Lagrangian neurodynamics for real-time error-backpropagation across cortical areas

D. Dold[1,2], W. Senn[2], A. F. Kungl[1,2], J. Sacramento[3], K. Schindler[4], J. Binas[5], Y. Bengio[5], M. A. Petrovici[1,2]

[1]Heidelberg University, Kirchhoff-Institute for Physics, [2]University of Bern, Department of Physiology, [3]Institute of Neuroinformatics, UZH / ETH Zurich, [4]Department of Neurology, Inselspital Bern, [5] MILA, University of Montreal.

The Manfred Stärk Foundation

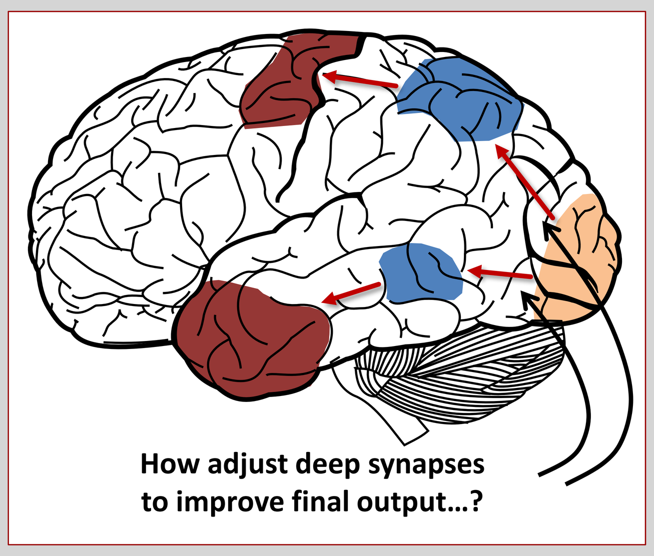
Abstract

The hierarchical structure of the cortex raises the question how plasticity in the brain is able to shape such a structure in the first place. The distant cousins of biological neurons, deep abstract neural networks, are commonly trained with the **backpropagation-of-errors algorithm (backprop)**, which solves the credit assignment problem for deep neural networks and is behind many of the recent achievements of deep learning. Despite its effectiveness in abstract neural networks, it remains **unclear whether backprop might represent a viable implementation of cortical plasticity**. Here, we present a new **theoretical framework that uses a least-action principle to derive a biologically plausible implementation of backprop**.

Backprop in the brain?

Whether the brain might use an optimization scheme like backprop to guide synaptic plasticity in deep hierarchical cortical areas is still an open question.

Here, we present a model that derives network dynamics from a **Lagrangian L**. Standard leaky integrator dynamics in u are obtained by requiring a least-action principle of L with respect to the **future discounted voltage $\tilde{u}$**.



Synaptic dynamics are defined to perform gradient descent on the same Lagrangian L. **The combined neurosynaptic dynamics lead to the emergence of backprop and learning as gradient descent on a cost function.**

Energy function

Cortex

$$E(u) = \sum_i \| u_i - W_i \varphi(u_{i-1}) \|^2 + \beta \cdot cost$$

teacher strength

prediction error $\bar{e}_i$

u_i $\quad W_i\varphi(u_{i-1})$

$$u = \tilde{u} - \tau\dot{\tilde{u}}$$
$$L = E(\tilde{u}, \dot{\tilde{u}})$$

$$\delta L = 0 \quad \& \quad \dot{W} = -\eta \cdot \nabla_W L$$

Theorem 1 *(real-time backprop)*

$$\bar{e}_i = W_{i+1}^{\mathrm{T}} \bar{e}_{i+1}, \qquad \dot{W}_i \propto \bar{e}_i \varphi(u_i)^{\mathrm{T}}$$

Theorem 2 *(real-time gradient descent)*

$$\frac{\mathrm{d}}{\mathrm{d}W_i} cost = \lim_{\beta \to 0} \frac{1}{\beta} \bar{e}_i^{\beta} \varphi^{\beta}(u_i)^{\mathrm{T}}$$

