Hardware Architecture of an FPGA-Cluster for Predictive Variability Benchmarking of RRAM-IMC-Based Neuromorphic K-SAT Solvers

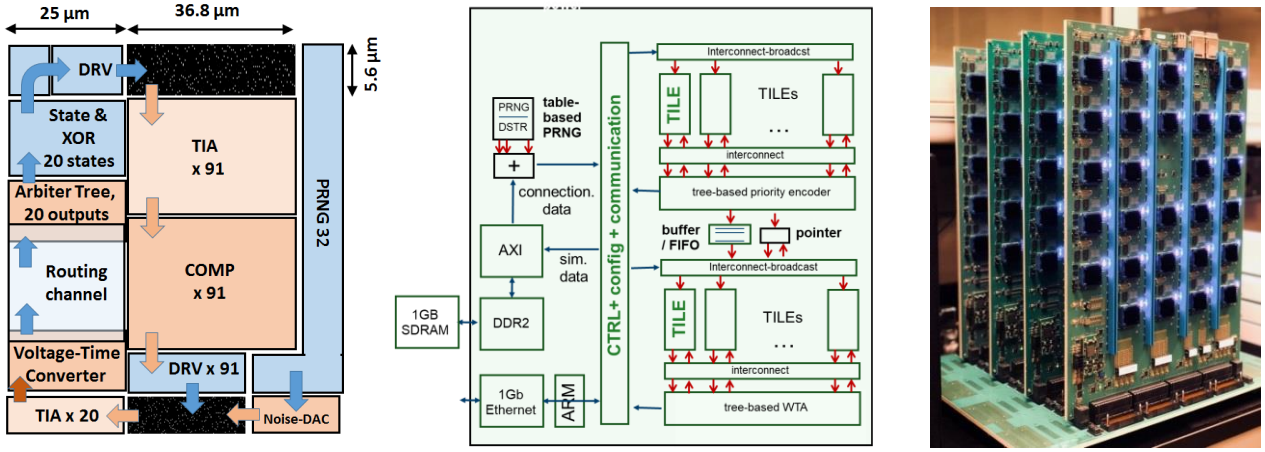
Arne Heittmann[(1)] and John Paul Strachan[(2)]

[(1)] PGI-15, Forschungszentrum Jülich, 52428 Jülich [(2)] PGI-14, Forschungszentrum Jülich, 52428 Jülich

The impact of variability on the performance of an RRAM-Computing-in-Memory-based architecture solving K-SAT unconstrained optimization problems is evaluated. Realistic CMOS variability parameters were extracted from a 28nm CMOS technology, while RRAM variability parameters were taken from measured devices reported in the literature. Importantly, whereas limited variability analysis typically is an essential part of the circuit design phase, the proposed modeling is going beyond low-level SPICE simulations and aims for scalable lightweight models which can be used for large-scale predictive benchmarking purposes.

The ultimate goal of this work is to capture the effects of statistical and transient component variability on the solution convergence of the target hardware [1], and to derive a specification for circuit design and algorithm definition regarding the RRAM programming. The evaluation of the effects of variability on the solver performance is realized by Monte Carlo sampling. For simulative acceleration, a dedicated multi-node FPGA-based digital compute-architecture (cf. Fig.1b,c) was elaborated, implemented, and used, which map the individual algorithmic steps of the K-SAT optimization to runtime-optimized and configurable hardware operations - and, in the case of digital peripheral circuits, a bit-exact behavior compared to the targeted dedicated hybrid RRAM/CMOS hardware.

It is shown by extensive Monte Carlo simulations that a naïve RRAM-based in-memory-compute (IMC) architecture (cf. Fig.1a) shows two different error characteristics with respect to (I) produce faulty solutions and (II) tendency for premature termination of iteration loops. A new variability compensation scheme is proposed, elaborated and evaluated for 3-SAT problems up to 200 variables and 860 clauses. It is shown that by application of the proposed compensation scheme errors of type (I) and (II) are completely eliminated.



***a. K*-Local Interactions In-Memory Accelerator (KLIMA)** **b. Prototype FPGA Architecture** **c. FPGA Cluster**

Figure 1: a. Floor plan of a hybrid RRAM-IMC solver, b. FPGA architecture, c. FPGA Cluster

[1] M. Hizzani *et al*., "Memristor-based hardware and algorithms for higher-order Hopfield optimization solver outperforming quadratic Ising machines," *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*, Singapore, Singapore, 2024, pp. 1-5, doi: 10.1109/ISCAS58744.2024.10558658.