationally challenging problems owing to its promise of outperforming conventional digital algorithms.

2.1.1. Ising Formulation of the Max-Cut problem

The Max-Cut problem can be mapped directly to the Ising Hamiltonian, and the solution to the problem is represented by the state s which minimizes H. The Max-Cut of a graph G(V, E) (V: vertices; E: edges) is a cut that divides G into two sets such that the number

of common edges between them is as large as possible; the number of common edges is the value of the Max-Cut.

To explain this mapping scheme, the vertices of a graph are divided into two sets- V1 and V2. Accordingly, all the edges in the graph are separated into two groups- those that connect vertices within V1 (W1), those within V2 (W2), and the cut set containing edges across V1 and V2 (Wcut). By convention, J_{ij} is considered to be the opposite of the weight of the edge between vertices i and j, i.e., $J_{ij} = -W_{ij}$. Then, W1+W2+Wcut $= \sum_{i,j} W_{ij} = -\sum_{i,j} J_{ij}$. We can write the Ising Hamiltonian as follows:

$$H = -\sum_{i,j} J_{ij} s_i s_j$$

$$= -\sum_{i,j} (-w_{ij}) s_i s_j$$

$$= \sum_{i,j \in V1} w_{ij}(+1).(+1) + \sum_{i,j \in V2} w_{ij}(-1).(-1) + \sum_{\substack{i \in V1, \\ j \in V2}} w_{ij}(+1).(-1)$$

$$= \sum_{i,j \in V1} w_{ij} + \sum_{i,j \in V2} w_{ij} - \sum_{\substack{i \in V1, \\ j \in V2}} w_{ij}$$

$$= (W1 + W2 - Wcut)$$

$$= \sum_{i,j} w_{ij} - Wcut - Wcut$$

$$= \sum_{i,j} w_{ij} - 2Wcut$$

So, Ising Hamiltonian = [sum of all weights] – 2×[cut size]

Therefore, when the Ising Hamiltonian is minimized, the cut size is maximized (corresponding to the Max-Cut solution). Due to this direct mapping between the Ising model and the max-cut problem, it is commonly used to benchmark the computational performance of the Ising machine.

2.1.2. Minimizing the Ising Hamiltonian Using Coupled Oscillators

Wang et al. in 2017 [47] were the first to develop the model showing the equivalence between coupled oscillator platform and the Ising model which is briefly discussed below. In order to map the Ising model to the coupled oscillator network, the spins (i, j) are represented by oscillators, the interaction co-efficient (J_{ij}) is represented by the coupling element and the spin state (+1 ↑, -1 ↓) is represented by the oscillator phases (0° or 180°). To binarize the oscillator phases, an external second harmonic injection signal (SHIL) is applied to all the oscillators in the network which induces each oscillator to settle to one of two stable phase-locked states (0° or 180°) [47][87]. It is to be noted that the external SHIL signal helps induce the bipartition (Fig. 2.1); without this signal, the oscillators exhibit a continuous phase ordering (as shown in prior works [70,88-89]).

The continuous-time dynamics of the phase difference between the oscillator and injection-locking signal can be described by [87][90],

$$\frac{d\left(\phi_i(t)\right)}{dt} = -K \sum_{j=1, j\neq i}^{n} J_{ij} \sin\left(\phi_i(t) - \phi_j(t)\right) - Ks. \sin\left(2\phi_i(t)\right) \tag{1}$$

where K is the coupling strength of the network, Ks is the strength of the SHIL signal, and J_{ij} sets the connectivity of the network.

The continuous-time dynamics of the network are dictated by the energy of the system, or the Lyapunov function $E(\phi)$, which is expressed as,

$$E\left(\phi(t)\right) = -K \sum_{i,j, j\neq i} J_{ij} \cos\left(\phi_i(t) - \phi_j(t)\right) - Ks. \cos\left(2\phi_i(t)\right) \tag{2}$$
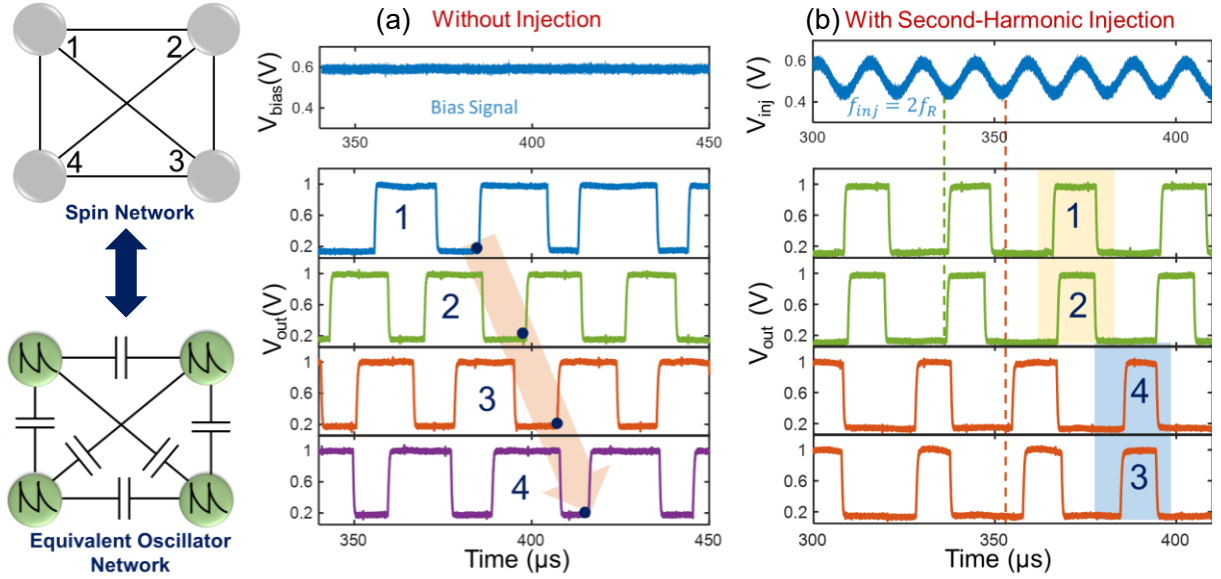
Fig. 2.1. Evolution of the oscillator phases (a) without and (b) with second-harmonic injection signal. Continuous phase ordering is observed in absence of SHIL whereas binarization of oscillator phases is observerd in presence of SHIL.

Now, differentiating, equation (2) with respect to φ,

$$\frac{dE\big(\phi(t)\big)}{d\phi_i} = -K\sum_{l=1,\,l\neq k}^{n} J_{kl}\frac{d}{d\phi_k}\cos\big(\phi_k(t)-\phi_l(t)\big) - K\sum_{l=1,\,l\neq k}^{n} J_{lk}\frac{d}{d\phi_k}\cos\big(\phi_l(t)-\phi_k(t)\big)$$

$$-Ks.\frac{d}{d\phi_k}\cos(2\phi_k(t))$$

$$= K\sum_{l=1,\,l\neq k}^{n} J_{kl}\sin\big(\phi_k(t)-\phi_l(t)\big) - K\sum_{l=1,\,l\neq k}^{n} J_{lk}\sin\big((\phi_l(t)-\phi_k(t)\big) + 2.\,Ks.\,\sin(2\phi_k(t))$$

$$= 2.\,K\sum_{l=1,\,l\neq k}^{n} J_{kl}\sin\big(\phi_k(t)-\phi_{lj}(t)\big) + 2.\,Ks.\,\sin\big(2\phi_k(t)\big)$$

$$= -2.\frac{d\big(\phi_k(t)\big)}{dt}$$

24

Therefore,

$$\frac{dE(\phi(t))}{dt} = \sum_{k=1}^{n} \left[\left(\frac{dE(\phi(t))}{d\phi_k(t)} \right) \cdot \left(\frac{d\phi_k(t)}{dt} \right) \right]$$

$$= \sum_{k=1}^{n} \left[\left(-2 \cdot \frac{d(\phi_k(t))}{dt} \right) \cdot \left(\frac{d\phi_k(t)}{dt} \right) \right]$$

$$= -2 \cdot \sum_{k=1}^{n} \left(\frac{d\phi_k}{dt} \right)^2 \leq 0$$

So, the energy function of the coupled oscillators under SHIL is naturally minimized over time. At points where every phase value (ϕ) is either 0° or 180°,

for $\Delta\phi = 0°$, $\cos\left(\phi_i(t) - \phi_j(t)\right) = s_i \cdot s_j = +1$ and $\cos(2\phi) = 1$ and

for $\Delta\phi = 180°$, $\cos\left(\phi_i(t) - \phi_j(t)\right) = s_i \cdot s_j = -1$ and $\cos(2\phi) = 1$

Therefore, (2) can be rewritten as,

$$E(\phi(t)) = -K \sum_{i,j,\, j \neq i} J_{ij} \cos\left(\phi_i(t) - \phi_j(t)\right) - n. \, Ks$$

$$= -K. \sum_{i,j} J_{ij} s_i s_j - n. \, Ks$$

which is equivalent to the Ising Hamiltonian with a constant offset. Therefore, as the coupled oscillator system under SHIL evolves to minimize its energy and reach the ground state, it minimizes the Ising Hamiltonian and functions as an Ising machine.

2.2 Hardware Implementation of Fully Reconfigurable Oscillator Ising Machine (Test Chip-1)
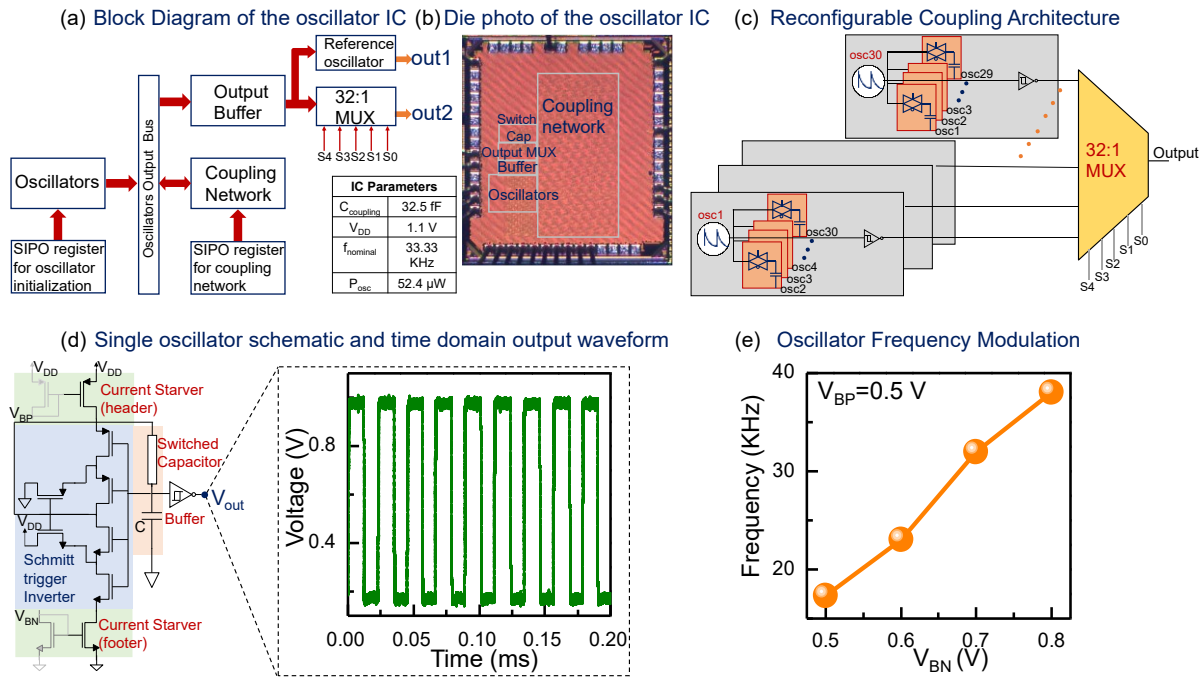
Fig. 2.2. (a) Block diagram; (b) Die photo of the IC consisting of coupled oscillators. (c) Schematic of the fully reconfigurable capacitive coupling scheme. The design facilitates 'all-to-all' connectivity which enables any arbitrary graph (up to 30 nodes) to be processed using the oscillators. (d) Circuit schematic and time domain output waveform of a single oscillator. (e) Programmability of the oscillator frequency using the current starver circuit.

The fully reconfigurable 30-oscillator Ising machine IC is fabricated using the bulk CMOS 65nm technology and occupies a total area of 1.44 mm^2. The measured power dissipation of the chip is 1.76 mW with each oscillator consuming 52.4 µW. Fig. 2.2a shows the block diagram of the IC which incorporates the following:

1. Oscillator. The basic computational unit of the hardware platform is the oscillator which is implemented using a Schmitt trigger inverter module along with negative RC feedback

(Fig. 2.2d). The total area occupied by the oscillator is approximately 28 μm^2. The oscillator implementation utilizes:

A. Switched-Capacitor: The feedback resistor (R_F = 402 MΩ) is implemented using a switched capacitor. The target resistance is achieved by applying two out-of-phase sinusoidal signals of frequency (f) 1.5 MHz onto the gates of the two MOSFETs that connect the capacitor (C_{gg} = 1.6 fF) to the input and output of the switched capacitor circuit; R_F = 1/(f*C_{gg}).

B. Output Buffer: The output of each oscillator is passed through an inverting hysteretic output buffer which digitizes the output while preserving the critical phase information (Fig. 2.2d). The binary output of the oscillator simplifies the measurement and read-out of the oscillator phases.

C. Current starver circuit: The operating parameters of the oscillator such as frequency can be controlled using the current starver circuits implemented at the header and the footer. The current starver helps modulate the charging and discharging current of the oscillator and therefore, the oscillation frequency as shown in Fig. 2.2e.

2. Reconfigurable Coupling Architecture. Fig. 2.2c shows a detailed view of the "all-to-all" reconfigurable coupling scheme. Each oscillator can be coupled to any other and all other oscillators through a T-gate and a coupling capacitor (C_C = 32.5 fF). The coupling elements are programmed according to the adjacency matrix of the input graph. When there is an edge between two vertices i.e. A_{ij} = A_{ji} = '1', the corresponding T-gate associated with the coupling capacitor C_{ij} is switched ON to facilitate capacitive coupling between vertices i and j.

3. I/O Ports

A. Serial-input parallel-output (SIPO) shift registers: Serial-input, parallel-output (SIPO) shift registers are used to program the oscillators (ON/OFF) and the coupling network.

B. 32:1 MUX: A 32:1 Multiplexer is used at the output to select and read the time domain output waveform of each oscillator in a serial fashion. The outputs are read using a PGLA (Pattern Generator Logic Analyzer).

2.3. Evaluation of the Computational Properties of the Oscillator Ising Machine
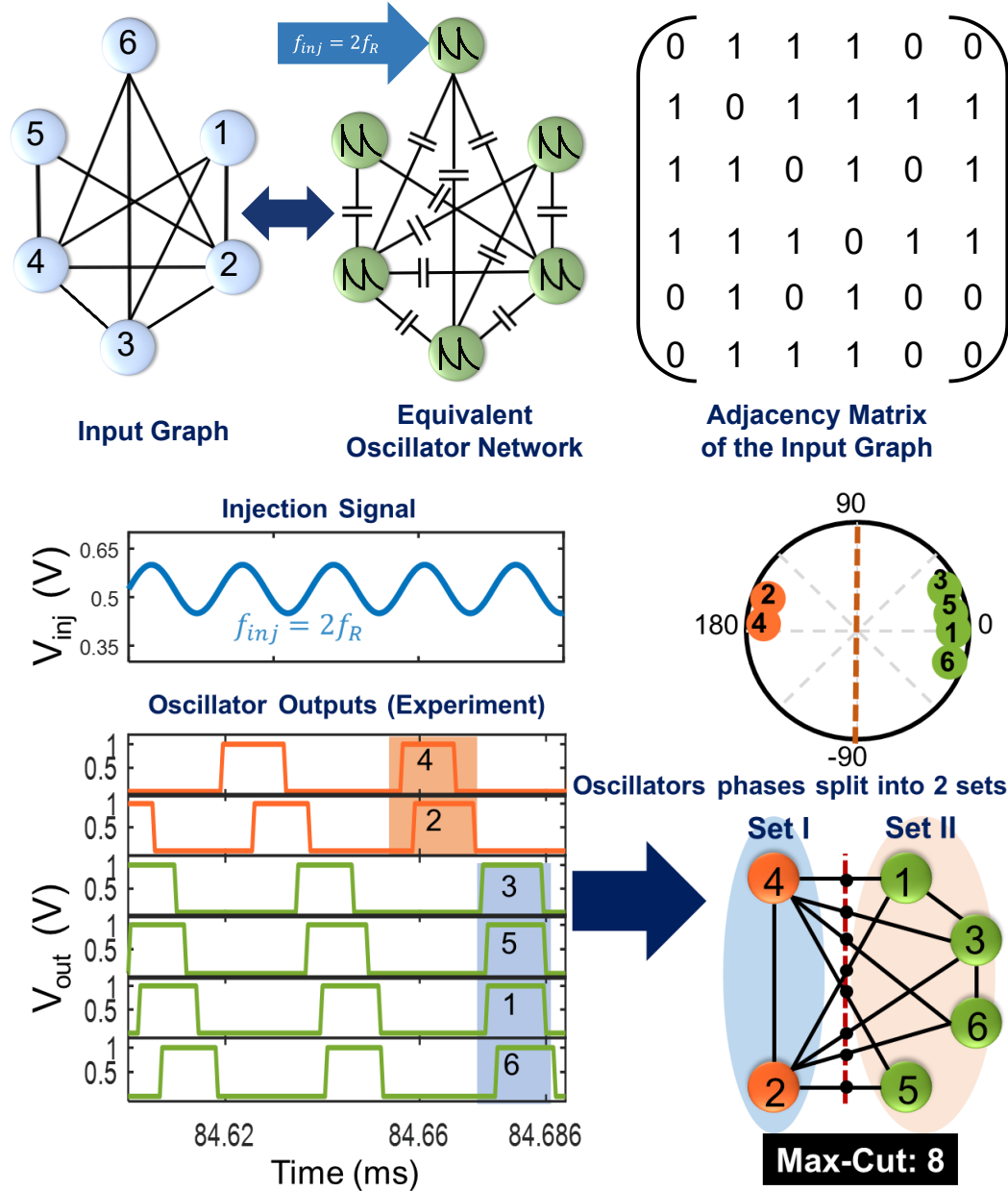
Fig. 2.3. Experimentally measured dynamics of the coupled oscillators (operating under the influence of a second-harmonic injection locking signal) for a representative 6 node graph. It can be observed from the polar plot that the oscillator phases exhibit a bi-partition which can be used to compute a high-quality Max-Cut of the graph.

2.3.1. Computing Max-Cut Using Oscillator Ising Machine

To compute the Max-Cut of a graph using the coupled oscillators, we start with the adjacency matrix A of the graph where A_{ij} indicates the presence (A_{ij}=1) or absence (A_{ij}=0) of an edge between node i and j of the graph. Since undirected graphs are considered here, A_{ij}=A_{ji}. Each node of the graph is mapped to an oscillator and every edge (represented by A_{ij}=A_{ji}) to a coupling capacitor; node ≡ oscillator; edge ≡ coupling capacitor; oscillator phase ≡ set (created by the cut) to which the node belongs. In the context of the proposed hardware, the number of rows (or columns) in A represents the number of oscillators required to process the graph, and A_{ij} is used to configure the corresponding coupling among the oscillators. The capacitors couple the oscillators negatively i.e. oscillators exhibit phase repulsion when capacitively coupled, and have a negative relationship to the edge weight. The matrix A is passed on to the SIPO registers to initialize a topologically similar oscillator network. The external injection signal which is applied to the current mirrors implemented at the header and footer of each oscillator constitutes a sinusoidal signal with a peak-to-peak amplitude of 150mV, DC offset (footer: 0.5 V, header: 0.3 V), and a frequency (f_{inj}) approximately twice the resonant frequency of the coupled circuit, i. e. $f_{inj} \cong 2f_R$. It is to be noted that the external sub-harmonic signal helps to induce the bi-partition i.e. 0º or 180º which corresponds to the two sub-sets created by the Max-Cut. Fig. 2.3 shows the experimentally measured oscillator outputs along with the corresponding phase plot for a representative 6-node graph. A bi-partition in the oscillator phases observed in the polar plot corresponds to the two subsets (Set I and II) created by the Max-Cut and the Max-Cut value can subsequently be computed by counting the number of common edges (=8 in this example) between the sets.

