Digital Object Identifier 10.1109/JXCDC.2023.3280846

Special Topic on Nontraditional Devices, Circuits, and Architectures for Energy-Efficient Computing

RECENTLY, novel applications in the space of artificial intelligence (AI) such as solving constraint optimization problems, probabilistic inferencing, contextual adaptation, and continual learning from noisy data are gaining momentum to address relevant real-world problems. A majority of these tasks are compute and/or memory intensive. While traditional deep learning has been fueled by the utilization of graphic processing units (GPUs) to accelerate algorithms primarily in the cloud, today we see a surge in the development of application/domain-specific integrated circuits and systems that aim at providing an order of magnitude improvement over traditional GPU-based approaches in terms of energy efficiency and latency. This growing branch of research taps into the realms of neuronal dynamics, collective computing using dynamical systems, harnessing stochasticity to enable probabilistic computing, and even draws inspiration from quantum computing. We envision such specialized application/domain-specific systems to perform complex tasks such as solving NP-hard optimization problems, performing reasoning and cognition in the presence of uncertainty with superior energy-efficiency (and/or area and latency improvements) compared to conventional GPU-based approaches and von Neumann computing using traditional silicon-based devices, circuits, and architectures. Of special interest is to utilize such nontraditional computing approaches to reduce the time to obtain solutions for computationally challenging problems that otherwise tend to grow exponentially with problem size. To support this vision, there needs to be fundamental advances in both nontraditional devices and circuits/architectures. Recent works have shown that novel circuit topologies and architectures involving non-Boolean, oscillatory, spiking, probabilistic, or quantum-inspired computing are more suited toward tackling applications such as solving constraint optimization problems, performing energy-based learning, performing Bayesian learning and inference, lifelong continual learning, and solving quantum-inspired applications such as Quantum Monte Carlo. A flurry of current research highlights that compared to traditional silicon-based devices, emerging nanodevices utilizing novel quantum materials such as complex oxides, ferroelectric materials, and spintronic materials can allow the realization of these novel circuits and architectures with lower foot-print area, higher energy efficiency, and lower latency.

The call for this Special Topic encouraged authors to report the recent research advances in the field of nontraditional devices, circuits, and architectures for energy-efficient computing. The call encouraged interdisciplinary research on the interaction and co-optimization of materials and devices as well as circuits and architecture. The possible topics of interest put forward were broad, encompassing topics such as utilizing emerging nanodevices involving novel quantum materials, such as phase-transition oxides, complex oxides, ferroelectric materials, magnetic/spintronic materials, and photonics, to perform novel non-Boolean computing, oscillatory or spiking neural networks, probabilistic computing, and quantum-inspired computing.

The following topics were specifically solicited:

- materials and devices that can enable dynamical systems, e.g., oscillatory Ising machines and probabilistic computing;
- materials and devices exhibiting neuronal dynamics enabling Bayesian and continual learning;
- materials and devices supporting quantum-inspired computing;
- integration of emerging technologies with silicon for building energy-efficiency systems;
- array-level demonstration and/or architecture-level design for nontraditional energy-efficient computing;
- co-optimization of hardware and algorithms for nontraditional energy-efficient computing;
- benchmarking simulators for nontraditional energy-efficient computing;
- new applications for nontraditional energy-efficient computing involving non-Boolean circuits, oscillatory or spiking neural networks, and probabilistic and quantum-inspired systems.

Here are the nine articles that have been accepted after the peer-review process.

1) Chowdhury et al. [A1] discuss a novel alternative to scaling transistors for general-purpose computing to augment transistor integration with emerging technologies, especially for domain-specific computing. In this article that spans over hardware, architecture, and algorithmic perspective, the authors describe a full-stack review of probabilistic computing with probabilistic-bit (p-bit) as a representative example of energy-efficient and domain-specific computing. The authors demonstrate how emerging nanoelectronics devices can be combined with existing CMOS ecosystems to build p-bits that can in turn be utilized for building energy-efficient probabilistic systems, tailored for probabilistic algorithms and applications.

