

A Probabilistic Self-annealing Compute Fabric based on 560 Hexagonally Coupled Ring Oscillators for Solving Combinatorial Optimization Problems

Ibrahim Ahmed, Po-Wei Chiu, and Chris H. Kim

Dept. of ECE, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA (Email: ahmed589@umn.edu)

Abstract

NP-hard combinatorial optimization problems (COPs) are very expensive to solve with traditional computers. COPs can be mapped to a coupled spin network where the ground state of the system is the solution. We propose a scalable truly coupled CMOS oscillator-based integrated system mimicking a spin network to solve COPs in hardware. Our simple latch-based coupling design finds solutions of max-cut problems with 85%-100% accuracy 10^4 - 10^6 times faster than commercial software running on a CPU. **Keywords:** Annealing processor, Ising machine, oscillator-based computation, max-cut.

Introduction

Combinatorial optimization problems (COPs), such as Boolean satisfiability, traveling salesman, max-cut, are a class of NP-hard problems that are intractable to solve using a traditional computer due to the extremely large search space. AI decision making, vehicle routing, VLSI layout optimization, network design, and many other modern applications can be modeled as COPs. A promising approach to solve COPs is mapping the problem to a network of spins [1] as shown in Fig. 1. The coupling dynamics minimizes the Ising Hamiltonian of the network and the final states of all spins is the solution. Previous implementations of spin networks require quantum devices operating at cryogenic temperatures [2], are based on digital logic without the coupling dynamics [3–4], or require special process [5–6]. We present a first-of-its-kind coupled ring oscillator (ROSC) based probabilistic compute engine for NP-hard COPs. A latch-based coupling of 560 ROSCs can produce satisfactory solutions of difficult max-cut problems 10^4 - 10^6 times faster than a commercial software running on a CPU.

Proposed Ising Fabric Design

Our proposed design has a 28×20 array of modular hexagonal unit cells representing the spin network, as shown in Fig. 2. We chose the hexagonal structure to maximize the number of neighbors per cell in 2D plane mimicking a spin-based system. We use nearest neighbor coupling where each cell is coupled with six of its neighbors. The coupled ROSCs oscillate in either the same-phase or opposite-phase with their neighbors. Hence, similar to the spin network, our system has two stable states. Fig. 3 shows the modular unit cell consisting ROSC, coupling and read blocks. Each ROSC can be synchronized with an on-chip clock. We use twice the frequency of ROSCs for the clock to allow two stable phases. The measured frequency of ROSC was 118 MHz at nominal VDD and room temperature. The coupling block consists of three latches that couples to three neighbors. The other three coupling connections come from the neighbors. We use “negative” phase coupling to demonstrate the max-cut problem [1]. The ROSCs and the latches are designed with local and global enable signals. Hence, any cell can be programmed to couple with any of its neighbors and latches can be enabled using the global enable signal. The read block measures the relative phase with its neighbor. Fig. 4 shows when the coupling latches are turned ON, the ROSCs are frequency locked with stable phases after exploring various minima. The latches are turned OFF repeatedly using the global latch enable signal for annealing the chip.

Probabilistic Exploration of Local Minima

We measure and reprogram the chip 100 times using the same graph shown in Fig. 5a. Fig. 5b shows the probability of a unit cell changing the state between iterations. Interestingly, Fig. 5c shows each iteration produces very similar results as max-cut can have multiple similar solutions. Hence, despite the solutions of different iterations are not the same, the quality of the result is surprisingly consistent. The distribution of Hamming distances between iterations, as shown in Fig. 5d, confirms the solutions are indeed widely different from each other, which proves the chip’s ability to converge to different local minima rather than finding deterministic solutions. The probabilistic nature of the chip and traversing through multiple local minima is a very crucial characteristic to solve COPs.

Measured COP Results

We map COPs with various levels of “difficulty” to the chip as shown in Fig. 6. We compare our results with a commercial COP software, LocalSolver, which is 10^4 - 10^6 times slower than our chip with roughly 4 orders of magnitude higher power consumption. The measured solutions for “easy” COPs have an accuracy between 98%-100%. Similarly, we mapped “difficult” COPs with various graph dimensions to the chip. For each dimension, we measured 150 problems. We compared the measured results with LocalSolver, as well as 1 million Monte-Carlo solutions sampled from the entire solution-space. The chip accuracy increases with annealing. We anneal the chip only three times to reduce delay. Fig. 7 shows the distribution of “difficult” max-cut results measured from the chip under nominal VDD and room temperature, and the sampled solutions normalized with software solution. The chip results are consistently better than the best solution from Monte-Carlo runs. For smaller graphs, such as 6×6 , the chip finds cut values within 95% of LocalSolver. For larger graph, such as 26×18 , the chip finds an average cut value within 86% of the software solution. A common heuristic finds max-cut solution within 88% of optimal result [7] indicating the chip solutions are satisfactory for practical applications.