2.3.2 Effect of Graph Size and Connectivity on the Solution Quality
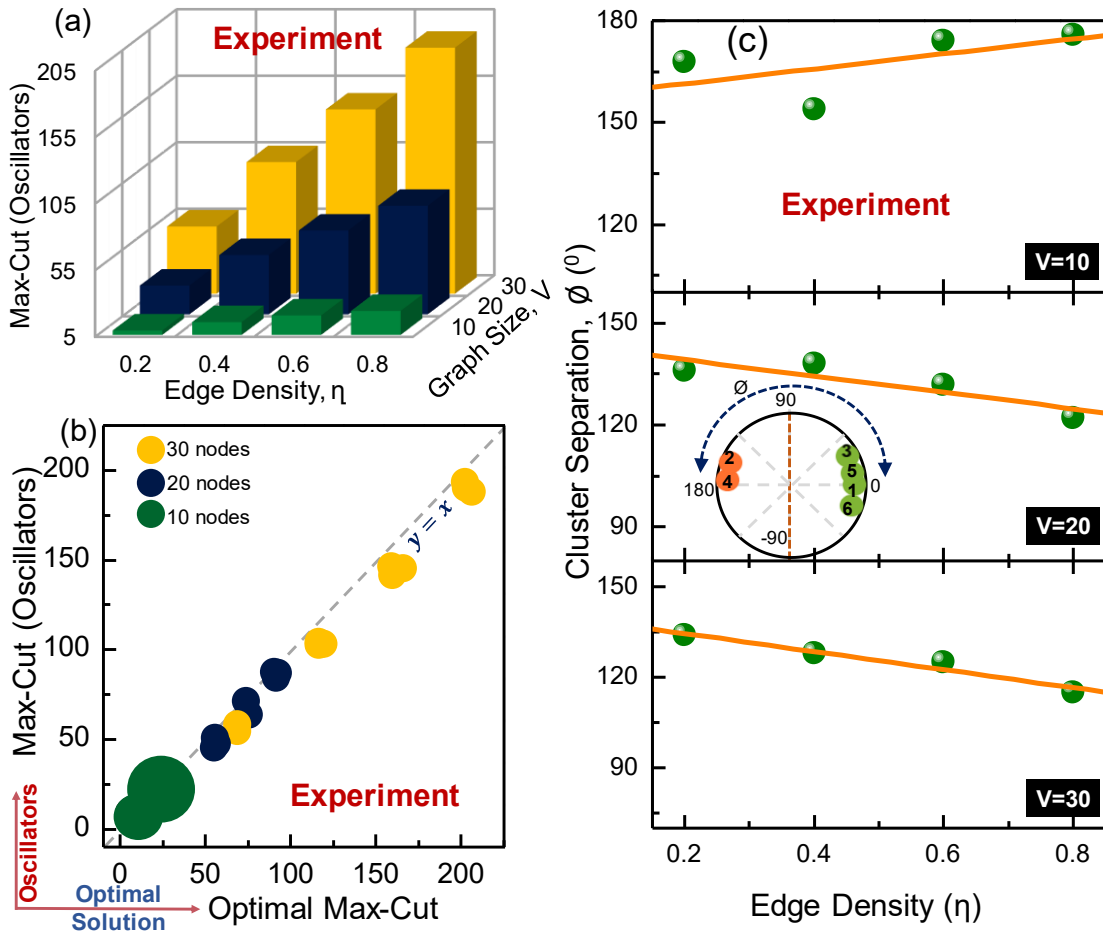
Fig. 2.4. (a) Bar plot showing the measured Max-Cut solutions for 36 randomly generated graph instances as a function of their size and edge density; (b) Bubble plot comparing (best case) Max-Cut solution obtained from the oscillators with the optimal Max-Cut of the graph; (c) Variation of cluster separation (i.e. angular separation between the two oscillator phases) with graph size and edge density.

The computational performance of the hardware is tested by generating random graph instances with V=10, 20, 30 nodes, and having a varied range edge density, η= 0.2, 0.4, 0.6, 0.8 (η is the ratio of the number of edges in the graph to the number of edges in an all-to-all connected graph of the same size); three graphs are tested for each combination of V and η (Fig. 2.4a) with each graph being measured 10 separate times (only non-planar graphs are considered here). It is evident that larger and denser graphs have larger Max-

Cuts, and consequently, are more challenging to solve. Fig. 2.4b shows a bubble plot comparing the value of the measured Max-Cut (best case) using the oscillators with the optimal solution obtained using the BiqMac solver developed by Rendl et al. [91], [92]. It can be observed that the solution to most of the analyzed graphs lies near- or on the identity line (y=x) although larger graphs tend to show higher deviations from the optimal solution. As measured, the oscillator solution is within 99% of the optimal solution in 12 of the 36 graphs. The larger value of the Max-Cut in the larger and denser graphs makes them challenging to solve. This property also manifests in the oscillator dynamics such as the cluster separation (defined as the difference between the mean phases of each cluster) shown in Fig. 2.4c. Larger and denser graphs show reduced phase separation (i.e. more deviation from the ideal 180° phase difference) in comparison to smaller and sparser graphs implying that the system finds it increasingly challenging to attain the global energy minima corresponding to the optimal Max-Cut solution.

To improve the Max-Cut solution obtained from the oscillators, a simple polynomial-time scheme based on local search is proposed as shown in the flow-chart in Fig. 2.5a. Using the Max-Cut solution computed from the oscillators, the scheme proceeds by moving nodes between the sets if and only if the move increases the value of the Max-Cut. This process is repeated until no more nodes can be found that can increase the value of the cut. The cumulative graph count distribution as a function of the distance from the optimal solution (i.e. difference between the Max-Cut solution obtained using the coupled oscillators and optimal Max-Cut) before and after post-processing shows the corresponding improvement in the solution for the experimentally measured graphs (Fig.

(a)

$G = (V,E), G_0 = (V_0, E_0), G_1 = (V_1, E_1)$

Put $i = 0, k = 0, k' = 1, G_{kex} = \emptyset, G_{k'ex} = \emptyset$

Find node v_{i+1} with degree $\Delta(G_k)$ & compute its degree with $G_{k'}$ ($\deg(v)$ in $G_{k'}$); $i = i + 1$

$\Delta(G_k) > \deg(v)$ in $G_{k'}$?

No

Yes

$V_{k'} = V_{k'} \cup \{v\}, \ V_k = V_k - \{v\}, \ G_{kex} = G_{kex} \cup \{v\}$

$\Delta(G_k) = minimum(\deg(v) \ in \ G_k)$?

No

Yes

Swap k and k'; $i = 0$

No

$k = 1$?

Yes

$G_{kex} = \emptyset$ & $G_{k'ex} = \emptyset$?

No

Yes

Find Max-cut by counting edges between G_k & $G_{k'}$

Max-Cut

- G_0, G_1 set partitions obtained from oscillators
- Degree: Number of edges connected to a node
- $\deg(v)$: Degree of vertex v
- $\Delta(G_k)$: Highest degree in G_k

(b) W/ & W/O post-Processing

W/O Post-Processing

W/ Post-Processing

Total 36 graphs

Experiment

Optimal

Cumulative count of graphs (#)

Distance from Optimal Solution

Fig. 2.5. (a) Flow chart for the local search-based post-processing scheme to improve the Max-Cut solution generated by the oscillators. (b) Deviation of the oscillator generated Max-Cut solution (w/ and w/o post-processing) from the optimal solution.

2.5b). The hybrid hardware-algorithm approach produces the optimal Max-Cut in 26 (~72%) of the 36 graphs. Moreover, the oscillators are also effective in solving challenging dense graphs where they produce optimal solutions in 7 out of the 9 measured graphs with edge density η=0.8.

2.4. Summary

In summary, the computational properties of coupled oscillators to solve the NP-Hard Max-Cut problem were experimentally investigated by developing a prototype IC of 30 oscillators with reconfigurable all-to-all coupling. The fully reconfigurable feature of the prototype oscillator Ising machine IC enables exploration of the effect of edge density

(sparsity and density of the graph) on the solution quality for varying problem sizes. The evaluation of the randomly generated graph instances over a varied range of connectivity experimentally reveals that as the problem size increases, the graphs with denser edge density yield lower solution accuracy. Moreover, to improve the solution quality, a polynomial time post-processing algorithm is developed and it shows that using minimal post-processing the solution accuracy can be significantly enhanced even in non-planar dense graphs. However, such lowering of the solution quality as we advance towards larger and denser graphs (even for a problem size of 30 nodes) preludes to the fact that the system scalability will be a critical factor in deciding the eventual success of oscillator-Ising machines- which is the main focal point of discussion in the next chapter.

Chapter 3

Scalability of the Oscillator Ising Machine

In this chapter, the key challenges regarding the scalability of oscillator Ising machines are investigated. To evaluate and ultimately circumvent the challenges of a scaled OIM system, the implementation of a 672 oscillator-based Ising machine IC with 30896 programmable coupling elements is performed. Such reconfigurability in the coupling scheme is achieved by employing a tile-based architecture, which facilitates >13x improvement in terms of connectivity (maximum degree) over prior OIM designs at scaled nodes [80,83]. Subsequently, the role of input graph properties (size & connectivity) on the computational performance (compute time & solution accuracy) is performed which enables experimental demonstration of the inherent trade-off for oscillator Ising machines, where achieving a significant speedup in compute time while maintaining high solution quality at scaled nodes becomes challenging. Consequently, a hybrid approach is developed that exploits the oscillator Ising machine to obtain near-optimal solutions (very fast), which are then subsequently improved using a simple local-search algorithm with minimum time penalty (elucidated in the following sections). Utilizing the hybrid approach, we experimentally demonstrate 3-100× improvement in experimentally measured time-to-compute at iso-solution quality, when compared to a digital algorithm (manifold optimization [93], executed on a 32-core processor with 256GB RAM). Finally, benchmarking of the proposed implementation with the digital algorithm is performed.

3.1 Hardware Implementation of Scaled Oscillator Ising Machine (Test Chip-2)

Die Photo of The OIM Chip   **Oscillator Circuit**   **Coupling Element**

(d) Computational Capability		
-# of oscillators (spins) measured	672	**Read out Interface**
Corresponding Max. # of Edges (coupling elements)	30896	1. **MUX**: Direct read-out of oscillator time-domain output.
Maximum degree of a vertex (max. # of coupling elements available to each oscillator)	111	2. **SPI** (Serial peripheral interface): Parallel read-out of all oscillator states. Used to obtain timing information

Fig.3.1 (a) Die photo of the scaled OIM IC. Architecture of the (b) CMOS Schmitt-trigger oscillator and the (c) coupling element; (d) Table showing computational capability of the oscillator Ising machine.

The CMOS oscillator-based scaled Ising machine IC (Fig. 3.1a) is realized using 672 oscillators implemented using a CMOS Schmitt trigger-based design with RC feedback, as shown in Fig. 3.1b. As entailed by the computational model [47], a second harmonic signal is injected into the oscillators through the header and the footer. The programmable coupling element between any two oscillators is realized using a series combination of a MOM (metal-oxide-metal) capacitor (5.2fF) in series with a transmission gate (T-gate) that can be used to turn the coupling ON/OFF (Fig. 3.1c). The computational
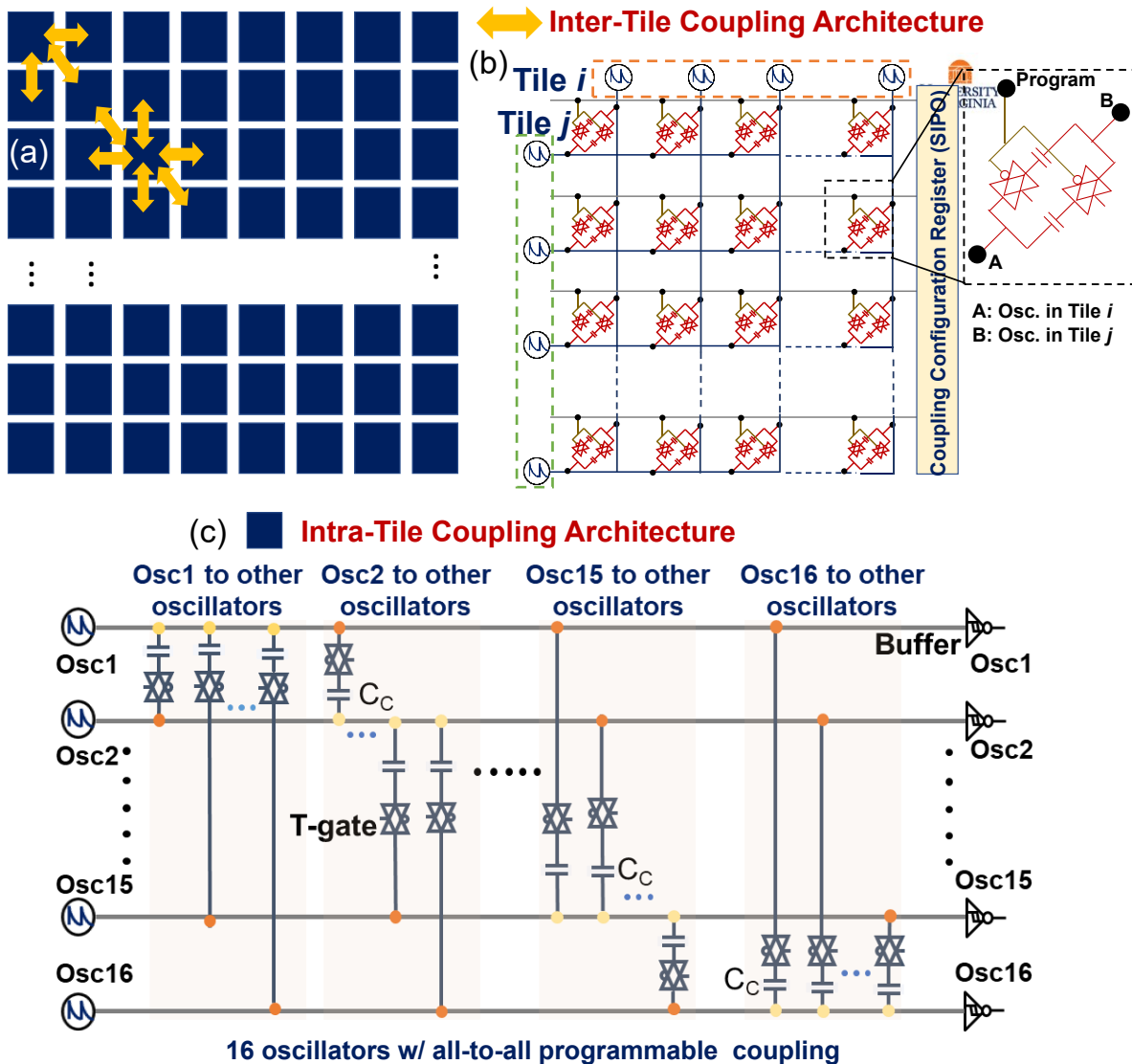
Fig.3.2 (a) Tile-based coupling architecture of the IC. (b) *Inter-tile*: Each oscillator in a tile is coupled to every other oscillator in 6 adjacent tiles i.e., all-to-all programmable coupling among one tile and six adjacent tiles (except at the edges). (c) *Intra-tile*: Each tile consists of 16 oscillators where each oscillator can be coupled to any and all other oscillators.

capabilities are summarized in Fig. 3.1d. The IC is fabricated using TSMC 65nm GP CMOS process. The input graph is mapped to the oscillator hardware using a SIPO (Serial Input Parallel Output) Register. Both MUX and SPI are available for read-out. The MUX enables direct observation of the time-domain output of the oscillators (in a serial fashion) whereas the SPI enables a parallel readout of the oscillators, required to obtain

timing information. To facilitate a scalable coupling scheme, a tiled architecture is adopted as shown in Fig. 3.2a. In the Intra-tile coupling, each tile consists of 16 oscillators (a total of 42 tiles) with all-to-all (programmable) coupling among the oscillators (Fig. 3.2c). *Inter-tile:* each oscillator in a tile is coupled to every other oscillator in 6 adjacent tiles i.e., there is all-to-all programmable coupling among all the oscillators in one tile and all other oscillators in six adjacent tiles (except at the edges) as shown in Fig. 3.2b. Consequently, the design allows us to measure graphs with up to 30,896 coupling elements with nodes having a maximum degree of 111.

It is to be noted that, as the system size is scaled up, implementation of the coupling network in OIM poses a significant challenge as the required coupling element increases quadratically with respect to the system size. For the 672 OIM implemented here, 225,456 ($=C(672,2)$) (non-directional) coupling elements would be needed to achieve full reconfigurability, which is not possible within a reasonable frame of area (2x2 mm^2 area was considered for this chip implementation). Moreover, since most practical graphs at scaled nodes [94] are locally very dense and globally sparse, we, therefore, adopt a tile-based coupling architecture that enables us to have densely connected oscillator clusters with relatively sparse connectivity amongst them to achieve an optimal trade-off between functionality and reconfigurability.

3.2 Accuracy vs. time-to-compute trade-off
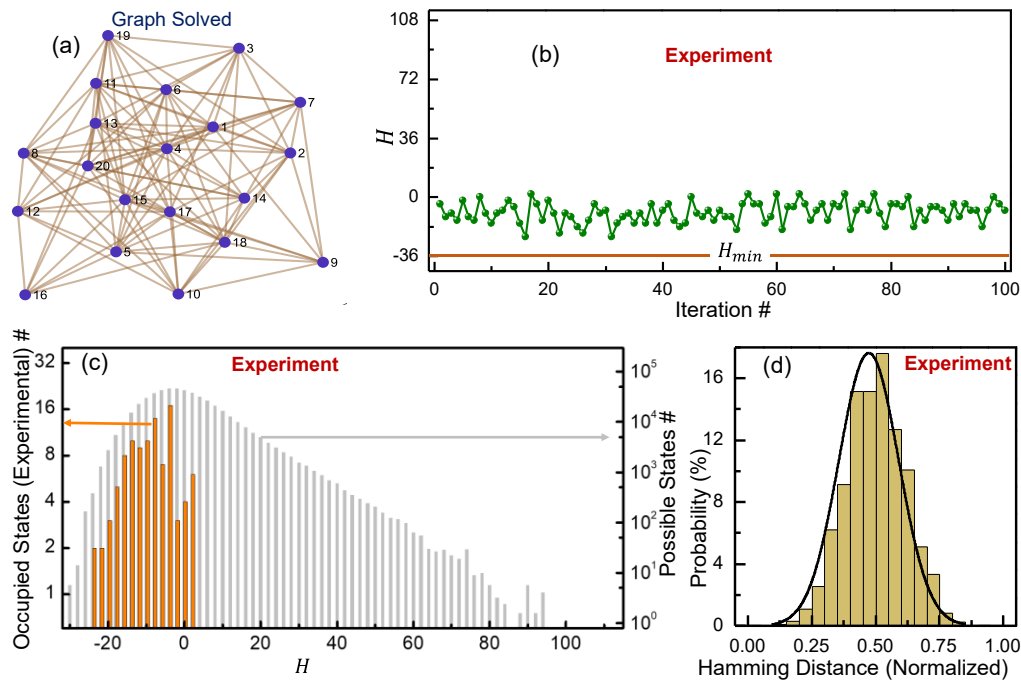
Fig. 3.3. (a) A representative network of 20 spins with randomly generated interactions (represented by edges); (b) Experimentally measured H over 100 separate trials (H_{min}=minimum/optimal H); (c) Distribution showing occupied energy states (represented by H) and their frequency (orange) compared to the complete solution space (grey) of the problem (=524,288 possible states); (d) Hamming distance distribution (normalized) between the experimentally measured solutions over 100 runs.

While the coupled oscillator system evolves through the high-dimensional solution space towards the ground state energy (corresponding to the optimal solution), it is likely to encounter many local minima where the system can get trapped, and subsequently, give rise to a sub-optimal solution. This is illustrated with the example of an oscillator network with 20 nodes and 114 edges (corresponding to the interactions among the oscillators) as shown in Fig. 3.3a. Fig. 3.3b shows the measured solutions (H) over 100 separate trials. Figure 3.3c compares the measured H attained by the system (and its frequency) with the entire combinatorial solution space i.e., H corresponding to all the possible spin assignments (grey in Fig. 3.3c) for the problem; there are 524,288 possible spin

assignments out of which only 7 correspond to the optimal solution. It is evident from Fig. 3.2c that the spin configurations measured using the system are sub-optimal with the best solution only equal to 92% of the optimal value. Moreover, it can also be observed that the system energy (proportional to H) at which the peak of the distribution of the measured solutions (over the 100 trials) occurs, coincides with the H value where the maximum number of local minima states occur. This indicates that the system - despite trying to minimize its energy- gets trapped in one of the many local minima of the phase space, consequently, giving rise to sub-optimal spin assignments. Furthermore, the Hamming distance among the measured spin assignments (Fig. 3.3d) is also computed to explore if there is a correlation among the solutions (generated in each trial). The resulting Hamming distance exhibits a Gaussian distribution implying that the solutions are different from each other and the exact local minima where the system gets trapped is random as the system evolves with a different trajectory during every run. This implies as the complexity of the solution space increases for scaled systems, the solution can get trapped in one of the many local minima resulting in an increasing sub-optimal nature of the solution. Therefore, there is an inherent solution quality vs compute-time trade-off for oscillator Ising machines, where achieving a significant speedup in compute-time while maintaining high solution quality becomes challenging.

