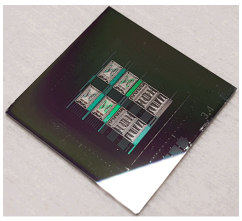

Bert Offrein, IBM Research

Photonic convolutional processors and optical crossbars for neural network training

Abstract: With the proliferation of ultrahigh-speed mobile networks and internet-connected devices, along with the rise of artificial intelligence (AI)1, the world is generating exponentially increasing amounts of data that need to be processed in a fast and efficient way. Highly parallelized, fast and scalable hardware is therefore becoming progressively more important. Here we demonstrate a computationally specific integrated photonic hardware accelerator (tensor core) that is capable of operating at speeds of trillions of multiply-accumulate operations per second 10^12 MAC operations per second or tera-MACs per second. The tensor core can be considered as the optical analogue of an application-specific integrated circuit (ASIC). It achieves parallelized photonic in-memory computing using phase-change-material memory arrays and photonic chip-based optical frequency combs (soliton microcombs). The computation is reduced to measuring the optical transmission of reconfigurable and non-resonant passive components and can operate at a bandwidth exceeding 14 gigahertz, limited only by the speed of the modulators and photodetectors. Given recent advances in hybrid integration of soliton microcombs at microwave line rates, ultralow-loss silicon nitride waveguides, and high-speed on-chip detectors and modulators, our approach provides a path towards full complementary metal–oxide–semiconductor (CMOS) wafer-scale integration of the photonic tensor core. Although we focus on convolutional processing, more generally our results indicate the potential of integrated photonics for parallel, fast, and efficient computational hardware in data-heavy AI applications such as autonomous driving, live video processing, and next-generation cloud computing services.

Speaker Biography: Bert Offrein received his Ph.D. degree in nonlinear integrated optics from the University of Twente (NL) in 1994. He then joined IBM Research - Zurich and contributed to establishing and commercializing adaptive integrated optical technology for DWDM networks. From 2004 to 2016, Bert Offrein was managing the photonics group, addressing optical interconnects for computing systems. Since 2016, he is leading the neuromorphic devices and systems group, focusing on novel hardware for neural networks. Bert Offrein is a principal research staff member at IBM Research and the co-author of over 150 publications and the co-inventor of more than 35 patents.