We compare prior works with our design in Fig. 8. The lack of details and accuracy statistics of previous proposals makes it difficult to fully compare various approaches. We measured 1000 problems with various difficulty levels in this work. Additionally, compared to quantum computers, our design does not require a special process and can work at room temperature. Fig. 9 shows the chip photo and summary.

Acknowledgement This work was supported in part by the Center for Probabilistic Spin Logic for Low-Energy Boolean and Non-Boolean Computing (CAPSL), one of the Nanoelectronic Computing Research (nCORE) Centers as task 2759.007, a Semiconductor Research Corporation (SRC) program sponsored by the NSF through ECCS 1739635.

References [1] A. Lucas, *Front. Phys.*, vol. 2, pp. 1-15, 2014. [2] M. Johnson, *Nature*, vol. 473, pp. 194-198, 2011. [3] M. Yamaoka, *JSSC*, vol. 51, pp. 303-309, 2016. [4] T. Takemoto, *ISSCC*, pp. 52-54, 2019. [5] D. Pierangeli, *PRL*, vol. 122, 213902, 2019. [6] S. Dutta, *IEDM*, pp. 911-914, 2019. [7] M. X., Goemans, *JACM*, vol. 42, pp. 1115-1145, 1995.

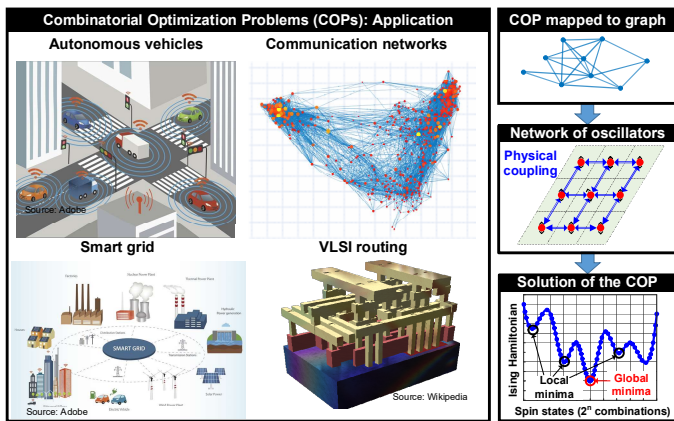


Fig. 1. (Left) Applications of COPs include autonomous vehicle routing, communication network design, smart grid design, VLSI layout optimization. (Right) COPs can be solved by a network of coupled oscillators where the global minima is the solution [1-4].

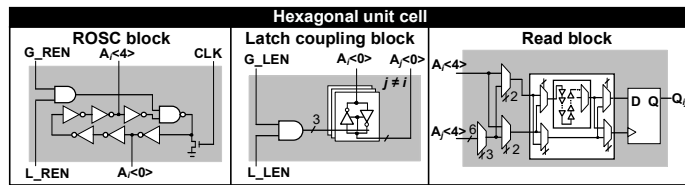


Fig. 3. Each modular cell consists of a ROSC, latch coupling block with global and local enabler, and read block to measure the neighboring cells.

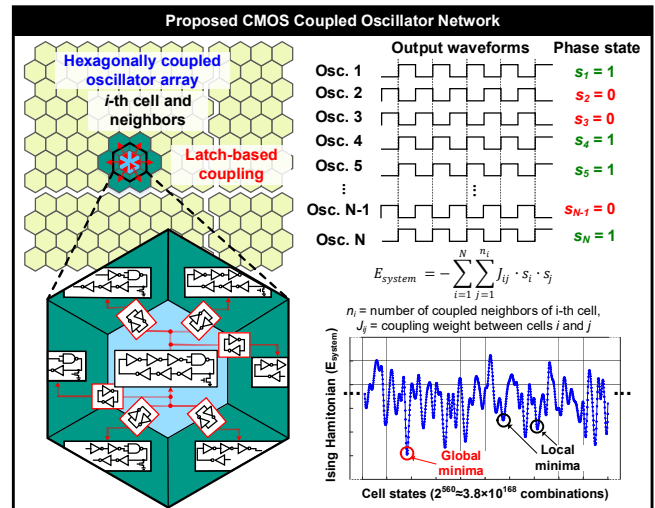


Fig. 2: (Left) Proposed 28×20 hexagonally coupled array. (Right) ROSCs have two stable states, $s_i = \{0,1\}$. The Ising Hamiltonian depends on the states of ROSCs and the coupling weights J_{ij} .

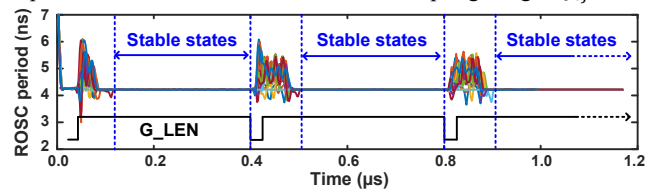


Fig. 4. Annealing: coupling is turned off for short period of time.