3.2.1 Evaluating the trade-off in scaled OIM (Test Chip-2)
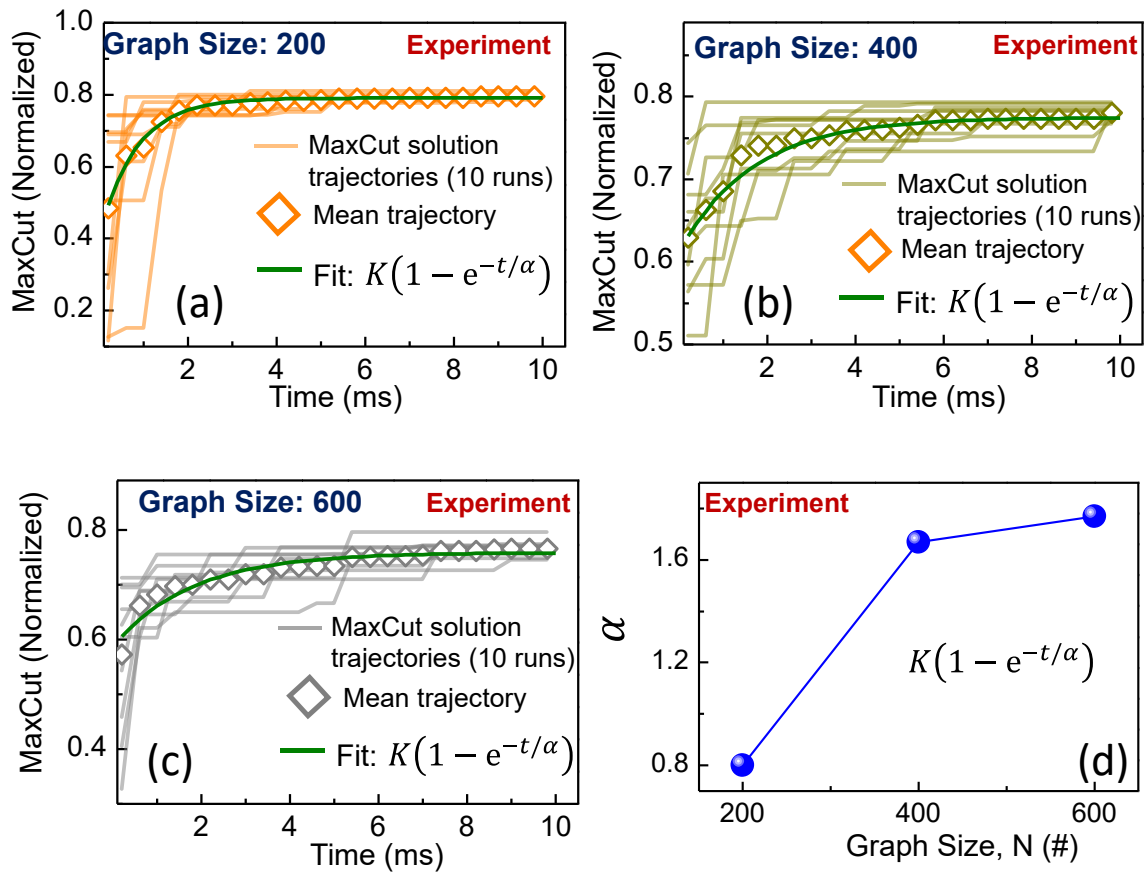
Fig. 3.4. Experimentally measured MaxCut solution vs. computation time for graphs with (a) N=200; (b) N=400; (c) N=600 nodes. 10 solution trajectories for oscillator Ising machine are shown in each case. Exponential behavior ($\propto (1 - \exp(-t/\alpha))$; similar to the charging of a capacitor) is observed. (d) Constant in the exponential function (α) as a function of graph size (N). α increases with N indicating that the convergence time required to obtain an optimal solution will degrade further in larger graphs.

In order to experimentally evaluate the solution accuracy vs. time-to-compute trade-off, the MaxCut solution trajectories as a function of time for varying graph sizes (N=200, 400, 600 nodes) are measured utilizing the scaled OIM hardware. Fig. 3.4(a-c) shows the representative measured MaxCut solution trajectories as a function of time for graph sizes of 200, 400, and 600 nodes, and for each instance, the solution trajectory was measured 10 separate times. It can be observed that, in all cases, the OIM shows an exponential

(capacitor charging-like) behavior wherein it converges to a near-optimal solution and saturates. The mean of the 10 separate measurements is fitted, represented by the diamond symbol in Fig. 3.4(a-c), to the equation $K(1 - exp(-t/\alpha))$. This implies that while the system will compute a near-optimal solution (~80%) very fast (within one "time constant" α), improving the solution from thereon will take a significant amount of time. Moreover, the coefficient α (similar to the time constant of a charging capacitor) increases with the size of the graph (Fig. 3.4d) indicating that the stand-alone Ising machine performance will degrade with size i.e., require exponentially longer time to converge. This can be attributed to the increasing complexity of the solution space for such a scaled-up system and the OIM is trapped in the local minima which consequently gives rise to a sub-optimal solution.

It is to be noted that, the OIM system might attain an improved solution if the measurement is kept on running above the timeframe considered here (10 ms). However, in such a case, the time penalty that the system will incur for improving the accuracy will be significant when compared to the digital algorithm. This creates a fundamental bottleneck for the (stand-alone) oscillator Ising machine for achieving a substantial speedup compared to digital algorithms while maintaining the same solution quality. Therefore, it is imperative to find an approach that can enhance the accuracy of the solution computed by the OIM hardware without acquiring significant compute time. To do so, a hardware-algorithm-based hybrid approach is developed and demonstrated as illustrated in the following sub-sections.

3.2.2 Hybrid Approach to overcome the trade-off
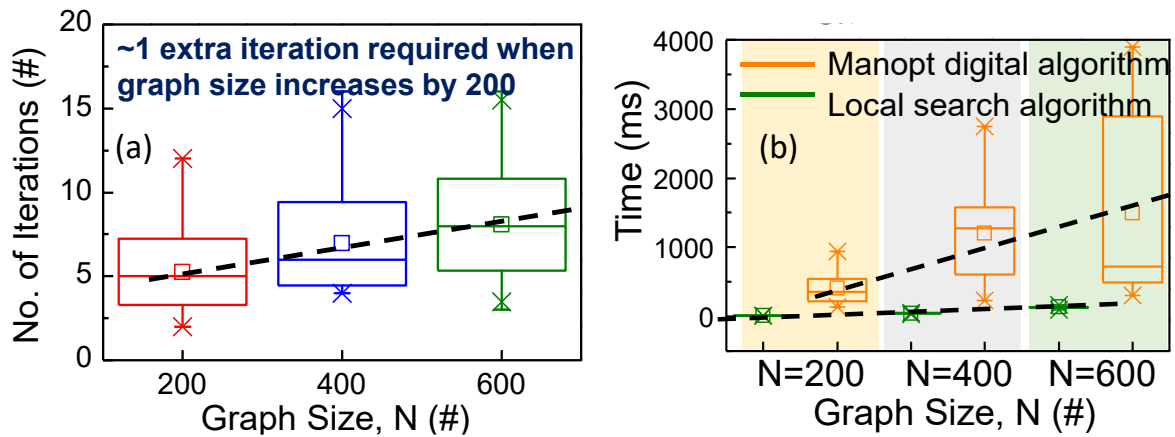
Fig.3.5. (a) No. of iteration required; and (b) computation time for the local search-based polynomial-time algorithm used in the hybrid approach; for comparison, the computation time for the manifold optimization algorithm, implemented on a workstation with 32-cores and 256 GB RAM, is also shown.

To circumvent this long tail in the solution trajectory of the oscillator Ising machine, a hybrid approach is proposed and implemented. Here, the oscillator Ising machine is utilized to compute a near-optimal solution (very fast), which is subsequently, improved using a simple polynomial-time algorithm based on local search [95]. The algorithm essentially exchanges nodes between the sets only if the move increases the MaxCut. This process is repeated until no more nodes can be moved. Fig. 3.5(a-b) shows the average number of iterations, and the corresponding time required for the post-processing scheme; ~1 extra iteration is required for every 200-node increase in the graph size. Moreover, the time required for this scheme is significantly smaller than that required for the digital algorithm. These characteristics enable our hybrid approach to improve the solution without a significant time penalty.

This is illustrated with the example graph solved in Fig. 3.6. The 600-node graph with 7366 edges, when mapped to a grid of 24x25 (=600) oscillators is designed such that the
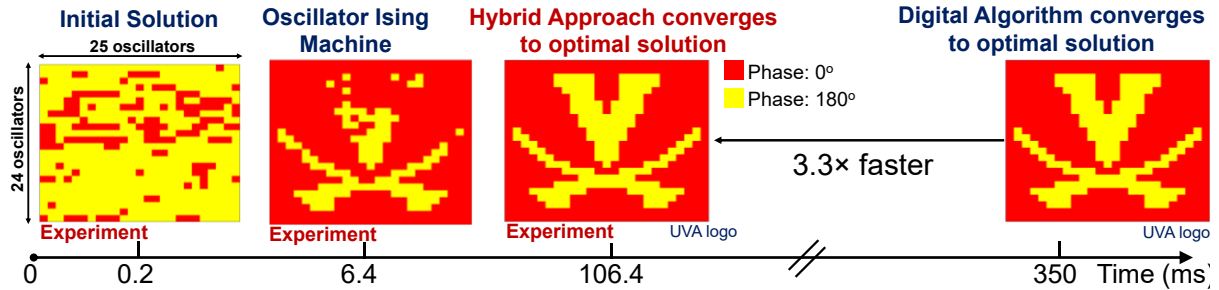
Fig. 3.6. Evolution of the oscillator phases while converging to the MaxCut solution. The MaxCut of the input graph is the UVA logo. It can be observed that the digital approach takes 3.3× longer to converge to the optimal solution. The oscillator-only approach is stuck in a sub-optimal solution even after the hybrid approach has converged to the optimal solution.

oscillator phases corresponding to the optimal MaxCut produce the UVA logo. Post-processing is performed on the solution measured from the oscillators at T=6.4ms. It can be observed that the hybrid approach can converge to the optimal solution 3.3× faster than the digital algorithm. We note that the speedup observed in this graph is at the lower end of the range (3-100×) observed in this work owing to the bipartite nature of the graph. Furthermore, while the oscillator-Ising machine converges to a near-optimal solution very fast (~6.4ms), it subsequently, remains stuck in a local minimum (sub-optimal solution) even after the hybrid approach has converged to the optimal solution.

Next, the effect of the graph size (N=200, 400, 600) and the density of edges on the performance and solution accuracy of the oscillator Ising machine (stand-alone and hybrid) as well as the digital algorithm (Fig. 3.7) are analyzed. A total of 60 graphs are measured; 20 graphs per graph size of varying edge density are measured 10 separate times. The density of edges is quantified using the average degree which represents the number of connections per oscillator; dense 600-node graphs with an average degree of 45 contain up to 27,000 coupling elements. It can be observed from Fig. 3.7(b-d) that the accuracy of the stand-alone oscillator Ising machine is sensitive to the
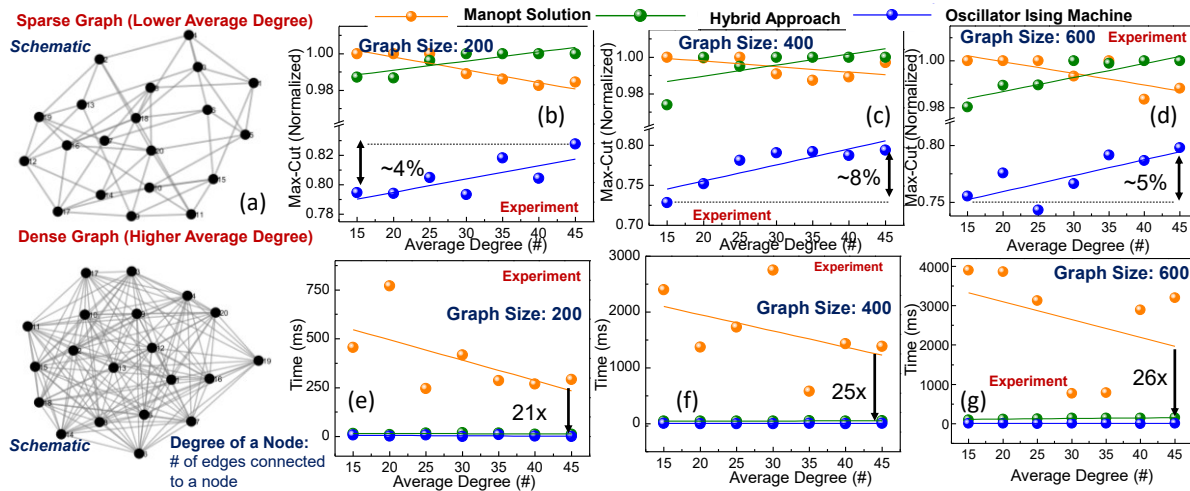
Fig. 3.7. (a) Illustration of a sparse & dense graph. (b-d) Accuracy; and (e-g) Computation time for the stand-alone and hybrid oscillator-based approach as well as digital algorithm as a function of the graph size (N=200-600) and connectivity (avg. degree: 15-45). The solution quality of the stand-alone oscillator-system is sensitive to the input graph connectivity. In contrast, the solution quality of the hybrid approach remains nearly constant. *The hybrid approach converges faster than the digital algorithm for graphs of all size & connectivity.*

sparsity of the graph with the solution accuracy varying by up to ∼ 8%. In contrast, it can be observed that the hybrid approach shows significantly less variation (~2%) in comparison to the stand-alone machine; similar behavior for accuracy is observed for the digital algorithm. Figs. 3.7e-g show the corresponding computation times. It can be observed that even with a time penalty associated with the local search, the hybrid approach facilitates a minimum 21× reduction in compute time compared to the digital algorithm while producing a similar quality of solutions.

3.3. Benchmarking
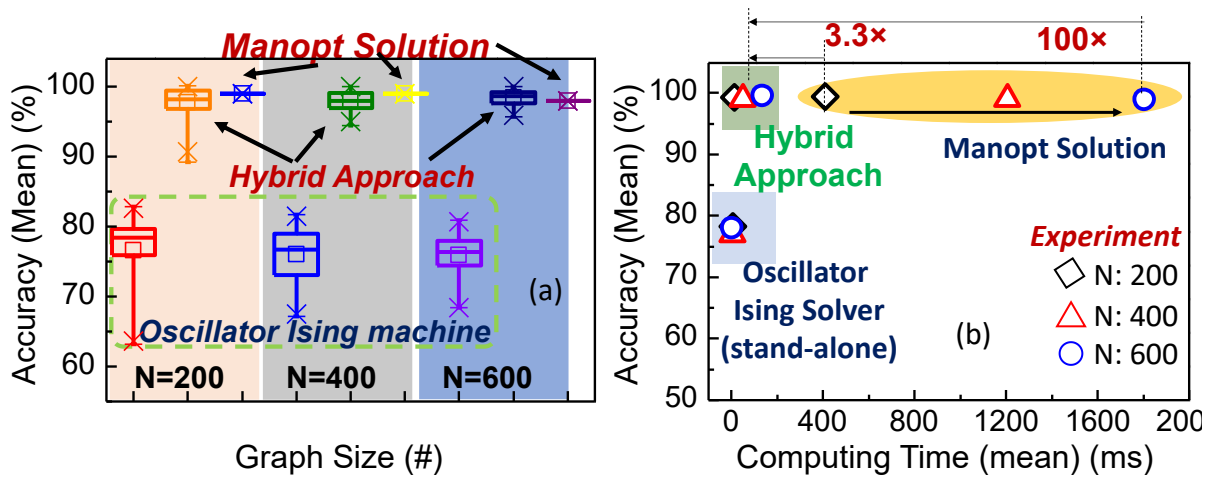
Fig. 3.8. (a) Comparison of the solution accuracy (mean) observed with oscillator-based stand-alone & hybrid approach, and the digital algorithm. (b) Comparison of the accuracy vs. compute time trade-off for the three approaches. The hybrid approach yields the best trade-off.

Finally, the benchmark of the proposed hybrid approach with the stand-alone oscillator Ising machine as well as the digital algorithm is presented. Fig. 3.8a shows that for all graph sizes (irrespective of connectivity), the hybrid approach produces better solutions than the stand-alone hardware, and is comparable to the digital solver. This is because of the exponential solution tail observed in the stand-alone Ising machine. The hybrid approach overcomes this using a simple local search algorithm. Fig. 3.8b compares the (mean) computation time vs. observed solution accuracy (mean) for the three approaches. It can be observed that: (i) the hybrid approach provides the best trade-off between performance and computation time. It produces solutions that have the same level of accuracy as a digital algorithm but require ~3-100× lesser time to compute. (ii) Unlike the digital algorithm where the (mean) solution time increases with graph size; the hybrid approach shows minimal sensitivity to the input graph properties.
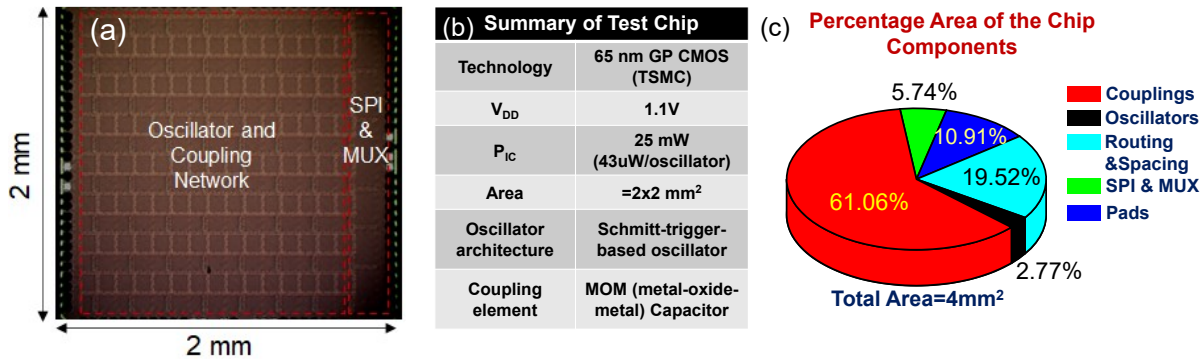
(d)	This work	VLSI'2020 [83]	IEDM'19 [82]	DAC'19 [80]	Sci Rep'19 [81]
Size & type of oscillator network (graph size)	672 (CMOS Schmitt-trigger)	560 (Ring Oscillators)	4 (Discrete IMT oscillators)	240 (Discrete LC oscillators)	4 (Discrete LC oscillators)
Coupling Architecture	Tiled coupling architecture -Intra-tile: All-to-all 16 oscillators -Inter-tile: Each tile is coupled to 6 other tiles (all-to-all coupling) **-Non-planar (NP-Hard)**	-Nearest neighbor Hexagonal coupling -Planar	All-to-all	Chimera architecture	All-to-all
# of Edges (undirected)	30,896	<3,360	6	1,200	6
Types of graphs evaluated	-Non-planar (NP-hard) -Sparse & dense graphs	-Planar (polynomial time solvable) -Sparse	Planar graphs (polynomial time solvable)	Sparse	Planar graphs (polynomial time solvable)
Maximum degree of graph	111 (Hardware)	6	3	~8	3
Power	43 µW/oscillator	41 µW/oscillator	20 µW/oscillator [59]	11 mW/oscillator*	11 mW/oscillator*
Technology	65nm CMOS (GP)	65nm CMOS (LP)	VO_2	LC oscillators with discrete coupling elements	LC oscillators with discrete coupling elements

*Power number was not reported in [80][81]. For benchmarking we considered power number of LC oscillator at 65nm technology node reported in [86].

Fig. 3.9. (a) Die photo, (b) summary of the test chip characteristics and (c) area breakdown for different components of the implemented scaled OIM IC. (d) Comparison of the capabilities of the hardware demonstrated in the present work with prior reports.

Fig. 3.9(a,b) show the die photo and the characteristics of the test chip. The area breakdown for different components of the OIM hardware is shown in Fig. 3.9c. A comparison of the present oscillator Ising hardware with prior works is shown in Fig. 3.9d. Earlier works have primarily focused on small prototypes while larger demonstrations

have used planar graphs with extremely sparse connectivity that is polynomial-time solvable. Consequently, performing a comprehensive comparison of the performance and trade-offs of the oscillator Ising machine with digital algorithms on NP-hard problems has been difficult. The oscillator hardware uniquely enables us to characterize the inherent trade-offs in oscillator Ising machines, and subsequently, design a hybrid hardware-software approach that produces high-quality solutions, significantly faster than digital approaches.

Moreover, from the comparison of the proposed hybrid approach with prior works (which are either limited to small graph sizes or have constrained connectivity), our implementation not only provides significant improvement in the connectivity that can be achieved (>13x) but also provides a pathway to exploit the time advantage of OIMs in large graphs of practical relevance.

3.4. Summary

In summary, the fundamental trade-off in solution accuracy and computation time in oscillator-Ising machines for solving NP-hard non-planar graphs (first of its kind) is experimentally revealed. Subsequently, we have demonstrated a hybrid approach, that exploits oscillator Ising machines and a simple local search algorithm, to reduce the computation time by 3-100× compared to the manifold optimization-based digital algorithm while yielding the same solution quality. Additionally, we address the challenge of implementing a scalable coupling scheme by adopting tile-based architecture that facilitates >13x enhancement in connectivity per oscillator. However, we note that our

evaluation of the oscillator Ising machine so far has been limited to solving the MaxCut problem. The solution to such a problem entails binary spin configuration (oscillator phase of 0° or 180° corresponding to spin state ± 1) and constitutes a small subset of the combinatorial optimization problems. Therefore, in the following chapter, we expand our evaluation of the OIM platform to solve a large subset of optimization problems, solutions of which constitute multi-valued spin configurations.

Chapter 4

Computational Models Based on Synchronized Oscillators for Solving Combinatorial Optimization Problems Beyond MaxCut

Prior work [47,80-83] as well as the evaluation of the OIM platform done so far in this dissertation, has primarily focused on formulating oscillator-based computational models for a relatively small subset of combinatorial problems whose solutions require binary spin states. For instance, while oscillator-based models are available for the MaxCut problem – a special instance of the Max-K-Cut (with K=2), no such models have been developed for the general instance of K. Therefore, the applicability of OIM for solving a broad range of problems whose solutions require multi-valued spin configurations remain unexplored. In this chapter, coupled oscillator-inspired computational models are developed and formulated for solving other archetypal combinatorial optimization problems, solutions of which are not limited to binary spin configurations but rather require multi-valued spin states. Namely, (a) the Max-K-Cut Problem: defined as the problem of dividing the nodes of a graph G (V, E) into K subsets such that the number of edges across the subsets is maximized; (b) the Graph Coloring Problem: defined as the problem of finding the minimum number of colors required such that every node in the graph is assigned one color, and no two connected nodes (i.e., nodes that share an edge) are assigned the same color; (c) the Traveling Salesman Problem (TSP): which asks the question: "Given a list of cities and the distances between each pair of cities, what is the shortest possible

route that visits each city exactly once and returns to the origin city?"; (d) the Hamiltonian Path (and Cycle) Problem: defined as the problem of identifying a path (if it exists) that visits every node in the graph exactly once. The Hamiltonian cycle is the same problem but imposes the additional constraint that the path must return to the node from where it originated, and (e) the Graph Partitioning Problem: defined as the challenge of dividing the vertices of a graph into two equal sets such that the number of common edges is minimized. Building on prior work by Wang *et al.* [47], the Lyapunov functions that map to the problems considered here are developed, and subsequently, the associated Kuramoto dynamics are constructed. It has been shown that besides phase partitioning/clustering (using external force functions), the relative phase ordering (in appropriately constructed Kuramoto dynamics) can be mapped to the solutions of combinatorial optimization problems such as the traveling salesman problem, and the Hamiltonian path problem, etc.

