Biologically inspired alternatives to backpropagation through time for learning in recurrent neural networks

Guillaume Bellec*, Franz Scherr*, Darjan Salaj, Elias Hajek, Robert Legenstein and Wolfgang Maass

Institute for Theoretical Computer Science, TU Graz

bellec@igi.tugraz.at

Human Brain Project

Abstract

1. Back-propagation through time (BPTT) is implausible and requires ample resources.
2. In "eligibility propagation" (*e-prop*) gradients are computed online without back-propagating through time.
3. It provides insights for improving truncated BPTT on hard machine learning problems with long time horizon.

E-prop: a new factorization of the error gradients

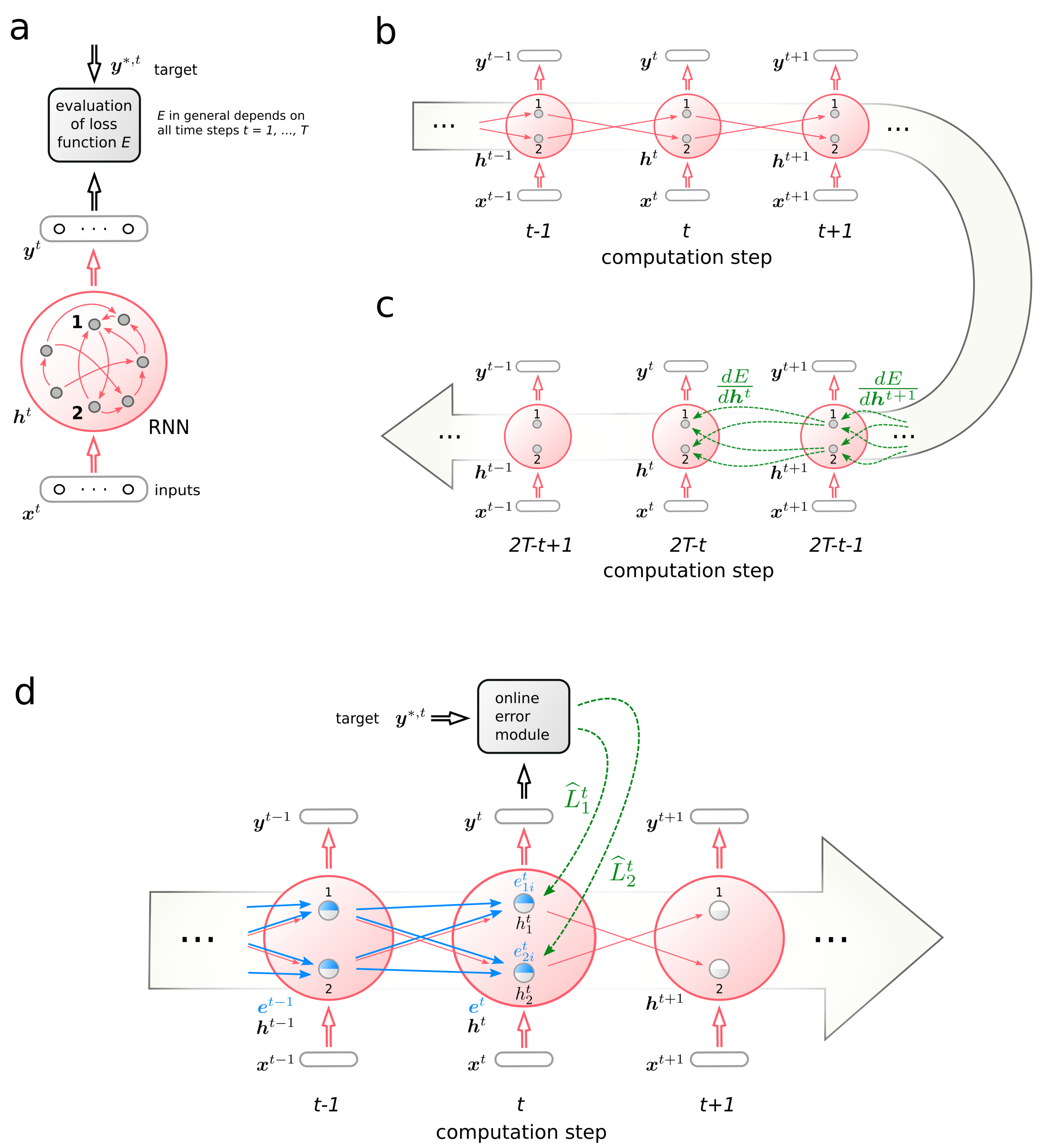

Figure 1: a The theory applies to a large family of RNN models (LSTMs and LIF are considered here). **b** Forward pass of BPTT. **c** Backward pass of BPTT. **d** Scheme of the computation targeted with *e-prop*.

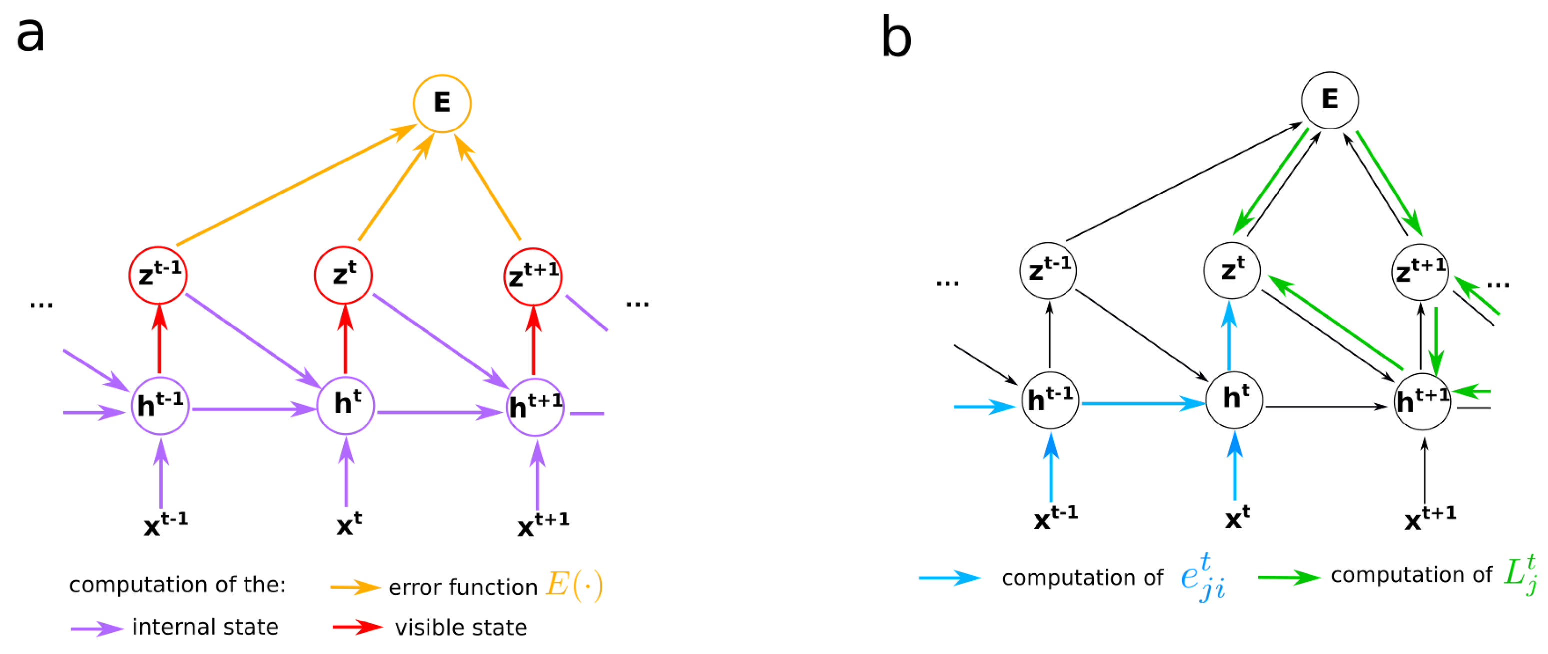

Figure 2: a Definition of the mathematical model. **c** Scheme of the gradient propagation in *e-prop*.

