Continuous learning in dendritic cortical microcircuits using Lagrangian mechanics

D. Dold[1,2], J. Sacramento[2], M. A. Petrovici[1,2], J. Binas[3], Y. Bengio[3], W. Senn[2]

[1]Heidelberg University, Kirchhoff-Institute for Physics, [2]University of Bern, Department of Physiology, [3]Université de Montréal, MILA

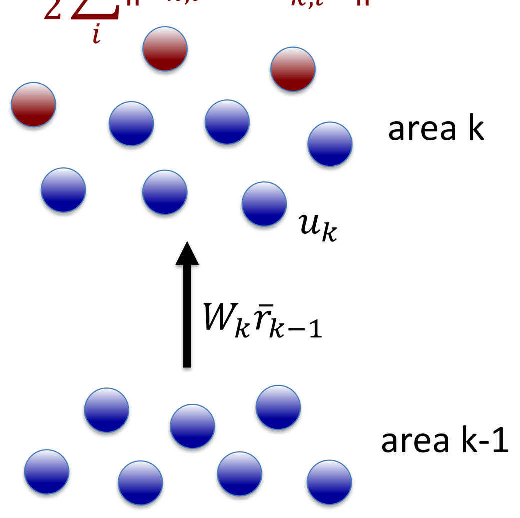
The Manfred Stärk Foundation

1. Energy-based models

Energy function encodes network:

$$E = \frac{1}{2}\sum_{k=1}^{N} \underbrace{\|u_k - W_k \bar{r}_{k-1}\|^2}_{\text{prediction error}} + \beta \underbrace{C}_{\text{cost function}}$$

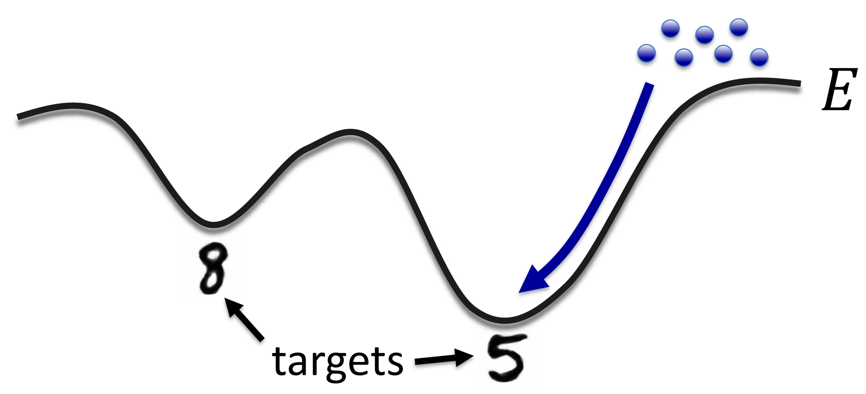
$$C = \frac{1}{2}\sum_i \|u_{k,i} - u_{k,i}^{trgt}\|^2$$



Standard -- dynamics from gradient descent:

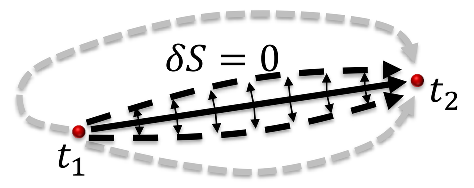
$$\dot{u} \sim -\nabla_u E$$

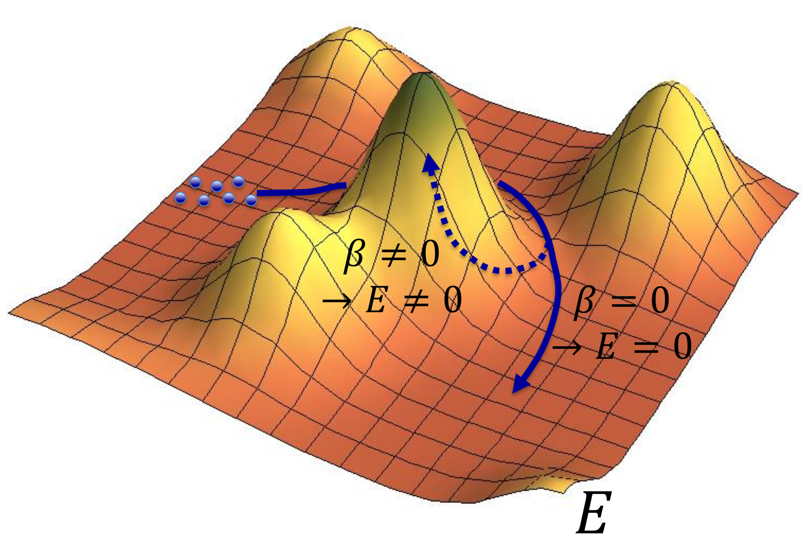
$$\dot{W} \sim -\nabla_W E$$



Here -- dynamics from Lagrangian mechanics:

$$\underbrace{S}_{\text{action}} = \int_{t_1}^{t_2} dt \underbrace{L(\tilde{u}, \dot{\tilde{u}})}_{\text{Lagrangian}}$$





$$\frac{\partial L}{\partial \tilde{u}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\tilde{u}}} = 0 \longrightarrow$$

no phases, continuous learning possible

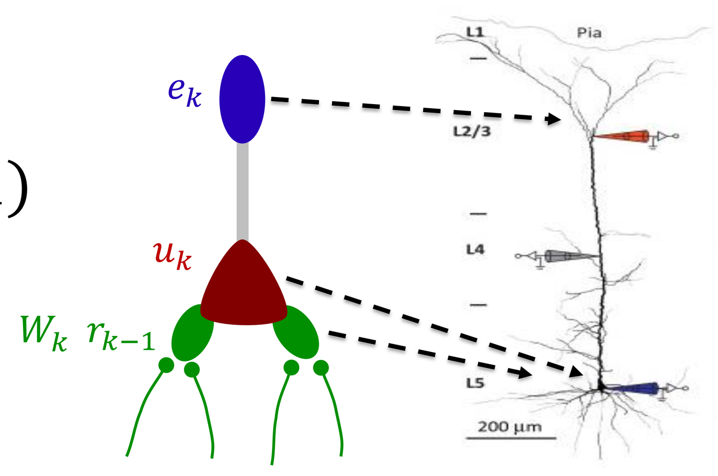
2. From Euler-Lagrange to neurons

Use future discounted voltage:

$$\tilde{u}(t) = \frac{1}{\tau}\int_t^\infty dt' u(t') e^{-\frac{t'-t}{\tau}} \xrightarrow{L(\tilde{u},\dot{\tilde{u}}) = -E(u)} \frac{\partial L}{\partial \tilde{u}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\tilde{u}}} = 0$$

Resulting neuronal dynamics:

$$\underbrace{\tau \dot{u}_k = -u_k}_{\text{somatic integration}} + \underbrace{W_k r_{k-1} + e_k}_{\text{basal and apical compartment}} \quad (1)$$

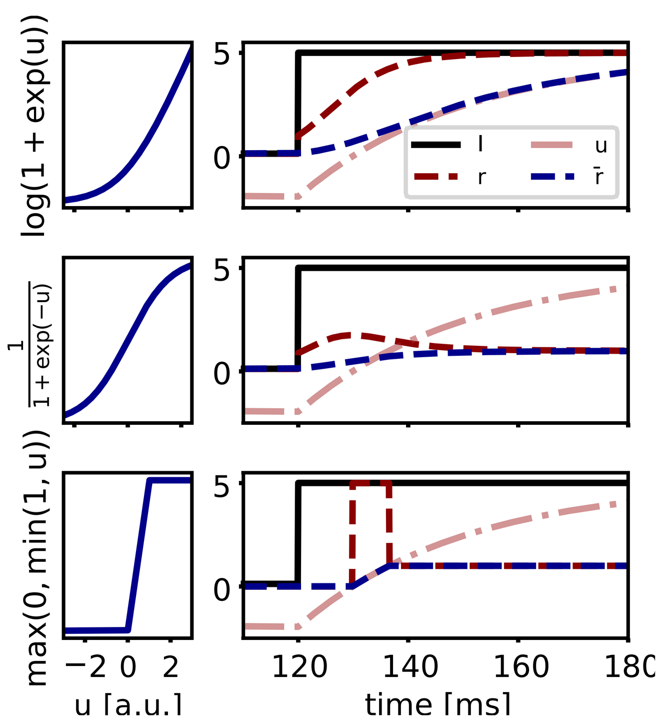


Advanced neuronal response:

$$r_k = \underbrace{\bar{r}_k + \tau \dot{\bar{r}}_k}_{\text{phase-advance of } \bar{r}_k = \varphi(u_k)}$$

$$\approx \bar{r}_k(t + \tau)$$

$$\approx \underbrace{m_\infty^3 h(u, \dot{u})}_{\text{sodium gating of HH neurons}}$$



