How can networks of spiking neurons wire themselves up for a specific computational task?

Wolfgang Maass

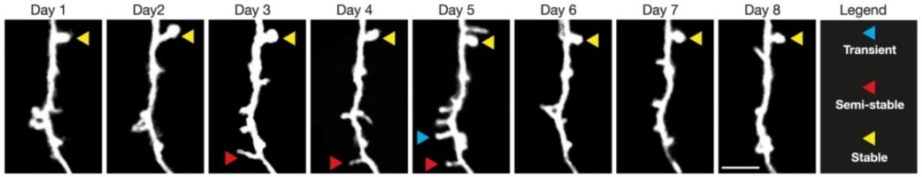
Institute for Theoretical Computer Science
Graz University of Technology, Austria

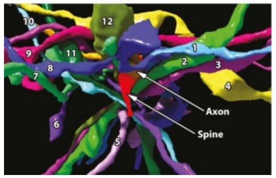
Experimental data suggest that common models for NN learning are incomplete, or even wrong

In the brain:

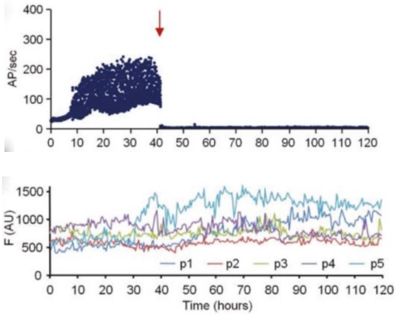
- Networks continuously rewire themselves

Svoboda Lab





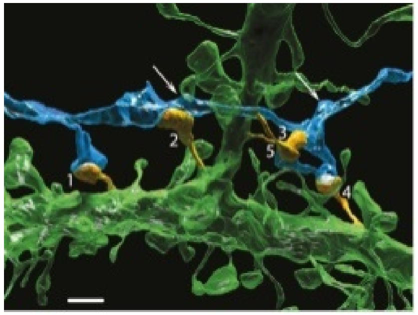
Lichtman Lab



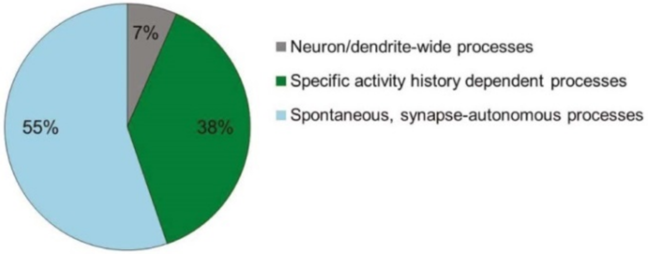
- This rewiring is based on an inherent **stochastic** component of synaptic plasticity, that even continues in the absence of neural activity

- STDP, Hebb, and other activity-dependent (deterministic) learning rules contribute at most 50% of the actual synaptic plasticity

estimated correlation of weigths of multiple synapses that connect the same axon with the same dendrite:
r = 0.23



Kasthuri et al., 2015



Dvorkin, Ziv, 2016

What do these biological data suggest for the design of continuously learning neuromorphic systems?

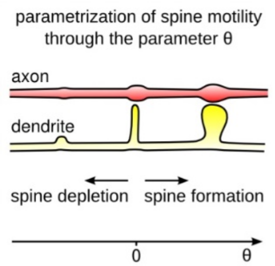
- We need to change our learning rules in order to accomodate rewiring

- We need new concepts and theories to design learning rules that integrate rewiring with synaptic plasticity in a goal-oriented manner

D. Kappel, R. Legenstein, S. Habenschuss, M. Hsieh and W. Maass
Reward-based self-conguration of neural circuits. 2017, in preparation.

A new conceptual and mathematical framework for integrating continuous rewiring into network plasticity

- We introduce a real-valued parameter θ_i for each potential synaptic connection i. This synaptic connections becomes functional when θ_i becomes positive, in which case $w_i = exp(\theta_i - \theta_0)$ is the synaptic weight (the exponential function provides a better fit to data)



- Plasticity of this potential synaptic connection is regulated by a **stochastic differential equations (SDE)** $$d\theta_i = \underbrace{\left(b \frac{\partial}{\partial \theta_i} \log p^*(\boldsymbol{\theta})\right)}_{drift} dt + \underbrace{\sqrt{2Tb} \cdot d\mathcal{W}_i}_{diffusion}$$

where $d\mathcal{W}_i$ denotes an infinitesimal step of a random walk, b = learning rate, T = temperature