In standard BPTT the gradients with respect to the recurrent weights are computed using the formula: $\frac{dE}{dW^{\mathrm{rec}}_{ji}} = \sum_t \frac{dE}{dh^t_j}\frac{\partial h^t_j}{\partial W^{\mathrm{rec}}_{ji}}$. We show that the same gradients can be re-written as follows

$$\frac{dE}{dW^{\mathrm{rec}}_{ji}} = \sum_t L^t_j e^t_{ji}, \tag{1}$$

where learning signals are defined by $L^t_j = \frac{dE}{dz^t_j}$ and the eligibility traces are computed using recursively defined eligibility vectors: $\epsilon^{t+1}_{ji} = \frac{\partial h^{t+1}_j}{\partial h^t_j} \cdot \epsilon^t_{ji} + \frac{\partial h^t_j}{\partial W^{\mathrm{rec}}_{ji}}$. Then the eligibility traces are $e^t_{ji} = \frac{\partial z^t_j}{\partial h^t_j}\epsilon^t_{ji}$

E-prop 1: direct feedback alignment in RNNs

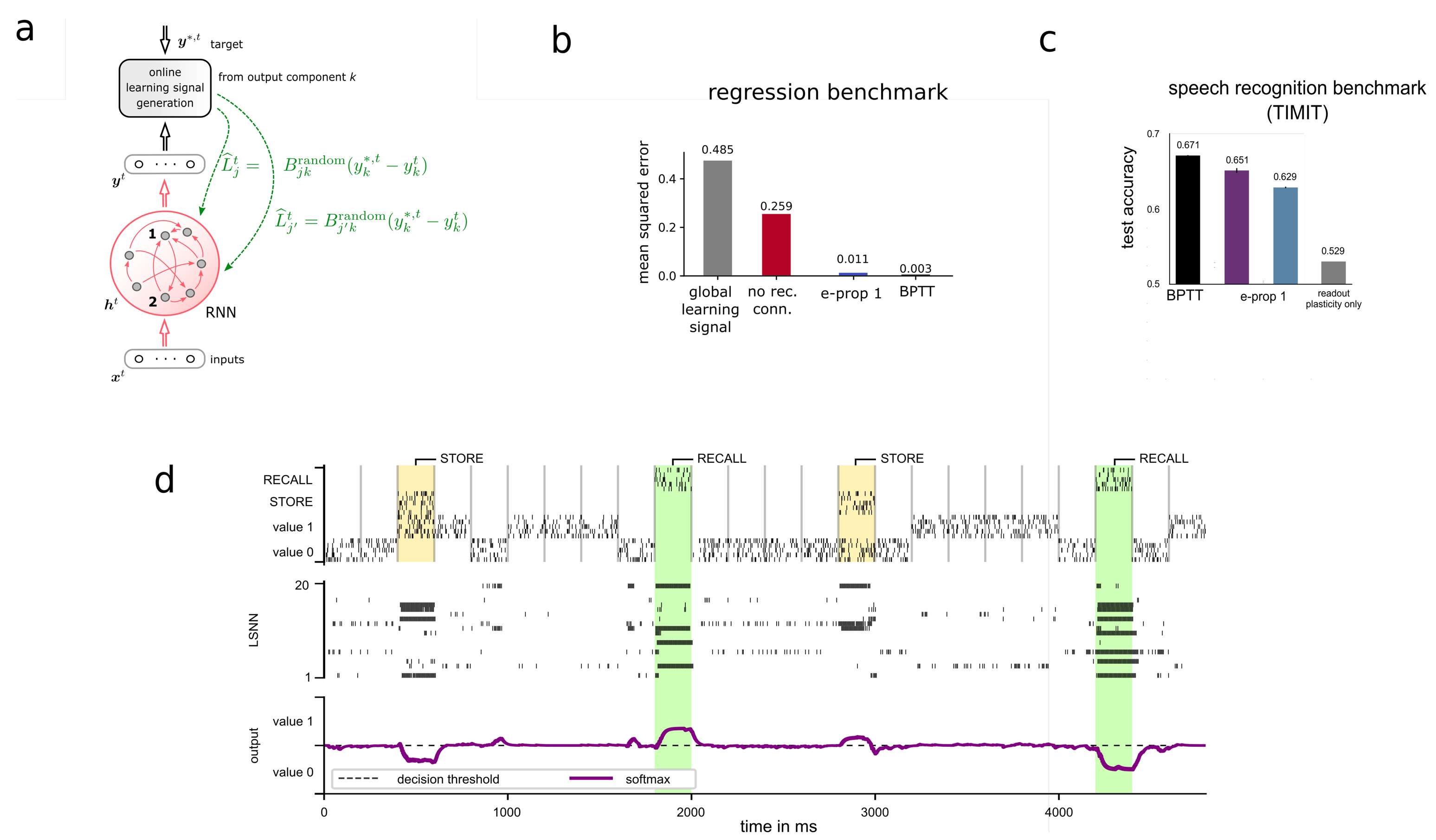

Figure 3: a Scheme of the simplest variant *e-prop 1*: the learning signals L^t_j is replaced by a random projection of the instantaneous error. The performance is reported on a regression task (**b**) and on a speech recognition task (textbfc). **d** Importantly, *e-prop* exploits the long "long short-term memory" of RNN models as the information remains in the eligibility traces as long as it does in the network itself.

