Ferroelectric-based Accelerators for Computationally Hard Problems

Mohammad Khairul Bashar
Electrical & Computer Engineering
University of Virginia
Charlottesville, USA
mb9mz@virginia.edu

Jaykumar Vaidya
Electrical & Computer Engineering
University of Virginia
Charlottesville, USA
jv9yy@virginia.edu

R S Surya Kanthi
Electrical & Computer Engineering
University of Virginia
Charlottesville, USA
sr2nk@virginia.edu

Chonghan Lee
Computer Science & Engineering
Pennsylvania State University
State College, USA
cvl5361@psu.edu

Feng Shi
Computer Science Department
University of California-Los Angeles,
Los Angeles CA, USA
shifeng@ucla.edu

Vijaykrishnan Narayanan
Computer Science & Engineering
Pennsylvania State University
State College, USA
vxn9@psu.edu

Nikhil Shukla
Electrical & Computer Engineering
University of Virginia
Charlottesville, USA
ns6pf@virginia.edu

ABSTRACT

Solving hard combinatorial optimization problems such as graph coloring efficiently continues to be an outstanding challenge for computing. Traditional digital computers typically entail an exponential increase in computing resources as the problem sizes increase. This makes larger problems of practical relevance intractable to compute, with subsequently adverse implications for a broad spectrum of ever-more relevant practical applications ranging from machine learning to electronic device automation (EDA). Here, we examine how analog coupled oscillators can enable area and energy-efficient methods to accelerate such problems. Further, we discuss how the implementation of such non-Boolean platforms can take advantage of emerging technologies such as scalable ferroelectrics.

CCS CONCEPTS

• Emerging Technologies • Models of Computation

KEYWORDS

Oscillators; synchronization; ferroelectrics

ACM ISBN 978-1-4503-8393-6/21/06...$15.00.
https://doi.org/10.1145/3453688.3461745

ACM Reference format:

M. K. Bashar, J. Vaidya, R. S. Surya, C. Lee, F. Shi, V. Narayanan, & N. Shukla. 2021. Ferroelectric-based Accelerators for Computationally Hard Problems. In *Proceedings of 2020 ACM* Great Lakes Symposium on VLSI 2021 *(GLVLSI'21), June 22-25, Virtual Event, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3453688.3461745

1 Introduction

Digital computers have been the backbone of modern computing, and have over the past five decades, enabled unprecedented performance improvement that has powered the information revolution. Problems are solved by creating their Boolean (1/0) abstractions which are mapped onto the underlying hardware platform, conventionally realized using CMOS-based digital switches. Hardware scaling, empowered by Moore's law, has enabled a doubling of the transistor (CMOS) density approximately every two years. Furthermore, the computational efficiency, has doubled every 1.57 years, as described by Koomey's law [1]. However, even with these tremendous strides in digital computation, there is still a large class of problems (known as NP-Hard) that are considered intractable to solve using digital machines [2]. Many combinatorial optimization problems (e.g., maximum cut (MaxCut), Boolean satisfiability, vertex cover) belong to this category, and typically require computing resources that scale exponentially with the input size of the problem [3].

We consider the Maximum Cut (MaxCut) problem as an illustrative example. Computing the MaxCut of a graph G(V,E) (V: vertices; E: edges) (unweighted graphs considered here i.e. edges only have a binary (1/0) value) entails finding a cut which divides G into two sets (S1 & S2) such that the number of common edges between them is as large as possible; the number of common edges is the MaxCut solution (Fig. 1a). Fig. 1b demonstrates the measured exponential increase in time to compute the optimal MaxCut in graphs of increasing size using an SDP (semi-definite programming)-based Branch & Bound digital algorithm (BiqMac) [4], [5].