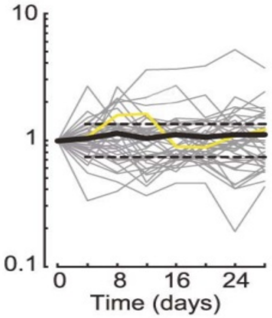
- The Fokker-Planck equation implies that $\frac{1}{Z}\boldsymbol{p}^*(\boldsymbol{\theta})^{\frac{1}{T}}$ is the resulting **stationary distribution** of the vector $\boldsymbol{\theta}$ of all these network parameters to which the stochastic system converges

  (but in general no convergence to a **particular** network configuration $\boldsymbol{\theta}$!).
- Hence the drift terms in the SDEs can „program“ a desired target performance into the network.
- I will focus on the case $\boldsymbol{p}^*(\boldsymbol{\theta}) \propto \boldsymbol{p_S}(\boldsymbol{\theta}) \cdot \boldsymbol{E}\,[total\ reward|\boldsymbol{\theta}]$

  where $p_S(\boldsymbol{\theta})$ is a prior that formalizes structural constraints (e.g., sparse connectivity).

A closer look at the SDE $d\theta_i = \left(b \frac{\partial}{\partial \theta_i} \log p^*(\boldsymbol{\theta})\right) dt + \sqrt{2Tb} \cdot d\mathcal{W}_i$

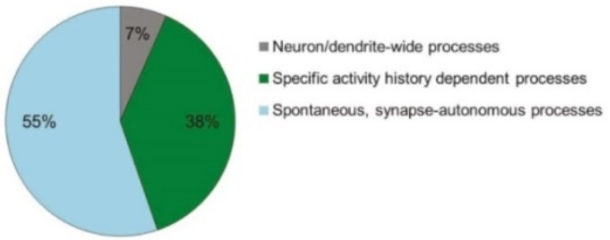
for $\boldsymbol{p}^*(\boldsymbol{\theta}) \propto \boldsymbol{p}_{\mathbf{S}}(\boldsymbol{\theta}) \cdot \boldsymbol{E}\,[total\ reward|\boldsymbol{\theta}]$



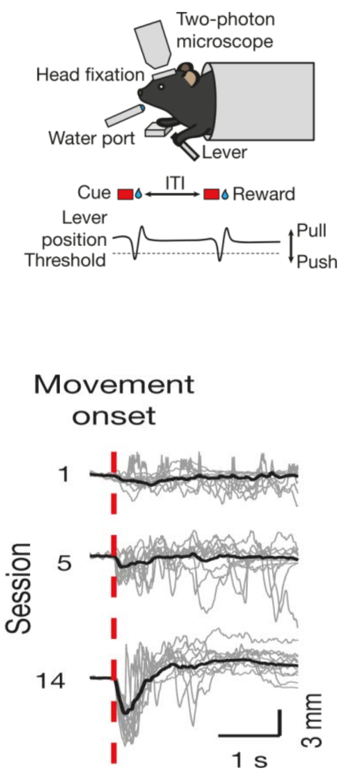
If one chooses a Gaussian for the prior $p_S(\theta)$, the derivative of its log models in conjunction with the diffusion term an Ornstein-Uhlenbeck process for θ_i i.e., for the log of the weight $w_i = exp(\theta_i - \theta_0)$, which fits data quite from the Lab of Noam ZIv quite well:

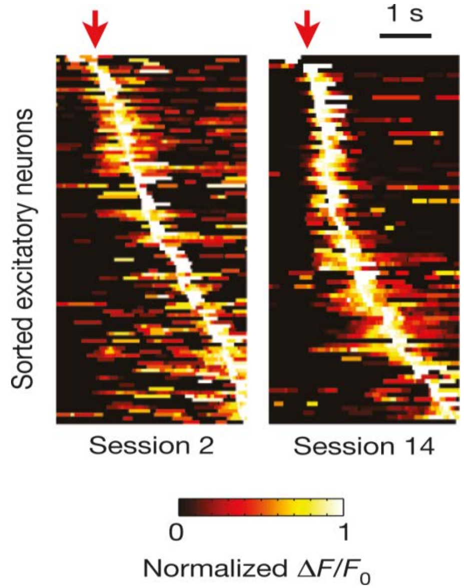
The derivative of the log of the second term $\boldsymbol{E}\,[total\ reward|\boldsymbol{\theta}]$ becomes significant only for θ_i >0. It then approximates standard rules for reward-gated STDP (with eligibility traces similar as reported in (Yagishita et al., Science 2014).