Additionally, these example problems also help illustrate the general principle and approach behind developing such models for other problems not covered here. It has been shown that the synchronized oscillator system dynamics can be engineered to solve different optimization problems by appropriately designing the coupling function and the external injection to each oscillator. Firstly, the formulation of the Max-K-Cut problem is discussed, which can be considered as a general case of the MaxCut problem.

4.1. Max-K-Cut

To develop the computational model for the oscillator-inspired Max-K-Cut solver, we first start with the description of the Ising model developed for MaxCut problem and illustrate how the oscillator dynamics can be used to solve the same. The MaxCut problem can be mapped to the Ising Hamiltonian H by representing each node of the corresponding graph by a spin $\{-1, +1\}$. Every edge between the nodes i and j is represented by the interaction coefficient J_{ij} among the spins; $J_{ij} = -1$ when an edge exists between the nodes, and $J_{ij} = 0$ when the edge is absent. Subsequently, solving the MaxCut problem is equivalent to minimizing the Ising Hamiltonian described as:

$$H = - \sum_{i,j=1,i<j}^{N} J_{ij} s_i s_j \qquad (4.1)$$

Wang *et. al* [47] considered the Lyapunov function defined in equation (4.2) and demonstrated that it is associated with the dynamics of a topologically equivalent (oscillator $\equiv$ spin; coupling element $\equiv$ interaction coefficient; oscillator phase $\equiv$ spin assignment) network of coupled oscillators, under second harmonic injection and is equivalent to H.

$$E(\phi(t)) = -C_1 \sum_{i,j=1,\,j\neq i}^{N} J_{ij} \cos(\Delta\phi_{ij}) - \sum_{i=1}^{N} C_{sync} \cos(2\phi_i(t)) \qquad (4.2)$$

where C_1 is the coupling strength, and C_{sync} modulates the strength of the second harmonic injection signal.

Furthermore, equation (4.2) being a Lyapunov function, the coupled oscillator phases, ϕ_I, will partition themselves in a way such that the system will tend to minimize $E(\phi(t))$ as

the system evolves over time. Consequently, this also minimizes H (Appendix 1). $E(\phi(t))$ is minimized at discrete points $\{0, \pi\}$ in the phase space which effectively correspond to the partitions/sets created by the MaxCut. This enables the oscillator network to directly compute the MaxCut solution as the system evolves towards the ground state energy.

Now this approach is extended to the Max-K-Cut problem. To aid the analysis for the Max-K-Cut problem, it is useful to recast the 'spins' that represent each set of the MaxCut as complex quantities, $re^{i\theta_k}$, where $r = 1$, $\theta_k = \frac{2\pi k}{K}$; $k = 1, 2, \dots, K$. For the MaxCut problem, where $K = 2$, it is evident that the possible spin configurations are $1e^{i\pi} = -1$ ($k = 1$) and $1e^{i(2\pi)} = 1$ ($k = 2$). To map the Max-K-Cut problem, we formulate the "spin" assignment for each of the K sets using the above scheme i.e., each set is assigned a value $re^{i\theta_k}$, where $r = 1, \theta_k = \frac{2\pi k}{K}$; $k = 1, 2 \dots, K$. It is to be noted that defining "spins" using this approach for $K > 2$ disengages it from the physical significance of a spin i.e., a spin with an assignment $1e^{i\left(\frac{2\pi}{3}\right)}$ ($k = 1, K = 3$) may not have physical relevance. However, we will continue to use the term here for continuity. It is also noteworthy though that while a complex spin has little physical significance, in oscillator networks, such assignments can be easily represented by the amplitude and phase. In fact, each set created by the Max-K-Cut will be represented by a specific oscillator phase, as discussed further.

The objective function describing the Max-K-Cut (modeled along the same lines as the Ising Hamiltonian) can be described as:

$$H_{Max-K-Cut}$$

$$= -\sum_{i,j,i<j} J_{ij}.\,Re\left(e^{i\left(\lim_{\sigma\to 0} \sum_{k=1}^{K-1} \left[\left((2k-1)\pi - \frac{2k\pi}{K}\right).e^{-\left(\frac{\left(\Delta\theta_{ij}-\frac{2k\pi}{K}\right)^2}{2\sigma^2}\right)} + \left(\frac{2k\pi}{K}-(2k-1)\pi\right).e^{-\left(\frac{\left(\Delta\theta_{ij}+\frac{2k\pi}{K}\right)^2}{2\sigma^2}\right)} \right] \right)} s_i s_j^* \right) \qquad (4.3)$$

where s_i, s_j represent the 'spin' assignments given by $re^{i\theta}$; $\Delta\theta_{ij} = \theta_i - \theta_j$, J_{ij} describes the connectivity among the nodes of the graph; $J_{ij} = -1(0)$, when an edge is present (absent) between nodes i and j. For simplicity, equation (4.3) can be re-written as:

$$H_{Max-K-Cut} = -\sum_{i,j,i<j}^{N} J_{ij}.\,Re\left(e^{if(\Delta\theta_{ij})} s_i s_j^* \right) \qquad (4.4)$$

where,

$$f(\Delta\theta_{ij}) = \lim_{\sigma\to 0} \sum_{k=1}^{K-1} \left[\left((2k-1)\pi - \frac{2k\pi}{K}\right).e^{-\left(\frac{\left(\Delta\theta_{ij} - \frac{2k\pi}{K}\right)^2}{2\sigma^2}\right)} \right.$$

$$\left. + \left(\frac{2k\pi}{K} - (2k-1)\pi\right).e^{-\left(\frac{\left(\Delta\theta_{ij} + \frac{2k\pi}{K}\right)^2}{2\sigma^2}\right)} \right] \qquad (4.5)$$

$f(\Delta\theta_{ij})$ is a 2π-periodic odd function. Equation (4.4) can also be expressed as:

$$H_{Max-K-Cut} = -\sum_{i,j,i<j}^{N} J_{ij} \cos\left(\Delta\theta_{ij} + f(\Delta\theta_{ij})\right) \tag{4.6}$$

$H_{Max-K-Cut}$ is designed such that $-J_{ij}\text{Re}\left(e^{if(\Delta\theta_{ij})}s_i s_j^*\right) = -1$ if and only if the nodes corresponding to an edge are in different partitions. For example, consider an edge whose corresponding nodes are placed in the sets with assignments $s_i = 1e^{i\left(\frac{4\pi}{3}\right)}$ (using $k_i = 2$), and $s_j^* = 1e^{-i\left(\frac{2\pi}{3}\right)}$ (using $k_j = 1$), respectively. For this edge, $f(\Delta\theta_{ij}) = \left(\frac{\pi}{3}\right)$; $e^{if(\Delta\theta_{ij})} = e^{i\left(\frac{\pi}{3}\right)}$, and thus, $-J_{ij}.\text{Re}\left(e^{if(\Delta\theta_{ij})}s_i s_j^*\right) = -1$. In contrast, if the edge is assigned to the same set (say, using $k_i = k_j = 1$), then $s_i = 1e^{i\left(\frac{2\pi}{3}\right)}$, $s_j^* = 1e^{-i\left(\frac{2\pi}{3}\right)}$ and $f(\Delta\theta_{ij}) = 0$; $e^{i(0)} = 1$. $-J_{ij}\text{Re}\left(e^{if(\Delta\theta_{ij})}s_i s_j^*\right)$ then evaluates to 1. Thus, computing the Max-K-Cut of the graph is equivalent to minimizing $H_{Max-K-Cut}$. Fig. 4.1 shows the spin assignment and $f(\Delta\theta_{ij})$ for various partitions K.

To emulate the minimization of $H_{Max-K-Cut}$ in a physical system, a coupled oscillator system with N oscillators is considered. For an oscillator in a coupled network, the phase evolution of that oscillator in the network can be described using the Gen-Adler's equation [97] as described by Wang *et al.* [47],

$$\frac{d\phi_i(t)}{dt} = \omega_i - \omega_{sync} + \omega_i \sum_{j=1,\ j\neq i}^{N} c_{ij}(\Delta\phi_{ij}) \tag{4.7}$$

where ω_I is the natural frequency of the individual oscillator, ω_{sync} is the frequency of the synchronized network and $c_{ij}(.)$ is a 2π-periodic function for the coupling among oscillators i and j. The oscillator network is designed to be topologically equivalent to the

# of Partitions (K)	Spin Assignment	$f(\Delta\theta_{ij})$
2 (Max-Cut)	$1e^{i\pi}(\equiv -1)$ $1e^{i(2\pi)}(\equiv 1)$	
3	$1e^{i\left(\frac{2\pi}{3}\right)}(\equiv -0.5 + i.\frac{\sqrt{3}}{2})$ $1e^{i\left(\frac{4\pi}{3}\right)}(\equiv -0.5 - i.\frac{\sqrt{3}}{2})$ $1e^{i\left(\frac{6\pi}{3}\right)}(\equiv 1)$	
4	$1e^{i\left(\frac{2\pi}{4}\right)}(\equiv +i)$ $1e^{i\left(\frac{4\pi}{4}\right)}(\equiv -1)$ $1e^{i\left(\frac{6\pi}{4}\right)}(\equiv -i)$ $1e^{i\left(\frac{8\pi}{4}\right)}(\equiv +1)$	
5	$1e^{i\left(\frac{2\pi}{5}\right)}(\equiv 0.3 + 0.95i)$ $1e^{i\left(\frac{4\pi}{5}\right)}(\equiv -0.8 + 0.58i)$ $1e^{i\left(\frac{6\pi}{5}\right)}(\equiv -0.8 - 0.58i)$ $1e^{i\left(\frac{8\pi}{5}\right)}(\equiv 0.3 - 0.95i)$ $1e^{i\left(\frac{10\pi}{5}\right)}(\equiv 1)$	

Fig. 4.1. Spin assignments and corresponding $f(\Delta\theta_{ij})$ for various values of K in the Max-K-Cut problem.

input graph i.e., oscillator ≡ node; coupling element ≡ edge. Thus, the coupling network can be described by the matrix (J_{ij}) of the graph.

Further, assuming that all the oscillators have the same frequency, equal to the synchronized frequency of the network [47], equation (4.7) evolves to:

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1, \ j\neq i}^{N} J_{ij} c_{ij}(\Delta\phi_{ij})$$

(4.8)

where C_1 is a positive constant that signifies the coupling strength among the oscillators.

Under the influence of the injection of the Kth harmonic signal to every oscillator in the system, equation (4.8) further evolves to:

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1, \ j\neq i}^{N} J_{ij} c_{ij}(\Delta\phi_{ij}) - C_{sync}\sin(K\phi_i(t))$$

(4.9)

where C_{sync} is a positive constant that describes the amplitude of the injected signal (K^{th} harmonic). In order to engineer the properties of the oscillator network for solving the Max-K-Cut problem, we carefully design the coupling function $c_{ij}(.)$ to be $\sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right)$, where $f(\Delta\phi_{ij})$ is derived from equation (4.5). The system dynamics can then be expressed as:

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1, \ j\neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) - C_{sync}\sin(K\phi_i(t))$$

(4.10)

Here, $f(\Delta\phi_{ij})$ can be considered as a coupling function as illustrated in Appendix 2. For a system with the dynamics as described in equation (4.10), the following Lyapunov function candidate is considered:

$$E(\phi(t)) = -\frac{KC_1}{2} \sum_{i,j,\, j \neq i}^{N} J_{ij} \cos\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) - \sum_{i=1}^{N} C_{sync} \cos(K\phi_i(t)) \tag{4.11}$$

To show that $E(\phi(t))$ is a Lyapunov function, we first analyze $\frac{dE(\phi(t))}{dt}$ which can be expressed as $\frac{dE(\phi(t))}{dt} = \frac{\partial E(\phi(t))}{\partial \phi_i(t)} \cdot \frac{d\phi_i(t)}{dt}$. It can be shown (detailed in Appendix 3) that

$$\frac{\partial E(\phi(t))}{\partial \phi_i(t)} = -K\frac{d\phi_i(t)}{dt} \tag{4.12}$$

Thus,

$$\frac{dE(\phi(t))}{dt} = \sum_{i=1}^{N} \left[\left(\frac{\partial E(\phi(t))}{\partial \phi_i(t)}\right) \cdot \left(\frac{d\phi_i(t)}{dt}\right) \right] \tag{4.13}$$

$$= -K\sum_{i=1}^{N} \left[\left(\frac{d\phi_i(t)}{dt}\right)^2 \right] \leq 0 \tag{4.14}$$

This indicates that the oscillator network described by the dynamics in equation (4.10) will aim to minimize the energy function in equation (4.11) as it evolves over time. $E(\phi(t))$ is minimized at the discrete phase points $\phi_k = \frac{2\pi k}{K}, k = 1, 2, \ldots, K$ which effectively correspond to the partitions / sets ($s_i = re^{i\theta_k}$, where $r = 1$, $\theta_k = \frac{2\pi k}{K}; k = 1, 2, \ldots, K$) created by the Max-K-Cut. It is to be noted that the minimization of $E(\phi(t))$ at the discrete phase points is facilitated by appropriately designing the injection/forcing function in the dynamics described in equation (4.9). The role of the forcing function is illustrated in Appendix 4. Equation (4.11) is equivalent to $H_{Max-K-Cut}$ in equation (4.6) with a constant offset which implies that the system will aim to compute the solution to the Max-K-Cut problem as well.
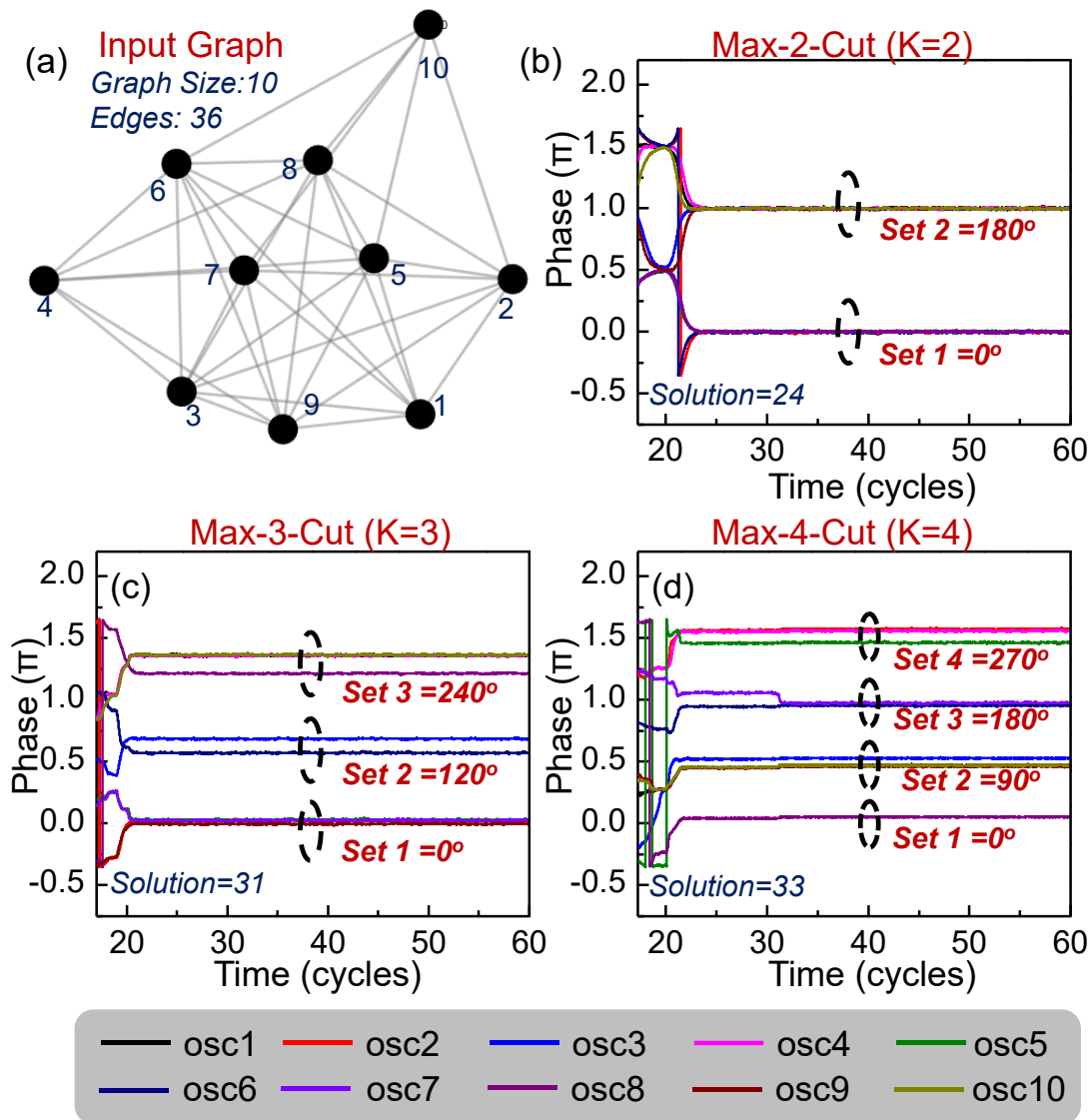
Fig. 4.2. (a) Illustrative 10 node graph; (b)-(d) Evolution of oscillator phases for different values of K (=2,3,4) in the Max-K-Cut problem.

Fig. 4.2 shows the phase partitions and the resulting Max-K-Cut (K= 2, 3, 4) solutions for an illustrative graph (with 10 nodes) using the oscillator-inspired model developed above. An annealing scheme similar to that used by Wang *et al.* [47], is used to help the system escape from local minima (Appendix 5). It can be observed that the oscillator phases,

representing the nodes of the graph, separate into K partitions to yield Max-K-Cut solutions. Further, using the oscillator-inspired solver developed above, larger graph instances (Appendix 6) from the G-Set dataset for K=2, 3, and 4 partitions are also evaluated. In each case, the system is simulated for 100 cycles and yields high-quality Max-K-Cut solutions. Next, the above approach is used to formulate oscillator-inspired computational models for other combinatorial optimization problems.

4.2. Graph Coloring
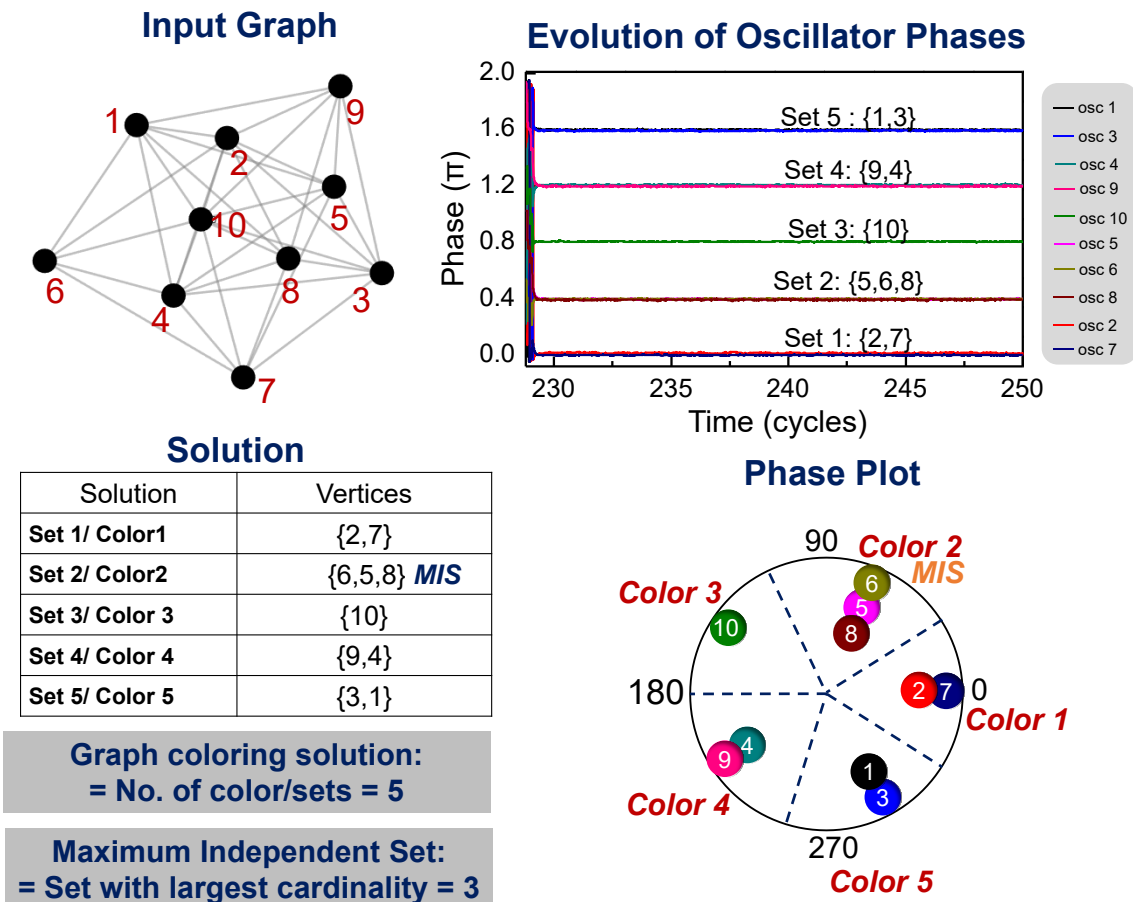
Fig. 4.3. A representative 10 node graph along with the corresponding phase partitions obtained using the oscillator-based computational model (K=5). The graph is 5 colorable since every edge is shared between (any) two sets i.e., there are no edges that lie entirely within one set. Further, the largest such set is the MIS (=3, optimal).