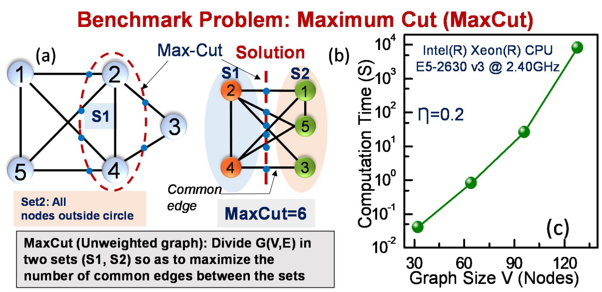

Figure 1: (a) The MaxCut problem and its solution illustrated using an example graph; (b) Time required to compute the optimal MaxCut as a function of graph size (V) using a Branch & Bound digital algorithm; an exponential increase is observed. The graphs are randomly generated instances having an edge density (η)=0.2

While the above algorithm calculates an optimal solution, even heuristic algorithms, which may be faster but do not guarantee optimal solutions, struggle when solving large problem instances. Such combinatorial optimization problems have important applications across a wide spectrum of areas ranging from VLSI design (interconnect routing) [6], [7], bioinformatics [8], medical imaging [9], protein folding [10], probabilistic reasoning [11], etc. Moreover, in certain applications such as deploying swarm intelligence [12], [13], autonomous driving [14], etc., such problems have to be solved not only in (close to) real-time but also in energy-constrained environments making energy efficiency of the computational platform equally important. Consequently, this has motivated the exploration of alternate computing paradigms to solve such problems.

Broadly, these alternatives can be classified as: (a) Quantum computing-based approaches [15]-[17]: the objective here is to overcome the fundamental hardness of the problem and provide an exponential speed up. This approach inherently relies on the idea that the speed up achieved with qubits will outweigh the additional energy (specifically, cryogenic cooling [18]) and space (area) overhead associated with quantum computers making it viable for at least for some applications. The alternate approach: (b) non-Von Neumann but classical systems, aim to use alternate (classical) computing models and hardware that may be more suitable for solving such problems. The objective here is that while the approach may not fundamentally overcome the theoretical hardness of the problem, it will still enable significantly better performance along with higher energy efficiency during runtime for a large number of problems. In fact, the energy efficiency (i.e., energy/compute) may even be better than that provided by quantum methods since most of the hardware implementations do not need cryogenic cooling. This may be particularly critical for applications that require performing such computation in energy-constrained environments, as discussed above.

One such non-Von Neumann approach is based on using dynamical systems [19], [20]. Since combinatorial optimization problems such as solving the MaxCut entail the maximization/minimization of an objective function in the combinatorial domain, the underlying concept here is to map the problem solution to the energy function of an appropriately designed dynamical system. Consequently, as the system evolves to minimize its energy and reach its ground state (≡ optimal solution), it will naturally compute the solution to the problem. The manifestation and effectiveness of this approach is evident in the physical world; real-time information processing is observed in natural dynamical systems such as the activation patterns of neural circuits, cellular signaling mechanisms, and information flow in social networks. Analytical solutions of such systems are rare and numerical simulations can be extremely computationally intensive. Yet the physical system "computes" its own dynamics in seconds.

One example of such a dynamical system is coupled oscillators- the focus of the present work. Coupled oscillators [21], in principle, represent a promising approach since oscillators can be made low-power, compact, room-temperature operation compatible, making them particularly attractive for applications such as special-purpose accelerators in heterogeneous computing platforms.

2 Oscillator-based computational models

The idea of analog computation with oscillators is not new. However, the field has been receiving increasing attention due to the advent of new hardware technologies (e.g., STT MRAM oscillators, phase change systems) that promise highly compact implementations, coupled with the enormous advances in process technology. Oscillator-based computational models have been proposed for a range of applications such as image processing (image segmentation- LEGION networks [22]-[25]; pattern recognition using time-delay encoding [26], frequency shift keying (FSK)) [27], [28], associative memory [29], oscillator-based neural networks [30], and even speech recognition. Romera *et al.* [31] demonstrated a 4-coupled spin-torque oscillator platform capable of classifying a pair of vowels from a speech sample. Furthermore, their application in combinatorial optimization has also received significant attention recently even though there were some early reports on their application in this area [32].

