



SRAM-based In-Memory Computing Hardware: Analog vs Digital and Macros to Microprocessors

Mingoo Seok

Associate Professor
Columbia University

BIO

Mingoo Seok is an associate professor of Electrical Engineering at Columbia University. He received his B.S. from Seoul National University, South Korea, in 2005 and his M.S. and Ph.D. degrees from the University of Michigan in 2007 and 2011, respectively, all in electrical engineering. His research interests are various aspects of VLSI circuits and architecture, including ultra-low-power integrated systems, cognitive and machine-learning computing, an adaptive technique for the process, voltage, temperature variations, transistor wear-out, integrated power management circuits, event-driven controls, and hybrid continuous and discrete computing. He won the 2015 NSF CAREER award and the 2019 Qualcomm Faculty Award. He is the technical program committee member for multiple conferences, including the IEEE International Solid-State Circuits Conference (ISSCC). In addition, He has been an IEEE SCS Distinguished Lecturer for Feb/2023-Feb/2025 and an associate editor for IEEE Transactions on Circuits and Systems Part I (TCAS-I) (2014-2016), IEEE Transactions on VLSI Systems (TVLSI) (2015-present), IEEE Solid-State Circuits Letter (SSCL) (2017-2022), and as a guest associate editor for IEEE Journal of Solid-State Circuits (JSSC) (2019).

ABSTRACT

In the last decade, SRAM-based in-memory computing (IMC) hardware has received significant research attention for its massive energy efficiency and performance boost. In this seminar, first, we will introduce two very recent macro prototypes that achieve state-of-the-art performance and energy efficiency yet leverage very different computing mechanisms. Specifically, one adopted analog-mixed-signal (AMC) computing mechanisms (capacitive coupling and charge sharing), whereas the other adopted a fully digital approach. After this macro-level introduction, we will present recent microprocessor prototypes employing IMC-based accelerators, which can perform on-chip inferences at high energy efficiency and low latency.

Friday, March 22, 2024 at 1:00 – 2:00 p.m.
Osborne Conference Room (ECSS 3.503)



Erik Jonsson School of Engineering
and Computer Science