Given a graph G(V, E) such that each node is to be assigned a specific color, the graph coloring problem entails finding the minimum number of colors required such that the no two connected nodes (i.e., nodes that share an edge) are assigned the same color. Nodes having the same color can be considered as a set that does not contain edges. Thus, the graph coloring problem can be formulated as finding the smallest K for which the sets created by the Max-K-Cut of the graph do not contain any edges within the set i.e., every edge in the graph are shared between (any) two sets. Solving the problem using this approach entails computing the Max-K-Cut of the graph for different values of K in ascending order, and determining the (smallest) K for which the sets created by the Max-K-Cut does not contain an edge internal to a set. The smallest value of K is the Chromatic number of the graph. This can be considered equivalent to solving the decision version of the graph coloring problem which evaluates if a graph is K-colorable i.e., can a graph be colored with K colors? It is noteworthy that Crnkic *et al.* [98] also proposed a similar oscillator-inspired model to solve the graph coloring problem but used a different formulation for the energy function. Some additional constraints on the value of K can be imposed by using graph properties e.g., the value of K will be at most $\Delta+1$, where Δ is the maximum degree of the graph (Brooks theorem [99]). Fig. 4.3 shows the coloring for an illustrative example for a graph (same as that considered in Fig. 4.2) considering K=5 where it can be observed that nodes split into 5 partitions. Moreover, none of the sets contain an edge within the set, implying that the graph is 5 colorable. In contrast, the partitions generated by the Max-4-Cut for the graph (Fig. 4.2d), which verifies if the graph is 4 colorable, shows that 2 edges lie within the set. Consequently, this implies that the

graph is not 4 colorable, and thus, 5 is the minimum number of colors required to color the graph (i.e., chromatic number).

4.3. Maximum Independent Set (MIS) and Maximum Clique (MC)

Each set obtained while computing the graph coloring problem (using the Max-K-Cut formulation) that does not contain an edge corresponds to an independent set of the graph. Thus, the largest independent set among them provides a good approximate solution for the MIS (Fig. 4.3) [88]. Additionally, using the relationship that the Maximum Clique of a graph is the MIS of its complement graph, the same approach can be used to compute the Maximum Clique of a graph by computing the MIS of the complement graph.

4.4. Traveling Salesman Problem (TSP)

Given a list of cities and the distances between each pair, the archetypal TSP is defined as the challenge of finding the shortest possible route that visits each city exactly once and returns to the city of origin. The objective function for the TSP can be formulated as minimizing H_{TSP}, where H_{TSP} is defined as:

$$H_{TSP} = -\sum_{i,j,i<j}^{N} J_{ij} \cdot Re\left(e^{if_{TSP}(\Delta\theta_{ij})}s_i s_j^*\right)$$
(4.15)

where s_i denotes the graph node (city), expressed as a complex quantity described earlier, and $J_{ij}=-D_{ij}$, where D_{ij} refers to the distance between nodes (cities) i and j. Here,

$$f_{TSP}(\Delta\theta_{ij}) = \lim_{\sigma\to 0} -\sum_{\gamma=1,N-1}\left(\frac{2\gamma\pi}{N}\cdot e^{-\left(\frac{\left(\Delta\theta_{ij}-\frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)} + \left(-\frac{2\gamma\pi}{N}\right)\cdot e^{-\left(\frac{\left(\Delta\theta_{ij}+\frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)}\right)$$

$$+ \lim_{\sigma\to 0}\sum_{k=2,k\neq N-1}^{N}\left(\left(\pi-\frac{2k\pi}{N}\right)e^{-\left(\frac{\left(\Delta\theta_{ij}-\frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)} + \left(\frac{2k\pi}{N}-\pi\right)\cdot e^{-\left(\frac{\left(\Delta\theta_{ij}+\frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)}\right)$$
(4.16)

$$H_{TSP} = -\sum_{i,j,i<j}^{N} J_{ij}\cos\left(\Delta\theta_{ij} + f_{TSP}(\Delta\theta_{ij})\right)$$
(4.17)

The TSP is equivalent to computing a phase ordering such that the sum of distances (D_{ij}) among the adjacent nodes (cities) in the phase ordering is as small as possible. To achieve this, $f_{TSP}(\Delta\theta_{ij})$ is designed such that $-J_{ij}Re\left(e^{if_{TSP}(\Delta\theta_{ij})}s_i s_j^*\right) = D_{ij}$ if the nodes are adjacent (i.e., $Re\left(e^{if_{TSP}(\Delta\theta_{ij})}s_i s_j^*\right)=1$), else $-J_{ij}Re\left(e^{if_{TSP}(\Delta\theta_{ij})}s_i s_j^*\right) = -D_{ij}$ (i.e., $Re\left(e^{if_{TSP}(\Delta\theta_{ij})}s_i s_j^*\right) = -1$); in other words, the system incurs a larger energy penalty for putting nodes with larger

distances adjacent to each other. Therefore, minimizing H_{TSP} minimizes the sum of distances among the adjacent nodes. Furthermore, the circular phase ordering ensures that dynamics not only compute the shortest route among the cities but also guarantee that one returns back to the original city after visiting each city only once. Along similar lines as the Max-K-Cut problem, the corresponding oscillator dynamics can be expressed as:
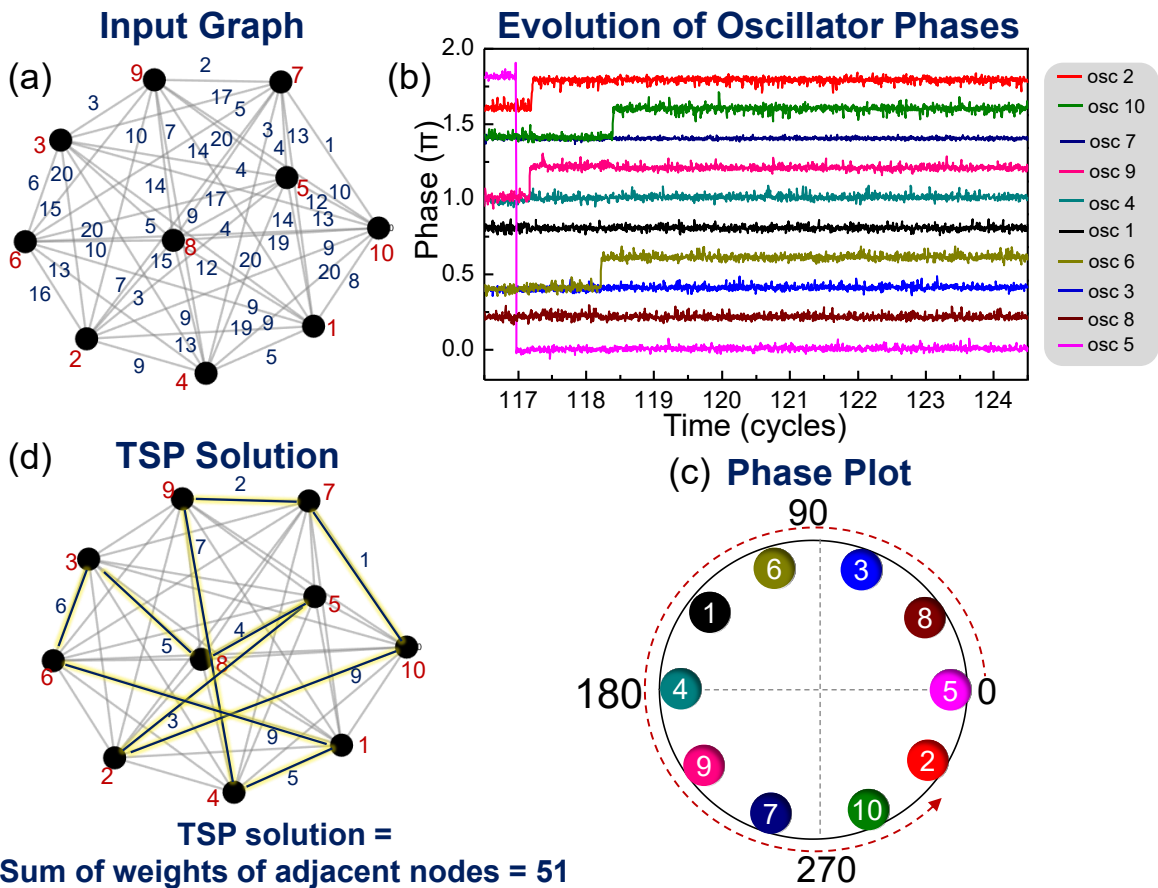
Fig. 4.4. (a) A representative 10 node graph; (b) oscillator phase evolution in a topologically equivalent oscillator system. (c) phase plot and the corresponding phase ordering that encodes the TSP solution. (d) TSP solution obtained from the phase ordering (=51; optimal for the above graph).

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1, j \neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f_{TSP}(\Delta\phi_{ij})\right) - C_{sync}\sin(N\phi_i(t)) \qquad (4.18)$$

As discussed earlier, the force term on the right-hand side ensures that the oscillators settle to a phase value of $\frac{2\pi k}{K}; k = 1, 2, \dots, N$. Similar to the Max-K-Cut analysis, the Lyapunov function candidate considered for these dynamics is:

$$E(\phi(t)) = -\frac{NC_1}{2} \sum_{i,j, j \neq i}^{N} J_{ij} \cos\left(\Delta\phi_{ij} + f_{TSP}(\Delta\phi_{ij})\right) - \sum_{i=1}^{N} C_{sync}\cos(N\phi_i(t)) \qquad (4.19)$$

Which can be shown to be minimized over time. Further, equation (4.19) is equivalent to H_{TSP} (with a constant offset), and thus, will be minimized over time. Fig. 4.4 shows a representative example for a 10 city TSP graph solved using the coupled oscillators.

4.5. Hamiltonian Cycle/Hamiltonian Path

The Hamiltonian path problem is defined as the problem of identifying a path (if it exists) that visits every node exactly once; computing the Hamiltonian cycle entails solving the same problem with the added constraint of returning back to the node of origin. In terms of the oscillator system dynamics, the objective here is to design a topologically equivalent coupled oscillator network that yields a phase ordering such that the number of edges among the adjacent nodes is maximized. Subsequently, if every node shares an edge with the adjacent nodes in the phase ordering, the resulting ordering represents the Hamiltonian cycle (and path); if only one edge is missing, then only a Hamiltonian path (but not cycle) exists. The Hamiltonian for the above problem can be formulated using the same approach used for TSP with the exception that the system is 'rewarded'

(in terms of energy) for bringing connected nodes with edges adjacent to each other. The objective function for the Hamiltonian cycle path can be expressed as:

$$H_{HC} = -\sum_{i,j,i<j}^{N} J_{ij}.\,\mathrm{Re}\left(e^{if_{HC}(\Delta\theta_{ij})}s_i s_j^*\right) \tag{4.20}$$

where,

$$f_{HC}(\Delta\theta_{ij}) = \lim_{\sigma\to 0} -\sum_{\gamma=1,N-1}\left(\left(\pi - \frac{2\gamma\pi}{N}\right).e^{-\left(\frac{\left(\Delta\theta_{ij}-\frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)} + \left(\frac{2\gamma\pi}{N} - \pi\right).e^{-\left(\frac{\left(\Delta\theta_{ij}+\frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)}\right)$$

$$+ \lim_{\sigma\to 0}\sum_{k=2,k\neq N-1}^{N}\left(\left(\frac{\pi}{2} - \frac{2\pi k}{N}\right)e^{-\left(\frac{\left(\Delta\theta_{ij}-\frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)} + \left(\frac{2\pi k}{N} - \frac{\pi}{2}\right).e^{-\left(\frac{\left(\Delta\theta_{ij}+\frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)}\right) \tag{4.21}$$

$$H_{HC} = -\sum_{i,j,i<j}^{N} J_{ij}\cos\left(\Delta\theta_{ij} + f_{HC}(\Delta\theta_{ij})\right) \tag{4.22}$$

It can be observed that $f_{HC}(\Delta\theta_{ij})$ has a similar form to $f_{TSP}(\Delta\theta_{ij})$ defined for TSP (equation (4.16)) with the exception that $\Delta\theta_{ij} + f_{HC}(\Delta\theta_{ij})$ converges to π when a pair of connected nodes are adjacent to each other (i.e., $\Delta\theta_{ij} = \pm\frac{2\pi}{N}$) in the phase space whereas $\Delta\theta_{ij} + f_{TSP}(\Delta\theta_{ij})$ converges to π when they are in non-adjacent positions (i.e., $\Delta\theta_{ij} = \pm\frac{2k\pi}{N}; 1 < k \leq N; k \neq 1, N-1$). In these respective (desired) scenarios, the system is 'rewarded' by lowering energy (equations (4.17) and (4.22)). Further, $\Delta\theta_{ij} + f_{HC}(\Delta\theta_{ij})$ (for the Hamiltonian cycle /path problem) converges to $\frac{\pi}{2}$ when the connected nodes are in non-adjacent
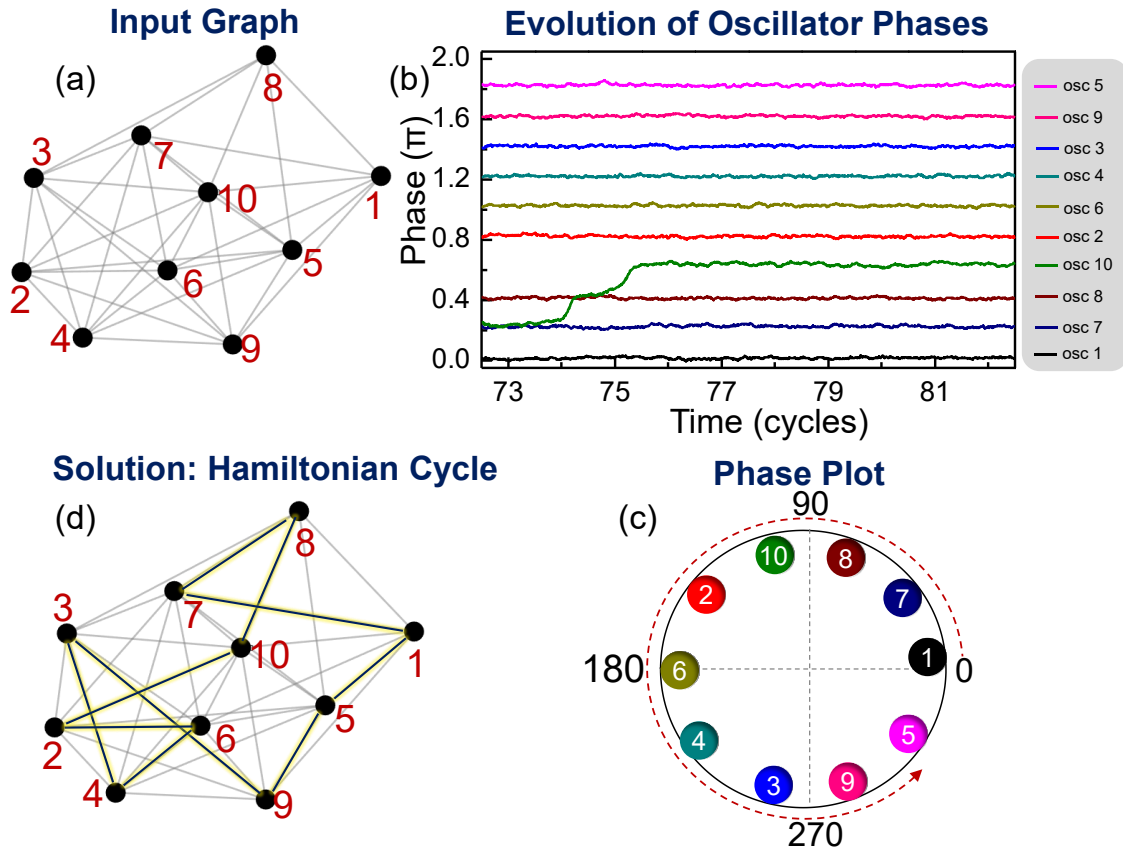
Fig.4.5. (a) A representative 10 node randomly instantiated graph; (b) phase dynamics of the equivalent coupled oscillator network for solving the Hamiltonian path/cycle problem; (c) corresponding ordering of phases, resulting in a Hamiltonian cycle as shown in (d).

positions (undesired case) whereas $\Delta\theta_{ij} + f_{TSP}(\Delta\theta_{ij})$ (for the TSP) converges to 0 when the nodes are in adjacent to each other in the phase space. In the case of the Hamiltonian cycle/path problem, this implies that the system does not lower energy when the connected nodes are in non-adjacent positions (i.e., system is rewarded only for placing nodes adjacent to each other) whereas in the case of TSP the system energy actually increases when the connected nodes are placed next to each other in the phase space. It is to be noted that similar to the TSP, $\Delta\theta_{ij} + f_{HC}(\Delta\theta_{ij})$ can be designed to converge to 0 (instead of $\frac{\pi}{2}$) by modifying the pre-factors for the gaussian distributions in the second term in the right-hand side of equation (4.21). This would impose an energy penalty similar

to TSP when a pair of connected nodes attains an undesirable configuration. Our motivation in choosing $\frac{\pi}{2}$ here was to illustrate how different energy functions can 'designed'. For both of the cases (i.e., $\Delta\theta_{ij} + f_{HC}(\Delta\theta_{ij})$ converging to $\frac{\pi}{2}$ or 0, when a pair of connected nodes are non-adjacent to each other), the equivalence between the optimal solution to the Hamiltonian cycle / path problem and the ground state of the system does not change, and the system will still continue to evolve towards a lower energy state in both the cases.

Similar to the TSP, the corresponding oscillator dynamics and the energy function of the system can be described by,

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1,j\neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f_{HC}(\Delta\phi_{ij})\right) - C_{sync}\sin(N\phi_i(t)) \qquad (4.23)$$

$$E(\phi(t)) = -\frac{NC_1}{2} \sum_{i,j,\, j\neq i}^{N} J_{ij} \cos\left(\Delta\phi_{ij} + f_{HC}(\Delta\phi_{ij})\right) - \sum_{i=1}^{N} C_{sync}\cos(N\phi_i(t)) \qquad (4.24)$$

Equation (4.24), which is equivalent to H_{HC} in equation (4.20) (with an offset), is minimized over time and in the process evolves towards the solution of the Hamiltonian cycle. Fig. 4.5 shows the Hamiltonian cycle computed on a demonstrative problem using the above approach.

4.6. Graph Partitioning
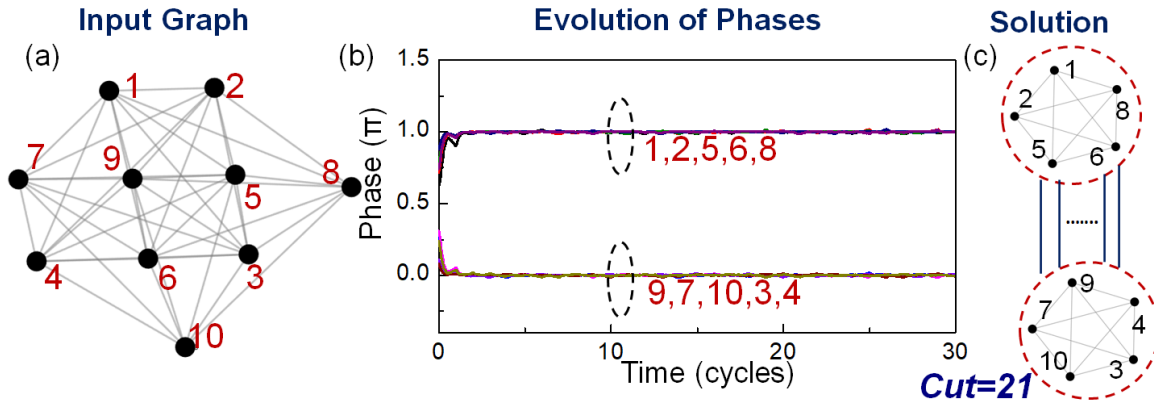
Fig. 4.6. (a) A representative 10 node graph and (b) phase dynamics of the coupled oscillators for the representative graph. (c) Corresponding graph partition solution of 21 (optimal).

The problem of dividing the nodes of a graph into two equal sets such that the number of common edges is minimized. For simplicity, here that the number of nodes (N) is assumed to be even. The objective function for the problem can be defined as:

$$H_{GP} = -\sum_{i,j,i<j}^{N} J_{ij} \left(A. \, \mathrm{Re}\left(\left(\prod_{n=1}^{N} s_n \right)^{\frac{1}{N}} . \, e^{-i\frac{\pi}{2}} \right) + B. \, \mathrm{Re}\left(s_i s_j^* \right) \right) \tag{4.25}$$

Here, $s_j = 1e^{i\theta_j}$, where $\theta_j \in \{0, \pi\}$. The first term on the RHS in equation (4.20) determines the relative reward (energy reduction) for dividing the set into two equal parts, and the second term determines the reward for minimizing the number of common edges. A (>0) and B (>0) are constants that determine the ratio of the two rewards and determine which constraint has preponderance in the optimization process. Further, $J_{ij} = +1(0)$ when an edge is present (absent).

To solve the problem using coupled oscillators, the objective can be expressed as designing coupled oscillator network where the oscillators exhibit a phase bi-partition $\{0, \pi\}$ such that each set has an equal number of oscillators, and the number of edges among the two clusters is small as possible. These properties can be realized using a system that exhibits the following dynamics:

$$\frac{d\phi_i(t)}{dt} = -\sum_{j=1, j\neq i}^{N} J_{ij}\left(2C_1\sin\left(\frac{1}{N}\sum_{n=1}^{N}\phi_n - \frac{\pi}{2}\right) + C_2\sin(\Delta\phi_{ij})\right) - C_{sync}\sin\left(2\phi_i(t)\right) \qquad (4.26)$$

with the corresponding Lyapunov function:

$$E(\phi(t)) = -\sum_{i,j=1, j\neq i}^{N} J_{ij}\left(2NC_1\cos\left(\frac{1}{N}\sum_{n=1}^{N}\phi_n - \frac{\pi}{2}\right) + C_2\cos(\Delta\phi_{ij})\right) \qquad (4.27)$$

$$-\sum_{i=1}^{N} C_{sync}\cos\left(2\phi_i(t)\right)$$

The solution to the graph partitioning problem then corresponds to the two sets of oscillators corresponding to a phase value of 0 and π, respectively, as shown in Fig. 4.6.