Several oscillator-based models have been proposed for solving combinatorial optimization problems. Wang *et. al.* [33] recently demonstrated that coupled oscillators, under second harmonic injection locking, can be used to minimize the Ising
njection locking, can be used to minimize the Ising
injection locking, can be used to minimize the Ising
ection locking, can be used to minimize the Ising Hamiltonian, $H =$

of the graph), under second harmonic injection, can be used to compute the MaxCut of a graph. The oscillators exhibit a phase bi-partition (0 & 180^{o}) that corresponds to the two sets created by the (Max-) Cut. Subsequently, this oscillator property has been demonstrated using various types of oscillators including emerging technologies such as insulator-metal transition oscillators [34], spin-torque oscillators [35], as well as using larger integrated systems [36], [37]. Figure 2a shows a representative graph and its equivalent oscillator network. Fig. 2b shows the corresponding oscillator outputs under the influence of a second harmonic signal. The resulting oscillator phases, shown in the phase plot in Fig. 2c, exhibit a phase bipartition (corresponding to S1 & S2) that can be used to compute the MaxCut of the graph.

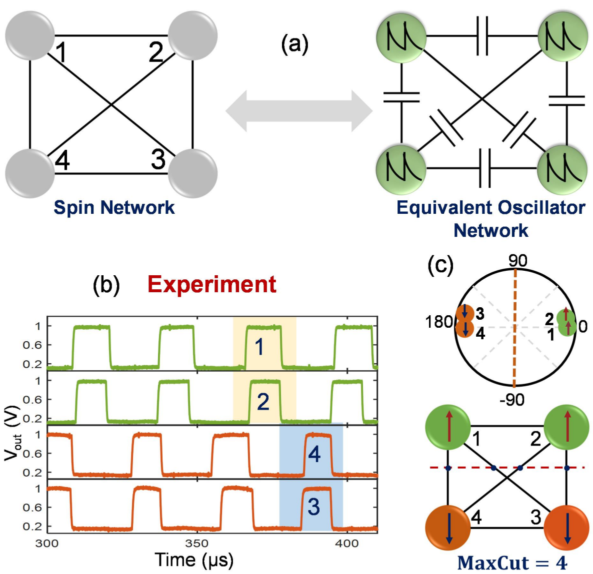

Figure 2: (a) A representative graph along with the corresponding topologically equivalent coupled oscillator network. (b) Experimentally measured time-domain output of the oscillators in the network under second harmonic injection. (c) Corresponding phase plot of the oscillators showing a phase bipartition with each set corresponding to a set created by the (Max-)Cut. An optimal solution (=4) is observed in this case.

Parihar *et al.* [38], also demonstrated the application of coupled relaxation oscillators in solving graph coloring– a prototypical NP-hard combinatorial optimization problem that entails computing the minimum number of colors (to be assigned to the nodes) such that no two nodes having a common edge are assigned the same color. They demonstrated that a topologically equivalent network of coupled oscillators (without any external injection signals) exhibits a unique phase ordering that can be used (with minimal post-processing) to solve the graph coloring problem [38], [39]. The underlying framework for this work lies in the equivalence between the eigenvalues of the coupled oscillator system in state space and those of the adjacency matrix of the graph being solved. Recently, the authors demonstrated this behavior in an integrated circuit of 30 oscillators with all-to-all reconfigurable coupling [40].

The application of this platform in solving other NP-complete problems such as computing the maximum independent set of a graph was shown [41], [42].

