

Session 14 - Emerging Device and Compute Technology - Neuromorphic Session II-Architecture Focus
Tuesday, December 10, 9:00 a.m.

Continental Ballroom 5

Co-Chairs: S. Ambrogio, IBM Research

J. Deng, Qualcomm

9:05 AM 14.1 Programmable Linear RAM: A New Flash Memory-based Memristor for Artificial Synapses and Its Application to Speech Recognition System

Shifan Gao, Guangjun Yang, Xiang Qiu, Chun Yang, Cheng Zhang, Binhan Li, Chao Gao, Hong Jiang, Zhexiong Wang, Jian Hu, Jun Xiao, Bo Zhang, Choong-Hyun Lee, Yi Zhao, Weiran Kong, Zhejiang University, Shanghai Huahong Grace Semiconductor Manufacturing Corporation, Flash Billion Semiconductor Co. Ltd, Peking University

In this work, a new type of flash memory-based memristor, named programmable linear random-access memory (PLRAM), is presented to store analog synaptic weights in a single Flash memory cell. With a linearity-aware design, an application of speech recognition is presented, with recognition accuracy higher than 90%.

9:30 AM 14.2 On-Chip Trainable 1.4M 6T2R PCM Synaptic Array with 1.6K Stochastic LIF Neurons for Spiking RBM

Masatoshi Ishii, Sangbum Kim, Scott Lewis, Atsuya Okazaki, Junka Okazawa, Megumi Ito, Malte J. Rasch, Wanki Kim, Akiyo Nomura, Uicheol Shin, , Kohji Hosokawa, Matthew BrightSky, Wilfried Haensch, IBM Research, Seoul National University, IBM Research - Tokyo,

Scalable six-transistor/two analog-weighted PCM resistor (6T2R) cell array integrating with fully-parallelized asynchronous stochastic leaky integrated-and-fire (LIF) neuron circuit demonstrated ultra low-power (8.95 pJ per synaptic operation) on-chip training and inference capabilities of restricted Boltzmann machine (RBM) using MNIST hand-written digit database.

9:55 AM 14.3 Fully Integrated Spiking Neural Network with Analog Neurons and RRAM Synapses

Alexandre Valentian, François Rummens, Elisa Vianello, Thomas Mesquida, Capucine Lecat-Mathieu de Boissac, Olivier Bichler, Carlo Reita, CEA-Leti, CEA-List

This paper presents, to the best of the authors' knowledge, the first complete integration of a Spiking Neural Network, combining analog spiking neurons and RRAM-based synapses. The test chip, fabricated in 130nm CMOS, shows well-controlled integration of synaptic currents and a high RRAM endurance to inference tasks (750M spikes sent).

10:20 AM 14.4 A Deep Neural Network Accelerator Based on Tiled RRAM Architecture

Xinxin Wang, Qiwein Wang, Seung hwan Lee, Fan-Hsuan Meng, Wei D. Lu, University of Michigan

Deep neural networks have been successfully mapped on an RRAM-based tiled in-memory computing (IMC) architecture. Effects of array size and quantized partial products (PPs) due to ADC precision constraints were analyzed. Methods were developed to solve these challenges and preserve DNN accuracies and IMC performance gains in the tiled architecture.

11:10 AM 14.5 On Designing Efficient and Reliable Nonvolatile Memory-Based Computing-In-Memory Accelerators (Invited)

Bonan Yang, Mengyun Liu, Yiran Chen, Krishnendu Chakrabarty, and Hai Li, Duke University

11:35 AM 14.6 Bayesian Neural Network Realization by Exploiting Inherent Stochastic Characteristics of Analog RRAM

Yudeng Lin, Qingtian Zhang, Jianshi Tang, Bin Gao, Chongxuan Li, Peng Yao, Zhengwu Liu, Jun Zhu, Jiwu Lu, Xiaobo Hu, He Qian, Huaqiang Wu, Tsinghua University, Hunan University, University of Notre Dame

For the first time, this paper develops a novel stochastic computing method by utilizing the inherent random noises of analog RRAM. The RRAM device can realize the function of sampling from a tunable probabilistic distribution required by BayNN. And this is the first demonstration work for BayNN with emerging devices.

12:00 PM 14.7 An Analog Neuro-Optimizer with Adaptable Annealing Based on 64×64 0T1R Crossbar Circuit

Mohammad Reza Mahmoodi, Hyungjin Kim, Zahra Fahimi, Hussein Nili, Leo Sedov, Valentin Polishchuk, Dmitri Strukov, University of California, Santa Barbara, Linkoping University

We demonstrate an analog neuro-optimization hardware, which supports various annealing techniques, using a crossbar circuit with 4096 passively-integrated analog-grade memristors. The hardware operation is successfully tested by experimentally solving weighted graph partitioning, maximum clique, vertex cover, and independent set problems, and observing good agreement with simulation results.

12:25 PM 14.8 A High-Speed and High-Reliability TRNG Based on Analog RRAM for IoT Security Application

Bohan Lin, Bin Gao, Yachuan Pang, Peng Yao, Dong Wu, Hu He, Jianshi Tang, He Qian, Huaqiang Wu, Tsinghua University

A novel True Random Number Generator (TRNG) based on analog RRAM is developed. The highest single-cell throughput >1 Mbit/sec is achieved among RRAM-based TRNGs with minimal circuit overhead. The TRNG performance shows excellent temperature stability, and the RRAM endurance issue is greatly relieved to meet the requirement for IoT application.