4.7. Summary

This work elucidates how natural energy minimization in oscilator-based dynamical systems can be used to develop computational models for solving combinatorial optimization problems. Moreover, the models developed in this work help expand the applications of this computing approach beyond the traditional combinatorial problems (e.g., MaxCut) considered in prior works, and thus, help bolster the case for exploring analog-inspired non-Boolean approaches to computationally intractable optimization

problems. However, it is to be noted that the scope of this dissertation has been primarily on solving problems with objective functions that have a quadratic degree i.e. an edge can connect to only two nodes. There is a larger class of problems (Boolean Satisfalibility, Integer Factorization, etc.) wherein the objective functions have a degree greater than two i.e an edge can connect to any number of nodes, and such data structures are represented using hypergraph. In the future, the oscillator-based computational model developed in this chapter to solve problems with the quadratic degree can be further extended to solving hypergraphs (degree>2) [100].

Chapter 5

Future Works

5.1. Towards a highly scalable Ising Machine Implementation
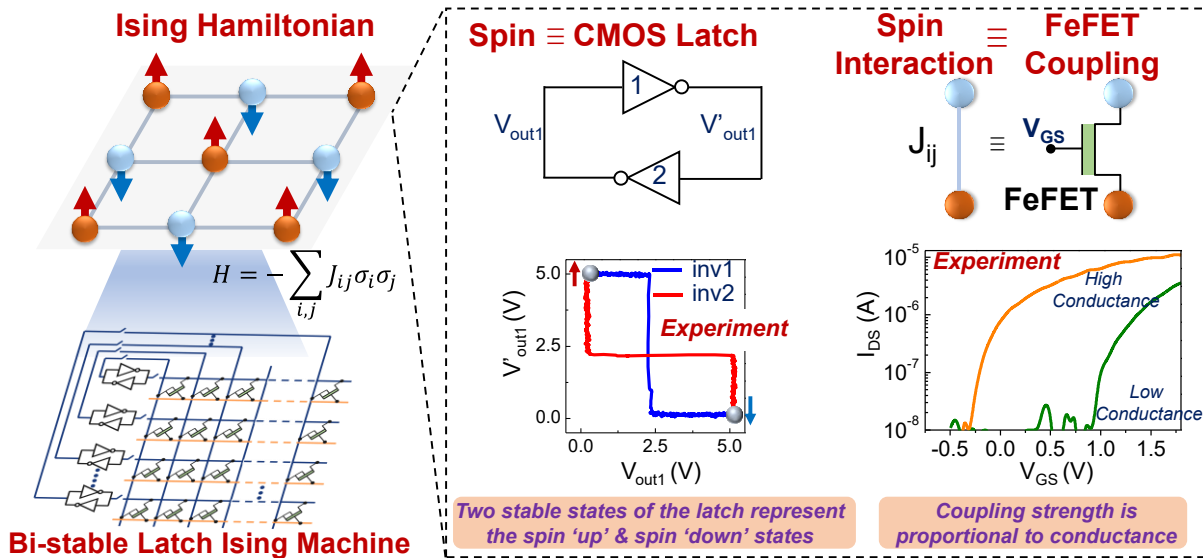
Fig. 5.1. Proposed design for the Ising machine using CMOS latches (cross-coupled inverters) as artificial Ising spins. The interaction among the spins is implemented using a CMOS-process compatible FeFET based array.

As illustrated in the earlier chapters, the intrinsic energy minimization in the highly interconnected system gives rise to rich spatio-temporal properties, which can subsequently be mapped to the solutions of many computationally intractable optimization problems. However, the highly interconnected nature of the system also poses a significant implementation and scalability challenge for Ising platforms (chapter 3). In fact, the number of coupling elements (representing edges) required for mapping an arbitrary graph scales up quadratically ($\sim N^2$) with the number of nodes in the graph, and consequently, scaling the system to large sizes continues to be a significant

challenge for most Ising machine designs. In chapter 3, it has been shown that the tile architecture-based coupling scheme can enable significant improvement in the connectivity that can be realized for a scaled-up system. In the future, this scaling work on Ising machines can be further extended to explore potential new avenues by developing novel hardware components that are not only compact but can also leverage the maturity of CMOS-process technology and integration. Fig. 5.1 shows an Ising machine platform that exploits the novel behavior of compact bi-stable CMOS-latches (cross-coupled inverters) as classical Ising spins interact through highly scalable and CMOS-process compatible ferroelectric-HfO_2-based Ferroelectric FETs (FeFETs) which act as coupling elements. Such an Ising machine can provide a significant area advantage as it facilitates a considerable reduction in the number of components not only for the realization of Ising spins (12T Schmitt-trigger oscillator to 4T bi-stable latch) but also for coupling network (1T-gate+1 bulky capacitor to 1T FeFET).

Proof-of-Concept Experiment: In the preliminary experiment, we demonstrate CMOS latches as highly scalable and compact Ising spins. To implement the coupling network, we aim to exploit the non-volatile behavior of CMOS-compatible FeFET memory arrays (in fact, the FeFETs used in this preliminary experiment are built in 28 nm high-κ metal gate technology platform) to implement the interaction among the spins (CMOS latches). Our choice of using the FeFET as the coupling element is motivated by the fact that FeFETs are compatible with CMOS process technology, provide a wide dynamic range for the resistance (coupling strength), and can be efficiently integrated and programmed
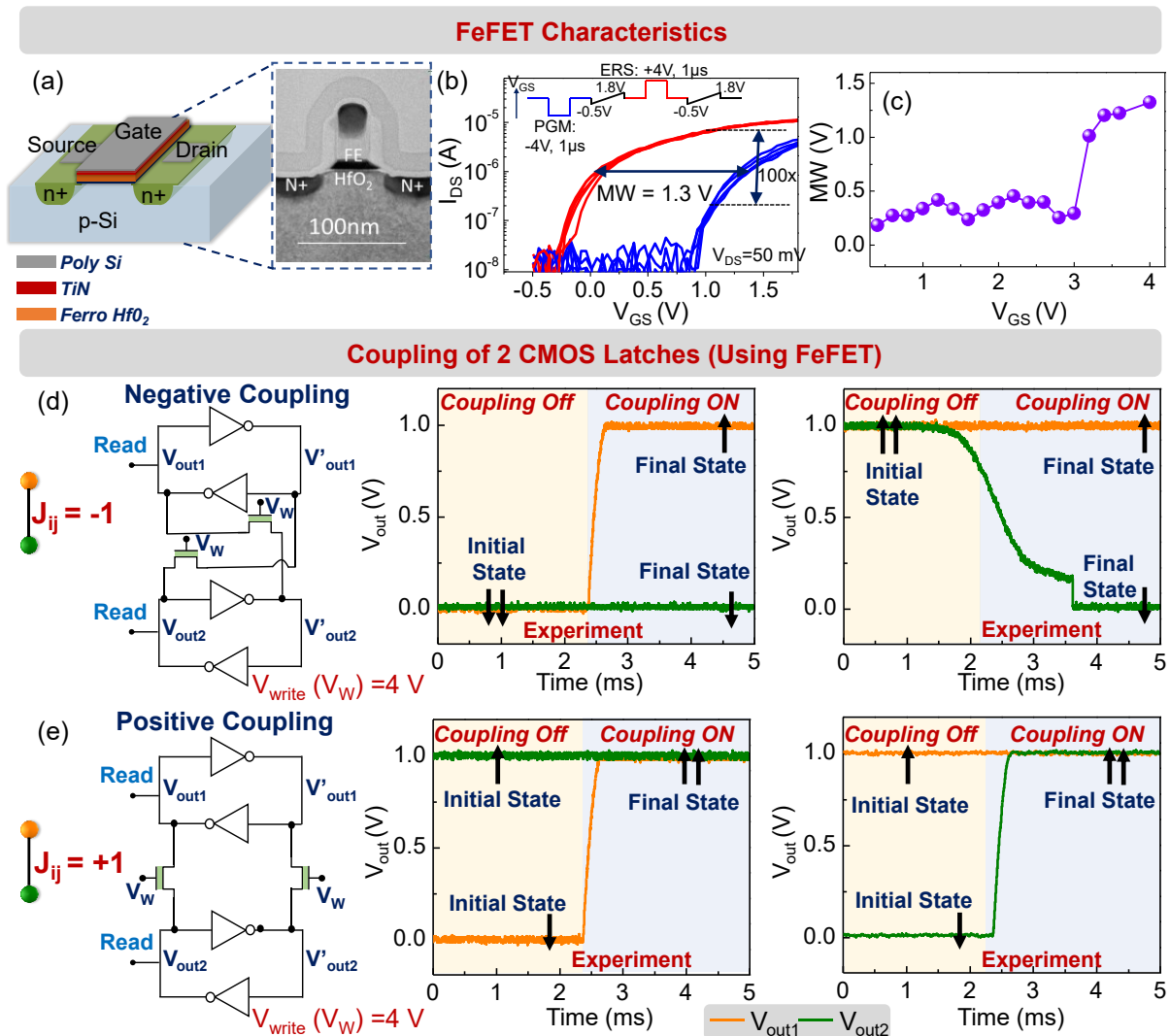
Fig. 5.2. FeFET coupled CMOS latches. (a) Schematic and TEM cross-section [102] of a 28 nm high-κ metal gate FeFET device. (b) I_{DS}–V_{GS} characteristics of the FeFET (W/L=0.5/0.5 μm) after program and erase pulses. (c) Evolution of memory window (MW) as a function of write voltage (V_{GS}). FeFET coupled two-latch system settles (d) out-of-phase and (e) in-phase when the coupling is negative (J_{ij}=-1) and positive (J_{ij}=+1), respectively.

in a scalable array that is required to map the spin interactions in the entire network. We envision that the tunable threshold voltage of the FeFET (Fig. 5.2a) would allow us to program the interaction between the latches; the low V_T (high conductance) state would correspond to J_{ij}=±1 whereas the high V_T (low conductance) state would correspond to J_{ij}=0.

Fig. 5.2b shows the experimentally measured transfer characteristics of the ferroelectric-HfO$_2$-based FeFETs used in this work; the devices are fabricated (see methods) on 28 nm high-κ metal gate technology platform, as shown in the cross-sectional TEM image [101][102]. It features a doped HfO$_2$ layer as the ferroelectric and SiO$_2$ as the interlayer. Detailed processing information is described elsewhere [102]. Fig. 5.2c shows the memory window vs. programming voltage characteristics for the FeFET. When a programming voltage of ±4 V is used to program the FeFET state, a 100× modulation in the current is obtained for V_{GS}= 1V.

We subsequently characterize the behavior of the FeFET as a programmable coupling element in a two-latch system. To evaluate this, the FeFETs are first programmed into the low V_T/high conductance state (J_{ij} =±1) using a programming pulse of magnitude +4V and a period of 1 µs. We test the interaction induced by the FeFETs among the latches by intentionally programming them into the 'incorrect' state, in order to observe the system evolve into the correct state i.e., when J_{ij}=-1 (+1), the latches are initialized into the same (opposite) states (0/V_{DD}), and subsequently, it is observed whether the system evolves to the correct state. During the initialization of the latches, the FeFETs are maintained at V_{GS}=-0.5V. This reduces the conductance of the FeFETs without affecting the threshold voltage. After the latches are initialized, the gate voltage is increased to V_{GS} = 1.5 V, and the corresponding dynamics of the FeFET-coupled CMOS latches are evaluated. We note that V_{GS}=1.5V was required to achieve the desired level of conductance from the FeFET-based coupling element (coupling strength) since the threshold voltage of the device has not been yet optimized for this demonstration. The gate voltage can be reduced to zero by appropriately adjusting the threshold as will be evaluated in the future.

The coupled two-latch system settles in-phase (out-of-phase) when the coupling is positive J_{ij}=+1 (negative, J_{ij}=-1), respectively.

5.2. Summary

The proposed implementation provides a potential pathway to realizing a compact and scalable Ising machine to solve computationally challenging problems such as MaxCut using CMOS-compatible components. In the future, evaluation of the scalability of Ising hardware based on the proposed architecture could be vital as the design is also well-positioned to take advantage of the maturity of the CMOS process technology to realize scaled implementations, making it a promising design approach for realizing high-performance application-specific accelerators for solving combinatorial optimization problems.

Chapter 6

Conclusion

The primary focus of this dissertation is to demonstrate and address the challenges to successfully realizing a low-power, CMOS-compatible, scalable, and reconfigurable coupled oscillator-based Ising machine. In order to achieve this objective, I start by designing a fully reconfigurable oscillator-based Ising machines (OIM) IC that can process graphs up to 30 nodes of arbitrary connectivity. The goal of this implementation was to understand the impact of graph connectivity on the solution quality computed by OIMs for non-planar NP-Hard graphs -which was missing in the prior literature. Utilizing the fully reconfigurable hardware platform, it has been experimentally shown that, as the problem size is increased, graphs with a higher edge density result in lower accuracy. Moreover, it is to be noted that to enable full reconfigurability, 870 coupling elements were required in the design, and the number of coupling elements (representing edges) for mapping an arbitrary graph scales quadratically ($\sim N^2$) with the number of nodes in the graph. This points to the fact that as the system size is increased in order to solve graphs with practical relevance two primary challenges evolve. One is mitigating the accuracy and compute time degradation as we increase the system size and the other is to circumvent the connectivity constraints in OIM implementation to realize non-planar graphs of varying connectivity.

To evaluate and ultimately address these limitations, next a 672 oscillator-based Ising machine IC (65 nm GP CMOS) with 30,896 programmable coupling elements is implemented. In order to facilitate reconfigurability in the connectivity of the non-planar

graphs that can be solved, a tile-based coupling architecture is developed in the design, which provides >13x enhancement over prior OIM demonstrations in terms of connectivity of each oscillator (maximum degree). Using this platform the fundamental trade-off in solution accuracy and computation time is revealed in oscillator-Ising machines for solving NP-hard non-planar graphs. Furthermore, we propose a hybrid approach utilizing which we experimentally demonstrate 3-100× improvement in experimentally measured time-to-compute at iso-solution quality, when compared to a digital algorithm.

However, prior work as well as the work in the dissertation so far has been limited to evaluating a relatively small subset of combinatorial problems whose solution entails binary spin configurations (such as the MaxCut problem). Therefore, in this dissertation, the applicability of oscillator Ising machines is furthered by solving broad set of combinatorial optimization problems whose solution entails multi-valued spin configurations namely, Max-K-Cut, TSP, Graph coloring, Maximum Independent Set, Hamiltonian Path/Cycle, Graph Partitioning, etc. Novel oscillator-based computational models are formulated which demonstrates that the synchronized oscillator system dynamics can be engineered to solve different optimization problems by appropriately designing the coupling function and the external injection to each oscillator. In the future, such dynamical system-based computational models can be extended to solving hypergraph problems.

Therefore, the coupled oscillator platform demonstrated here facilitates a scalable non-Boolean approach to problems that are considered computationally hard to solve on digital platforms. Our work demonstrates, using experiment and simulation, that the rich spatio-temporal dynamics of the coupled oscillators can be leveraged to compute (-near)

optimal solutions to a varied range of challenging optimization problems. Our work marks an important step towards enabling application-specific analog computing platforms to solve computationally hard problems.

Appendices

Appendix 1:

Here, we illustrate that minimizing the Lyapunov function for the MaxCut problem in equation (4.2) is equivalent to minimizing H in equation (4.1), first shown by Wang *et al.* [47]. Similar analysis can be extended to the other combinatorial optimization problems considered here.

The Lyapunov function for the oscillator system to solve the MaxCut is given by:

$$E\big(\phi(t)\big) = -C_1 \sum_{i,j=1,\, j\neq i}^{N} J_{ij} \cos\big(\Delta\phi_{ij}\big) - \sum_{i=1}^{N} C_{sync}\cos\big(2\phi_i(t)\big) \qquad (A1.1)$$

where, $C_1 > 0$, $C_{sync} > 0$; $J_{ij} = -1$.

Equation (A1.1) (same as equation (4.2)) will achieve a minimum when $\cos\big(\Delta\phi_{ij}\big) = -1$ (resulting from the first term on the right-hand side) and $\cos\big(2\phi_i(t)\big) = 1$ (resulting from the second term on the right-hand side of the equation).

These conditions are satisfied (facilitating E to achieve a minimum) when $\phi_i = 0$ or π. At these discrete phase points, equation (A1.1) can be expressed as:

$$E\big(\phi(t)\big) = -C_1 \sum_{i,j=1,\, j\neq i}^{N} J_{ij} \cos\big(\Delta\phi_{ij}\big) - NC_{sync} \qquad (A1.2)$$

since $\cos\big(2\phi_i(t)\big)=1$ for $\phi_i=0$ or π.

Further, by mapping every spin $s_i = +1$ to $\phi_i = 0$ (π), and $s_i = -1$ to $\phi_i = \pi$ (0), it can be observed that $\cos\big(\Delta\phi_{ij}\big) = -1,$ when the nodes lie in the opposite set and $\cos\big(\Delta\phi_{ij}\big) =$

$+1$ when the nodes lie in the same set. This is similar to $s_i s_j = -1 (+1)$ when the spins have opposite (same) alignment, respectively. Furthermore, the phase configuration $[\phi_1, \phi_2, \phi_3, \ldots \phi_N]$ (where $\phi_i = 0$ or π) which minimizes $E(\phi(t))$ will correspond to the spin configuration $[s_1, s_2, s_3 \ldots, s_N]$ that minimizes the Ising Hamiltonian H given by:

$$H = -\sum_{i,j=1,i<j}^{N} J_{ij} s_i s_j \qquad (A1.3)$$

In fact, with $C_1 = \frac{1}{2}$, equation (A1.2) is equivalent to the Ising Hamiltonian with a constant offset.

$$E(\phi(t)) = -\sum_{i,j=1,i<j}^{N} J_{ij} \cos(\Delta\phi_{ij}) - NC_{sync} \qquad (A1.4)$$

Here, $\frac{1}{2}$ is used since the spin interaction in the Hamiltonian is represented by two terms (J_{ij}, J_{ji}) in the Lyapunov function.

Appendix 2:

Here, we illustrate how $f(\Delta\phi_{ij})$ can essentially be considered as a coupling function. Equation (4.10) can be expanded and rewritten as:

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1,i\neq j}^{N} \left[\mathbf{J_{ij} cos}\left(\mathbf{f}(\mathbf{\Delta\phi_{ij}})\right) . \sin(\Delta\phi_{ij}) + \left(\mathbf{J_{ij} sin}\left(\mathbf{f}(\mathbf{\Delta\phi_{ij}})\right) \right) . \cos(\Delta\phi_{ij}) \right]$$
$$- C_{sync} \sin(K\phi_i) \qquad (A2.1)$$

Here, $J_{ij}\cos\left(f\left(\Delta\phi_{ij}\right)\right)$ and $J_{ij}\sin\left(f\left(\Delta\phi_{ij}\right)\right)$ can be considered as coupling functions between the quadrature outputs of the oscillator.

Appendix 3: Proof to show $\frac{\partial E(\phi(t))}{\partial\phi_i(t)} = -K\frac{d\phi_i(t)}{dt}$

$$E(\phi(t)) = -\frac{KC_1}{2}\sum_{i,j,\, j\neq i}^{N} J_{ij}\cos\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) - \sum_{i=1}^{N} C_{sync}\cos(K\phi_i(t)) \tag{A3.1}$$

$$\frac{\partial E(\phi(t))}{\partial\phi_i(t)} = -\frac{KC_1}{2}\sum_{j=1,\, j\neq i}^{N} J_{ij}\frac{\partial}{\partial\phi_i(t)}\cos\left(\phi_i - \phi_j + f(\phi_i - \phi_j)\right)$$

$$-\frac{KC_1}{2}\sum_{j=1,\, j\neq i}^{N} J_{ji}\frac{\partial}{\partial\phi_i(t)}\cos\left(\phi_j - \phi_i + f(\phi_j - \phi_i)\right) - C_{sync}\frac{\partial}{\partial\phi_i(t)}\cos(K\phi_i) \tag{A3.2}$$

$$\frac{\partial E(\phi(t))}{\partial\phi_i(t)} = C_2\sum_{j=1,\, j\neq i}^{N} J_{ij}\sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right)\cdot\left(1 + \frac{\partial}{\partial\phi_i(t)}f(\Delta\phi_{ij})\right)$$

$$+C_2\sum_{j=1,\, j\neq i}^{N} J_{ji}\sin\left(\Delta\phi_{ji} + f(\Delta\phi_{ji})\right)\cdot\left(-1 + \frac{\partial}{\partial\phi_i(t)}f(\Delta\phi_{ji})\right) + C_{sync}K.\sin(K\phi_i) \tag{A3.3}$$

$$\frac{\partial}{\partial\phi_i(t)}f(\Delta\phi_{ij}) = \lim_{\sigma\to 0}\sum_{k=1}^{K-1}\left(-\left((2k-1)\pi - \frac{2k\pi}{K}\right).e^{-\left(\frac{\left(\phi_i - \phi_j - \frac{2k\pi}{K}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_i - \phi_j - \frac{2k\pi}{K}\right)}{2\sigma^2}\right)\right.$$

$$\left.-\left(\frac{2k\pi}{K} - (2k-1)\pi\right).e^{-\left(\frac{\left(\phi_i - \phi_j + \frac{2k\pi}{K}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_i - \phi_j + \frac{2k\pi}{K}\right)}{2\sigma^2}\right)\right) \tag{A3.4}$$

$$\frac{\partial}{\partial \phi_i(t)} f(\Delta\phi_{ji}) = \lim_{\sigma \to 0} \sum_{k=1}^{K-1} \left((2k-1)\pi - \frac{2k\pi}{K} \right) . e^{-\left(\frac{\left(\phi_j - \phi_i - \frac{2k\pi}{K} \right)^2}{2\sigma^2} \right)} . \left(\frac{2\left(\phi_j - \phi_i - \frac{2k\pi}{K} \right)}{2\sigma^2} \right) \qquad (A3.5)$$

$$+ \left(\frac{2k\pi}{K} - (2k-1)\pi \right) . e^{-\left(\frac{\left(\phi_j - \phi_i + \frac{2k\pi}{K} \right)^2}{2\sigma^2} \right)} . \left(\frac{2\left(\phi_j - \phi_i + \frac{2k\pi}{K} \right)}{2\sigma^2} \right)$$

Using the relation that, $\lim_{\sigma \to 0} \dfrac{e^{-\alpha^2/\sigma^2}}{\sigma^2} = 0$ in equations (A3.4) and (A3.5)

$$\frac{\partial}{\partial \phi_i(t)} f(\Delta\phi_{ij}) = \frac{\partial}{\partial \phi_i(t)} f(\Delta\phi_{ji}) = 0 \qquad (A3.6)$$

Similarly, $\frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ij}) = \frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ji})$ relevant to solving the TSP can be shown to be

equal to 0 as illustrated further on.

