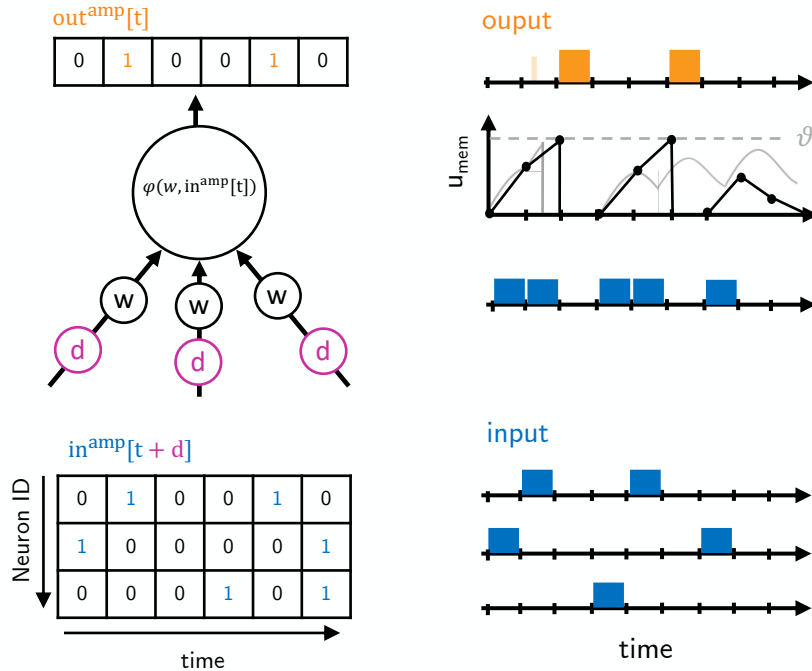


DelGrad: Exact event-based gradients in spiking networks for training delays and weights

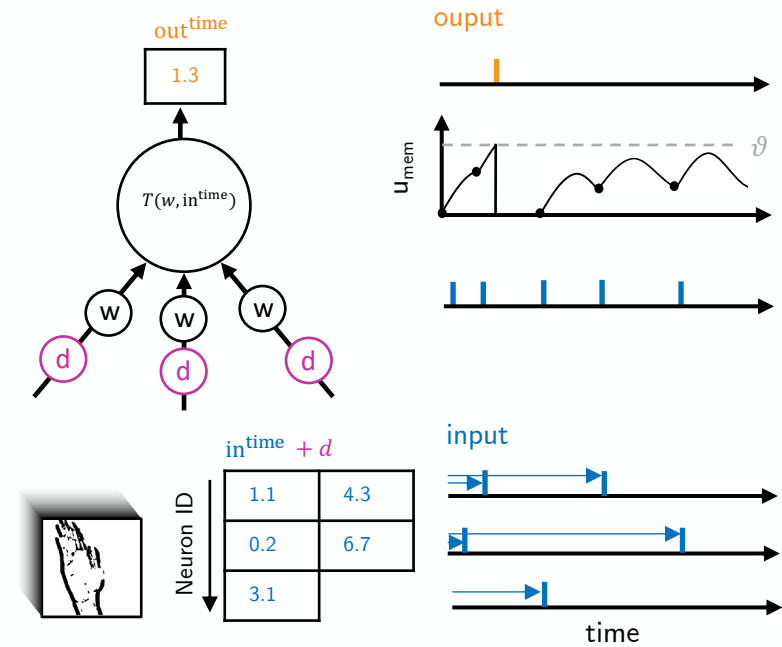
Julian Goltz, Jimmy Weber, Laura Kriener, Sebastian Billaudelle, Peter Lake, Johannes Schemmel, Melika Payvand and Mihai A. Petrovici.



Time step framework [1]



Event-based framework [2]



Gradients for the delays

$$\frac{\partial \mathcal{L}}{\partial d} = \frac{\partial \mathcal{L}}{\partial out^{amp}[t]} \frac{\partial out^{amp}[t]}{\partial in^{amp}[t+d]} \frac{\partial in^{amp}[t+d]}{\partial d}$$

Surrogate Gradient	Non differentiable
--------------------	--------------------

- Exact at forward (precise timings)
- Exact at backward (differentiable $\Rightarrow$ exact gradients).
- Complexity scales with events sparsity.
- Fewer observables: only the spike timings (HW-friendly).
- Natural training of **delays**

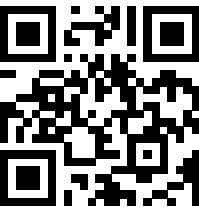
Gradients for the delays

$$\frac{\partial \mathcal{L}}{\partial d} = \frac{\partial \mathcal{L}}{\partial out^{time}[t]} \frac{\partial out^{time}[t]}{\partial (in^{time} + d)} \frac{\partial (in^{time} + d)}{\partial d}$$