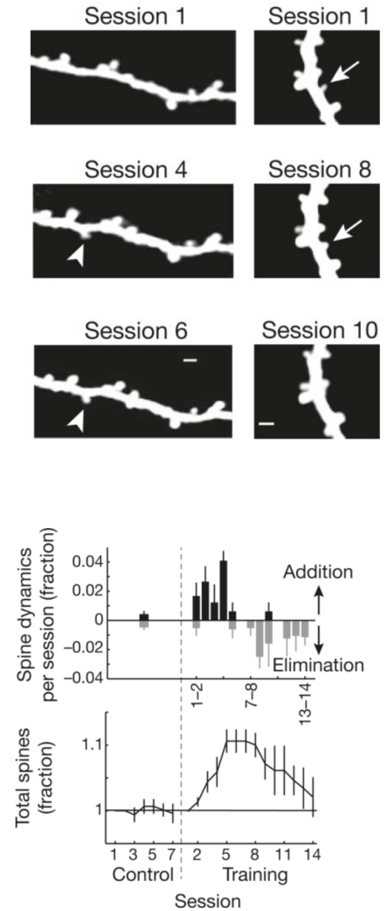
If the temperature T is sufficiently large, this model reproduces the experimentally found strong contribution of activity-independent synaptic plasticity



We have selected the following biological paradigm for a first test of this new model for network plasticity

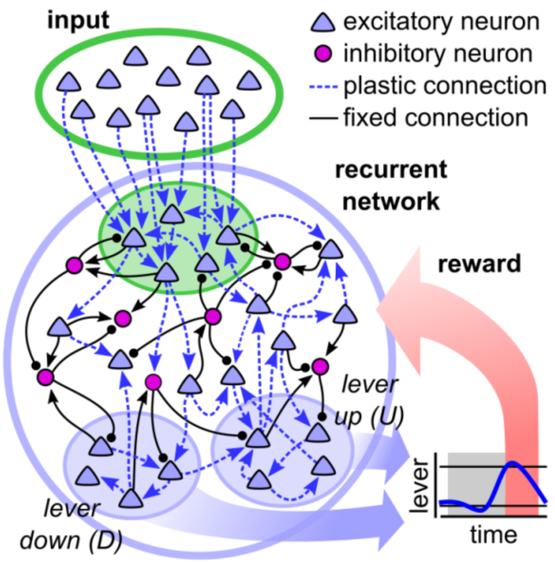


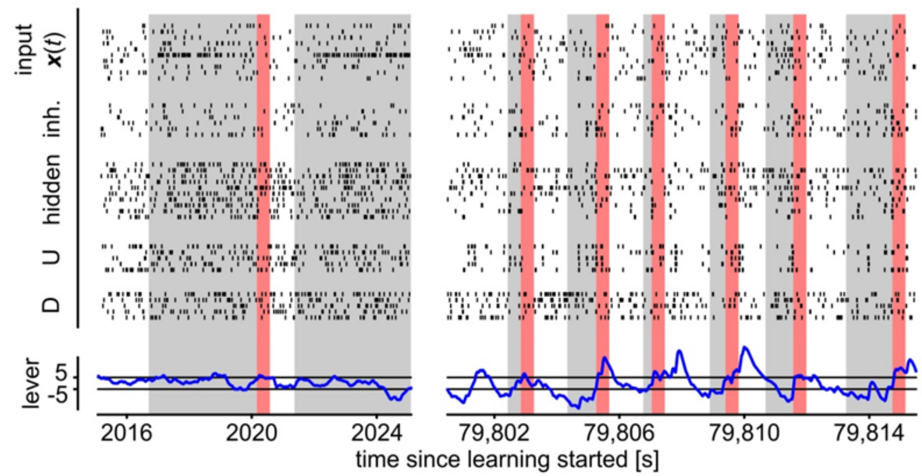


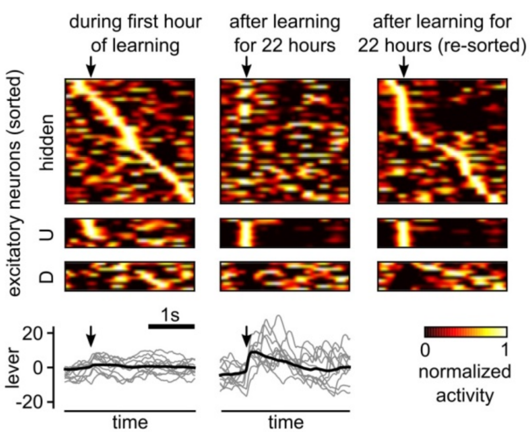


Emergence of reproducible spatiotemporal activity during motor learning;
Andrew J. Peters, Simon X. Chen & Takaki Komiyama; Nature(510) 2014