E-prop 2: One shot learning via Learning-to-learn

For a given family $\mathcal{F}$ of tasks, the outer loop optimization trains the parameters of an error module and the initialization W^{init} of the neural network, such that the main network learns to solve any task $\mathcal{C}$ of $\mathcal{F}$ after one parameter update. The training over tasks of $\mathcal{F}$ called the outer loop, the parameter update between the demonstration and the test trial is the inner loop.

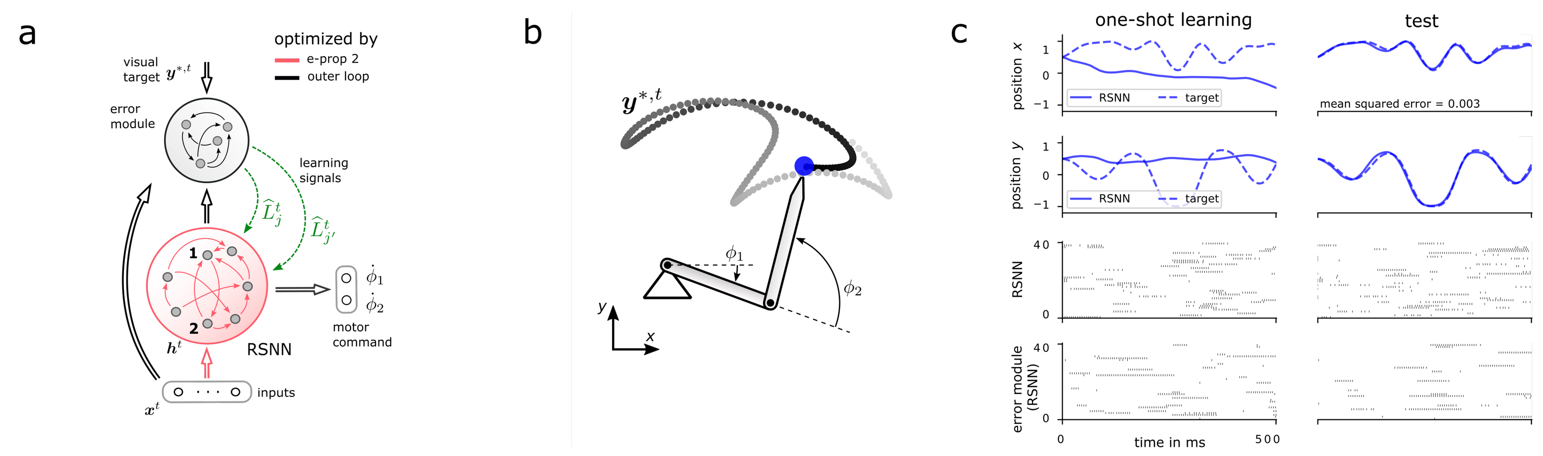

Figure 4: a Network architecture, the learning signal is provided by a separate error module. The error module is trained on the outer-loop. **b** The movement had to be performed by a two joint arm, and the target trajectory is two dimensional trajectory generated randomly. **c** Network activity after outer loop training. The activity is displayed on the first and single presentation of the motion to be learnt (left) and the test trial (right).

