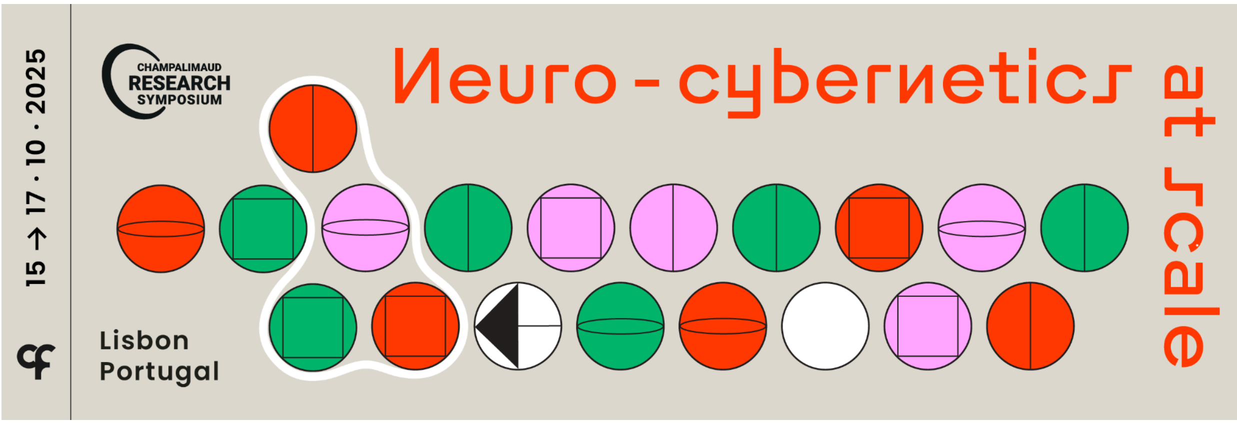

Second-order forward-mode optimization of recurrent neural networks for neuroscience

Youjing Yu, Rui Xia, Qingxi Ma, Máté Lengyel, Guillaume Hennequin
1. University of Cambridge, Cambridge, UK

Training recurrent neural networks (RNNs) to perform neuroscience tasks can be challenging. Unlike in machine learning where any architectural modification of an RNN (e.g. GRU or LSTM) is acceptable if it facilitates training, the RNN models trained as models of brain dynamics are subject to plausibility constraints that fundamentally exclude the usual machine learning hacks. The "vanilla" RNNs commonly used in computational neuroscience find themselves plagued by ill-conditioned loss surfaces that complicate training and significantly hinder our capacity to investigate the brain dynamics underlying complex tasks. Moreover, some tasks may require very long time horizons which backpropagation cannot handle given typical GPU memory limits. In earlier work [Ref 1], we developed SOFO, a second-order random subspace optimizer that efficiently explores loss surfaces without requiring backpropagation. Instead, SOFO relies on parallelizable batched forward-mode differentiation, yielding constant memory cost over time. By exploiting second-order curvature, SOFO significantly outperforms Adam on various RNN tasks. In more recent work, we improve SOFO by using curvature-guided navigation of the parameter space, replacing random subspace sampling in each iteration. We show that the improved SOFO converges faster than [Ref 1] across tasks including a delayed addition task [Ref 2], a challenging double-reach motor task, and a meta continual-learning task using a Hebbian learning rule. By accelerating and scaling the training of biologically grounded network models, SOFO greatly facilitates research into neural networks modeling behavioral tasks.

biologically plausible recurrent neural networks; behaviour modelling; optimization; memory-efficient training; second-order method