Substituting equation (A3.6) into (A3.3) and using $\sin(x) = -\sin(-x)$ & $J_{ij} = J_{ji}$ we get,

$$\frac{\partial E(\phi(t))}{\partial \phi_i(t)} = 2C_2 \sum_{j=1, j \neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij}) \right) + C_{sync} K. \sin(K\phi_i) \qquad (A3.7)$$

$$\frac{\partial E(\phi(t))}{\partial \phi_i(t)} = 2.\frac{K.C_1}{2} \sum_{j=1, j \neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij}) \right) + C_{sync} K. \sin(K\phi_i) \qquad (A3.8)$$

$$\frac{\partial E(\phi(t))}{\partial \phi_i(t)} = KC_1 \sum_{j=1, j \neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f\left(\Delta\phi_{ij} \right) \right) + C_{sync} K. \sin(K\phi_i) = -K.\frac{d\phi_i(t)}{dt} \qquad (A3.9)$$

$$\frac{dE(\phi(t))}{dt} = \sum_{i=1}^{N}\left[\left(\frac{\partial E(\phi(t))}{\partial \phi_i(t)}\right).\left(\frac{d\phi_i(t)}{dt}\right)\right] = -K\sum_{i=1}^{N}\left[\left(\frac{d\phi_i(t)}{dt}\right)^2\right] \le 0 \qquad \text{(A3.10)}$$

To show $\frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ij}) = \frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ji}) = 0$.

$\frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ij})$

$$
\begin{aligned}
= \lim_{\sigma\to 0} -\sum_{\gamma=1,N-1} & \left(\left(-\frac{2\gamma\pi}{N}\right).e^{-\left(\frac{\left(\phi_i - \phi_j - \frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_i - \phi_j - \frac{2\gamma\pi}{N}\right)}{2\sigma^2}\right)\right. \\
& \left. + \left(\frac{2\gamma\pi}{N}\right).e^{-\left(\frac{\left(\phi_i - \phi_j + \frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_i - \phi_j + \frac{2\gamma\pi}{N}\right)}{2\sigma^2}\right)\right) \\
+ \lim_{\sigma\to 0} \sum_{k=2;k\neq N-1}^{N} & \left(-\left(\pi - \frac{2k\pi}{N}\right).e^{-\left(\frac{\left(\phi_i - \phi_j - \frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_i - \phi_j - \frac{2k\pi}{N}\right)}{2\sigma^2}\right)\right. \\
& \left. - \left(\frac{2k\pi}{N} - \pi\right).e^{-\left(\frac{\left(\phi_i - \phi_j + \frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_i - \phi_j + \frac{2k\pi}{N}\right)}{2\sigma^2}\right)\right)
\end{aligned}
$$

(A3.11)

$$\frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ji})$$

(A3.12)

$$= \lim_{\sigma \to 0} - \sum_{\gamma=1,N-1} \left(\left(\frac{2\gamma\pi}{N}\right).e^{-\left(\frac{\left(\phi_j - \phi_i - \frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_j - \phi_i - \frac{2\gamma\pi}{N}\right)}{2\sigma^2}\right) \right.$$

$$\left. + \left(-\frac{2\gamma\pi}{N}\right).e^{-\left(\frac{\left(\phi_j - \phi_i + \frac{2\gamma\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_j - \phi_i + \frac{2\gamma\pi}{N}\right)}{2\sigma^2}\right) \right)$$

$$+ \lim_{\sigma \to 0} \sum_{k=2}^{N} \left(\left(\pi - \frac{2k\pi}{N}\right).e^{-\left(\frac{\left(\phi_j - \phi_i - \frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_j - \phi_i - \frac{2k\pi}{N}\right)}{2\sigma^2}\right) \right.$$

$$\left. + \left(\frac{2k\pi}{N} - \pi\right).e^{-\left(\frac{\left(\phi_j - \phi_i + \frac{2k\pi}{N}\right)^2}{2\sigma^2}\right)}.\left(\frac{2\left(\phi_j - \phi_i + \frac{2k\pi}{N}\right)}{2\sigma^2}\right) \right)$$

Using the relation that, $\lim_{\sigma \to 0} \frac{e^{-\alpha^2/\sigma^2}}{\sigma^2} = 0$ in equations (A3.11) and (A3.12)

$$\frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ij}) = \frac{\partial}{\partial \phi_i(t)} f_{TSP}(\Delta\phi_{ji}) = 0$$

(A3.13)

Substituting equation (A3.13) into (A3.3), and using a similar approach as described above, it can be shown that the energy function for the TSP (equation (4.19) in main text) is a Lyapunov function. The approach can be used to show that $\frac{\partial}{\partial \phi_i(t)} f_{HC}(\Delta\phi_{ij}) = \frac{\partial}{\partial \phi_i(t)} f_{HC}(\Delta\phi_{ji}) = 0$, and the corresponding energy function for the Hamiltonian cycle/path (equation (24) in the main text) is a Lyapunov function.

Appendix 4:

Here, we elucidate the role of the injection term/ external forcing function in discretizing the phases for the general case of the Max-K-Cut problem. We first consider the system dynamics for solving the Max-K-Cut problem *without the external injection*:

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1,\ j\neq i}^{N} J_{ij} \sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) \tag{A4.1}$$

The corresponding Lyapunov function for this system is given by:

$$E(\phi(t)) = -\frac{KC_1}{2} \sum_{i,j,\ j\neq i}^{N} J_{ij} \cos\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) \tag{A4.2}$$

Equation (A4.2) achieves a minimum when the phase difference is $\frac{2\pi k}{K}$, $k = 1,2,\dots,K$. We also note that $f(\Delta\phi_{ij})$ in the first term is specifically designed such that the $\cos\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right)$ term equals -1 (minimum value for the cos(.) function) when $\Delta\phi_{ij} = \frac{2\pi k}{K}$ (for MaxCut, $f(\Delta\phi_{ij})=0$ since the cos(.) naturally achieves a minimum at a phase difference of π).

While this minimum is attained when $\Delta\phi_{ij} = \frac{2\pi k}{K}$, ϕ_i and ϕ_j can assume analog values. For instance, in the case of the Max-3-Cut problem, $\phi_i = 0$ and $\phi_j = \frac{2\pi}{3}$, and $\phi_i = \frac{\pi}{3}$ and $\phi_j = \pi$ are both probable solutions since $\Delta\phi_{ij} = \frac{2\pi}{3}$ in both cases ($f(\Delta\phi_{ij}) = \frac{\pi}{3}$ in both cases). Consequently, the oscillators exhibit a continuous distribution of phases, that cannot be directly mapped to the K-sets created by the Max-K-Cut.

In contrast, when the force term is considered, the resulting system dynamics are described by,

$$\frac{d\phi_i(t)}{dt} = -C_1 \sum_{j=1,\ j\neq i}^{N} J_{ij} \ \sin\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) - C_{sync}\sin(K\phi_i(t)) \qquad \text{(A4.3)}$$

for which the corresponding Lyapunov function is given by,

$$E(\phi(t)) = -\frac{KC_1}{2} \sum_{i,j,\ j\neq i}^{N} J_{ij} \cos\left(\Delta\phi_{ij} + f(\Delta\phi_{ij})\right) - \sum_{i=1}^{N} C_{sync}\cos(K\phi_i(t)) \qquad \text{(A4.4)}$$

In order for Equation (A4.4) to achieve a minimum, not only does $\Delta\phi_{ij} + f(\Delta\phi_{ij}) = (2m + 1)\,\pi$ has to hold true (condition arises from the first term), but ϕ_i and ϕ_j have to also equal $\frac{2\pi k}{K}$. In other words, the forcing term discretizes the phases to K points. Subsequently, these phase points can be directly mapped to the K sets created by the Max-K-Cut problem.

Appendix 5:

The strength of the coupling among the oscillators (C_1) is increased linearly given by $C_1 = 1 + t\frac{A-1}{T}$; where T is the total simulation time (equivalent to 100 oscillation cycles), and A is a constant which is chosen such that the maximum value of $C_1 = A$ (at $t = T$). This implementation is inspired by prior work by Wang et al. [47], wherein increasing C_1 linearly was effectively used as an annealing schedule that helped the system better escape from local minima, and thus, facilitated improved solution quality. We observe
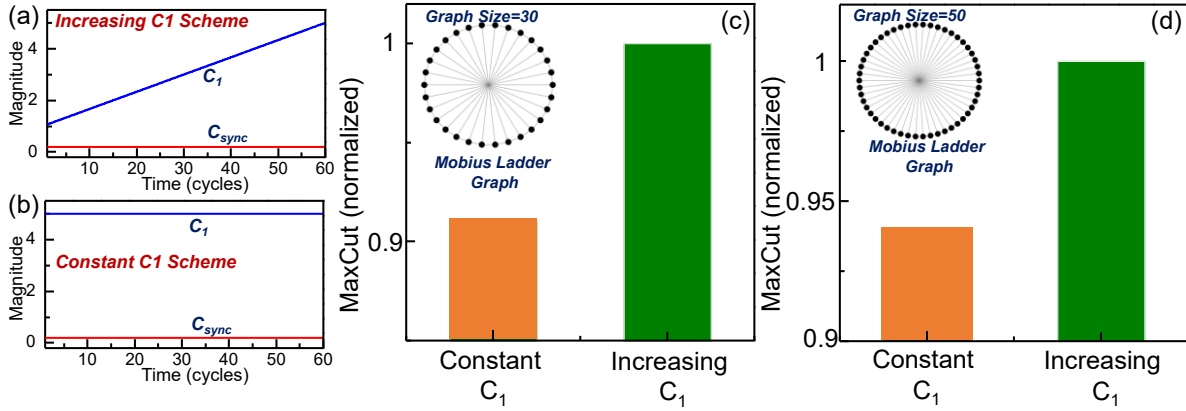
Fig. A1. Impact of C_1 evolution on the MaxCut solution quality. (a)(b) Two schemes for evolution of the coupling constants considered in the example. C_1 increases linearly in (a), while it is constant in (b); C_{sync} is constant in both the cases. (c)(d) Observed MaxCut solutions (normalized to optimal solution) for two representative Mobius ladder graphs of size 30 and 50 nodes, respectively. It can be observed empirically that linearly increasing C_1 produces improved solutions.

similar behavior as shown in the representative graph (Fig. A1), wherein increasing the coupling strength linearly (while maintaining C_1 below a critical threshold) produced better solutions compared to maintaining a constant coupling strength.

However, it must be noted that considering the negative nature of coupling, C_1 cannot be made indefinitely large and must be below a certain threshold. In fact, we observe that the $\frac{C_1}{C_{sync}}$ ratio (i.e., coupling strength among the oscillator nodes relative to the coupling constant for external forcing function) is critical to achieving the desired functionality, and the system clusters get destabilized if the $\frac{C_1}{C_{sync}}$ ratio is above a certain threshold. This is illustrated for a representative graph in Fig. A2 where it can be observed that the clusters start to destabilize for $\frac{C_1}{C_{sync}} > 70$. We note that the exact value of the threshold will depend on the properties of the graph such as size and connectivity as well as the computing problem.
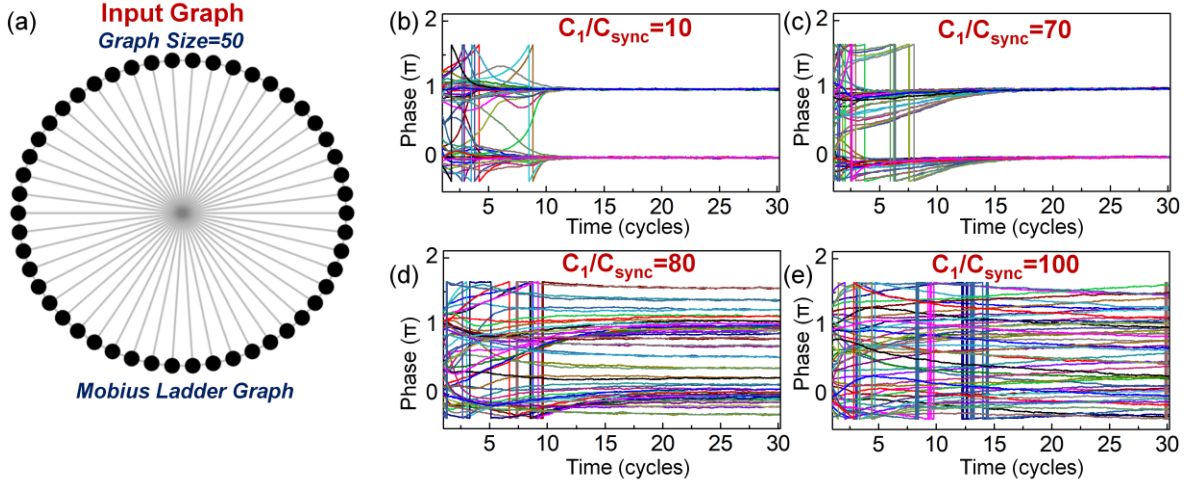
Fig. A2. (a) Mobius ladder graph considered in the illustrative example. (b)-(e) Evolution of oscillator phases for different C_1/C_{sync} ratios (=10, 70, 80, 100, respectively). C_1: coupling strength among the oscillators; C_{sync}: Coupling strength of external injection signal / force function. It can be observed that C_1 (relative to C_{sync}) must remain below a critical threshold in order to observe the phase clustering.

We also note that besides C_1, C_{sync} the standard deviation σ of the gaussian distributions must be carefully designed for the problems considered in this work. While force function ensures that certain angles $\left(\frac{2k\pi}{N}; k = 1, 2, 3..N\right)$ are energetically favored, the system, while evolving towards this low energy configuration, may assume other phase angles not equal to $\frac{2k\pi}{N}$. It is important to make sure that such angles do not help the system evolve towards a lower energy configuration. This can be ensured by appropriately choosing σ as well as adding additional gaussian terms in the function $f\left(\Delta\theta_{ij}\right)$ designed to penalize the system (i.e., increase energy) for such configurations. For instance, in the case of the Hamiltonian cycle, gaussian functions can be added to ensure that the system energy is increased for all angles not equal to the $\pm\frac{2\pi}{N}$ (instead of just considering an energy penalty for $\frac{2k\pi}{N}; k \neq 1, N - 1$.

Appendix 6:

Input Graph	No. of Partitions (K)	Evolutions of oscillator phases
	Max-2-Cut (K=2) *Solution : 1561*	
	Max-3-Cut (K=3) *Solution : 2040*	
Graph Size: 100 *No. of Edges: 2970*	**Max-4-Cut (K=4)** *Solution : 2279*	
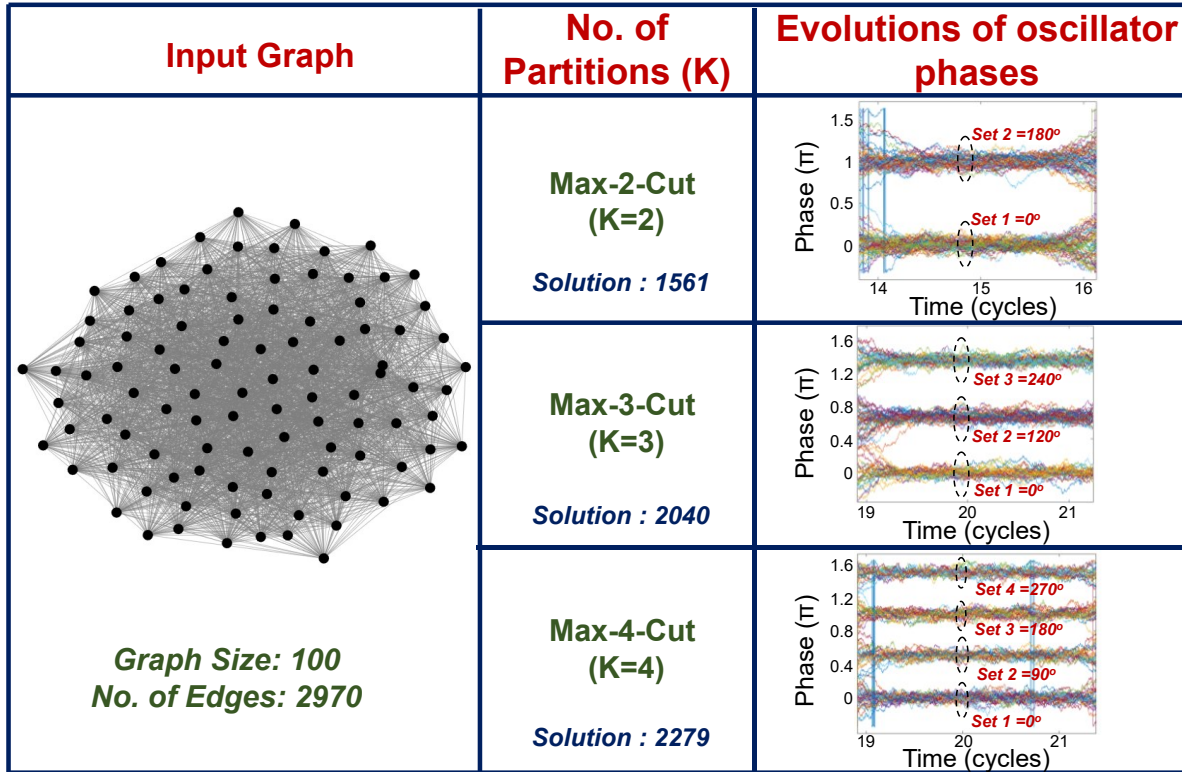
Fig. A3. A representative 100 node (randomly instantiated) graph and the evolution of oscillator phases corresponding to the solution of the Max-K-Cut problem (K=2, 3, 4).

Graph	Nodes (V)	Edges (E)	Oscillator Solution (Max-2-cut)	Solution Quality (Max-2-cut)	Oscillator Solution (Max-3-cut)	Solution Quality (Max-3-cut)	Oscillator Solution (Max-4-cut)	Solution Quality (Max-4-cut)
G1	800	19176	11624	100%	15032	99.1%	16166	96.2%
G2	800	19176	11606	99.8%	14878	98.1%	16197	96.4%
G3	800	19176	11613	99.9%	14852	97.88%	16204	96.4%
G4	800	19176	11635	99.9%	14894	98.1%	16208	96.4%
G5	800	19176	11570	99.4%	14915	98.2%	16211	96.4%

Fig. A4. Graph instances from the G-Set solved using the oscillator-based computational model; the oscillators compute yields high quality Max-K-Cut solutions within >96% of the solutions obtained by F. Ma et al. [103]. Here, $\text{Solution Quality} = \frac{\text{OScillator Solution}}{\text{Best Known Solution}} \times 100\%$.

Publications

Journal Articles:

1. **<u>A. Mallick</u>**, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, ''Using synchronized oscillators to compute the maximum independent set,'' *Nature Communications*., vol. 11, no. 1, pp. 1–7, Dec. 2020.

2. **<u>A. Mallick</u>**[*], M. K. Bashar[*], D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, ''Experimental demonstration of a reconfigurable coupled oscillator platform to solve the max-cut problem,'' IEEE J. Explor. Solid State Comput. Devices Circuits, vol. 6, pp. 116–121, 2020. *equal contribution*

3. **<u>A. Mallick</u>**, M. K. Bashar, Z. Lin, and N. Shukla, "Computational Models based on Synchronized Oscillators for Solving Combinatorial Optimization Problems," Phys. Rev. Applied 17, 064064.

4. **<u>A. Mallick</u>**, Z. Zhao, M. K. Bashar, S. Alam, M.M. Islam, Y. Xiao, Y. Xu, A. Aziz, V. Narayanan, K. Ni and N. Shukla, "CMOS-Compatible Ising Machines built using Bistable Latches Coupled through Ferroelectric Transistor Arrays," arXiv preprint ArXiv: 2205.14729 (2022).

5. **<u>A. Mallick</u>**, and N. Shukla*.,* "Evaluation of Bulk and SOI FeFET Architecture for Non-Volatile Memory Applications," in *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 425-429, 2019.

6. M. K. Bashar, **<u>A. Mallick,</u>** and N. Shukla*,* "Experimental Investigation of the Dynamics of Coupled Oscillators as Ising Machines," in *IEEE Access*, vol. 9, pp. 148184-148190, 2021, doi: 10.1109/ACCESS.2021.3124808.

7. M. K. Bashar, **A. Mallick**, A. W. Ghosh, and N. Shukla, "Dynamical system-based computational models for solving combinatorial optimization on hypergraphs" *arXiv preprint arXiv:2207.09618 (2022)*.