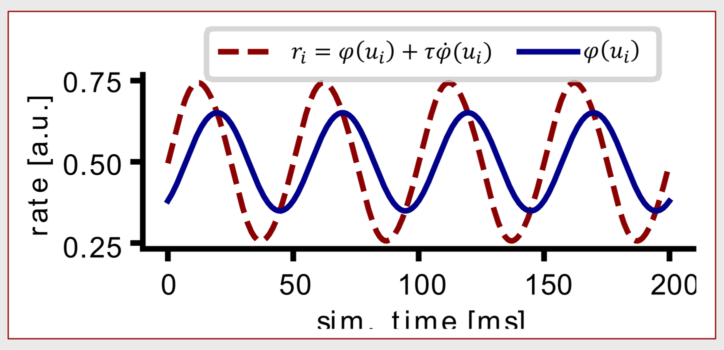
Lagrange formalism

Backprop

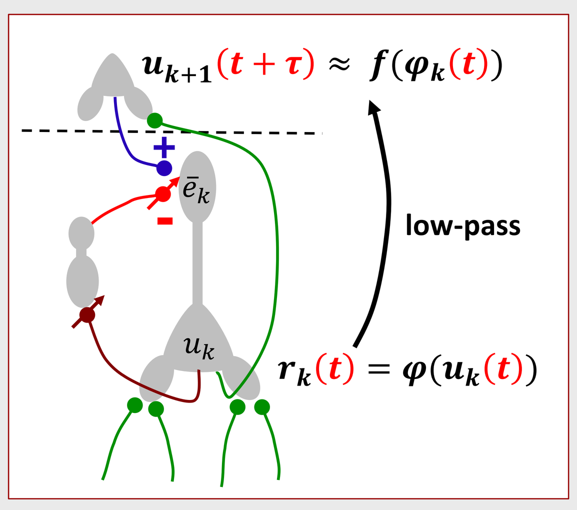
Advanced response compensates delays

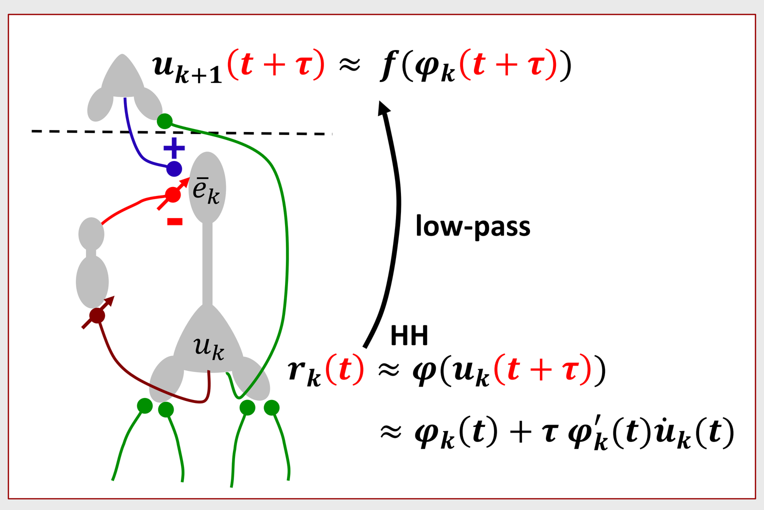
In this model, **neurons communicate with an advanced (look-ahead) rate** that takes the time-derivative of the membrane potential into account:

$$r_i(t) = \varphi(u_i(t)) + \tau\varphi'(u_i(t))\dot{u}_i(t) \approx \varphi(u_i(t+\tau)) .$$