3 Leveraging emerging technologies

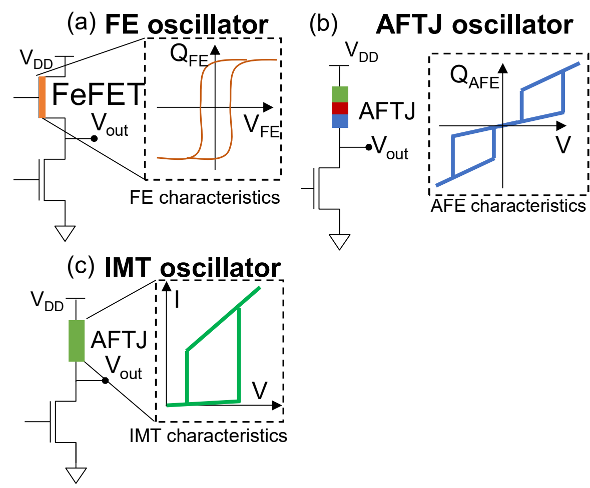

Figure 3: Illustrative examples of emerging technology-based oscillator designs. (a) Ferroelectric (FET)-based oscillator, proposed by [43]; (b) Anti-ferroelectric tunnel junction-based oscillator proposed by [44]; (c) insulator-metal transition-based oscillator

While the oscillator-based computational models promise novel and potentially more efficient algorithmic pathways to solve combinatorial optimization, the ultimate efficiency of such an approach will also depend on the area-, energy- efficiency and performance of the underlying hardware implementation.

This inspires the exploration of emerging hardware technologies that may be more amenable to such implementations than conventional CMOS-based designs which are primarily optimized as digital switches. Beyond silicon materials systems that naturally exhibit characteristics such as non-linear charge vs. voltage properties (ferroelectrics), non-linear current vs. voltage characteristics (spin transfer torque devices, insulator-metal transition materials such as VO_2, NbO_2), hysteresis - typically absent in conventional silicon, can offer more energy and area efficient pathways to implement the required hardware primitives. For instance, in contrast to the (minimum) 6 transistors (6T) required for implementing a ring oscillator in CMOS technology, many of the above technologies facilitate a significantly compact 1T1R implementation; examples include, VO_2 & NbO_2-based insulator-metal transition oscillators, spin-transfer torque (STT) oscillators, ferroelectric and anti-ferroelectric oscillators.

In particular, with the recent discovery of a ferroelectric phase in doped-HfO_2 (e.g., $Hf_xZr_{1-x}O_2$; *f*-HfO_2), there has been extensive interest in exploring this material system for such applications. Besides the inherent compatibility with CMOS process technology, *f*-HfO_2 is highly scalable, making it suitable for scaled technology nodes. 1nm thick films have been experimentally demonstrated [46] in alignment with theoretical predictions [47] and there is no known limit to lateral scaling. Furthermore, its response can not only be modulated between that of a FE and AFE, but its polarization can also be varied over a wide range using techniques (e.g., composition & doping) that are largely compatible with existing process technology. Besides the above materials-centric advantages provided by *f*-HfO_2, ferroelectric devices such as FeFETs and FTJs can extremely exhibit energy efficient [48], fast, and non-filamentary (unlike VO_2, NbO_2) switching that can be tuned between volatile, non-volatile, and oscillatory operation (Fig. 3). Oscillator designs based on ferroelectrics [43], and anti-ferroelectric tunnel junctions [44] have been recently shown. Additionally, these emerging technologies can potentially provide efficient routes to implement the interaction/coupling among the oscillators. For instance, the cross-point memory architecture implemented using resistive [49] and ferroelectric memory [50], [51] provides a natural substrate for mapping the adjacency matrix of a graph.

In summary, computational models & systems based on coupled oscillators, co-designed and optimized with emerging technologies such as ferroelectrics, can provide energy, area, and performance efficient pathways to accelerate computationally hard problems.

ACKNOWLEDGMENTS

This work was supported in part by Semiconductor Research Corporation (SRC) under Grant 2841.001, in part by the National Science Foundation (NSF) under Grant ECCS-1807551, NSF-2008365, NSF - 1914730.