Conference Proceedings:

8. **A. Mallick**, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, and N. Shukla, "Overcoming the Accuracy vs. Performance Trade-off in Oscillator Ising Machines," in *2021 IEEE International Electron Devices Meeting (IEDM)*, pp. 40.2.1-40.2, 2021.

9. **A. Mallick**, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla "Graph Coloring using Coupled Oscillator-based Dynamical Systems," in IEEE ISCAS, 2021.

10. **A. Mallick,** M.K. Bashar, and N. Shukla "Oscillator-based Dynamical Computing Platforms to Solve Combinatorial Optimization," in *6th IEEE Electron Devices Technology & Manufacturing Conference (EDTM)*, 2022, pp. 125-127.

11. M. K. Bashar, R. Hrdy, **A. Mallick**, F. Farnoud Hassanzadeh and N. Shukla, "Solving the Maximum Independent Set Problem using Coupled Relaxation Oscillators," in *2019 Device Research Conference (DRC)*, Ann Arbor, MI, USA, pp. 187-188, 2019.

Bibliography

1. J. Von Neumann, "First Draft of a Report on the EDVAC," archived from the original (1945) PDF on March 14, 2013, retrieved August 24, 2011.

2. M. A. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," in Proceedings of the London Mathematical Society, vol. 42, pp. 230–265, published 1937.

3. N. Stern, "John von Neumann's Influence on Electronic Digital Computing," in Annals of the History of Computing, vol. 2, no. 4, pp. 349-362, Oct.-Dec. 1980.

4. W. Aspray, "John von Neumann and the origins of modern computing," MIT Press,1990.

5. J. Welser, J. W. Pitera and C. Goldberg, "Future Computing Hardware for AI, " in IEEE International Electron Devices Meeting (IEDM), pp. 1.3.1-1.3.6, 2018.

6. P. Kidwell, "Landmarks in Digital Computing: A Smithsonian Pictorial History [Reviews]," in IEEE Annals of the History of Computing, vol. 17, no. 3, pp. 82-, Fall 1995.

7. P. Chen, L. Wu, S.Liu, and I. Rajapakse, "Fast Incremental von Neumann Graph Entropy Computation: Theory, Algorithm, and Applications," in Proceedings of the 36th International Conference on Machine Learning, PMLR 97:1091-1101, 2019.

8. T. Zanotti, F. M. Puglisi, and P. Pavan, "Energy-efficient non-von neumann computing architecture supporting multiple computing paradigms for logic and binarized neural networks," Journal of Low Power Electronics and Applications, 11(3), p.29, 2021.

9. S. Roy, M. Shrivastava, C. V. Pandey, S. K. Nayak, and U. Rawat, "IEVCA: An efficient image encryption technique for IoT applications using 2-D Von-Neumann cellular automata,"Multimedia Tools and Applications, vol. 80(21), pp.31529-31567, 2021.

10. Y.G. Bao, and S. Wang, "Labeled von Neumann architecture for software-defined cloud," Journal of Computer Science and Technology, 32(2), pp.219-223, 2017.

11. W. Aspray, "The transformation of numerical analysis by the computer: An example from the work of John von Neumann," In Institutions and Applications, Academic Press, pp. 306-322, 1989.

12. F. Neumann, & C. Witt, "Combinatorial optimization and computational complexity", in Bioinspired computation in combinatorial optimization, pp.9-19, Springer, Berlin, Heidelberg, 2010.

13. D. Du, and P. M. Pardalos, "Handbook of combinatorial optimization," Springer Science & Business Media, vol.4, 1998.

14. I. Averbakh, "On the complexity of a class of combinatorial optimization problems with uncertainty," Mathematical Programming, vol. 90(2), pp.263-272, 2001.

15. G. J. Woeginger, "Exact algorithms for NP-hard problems: A survey. In Combinatorial optimization—eureka, you shrink!, pp. 185-207. Springer, Berlin, Heidelberg, 2003.

16. Arora, S., "The approximability of NP-hard problems," in Proceedings of the 13th annual ACM symposium on Theory of computing, pp. 337-348, 1998.

17. B. Shahriari, K. Swersky, Z. Wang et. al, "Taking the human out of the loop: A review of Bayesian optimization", Proceedings of the IEEE **104**, 148-175, 2015.

18. B. Tosuni, "Graph Coloring Problems in Modern Computer Science", European Journal of Interdisciplinary Studies **1**, 87-95, 2015.

19. B. Padmanabhan, & A. Tuzhilin, "On the use of optimization for data mining: Theoretical interactions and eCRM opportunities," Management Science 49, 1327-1343, 2003.

20. J. H. Greenberg, W. E. Hart, & G. Lancia, "Opportunities for combinatorial optimization in computational biology. INFORMS Journal on Computing **16**, 211-231, 2004.

21. J. J. Fallas-Monge, J. Chavarría-Molina, & G. Alpízar-Brenes, "Combinatorial metaheuristics applied to infectious disease models", in 2016 IEEE 36th Central American and Panama Convention (CONCAPAN XXXVI), pp. 1-6 (IEEE, 2016).

22. S. Siddhaye, K. Camarda, M. Southard, & E. Topp, "Pharmaceutical product design using combinatorial optimization", Computers & Chemical Engineering 28, 425-434, 2004.

23. N. Rhodes, P. Willett, A. Calvet, J. B. Dunbar, & C. Humblet, "CLIP: similarity searching of 3D databases using clique detection," Journal of chemical information and computer sciences 43, 443-448, 2003.

24. D. B. Kell, "Scientific discovery as a combinatorial optimisation problem: how best to navigate the landscape of possible experiments?," Bioessays 34, 236-244 (2012).

25. C. S. Calude, "Unconventional computing: A brief subjective history," in Advances in Unconventional Computing, 855-864, Springer, Cham, 2017.

26. A. Chen, S. Datta, X. S. Hu, M. T. Niemier, T. Š. Rosing, & J. J. Yang, "A survey on architecture advances enabled by emerging beyond-cmos technologies" IEEE Design & Test 36, 46-68, 2019.

27. J. Backus, "Can programming be liberated from the von Neumann style? A functional style and its algebra of programs." Communications of the ACM **21**, 613-641, 1978.

28. D. Efnusheva, A. Cholakoska, and A. Tentov, "A survey of different approaches for overcoming the processor-memory bottleneck," International Journal of Computer Science and Information Technology, 9(2), pp.151-163, 2017.

29. I. I.Arikpo, F. U. Ogban, and I. E. Eteng, "Von neumann architecture and modern computers," *Global Journal of Mathematical Sciences* 6, no. 2 (2007): 97-103.

30. A. Laird, "The Von Neumann architecture topic paper# 3," Computer Science, 319, pp.360-8771.

31. P. M. Kogge, "Function-based computing and parallelism: A review. Parallel computing," 2(3), pp.243-253, 1985.

32. X. Zou, S. Xu, X. Chen, L. Yan, and Y. Han, "Breaking the von Neumann bottleneck: architecture-level processing-in-memory technology," Science China Information Sciences, 64(6), pp.1-10, 2021.

33. W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, , H. Ohno, and S. Datta, "Integer factorization using stochastic magnetic tunnel junctions," Nature 573, 390–393, 2019.

34. K. Y. Camsari, R. Faria, B. M. Sutton, & S. Datta, "Stochastic p-bits for invertible logic," Phys. Rev. X 7, 031014, 2017.

35. D. Pierangeli, G. Marcucci, & C. Conti, "Large-scale photonic Ising machine by spatial light modulation," Phys. Rev. Lett. 122, 213902, 2019.

36. R. Andrawis & K. Roy, "Antiferroelectric Tunnel Junctions as Energy-Efficient Coupled Oscillators: Modeling, Analysis, and Application to Solving Combinatorial Optimization Problems" IEEE Transactions on Electron Devices, **67** (7), pp. 2974-2980, (2020).

37. C. Roques-Carmes, Y. Shen, C. Zanoci, M. Prabhu, F. Atieh, L. Jing, T. Dubček, *et al.*, "Heuristic recurrent algorithms for photonic Ising machines. Nat. communications 11, 1–8 (2020).

38. F. Cai, S. Kumar, T. V. Vaerenbergh, X. Sheng, R. Liu, C. Li, Z. Liu *et al.*,"Power-efficient combinatorial optimization using intrinsic noise in memristor hopfield neural networks," Nat. Electron. 3, 409–418 (2020).

39. M. N. Bojnordi,. & E. Ipek, "Memristive boltzmann machine: A hardware accelerator for combinatorial optimization and deep learning," in 2016 IEEE International Symposium on High Performance Computer Architecture (HPCA), 1–13 (IEEE, 2016).

40. B. Kiraly, E. J. Knol, W. M. van Weerdenburg, H. J. Kappen, & A. A. Khajetoorians, "An atomic Boltzmann machine capable of self-adaption," Nat. Nanotechnol. 1–7 (2021).

41. T. Byrnes, S. Koyama, K. Yan, & Y. Yamamoto, "Neural networks using two-component Bose-Einstein condensates," Sci. reports 3, 1–7 (2013).

42. K. Mizushima, H. Goto, H. & R. Sato, "Large-scale Ising-machines composed of magnetic neurons," Appl. Phys. Lett. 111, 172406 (2017).

43. D. Reis, A. F. Laguna, M. Niemier, and X. S. Hu, "Attention-in-Memory for Few-Shot Learning with Configurable Ferroelectric FET Arrays," in Proceedings of the 26th Asia and South Pacific Design Automation Conference, pp. 49-54 (2021).

44. E. Ising, "Beitrag zur theorie des ferro-und paramagnetismus." PhD diss., Grefe & Tiedemann,1924.

45. G. B. Stephen, "History of the Lenz-Ising Model. Rev. Mod. Phys.," 39:883–893, Oct 1967

46. A. Lucas, "Ising formulations of many NP problems." Frontiers in physics 2, 5 (2014).

47. T. Wang, & J. Roychowdhury, "OIM: Oscillator-based Ising machines for solving combinatorial optimisation problems" In International Conference on Unconventional Computation and Natural Computation, 232-256, Springer, Cham, 2019.

48. M. W. Johnson, M. H. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris et al. et al., "Quantum annealing with manufactured spins," Nature, vol. 473, no. 7346, pp. 194–198, May 2011.

49. A. D. King and C. C. McGeoch, "Algorithm engineering for a quantum annealing platform," 2014, arXiv:1410.2628. [Online]. Available: http://arxiv.org/abs/1410.2628

50. V. N. Smelyanskiy et al., "A near-term quantum computing approach for hard computational problems in space exploration," 2012

51. A. Das, A. & B. K. Chakrabarti, "Colloquium: Quantum annealing and analog quantum computation," Rev. Mod. Phys. 80, 1061–1081 (2008).

52. P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, & W. D. Oliver, "Perspectives of quantum annealing: Methods and implementations," Reports on Prog. Phys. 83, 054401 (2020)

53. T. Albash, & D. A. Lidar, "Adiabatic quantum computation," Rev. Mod. Phys. 90, 015002 (2018).

54. S. Utsunomiya, K. Takata,. & Y. Yamamoto, "Mapping of Ising models onto injection-locked laser systems," Opt. express 19, 18091–18108 (2011).

55. A. Marandi, Z. Wang, K. Takata, R. L. Byer, & Y. Yamamoto, "Network of time-multiplexed optical parametric oscillators as a coherent Ising machine," Nat. Photonics 8, 937–942 (2014).

56. Z. Wang, A. Marandi, K. Wen, R. L. Byer, and Y. Yamamoto, "Coherent ising machine based on degenerate optical parametric oscillators," Phys. Rev. A, Gen. Phys., vol. 88, no. 6, Dec. 2013, Art. no. 063853.

57. P. L. Mcmahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki et al., "A fully programmable 100-spin coherent Ising machine with all-to-all connections," Science, vol. 354, no. 6312, pp. 614–617, Nov. 2016.

58. Y. Haribara, S. Utsunomiya, and Y. Yamamoto, "A coherent ising machine for MAX-CUT problems: Performance evaluation against semidefinite programming and simulated annealing," Lect. Notes Phys., vol. 911, pp. 251–262, Oct. 2016.

59. T. Inagaki, Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi *et al.*, "A coherent Ising machine for 2000-node optimization problems," Science 354, 603–606 (2016).

60. M. Yamaoka, C. Yoshimura, M. Hayashi, T. Okuyama, H. Aoki, and H. Mizuno, "A 20k-spin ising chip to solve combinatorial optimization problems with CMOS annealing," IEEE J. Solid-State Circuits, vol. 51, no. 1, pp. 303–309, Jan. 2016.

61. T. Takemoto, M. Hayashi, C. Yoshimura, and M. Yamaoka, "2.6 a 2×30kspin multichip scalable annealing processor based on a Processing-inmemory approach for solving large-scale combinatorial optimization problems," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2019, pp. 52–54.

62. Y. Su, H. Kim, and B. Kim, "31.2 CIM-spin: A 0.5-to-1.2 V scalable annealing processor using digital compute-in-memory spin operators and register-based spins for combinatorial optimization problems," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2020, pp. 480–482.

63. K. Yamamoto et al., "7.3 STATICA: A 512-spin 0.25 M-weight fulldigital annealing processor with a near-memory all-spin-updates-at-once architecture for combinatorial optimization with complete spin-spin interactions," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, pp. 138–140, IEEE, 2020.

64. D-Wave Announces D-Wave 2000Q Quantum Computer and First System Order | D-Wave Systems". www.dwavesys.com. Retrieved 2017-01-25.

65. D-Wave Systems, PDF, 01-2017, http://www.dwavesys.com/sites/default/files/D-Wave%202000Q%20Tech%20Collateral_0117F.pdf

66. R. Hamerly, T. Inagaki, P. L. McMahon, D. Venturelli, A. Marandi, T. Onodera, E. Ng et al., "Experimental investigation of performance differences between coherent Ising machines and a quantum annealer," Sci. Adv., vol. 5, no. 5, May 2019.

67. G. Csaba and W. Porod, "Coupled oscillators for computing: A review and perspective," Appl. Phys. Rev., vol. 7, no. 1, Mar. 2020, Art. no. 011302.

68. D. E. Nikonov, P. Kurahashi, J. S. Ayers, H.-J. Lee, Y. Fan, and I. A. Young, "A coupled CMOS oscillator array for 8ns and 55pJ inference in convolutional neural networks," 2019, arXiv:1910.11803. [Online]. Available: http://arxiv.org/abs/1910.11803

69. Z. Hull, D. Chiarulli, S. Levitan, G. Csaba, W. Porod, M. Pufall, W. Rippar *et al.*, "Computation with coupled oscillators in an image processing pipeline," in Proc. 15th Int. Workshop Cellular Nanosc. Netw. Appl., Aug. 2016, pp. 1–2.

70. A. Parihar, N. Shukla, M. Jerry, S. Datta, and A. Raychowdhury, "Vertex coloring of graphs via phase dynamics of coupled oscillatory networks," Sci. Rep., vol. 7, no. 1, Apr. 2017, Art. no. 911.

71. D. E. Nikonov, I. A. Young, and G. I. Bourianoff, "Convolutional networks for image processing by coupled oscillator arrays," 2014, arXiv:1409.4469. [Online]. Available: http://arxiv.org/abs/1409.4469

72. F. C. Hoppensteadt and E. M. Izhikevich, "Synaptic organizations and dynamical properties of weakly connected neural oscillators," Biol. Cybern., vol. 75, no. 2, pp. 117–127, Aug. 1996.

73. J. Torrejon et al., "Neuromorphic computing with nanoscale spintronic oscillators," Nature, vol. 547, no. 7664, pp. 428–431, Jul. 2017.

74. G. Csaba, A. Papp, W. Porod, and R. Yeniceri, "Non-Boolean computing based on linear waves and oscillators," in Proc. 45th Eur. Solid State Device Res. Conf. (ESSDERC), Sep. 2015, pp. 101–104.

75. D. E. Nikonov et al., "Coupled-oscillator associative memory array operation for pattern recognition," IEEE J. Explor. Solid-State Comput. Devices Circuits, vol. 1, pp. 85–93, 2015.

76. S. P. Levitan et al., "Associative processing with coupled oscillators," in Proc. Int. Symp. Low Power Electron. Des. (ISLPED), Sep. 2013, p. 235.

77. G. Csaba, A. Raychowdhury, S. Datta, and W. Porod, "Computing with coupled oscillators: Theory, devices, and applications," in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), 2018, pp. 1–5.

78. S. P. Levitan, Y. Fang, D. H. Dash, T. Shibata, D. E. Nikonov, and G. I. Bourianoff, "Non-Boolean associative architectures based on nanooscillators," in Proc. 13th Int. Workshop Cellular Nanosc. Netw. Appl., Aug. 2012, pp. 1–6.

79. G. Csaba and W. Porod, "Computational study of spin-torque oscillator interactions for non-Boolean computing applications," IEEE Trans. Magn., vol. 49, no. 7, pp. 4447–4451, Jul. 2013.

80. T. Wang, L. Wu, and J. Roychowdhury, "New computational results and hardware prototypes for oscillator-based ising machines," in Proc. 56th Annu. Des. Autom. Conf., Jun. 2019, pp. 1–2.

81. J. Chou, S. Bramhavar, S. Ghosh, and W. Herzog, "Analog coupled oscillator based weighted ising machine," Sci. Rep., vol. 9, no. 1, Oct. 2019, Art. no. 14786.

82. S. Dutta, A. Khanna, J. Gomez, K. Ni, Z. Toroczkai, and S. Datta, "Experimental demonstration of phase transition nano-oscillator based ising machine," in IEDM Tech. Dig., Dec. 2019, pp. 911–914.

83. I. Ahmed, P.-W. Chiu, and C. H. Kim, "A probabilistic self-annealing compute fabric based on 560 hexagonally coupled ring oscillators for solving combinatorial optimization problems," in Proc. IEEE Symp. VLSI Circuits, 2020, pp. 1–2.

84. F. Hadlock, "Finding a maximum cut of a planar graph in polynomial time," SIAM Journal on Computing, vol. 4, pp.221-225, 1975.

85. M. Babaeian, D. T. Nguyen, V. Demir, M. Akbulut, P. A. Blanche, Y. Kaneda, S. Guha, M. A. Neifeld, and N. Peyghambarian, "A single shot coherent Ising machine based on a network of injection-locked multicore fiber lasers",Nature communications, 10(1), pp.1-11,2019.

86. S. Dal Toso, A. Bevilacqua, M. Tiebout, N. Da Dalt, A. Gerosa, and A. Neviani, "A 0.06 mm^2 11 mW local oscillator for the GSM standard in 65 nm CMOS", IEEE Journal of Solid-State Circuits **45**, 1295-1304 (2010).

87. A. Neogy and J. Roychowdhury, "Analysis and design of sub-harmonically injection locked oscillators," 2012 Design, Automation & Test in Europe Conference & Exhibition (DATE), 2012, pp. 1209-1214, doi: 10.1109/DATE.2012.6176677

88. A. Mallick, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, "Using synchronized oscillators to compute the maximum independent set", Nature Communications 11, 4689 (2020).

89. A. Mallick, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, "Graph Coloring using Coupled Oscillator-based Dynamical Systems," IEEE ISCAS, 2021.

90. J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort and R. Spigler, "The Kuramoto Model: A Simple Paradigm for Synchronization Phenomena," Reviews of Modern Physics, 77(1):137, 2005.

91. F. Rendl, G. Rinaldi, and A. Wiegele, ''Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations,'' Math. Program., vol. 121, no. 2, pp. 307–335, Feb. 2010

92. Biq Mac Solver—Binary Quadratic and Max Cut Solver. Accessed: Aug. 1, 2020. [Online]. Available: http://biqmac.uni-klu.ac.at

93. N. Boumal, B. Mishra, P. A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," The Journal of Machine Learning Research, vol. 15, 2014.

94. G. Melancon, ''Just how dense are dense graphs in the real world?: A methodological note,'' in Proc. AVI Workshop BEyond Time Errors Novel Eval. Methods Inf. Visualizat., 2006, pp. 1–7.

95. M. K. Bashar*, A. Mallick*, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, "Experimental Demonstration of a Reconfigurable Coupled Oscillator Platform to Solve the Max-Cut Problem," IEEE JXCDC, 2020. *equal contribution

96. S. Dutta, A. Khanna, A. S. Assoa, H. Paik, D. G. Schlom, Z. Toroczkai, A. Raychowdhury, and S. Datta, "An Ising Hamiltonian solver based on coupled stochastic phase-transition nano-oscillators", Nat Electron 4, 502–512 (2021).

97. P. Bhansali and J. Roychowdhury, in Proceedings of the IEEE ASP-DAC (2009), p. 522

98. A. Crnkić, J. Povh, V. Jaćimović, and Z. Levnajić, Collective dynamics of phase repulsive oscillators solve graph coloring problem, Chaos: An Interdisciplinary Journal of Nonlinear Science 30, 033128 (2020).

99. R. Brooks, On coloring the nodes of a network, Math. Proc. Camb. Philos. Soc. 37, 194 (1941).

100. M. K. Bashar, A. Mallick, A. W. Ghosh, and N. Shukla, "Dynamical system-based computational models for solving combinatorial optimization on hypergraphs" arXiv preprint arXiv:2207.09618 (2022).

101. M. Trentzsch, S. Flachowsky, R. Richter, J. Paul, B. Reimer, D. Utess, S. Jansen et al. "A 28 nm HKMG super low power embedded NVM technology based on ferroelectric FETs," in 2016 IEEE International Electron Devices Meeting (IEDM) 11–5 (IEEE, 2016).

102. S. Deng, G. Yin, W. Chakraborty, S. Dutta, S. Datta, X. Li, and K. Ni, "A Comprehensive Model for Ferroelectric FET Capturing the Key Behaviors: Scalability, Variation, Stochasticity, and Accumulation," in 2020 IEEE Symposium on VLSI Technology, 1-2 (2020).

103. F. Ma and J. K. Hao, A multiple search operator heuristic for the Max-K-Cut problem, Ann. Oper. Res. 248, 365 (2017).