2) Dynamical systems using coupled oscillator networks serve as a novel pathway toward performing non-Boolean computing. Bashar et al. [A2] discuss how the inherent energy minimization of such a system can be utilized for solving combinatorial optimization problems. The authors showcase two oscillator-inspired dynamical systems to solve a quintessential computationally intractable problem—Boolean satisfiability (SAT). The authors engineer the system dynamics such that they facilitate solutions to two different flavors of the SAT problem. The authors show techniques to formulate the dynamical system tailored for computing a 3-SAT problem and then demonstrate how the dynamics map to the solution of the Max-NAE-3-SAT problem.
3) In another article, Bashar et al. [A3] emphasize the need to construct dynamical systems that can go beyond the traditional quadratic degree of interaction and can be used for minimizing objective functions with higher degrees. The authors develop dynamical system-inspired computational models that can be used to define ''energy functions'' for hyper-graph-based combinatorial optimization problems ranging from Boolean SAT to integer factorization problems.
4) While stochastic computing shows promise for performing complex functions with large amounts of data, the underlying hardware needed to generate random bit-streams using conventional CMOS is area and delay expensive. Zink et al. [A4] have proposed a novel approach for embedding stochastic bit generation and processing in a computational random access memory. The authors show that such a computational memory is resilient and low cost and has wide applications for image processing, Bayesian inferencing, and Bayesian belief networks.
5) Xie et al. [A5] explore the utilization of resistive random access memory (RRAM) for compute-in-memory (CIM) applications. The authors propose an analog-oriented RRAM array/read circuits co-design technique to solve the tradeoffs between distortion, power consumption, throughput, and chip area. While previous works on RRAM-based CIM utilized access transistors as a switch, here the authors investigate a new perspective that utilizes access transistors as common-gate current buffers which reduce the operating current and amplify the output impedance. Furthermore, the authors also introduce the idea of in-ADC computing (IAC) to reduce peripheral circuit complexity.
6) In-memory computing (IMC) shows exceptional throughput and energy efficiency for performing numerous tasks pertaining to modern-day machine learning applications. Usually, IMC is used by coupling crosspoint memory devices in open-loop matrix-vector-multiplication and closed-loop inverse-matrix-multiplication (IMVM). However, with each application demanding a different circuit topology, building a reconfigurable and general-purpose IMC system becomes a problem. Mannocci and Ielmini [A6] have proposed a generalized closed-loop IMVM circuit that is capable of performing any linear matrix operation by appropriate memory remapping. The proposed circuit represents an ideal candidate for general-purpose accelerators of machine learning.
7) Jacob et al. [A7] propose a novel nonvolatile compute-in-memory macro using voltage-controlled magnetic tunnel junction (VC-MTJ) memory and an in situ magnetic-to-digital conversion (MDC). The authors show that VC-MTJ provides almost 10× lower write energy and switching time compared to conventional STT-MRAM. The in-situ MDC embedded inside each VC-MRAM row converts magnetically stored weight information to CMOS logic levels and enables switched-capacitor-based multiply-accumulate (MAC) operation. The authors show that the proposed approach demonstrates accuracy comparable to state-of-the-art SRAM-CIM while providing 1.5× higher energy efficiency and 2× higher density.
8) Zogbi et al. [A8] propose a novel magnetic domain wall-based logic, by combining domain wall-magnetic tunnel junction (DW-MTJ) and voltage-controlled magnetic anisotropy (VCMA) effect to improve the reliability of logic concatenation. The authors simulate a systolic array of DW-MTJ for multiply-and-accumulate (MAC) operation which shows comparable throughput and efficiency to state-of-the-art CMOS, while being radiation hard. This has implications for designing DW-based logic for harsh environment electronics.
9) Toward addressing the call for efficient computing, Kumar et al. [A9] propose a novel design of dynamic logic circuits using a fully depleted silicon-on-insulator (FD-SOI) technology. Using calibrated models, the authors show improvements in transistor count, propagation delay, power, and power-delay product compared to the conventional designs, with reduced impact from variation and charge sharing effect.

Sourav Dutta
Components Research
Intel Corporation
Hillsboro, OR 97124 USA
e-mail: sourav.dutta@intel.com

Punyashloka Debashis
Components Research
Intel Corporation
Hillsboro, OR 97124 USA
e-mail: punyashloka.debashis@intel.com

AMIR KHOSROWSHAHI
Components Research
Intel Corporation
Hillsboro, OR 97124 USA
e-mail: amir.khosrowshahi@intel.com

APPENDIX: RELATED ARTICLES

[A1] S. Chowdhury et al., "A full-stack view of probabilistic computing with p-bits: Devices, architectures, and algorithms," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 1–11, Jun. 2023, doi: 10.1109/JXCDC.2023.3256981.

[A2] M. K. Bashar, Z. Lin, and N. Shukla, "Oscillator-inspired dynamical systems to solve Boolean satisfiability," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 12–20, Jun. 2023, doi: 10.1109/JXCDC.2023.3241045.

[A3] M. K. Bashar, A. Mallick, A. W. Ghosh, and N. Shukla, "Dynamical system-based computational models for solving combinatorial optimization on hypergraphs," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 21–28, Jun. 2023, doi: 10.1109/JXCDC.2023.3235113.

[A4] B. R. Zink et al., "A stochastic computing scheme of embedding random bit generation and processing in computational random access memory (SC-CRAM)," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 29–37, Jun. 2023, doi: 10.1109/JXCDC.2023.3266136.

[A5] T. Xie, S. Yu, and S. Li, "A high-parallelism RRAM-based compute-in-memory macro with intrinsic impedance boosting and in-ADC computing," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 38–46, Jun. 2023, doi: 10.1109/JXCDC.2023.3255788.

[A6] P. Mannocci and D. Ielmini, "A generalized block-matrix circuit for closed-loop analogue in-memory computing," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 47–55, Jun. 2023, doi: 10.1109/JXCDC.2023.3265803.

[A7] V. K. Jacob, J. Yang, H. He, P. Gupta, K. L. Wang, and S. Pamarti, "A nonvolatile compute-in-memory macro using voltage-controlled MRAM and in-situ magnetic-to-digital converter," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 56–64, Jun. 2023, doi: 10.1109/JXCDC.2023.3258431.

[A8] N. Zogbi et al., "Parallel matrix multiplication using voltage-controlled magnetic anisotropy domain wall logic," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 65–73, Jun. 2023, doi: 10.1109/JXCDC.2023.3266441.

[A9] S. Kumar, S. Chatterjee, C. K. Dabhi, Y. S. Chauhan, and H. Amrouch, "Nontraditional design of dynamic logics using FDSOI for ultra-efficient computing," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 9, no. 1, pp. 74–82, Jun. 2023, doi: 10.1109/JXCDC.2023.3269141.