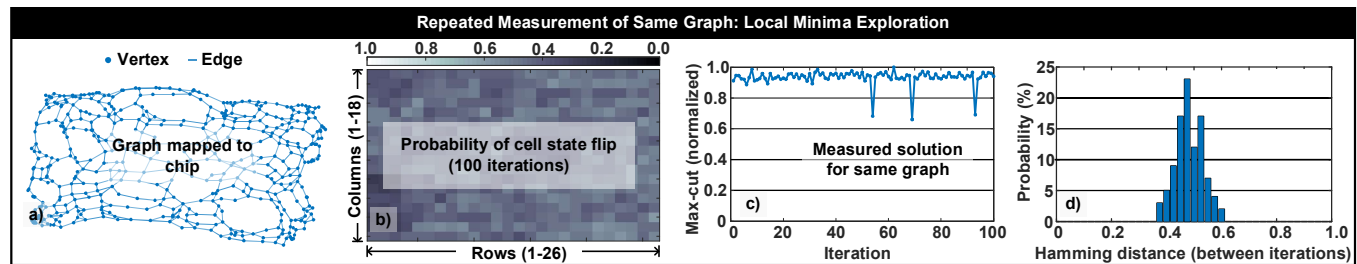


Fig. 5. a) A random graph is measured and reprogrammed 100 times. b) Unit cell state may flip between iterations, while the results remain similar as shown in c). d) Hamming distance of the solutions are different proving the chip is randomly exploring different minima.

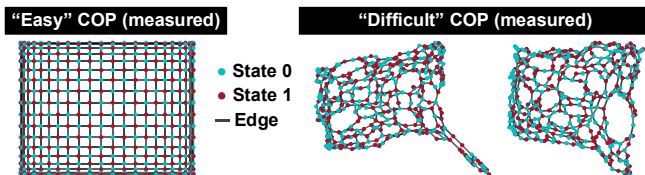


Fig. 6. Example of various graphs mapped and solved using chip.

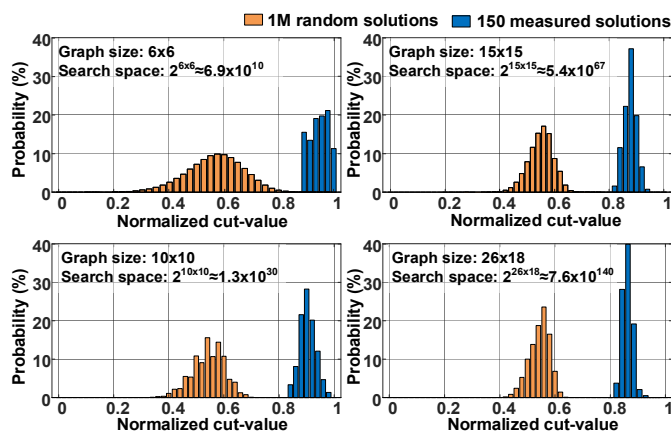
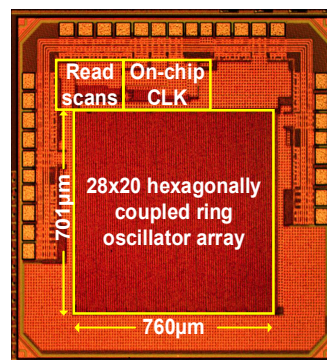


Fig. 7. Normalized max-cut distributions for 150 "Difficult" COPs. Measured results are compared with 1 million randomly sampled solutions from whole solution-space for each specific graph.

	This work	PRL '19 [5]	IEDM '19 [6]	D-Wave
Architecture	Ring oscillators	Optical LASER	IMT Oscillators	Qubits
Requires special process	No	Yes	Yes	Yes
Annealing method	Oscillator interaction	Spatial light modulation	IMT stochasticity	Quantum interaction
Technology	65nm CMOS	Photonics	IMT device	Superconductor
Operating temperature	Room temp.	Room temp.	Room temp.	-273.14°C
Total measured solutions	1000	16	1	N/A
Peak power	23mW	Not reported	Not reported	25KW
Delay	1μs - 10μs	1000 iterations	1ms (4 spins)	N/A
Measured accuracy	Easy: 98-100%* Difficult: 82-100%*	>95% (87% cases)	97.6% (4 spins)	N/A

*Compared with commercial optimization solver solution

Fig. 8. Comparison with prior works with true coupling dynamics.



Application	Combinatorial optimization problems
Process	65nm LP CMOS
Architecture	ROSC, latch based coupling, self-annealing
Voltage	1.0V
Area	Chip: 1.44mm ²
	Core: 0.53mm ²
	Unit cell: 0.00095mm ²
Peak power	23mW
Power per cell	41μW

Fig. 9. Die photo and summary of the 65nm test chip.