3. Learning with dendrites

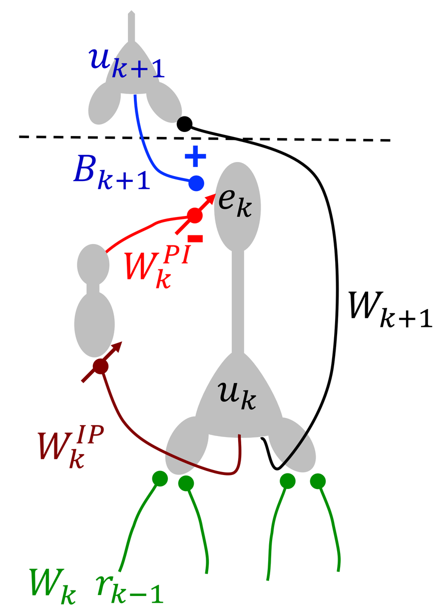
Prediction error encoded in apical dendrites:

$$\bar{e}_k = \bar{r}'_k \cdot W_{k+1}^T (u_{k+1} - W_{k+1}\bar{r}_k)$$

$$\sim W_{k+1}^T u_{k+1} - W_{k+1}^T W_{k+1} \bar{r}_k$$

$$\sim \underbrace{B_{k+1} u_{k+1}}_{\text{top-down feedback}} - \underbrace{W_k^{PI} W_k^{IP} \bar{r}_k}_{\text{bottom-up prediction}}$$

⟶ **nudges u_k away from basal voltage**



Basal prediction of soma drives plasticity:

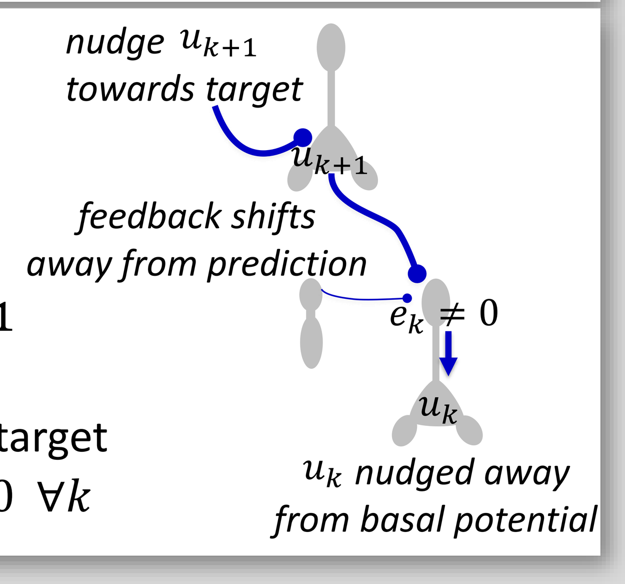
$$\dot{W} \sim -\nabla_W E \longrightarrow \dot{W}_k \sim \underbrace{(u_k - W_k \bar{r}_{k-1})}_{\text{Urbanczik-Senn (basal prediction of soma)}} \bar{r}_{k-1}^T \quad (2)$$

Backpropagation of errors:

combining (1) and (2): $\dot{W}_k \sim \bar{e}_k \bar{r}_{k-1}^T$

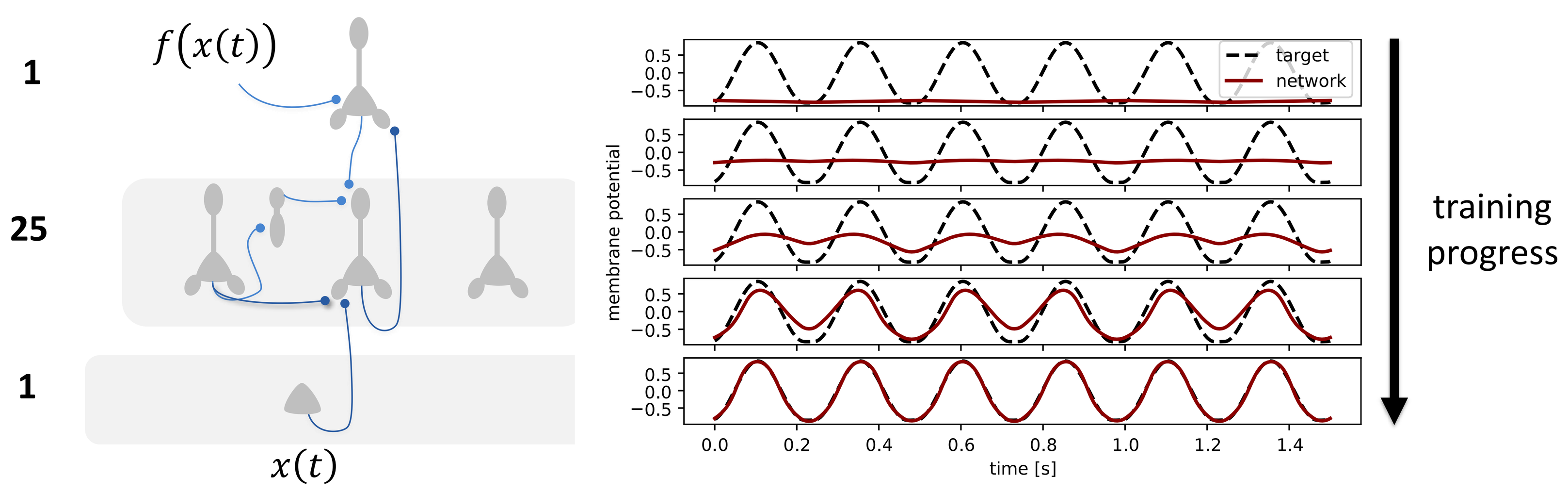
with $\bar{e}_k = \bar{r}'_k \cdot W_{k+1}^T \bar{e}_{k+1}$

- **errors** introduced by **nudging** neurons towards target
- no nudging ⟶ no errors ⟶ $\dot{W}_k = 0 \;\; \forall k$

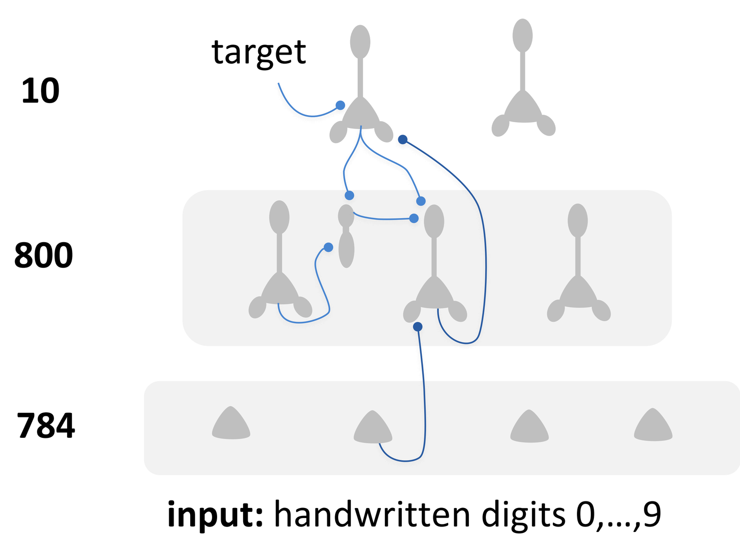


4. Time-continuous learning

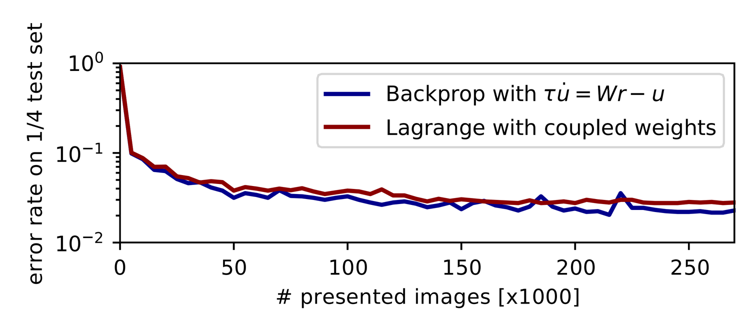
Learning a simple nonlinear continuous mapping:



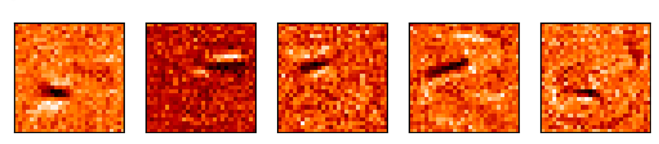
Learning handwritten digits:



Performance:



Receptive fields:



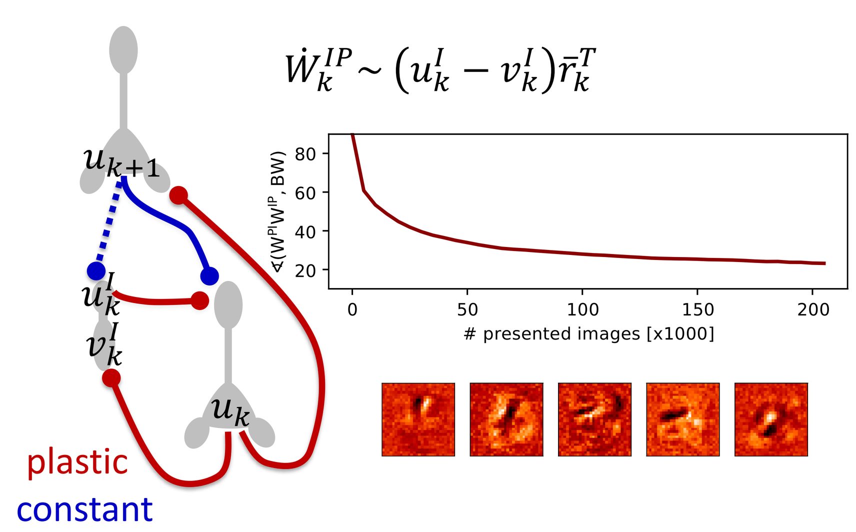
Learning lateral weights:

from theory: coupled forward and backward weights

deviation: weights align during training *(work in progress)*

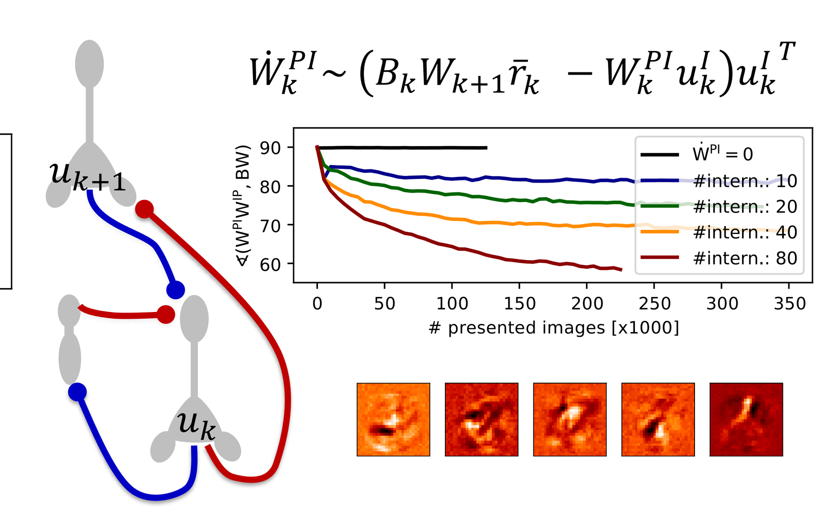
nudging of interneurons:

$$\dot{W}_k^{IP} \sim (u_k^I - v_k^I)\bar{r}_k^T$$



no nudging of interneurons:

$$\dot{W}_k^{PI} \sim (B_k W_{k+1}\bar{r}_k - W_k^{PI} u_k^I) u_k^{I\,T}$$



This work has received funding from the European Union Horizon 2020 Framework Programme under grant agreement 720270 and 785907 (HBP) as well as the Manfred Stärk Foundation. Calculations using GPUs were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern. Furthermore, the authors acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no INST 39/963-1 FUGG.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533.
LeCun, Y. A (1988). Theoretical Framework for Back-Propagation. Proceedings of the 1988 Connectionist Models Summer School, 21{28.
Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extraclassical receptive-field effects. Nature neuroscience, 2(1), 79.
Scellier, B., & Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. Frontiers in computational neuroscience, 11, 24.
Sacramento, J., Costa, R. P., Bengio, Y.,& Senn, W. (2017). Dendritic error backpropagation in deep cortical microcircuits. arXiv preprint arXiv:1801.00062.

Email contact: dodo@kip.uni-heidelberg.de

UBELIX