REFERENCES

[1] J. Koomey *et al.* 2010. Implications of historical trends in the electrical efficiency of computing. *IEEE Annals of the History of Computing*, 33.3, 46-54.
[2] C. S. Calude. 2017. Unconventional computing: A brief subjective history. In *Advances in Unconventional Computing*, Springer, Cham, 855-864.
[3] F. Neumann, and C. Witt. 2010. Combinatorial optimization and computational complexity. *Bioinspired computation in combinatorial optimization*, Springer, Berlin, Heidelberg, 9-19.
[4] F. Rendl *et al.* 2010. Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations. *Mathematical Programming*, 121.2, 307-335.
[5] Biq Mac Solver - Binary quadratic and Max cut Solver. Retrieved May 1, 2021. http://biqmac.uni-klu.ac.at/
[6] F. Liers *et al.* 2011. Via Minimization in VLSI Chip Design Application of a Planar Max-Cut Algorith. http://e-archive.informatik.unikoeln.de/630/
[7] S. Held *et al.* 2011. Combinatorial Optimization in VLSI Design. *Combinatorial Optimization-Methods and Applications*, 31, 33-96.
[8] H. J. Greenberg *et al.* 2004. Opportunities for combinatorial optimization in computational biology. *INFORMS Journal on Computing*, 16.3, 211-231.
[9] P. L. Hammer, and P. O. Bonates. 2006. Logical analysis of data—An overview: From combinatorial optimization to medical applications. *Annals of Operations Research*, 148.1, 203-225.
[10] A. S. Fraenkel. 1993. Complexity of protein folding. *Bulletin of Mathematical Biology*, 55.6, 1199-1210.
[11] G. F. Cooper. 1990. The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, 42.2-3, 393-405.
[12] T. Krausburg. 2019. Constrained Coalition Formation among Heterogeneous Agents for the Multi-Agent Programming Contest. In *ICAART (1)*.
[13] E. Demirovic *et al.* 2018. Recoverable team formation: Building teams resilient to change. Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems.
[14] W. Xiao *et al.* 2021. Rule-based optimal control for autonomous driving. https://arxiv.org/abs/2101.05709
[15] A. Steane. 1998. Quantum computing. *Reports on Progress in Physics*, 61.2, 117.
[16] A. B. Finnila *et al.* 1994. Quantum annealing: A new method for minimizing multidimensional functions. *Chemical Physics Letters*, 219.5-6, 343-348.
[17] S. Boixo *et al.* 2013. Experimental signature of programmable quantum annealing. *Nature Communications*, l 4.1, 1-8.
[18] T. Albash *et al.* 2015. Consistency tests of classical and quantum models for a quantum annealer. *Physical Review A*, 91.4, 042314.
[19] J. J. Hopfield. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79, 2554-2558.
[20] H. T. Siegelmann, and S. Fishman. 1998. Analog computation with dynamical systems. *Physica D: Nonlinear Phenomena*, 120, 214-235.
[21] G. Csaba, and W. Porod. 2020. Coupled oscillators for computing: A review and perspective. *Applied Physics Reviews*, 7, 011302.
[22] D. Wang, and D. Terman. 1995. Locally excitatory globally inhibitory oscillator networks. *IEEE Transactions on Neural Networks*, 6.1, 283-286.
[23] N. Shareef *et al.* 1999. Segmentation of medical images using LEGION. *IEEE Transactions on Medical Imaging*, 18.1, 74-91.
[24] J. Yuan *et al.* 2011. LEGION-based automatic road extraction from satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 49.11, 4528-4538.
[25] Y. Fang *et al.* 2014. Image segmentation using frequency locking of coupled oscillators. In *14th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA)*, IEEE.
[26] E. Corti *et al.* 2020. Time-Delay Encoded Image Recognition in a Network of Resistively Coupled VO_2 on Si Oscillators. *IEEE Electron Device Letters*, 41, 629-632.
[27] D. E. Nikonov *et al.* 2015. Coupled-oscillator associative memory array operation for pattern recognition. *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, 1, 85-93.
