Open questions on delay optimization in Spiking Neural Networks

Alexandre Queant [1, 2, 3], Benoit R Cottereau [1, 2, 3] and Timothée Masquelier [1, 2]

[1] CerCo, CNRS UMR 5549, Toulouse, France, [2] Université de Toulouse, France, [3] IPAL, CNRS IRL 2955, Singapore

While surrogate gradient learning (SGL) has enabled the training of deep and recurrent spiking neural networks (SNNs) with performances approaching those of standard deep learning [1], the optimization of these networks remains largely focused on three aspects: synaptic weight optimization, as in regular deep learning; spiking neuron model development and tuning; and structural or architectural exploration. Another form of optimization has gained traction in recent years, which directly learns how neurons should transmit information in time. In SNNs, each connection is characterized not only by a weight but also by a transmission delay, and both can be considered as coupled parameters for optimization. Theoretical works have long suggested that trainable delays drastically enhance expressivity [2], but practical methods for learning them have only emerged recently and remain mostly limited to feedforward architectures [3–9]. We recently introduced DelRec [10], an SGL-based method to jointly optimize axonal or synaptic delays along with synaptic weights in recurrent SNNs (RSNNs) via backpropagation through time. DelRec uses a differentiable interpolation scheme to handle non-integer delays during training, is neuron-model agnostic, and achieved state-of-the-art results on challenging temporal benchmarks. The availability of recurrent delay learning immediately raises new questions. What can be modeled with such delay-equipped recurrent networks that could not be modeled before? Can we identify tasks where delay learning is not merely helpful but predominantly necessary for good performance? More broadly, learning such a fundamental parameter, one that directly shapes the temporal behavior of a network, opens research directions that are yet to be explored, for both feedforward delays and recurrent delays. How do recurrent and feedforward delays interact when both are learned? Is learning delays always the best trade-off compared to having fixed delays, increasing network size or enriching neuron dynamics? To what degree can delay learning lead to sparser, more parameter-efficient models? These remain open questions of direct relevance to the neuromorphic community, where programmable delays are natively supported by event-driven hardware and where efficiency is the central concern. We will present the DelRec framework along with preliminary results on these directions.

[1] E. O. Neftci, H. Mostafa, F. Zenke, IEEE Signal Process. Mag., 2019.

[2] W. Maass, M. Schmitt, Inf. Comput., 1999.

[3] S. B. Shrestha, G. Orchard, arXiv, 1810.08646, 2018.

[4] P. Sun, Y. Chua, P. Devos, D. Botteldooren, Front. Neurosci., 2023.

[5] I. Hammouamri, I. Khalfaoui-Hassani, T. Masquelier, ICLR, 2024.

[6] L. Deckers et al., Front. Neurosci., 2024.

[7] J. Göltz, J. Weber, L. Kriener et al., Nat. Commun., 2025.

[8] B. Mészáros, J. C. Knight, T. Nowotny, Nat. Commun., 2025.

[9] L. Vassallo, N. Taherinejad, arXiv, 2601.00668, 2026.

[10] A. Queant, U. Rançon, B. R. Cottereau, T. Masquelier, arXiv, 2509.24852, 2025.