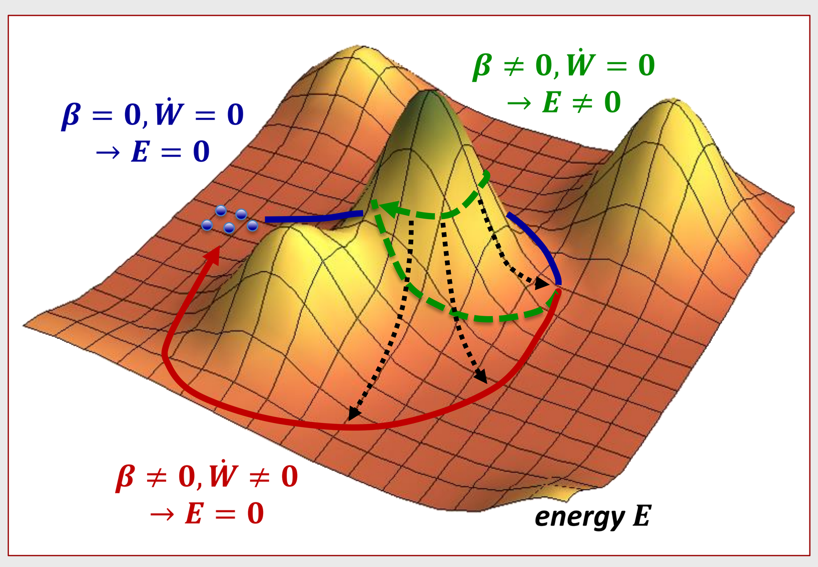


Such an advancing mechanism might be implemented by the spiking mechanism, e.g., sodium gating of the Hodgkin-Huxley (HH) model. In the apical dendrites, similar mechanisms might advance the error signal as well.





The resulting dynamics become intuitive when we look at the Lagrange (or, more commonly, Energy) picture:
Without nudging, the network traverses states with $E = 0$. Only if we nudge, the network starts moving through a regime $E \neq 0$, while plasticity pulls the network back into a low energy regime.

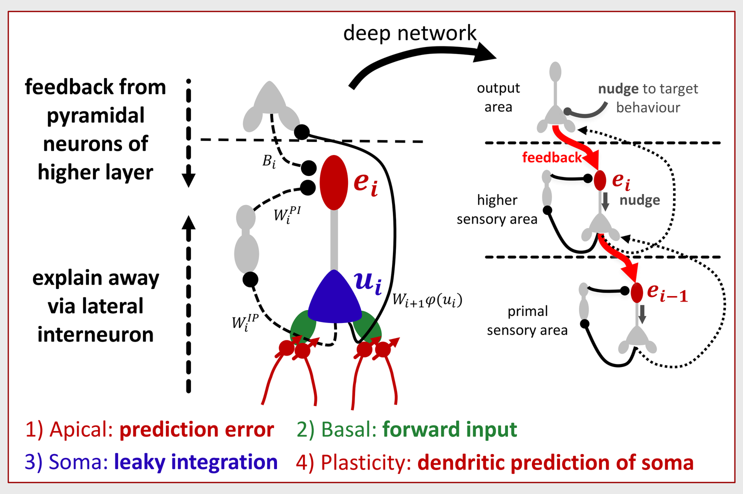


Cortical circuitry implements backprop

The derived neurodynamics can be interpreted as pyramidal neurons integrating basal ($W_i r_{i-1}$) and apical (e_i) input at the soma:

$$\tau\dot{u}_i = -u_i + W_i r_{i-1} + e_i .$$

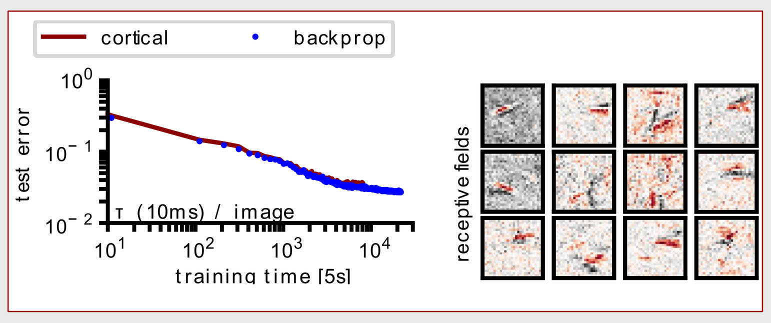
The apical potential encodes a prediction error that is calculated locally via lateral interneuron circuits that try to explain away feedback coming from higher areas.