E-prop 3: Boosting truncated BPTT with eligibility traces

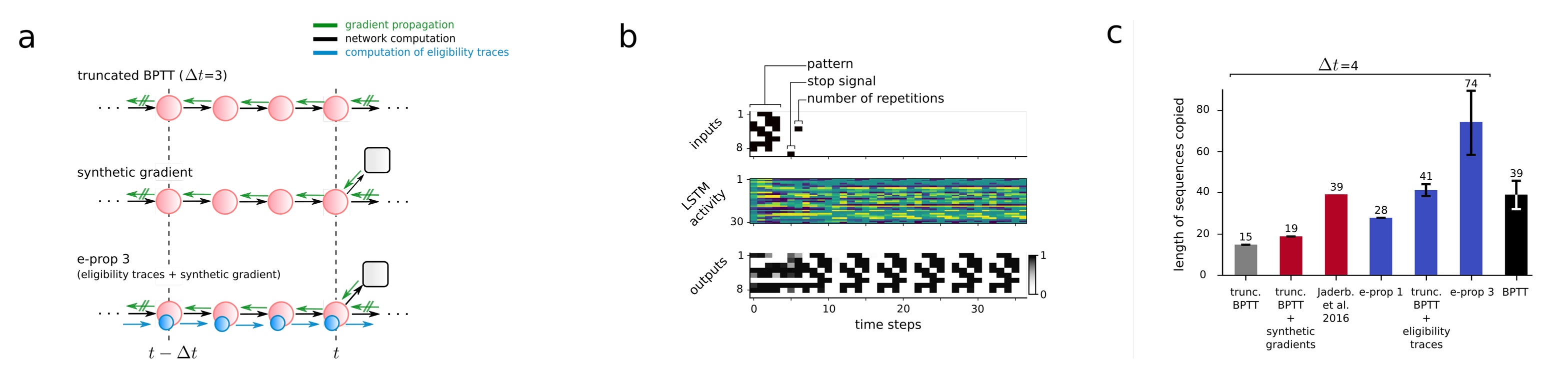

Figure 5: *E-prop 3* improves truncated BPTT. a *E-prop 3* extends the time horizon towards the past and the future using eligibility traces and synthetic gradients. **b** A trial of the copy repeat task. **c** Performance on the copy-repeat task.

Conclusions

1. In [2] we provide a new theory for learning in RNNs.
2. For biophysical neuron models, the learning results in learning rule compatible with the experimental data on synaptic plasticity. For a review see [3].
3. In [1], it was already shown that spiking neuron can achieve performance comparable to that of LSTMs when trained with BPTT. It seems that learning rules that can operate online on a neuromorphic hardware such as Loihi with in-built plasticity can achieve comparable performance.

Acknowledgements

This research was supported by the Human Brain Project of the European Union, Grant agreement No. 785907

References

[1] Guillaume Bellec, Darjan Salaj, Anand Subramoney, Robert Legenstein, and Wolfgang Maass. Long short-term memory and learning-to-learn in networks of spiking neurons. In *Advances in Neural Information Processing Systems*, pages 787–797, 2018.

[2] Guillaume Bellec, Franz Scherr, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. Biologically inspired alternatives to backpropagation through time for learning in recurrent neural nets. *arXiv preprint arXiv:1901.09049*, 2019.

[3] Wulfram Gerstner, Marco Lehmann, Vasiliki Liakoni, Dane Corneil, and Johanni Brea. Eligibility traces and plasticity on behavioral time scales: experimental support of neohebbian three-factor learning rules. *Frontiers in neural circuits*, 12, 2018.