Physical AI with analog memristors for Edge Computing

Heidemarie Krüger[1,2,3]

[1] Leibniz-IPHT e.V., Albert-Einstein Str. 9, 07745 Jena, Germany

[2]FSU Jena, Helmholtzweg 3, 07743 Jena, Germany

[3]TECHiFAB GmbH, Bautzner Landstr. 45, 01454 Radeberg, Germany

heidemarie.krueger@leibniz-ipht.de

The AI accelerators for edge sensing and edge computing currently available on the market are based on the classic von Neumann architecture, process digital data, consume more than 90% of their energy for data exchange between the processor and memory unit, and are not real-time capable. In comparison, AI accelerators with the memory unit located very close to the processor exhibit 30% reduced energy consumption and lower latency.

However, addressing the problem of the exponential global increase in data volumes and the associated computing power requirements, as well as the increasing complexity of AI training algorithms, necessitates further fundamental breakthroughs in AI accelerator hardware. For example, new AI accelerators utilize ReRAM memristor Xbars. Such memristors are reconfigurable, non-volatile memory cells that can be used, for instance, to simulate the function of weights in neural networks.

The analog BFO memristor presented here [1-5] is one of the few memristors with continuous memory and time-separated read and write operations for data storage and processing in the same memristor cell [6]. It is discussed why BFO memristors are promising for low-power, low-latency AI accelerators, e.g., for autonomous driving, surveillance, and robotics.

1) Y. Shuai, S. Zhou, D. Bürger, M. Helm, H. Schmidt, J. Appl. Phys. 109, 124117 (2011).
2) T. You, Y. Shuai, W. Luo, N. Du, D. Bürger, I. Skorupa, R. Hübner, S. Henker, C. G. Mayr, R. Schüffny, T. Mikolajick, O.G. Schmidt, H. Schmidt, Adv. Funct. Mat. 24, 3357-3365 (2014).
3) N. Du, M. Kiani, C.G. Mayr, T. You, D. Bürger, I. Skorupa, O.G. Schmidt, H. Schmidt, Front. Neurosci. 9, 227 (2015).
4) N. Du, X. Zhao, Z. Chen, B. Choubey, M. Di Ventra, I. Skorupa, D. Bürger, H. Schmidt, Front. Neurosci. 15, 660894 (2021).
5) S. V. Vegesna, V. R. Rayapati, H. Schmidt, Phys. Rev. Applied **22**, 034028 (2024).
6) H. Schmidt, J. Appl. Phys. 135, 200902 (2024).