[28] D. E. Nikonov et al. 2019. A Coupled CMOS Oscillator Array for 8ns and 55pJ Inference in Convolutional Neural Networks. https://arxiv.org/abs/1910.11803
[29] N. Shukla *et al.* 2014. Pairwise coupled hybrid vanadium dioxide-MOSFET (HVFET) oscillators for non-boolean associative computing. *In 2014 IEEE International Electron Devices Meeting*, IEEE.
[30] H. D. Sompolinsky. 1990. Global processing of visual stimuli in a neural network of coupled oscillators. *Proceedings of the National Academy of Sciences*, 87.18, 7200-7204.
[31] M. Romera *et al.* 2018. Vowel recognition with four coupled spin-torque nano-oscillators. *Nature*, 563.7730, 230-234.
[32] C. W. Wu. 1998. Graph coloring via synchronization of coupled oscillators. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 45.9, 974-978.
[33] T. Wang, and J. Roychowdhury. 2019. OIM: Oscillator-based Ising machines for solving combinatorial optimisation problems. In *International Conference on Unconventional Computation and Natural Computation*, 232-256, Springer, Cham.
[34] S. Dutta *et al.* 2019. Experimental demonstration of phase transition nano-oscillator based ising machine. In *IEEE International Electron Devices Meeting (IEDM)*, IEEE.
[35] D. I. Albertsson *et al.* 2021. Ultrafast Ising Machines using spin torque nano-oscillators. *Applied Physics Letters*, 118.11, 112404.
[36] I. Ahmed *et al.* 2020. A Probabilistic Self-Annealing Compute Fabric Based on 560 Hexagonally Coupled Ring Oscillators for Solving Combinatorial Optimization Problems. In *2020 IEEE Symposium on VLSI Circuits*, IEEE.
[37] M. K. Bashar *et al.* 2020. Experimental Demonstration of a Reconfigurable Coupled Oscillator Platform to Solve the Max-Cut Problem. In *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, 6, 116-121.
[38] A. Parihar *et al.* 2017. Vertex coloring of graphs via phase dynamics of coupled oscillatory networks. *Scientific Reports*, 7, 1-11.
[39] X. Ren *et al.* 2021. Cardiac Muscle Cell-Based Coupled Oscillator Network for Collective Computing. *Advanced Intelligent Systems*, 3, 2000253.
[40] A. Mallick *et al.* 2021. Graph Coloring using Coupled Oscillator-based Dynamical Systems. In *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1-5.
[41] A. Mallick *et al.* 2020. Using synchronized oscillators to compute the maximum independent set. *Nature Communications*, 11, 1-7.
[42] M. K. Bashar *et al.* 2019. Solving the Maximum Independent Set Problem using Coupled Relaxation Oscillators. In *2019 Device Research Conference (DRC)*, 187-188.
[43] Z. Wang *et al.* 2017. Ferroelectric oscillators and their coupled networks. *IEEE Electron Device Letters*, 38.11, 1614-1617.
[44] R. Andrawis and K. Roy. 2020. Antiferroelectric Tunnel Junctions as Energy-Efficient Coupled Oscillators: Modeling, Analysis, and Application to Solving Combinatorial Optimization Problems. *IEEE Transactions on Electron Devices*, 67, 2974-2980.
[45] N. Shukla *et al.* 2014. Synchronized charge oscillations in correlated electron systems. *Scientific Reports*, 4.1, 1-6.

[46] S. S. Cheema *et al.* 2020. Enhanced ferroelectricity in ultrathin films grown directly on silicon. *Nature*, 580.7804, 478-482.
[47] H. J. Lee *et al.* 2020. Scale-free ferroelectricity induced by flat phonon bands in HfO2. *Science*, 369.6509, 1343-1347.
[48] A. Keshavarzi *et al.* 2020. FerroElectronics for Edge Intelligence. *IEEE Micro*, 40.6, 33-48.
[49] S. H. Jo *et al.* 2014. 3D-stackable crossbar resistive memory based on field assisted superlinear threshold (FAST) selector. In *2014 IEEE international electron devices meetin*g, IEEE.
[50] W. Chen. 2019. Selector-free cross-point memory architecture based on ferroelectric MFM capacitors. In *2019 IEEE 11th International Memory Workshop (IMW)*, IEEE.
[51] H. S. Lee, *et al.* 2015. Ferroelectric tunnel junction for dense cross-point arrays. ACS applied materials & interfaces, 7.40, 22348-22354.