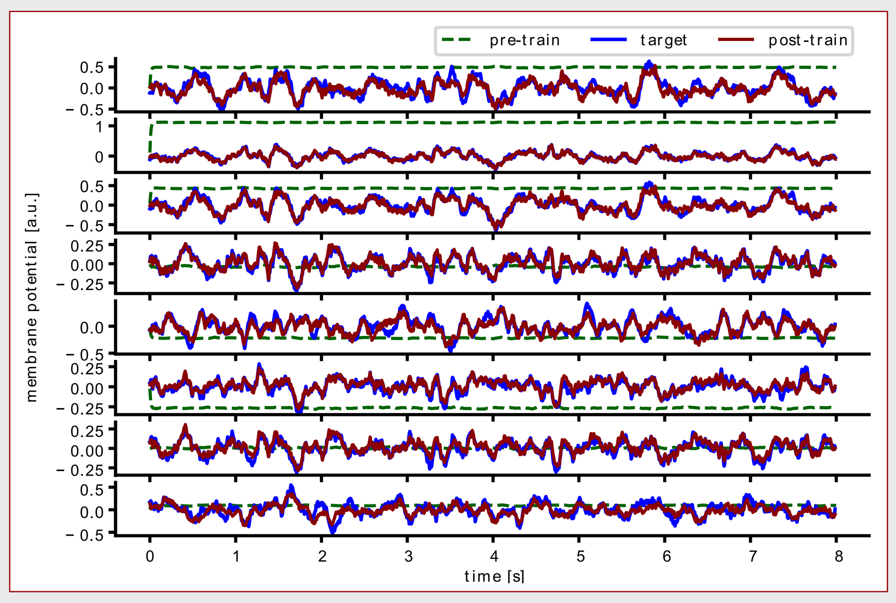
These errors nudge the soma, becoming accessible to the plasticity rule driving the forward projections W_i, which can be interpreted as a voltage-based version of the Urbanczik-Senn rule (dendritic prediction of somatic activity):

$$\dot{W}_i \propto [\underbrace{u_i - W_i\varphi(u_{i-1})}_{\bar{e}_i}]\varphi^T(u_{i-1}).$$

Errors are propagated backward through the network via feedback connections while sensory information is propagated forward. **Neurons minimize these local prediction errors e_i, which in turn reduces a global cost function.**



This is demonstrated here for two tasks: Learning MNIST continuously and learning to reproduce a (continuous) human iEEG signal.



No weight coupling: Learning the microcircuit

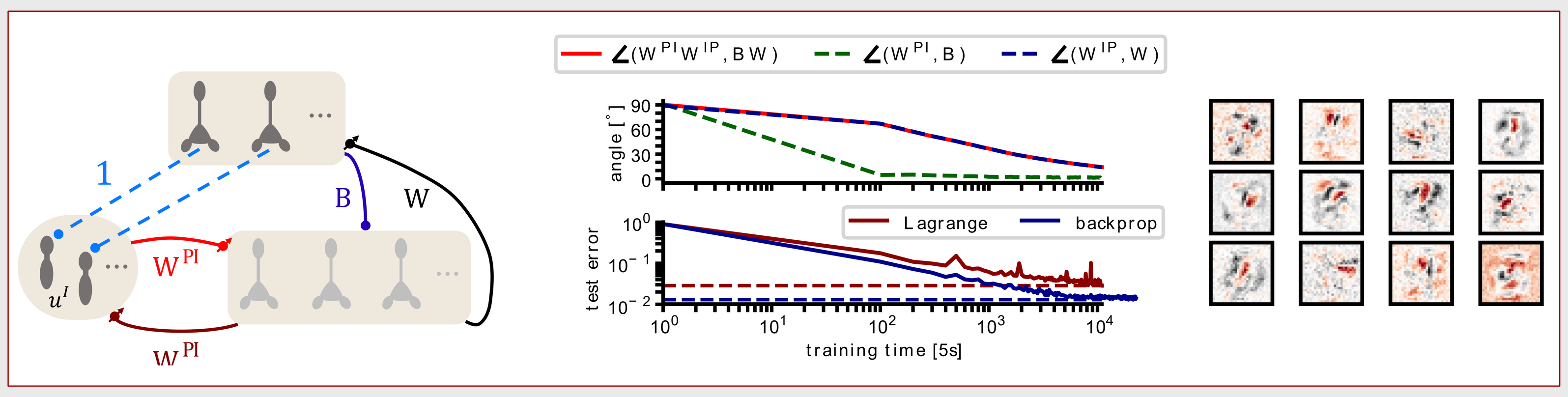
The microcircuit can be trained to avoid weight coupling. In this case, the interneurons are also nudged (either one-to-one or mixed) by the higher cortical area. The pyramidal-to-interneuron weights learn to mimic the top-down feedback

