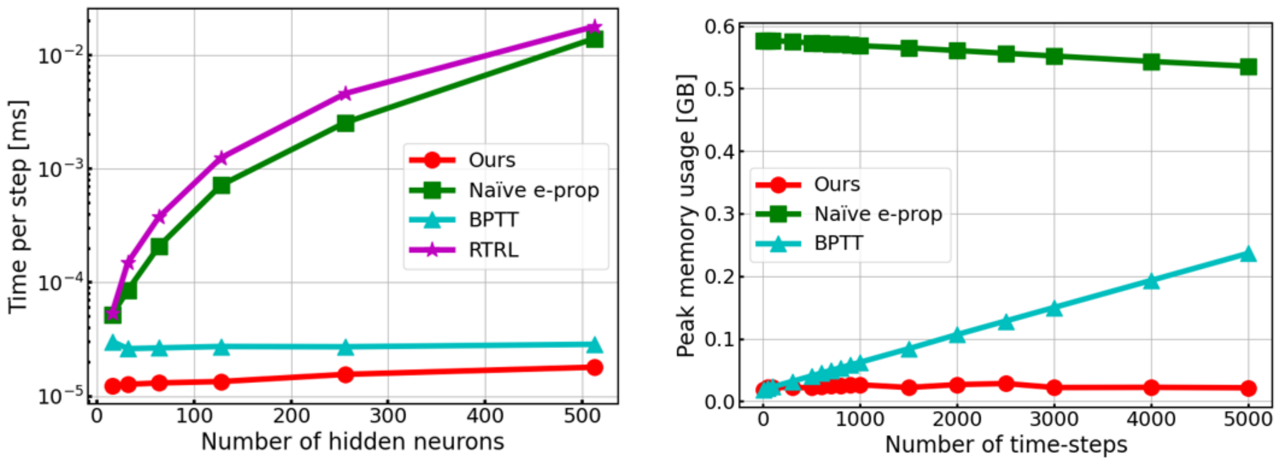
A Truly Sparse and General Implementation of Gradient-Based Synaptic Plasticity

J. Lohoff[1,2], A. Kaya[1,2], F. Assmuth[1,2], E. Neftci[1,2]

[1] Dept. of Electrical Engineering, RWTH Aachen, [2] PGI-15, Forschungszentrum Jülich

Online synaptic plasticity rules derived from gradient descent achieve high accuracy on a wide range of practical tasks. However, their software implementation often requires tediously hand-derived gradients or using gradient backpropagation which sacrifices the online capability of the rules. In this work, we present a custom automatic differentiation (AD) pipeline for sparse and online implementation of gradient-based synaptic plasticity rules that generalizes to arbitrary neuron models. Our work combines the programming ease of backpropagation-type methods for forward AD while being memory-efficient. To achieve this, we exploit the advantageous compute and memory scaling of online synaptic plasticity by providing an inherently sparse implementation of AD where expensive tensor contractions are replaced with simple element-wise multiplications if the tensors are diagonal. Gradient-based synaptic plasticity rules such as eligibility propagation (e-prop) have exactly this property and thus profit immensely from this feature. We demonstrate the alignment of our gradients with respect to gradient backpropagation on an synthetic task where e-prop gradients are exact, as well as audio speech classification benchmarks. We demonstrate how memory utilization scales with network size without dependence on the sequence length, as expected from forward AD methods.



(a) Evaluation time of a single step for changing number of hidden neurons with 1000 training example time-steps for our e-prop implementation with Graphax, BPTT, naïve e-prop, and RTRL.

(b) Peak device memory usage for 128 hidden neurons with changing training example time-steps for our e-prop implementation with Graphax, BPTT, and naïve e-prop.

Figure 1: Performance comparison of e-prop implementations.

[1] J. Lohoff, A. Kaya, F. Assmuth, E. Neftci. *A truly sparse implementation of gradient-based plasticity*. Neuro-Inspired Computational Elements, 2025.

[2] J. Lohoff, E. Neftci. *Optimizing Automatic Differentiation with Deep Reinforcement Learning*. NeurIPS, 2024.

[3] Guillaume Bellec, Franz Scherr, Anand Subramoney, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. *A solution to the learning dilemma for recurrent networks of spiking neurons*. Nature communications, 11, 2020. Authors, Abbreviated Journal title, Volume (Issue), Page(s), Year.