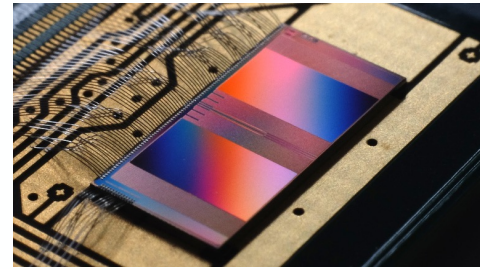
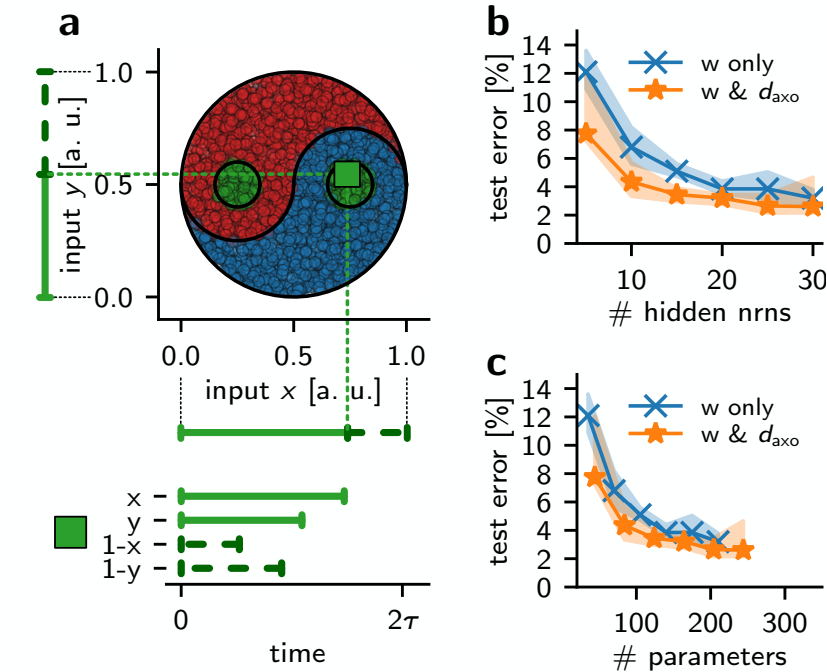
Trivial: **=1**

DelGrad: Exact event-based gradients in spiking networks for training delays and weights

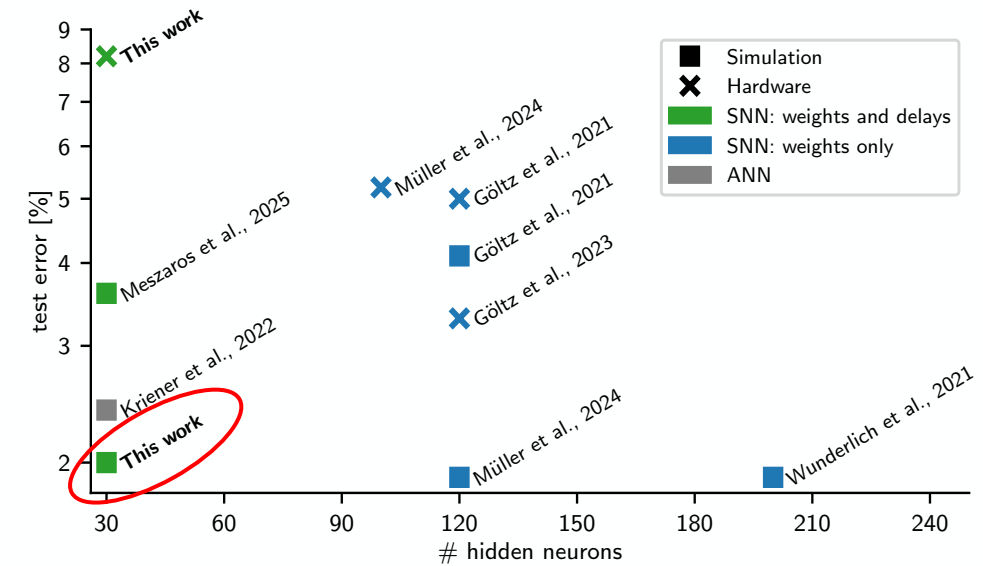
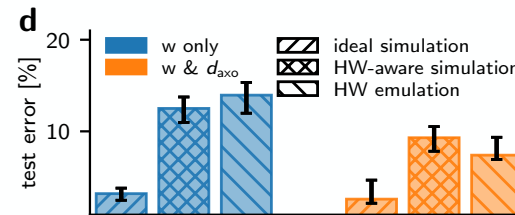
Julian Göltz, Jimmy Weber, Laura Kriener, Sebastian Billaudelle, Peter Lake, Johannes Schemmel, Melika Payvand and Mihai A. Petrovici.



Results



BrainScaleS-2 [4]



- Assessed on the Yin-Yang Dataset [3]
- Better performance with reduced number of parameters

- Demonstrate successful training with HW in-the-loop

- Outperform State-of-the-art on the Yin-Yang Dataset