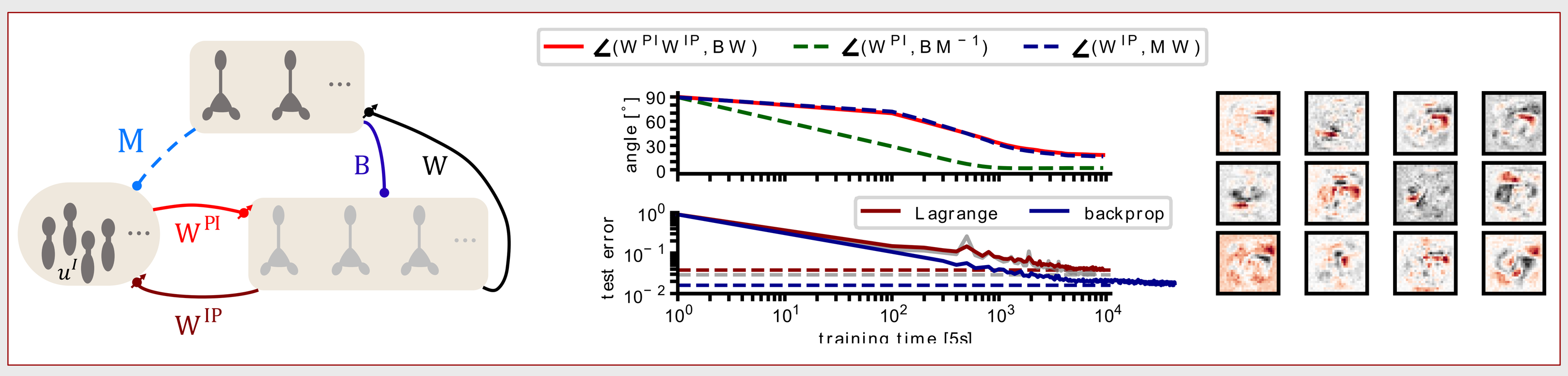
$$\dot{W}_i^{IP} \propto \left(u_i^I - W^{IP}\varphi(u_i)\right)\varphi(u_i)^{\mathrm{T}} ,$$

while the interneuron-to-pyramidal weights learn to undo the mixing as well as compensate for the feedback through B:

$$\dot{W}_i^{PI} \propto (B_i u_{i+1} - W^{PI}u_i^I)(u_i^I)^{\mathrm{T}} .$$

While training the forward weights, the interneuron weights learn to align accordingly.





Conclusion

Why backprop?[1]
- widely and successfully used in many deep learning applications
- very simple learning rule

Backprop in the brain?[2,3,4,5,6]
- recently, several papers proposed different models of how backprop might be implemented in the brain

Cortical circuits implement backprop[5,6]
- possible by employing dendrites, feedback and cortical circuitry
- learning rule itself local and biologically plausible / interpretable

Linking biology and deep learning[6]
- using the Lagrange framework, we have a clear link from abstract theory to the biophysical implementation

[1]LeCun, Yann, Bengio, Yoshua, & Hinton, Geoffrey. **Deep learning**. Nature (2015).
[2]Whittington, James C., & Bogacz, Rafal. **An approximation of the error backpropagation algorithm in a predictive coding network with local Hebbian synaptic plasticity**. Neural computation (2017).
[3]Scellier, Benjamin, & Bengio, Yoshua. **Equilibrium propagation: Bridging the gap between energy-based models and backpropagation**. Frontiers in computational neuroscience (2017).
[4]Guerguiev, Jordan, et al. **Towards deep learning with segregated dendrites**. Elife (2017).
[5]Sacramento, João, et al. **Dendritic cortical microcircuits approximate the backpropagation algorithm**. NeurIPS (2018).
[6]Dold, Dominik, et al. **Lagrangian dynamics of dendritic microcircuits enables real-time backpropagation of errors**. Cosyne Abstracts (2019).

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

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