Our model qualitatively reproduces the experimental data

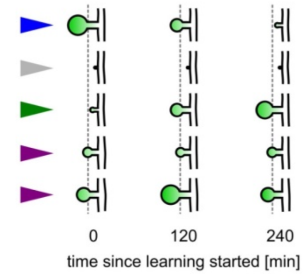






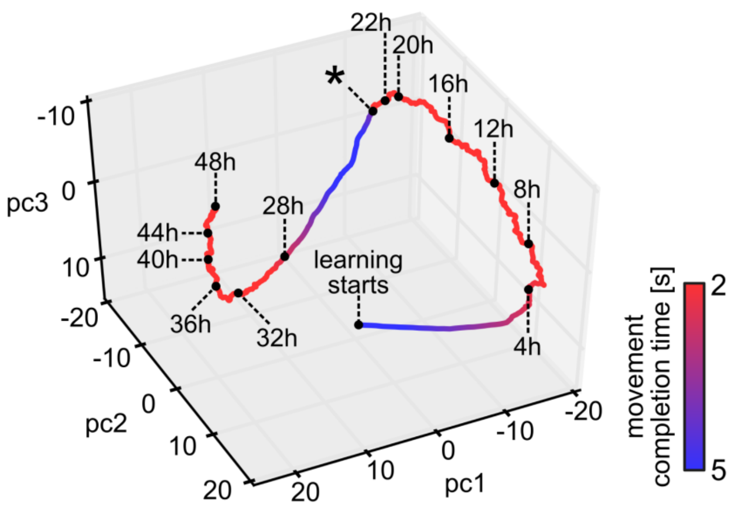
A stereotypical assembly dynamics ermerges during learning

some spines vanish, new ones emerge, and some of them become stable

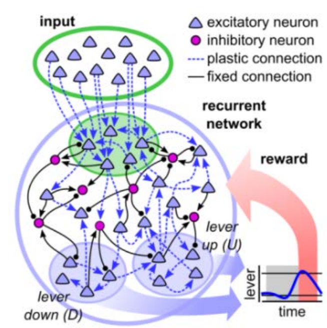


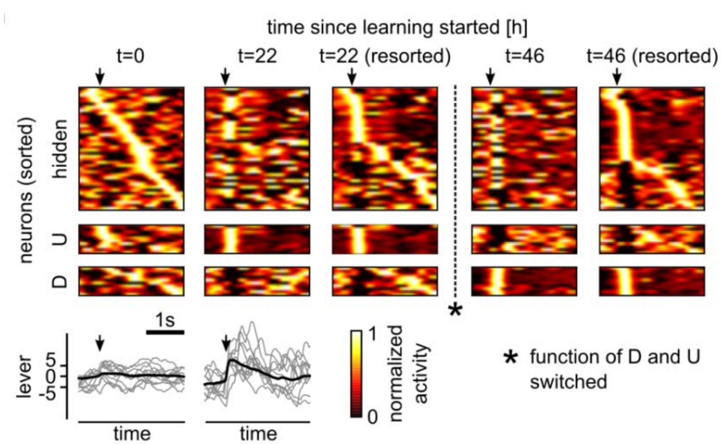
In addition, our model introduces lifelong learning capability into the neural network

The parameter vector $\boldsymbol{\theta}$ keeps moving even after good performance has been reached within some low-D manifold (red color indicates good performance).

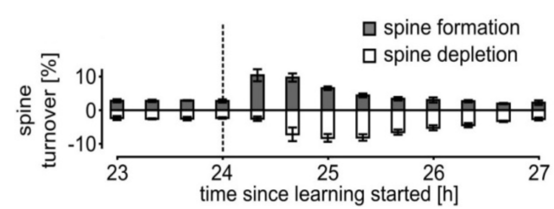


Functional benefit of ongoing stochastic parameter dynamics:
Immediate and automatic compensation for a drastic network perturbation: Switch of function of the populations U and D after 24h





This switch gives rise to a reorganization of network connections, and of the assembly dynamics, similar as observed in the biological data



A note on the learning (compensation-) speed of the model

This speed can be tuned by

--starting with a suitable network scaffold (e.g., reflecting genetically encoded aspects of brain networks)

--choosing (and adapting) suitable priors

--optimizing the sampling process (e.g., Hamiltonian, rather than Langevin)

--modulating the temperature T (like in simulated annealing)

Resulting new perspective of network learning from a more abstract perspective

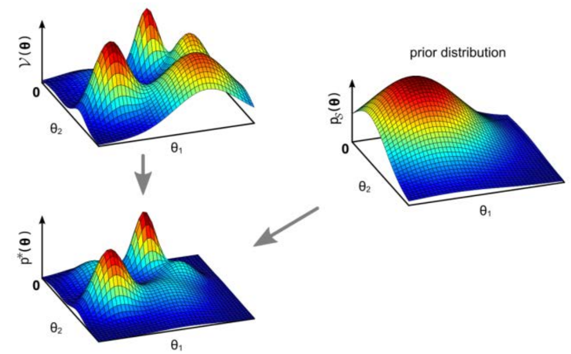
1. We arrive at a **Bayesian** model for network plasticity, where a prior (encoding e.g. structural constraints, innate knowledge, previously learned information....) modulates network plasticity

2. Gradient ascent in network network fitness is replaced by **stochastic sampling from a posterior distribution**

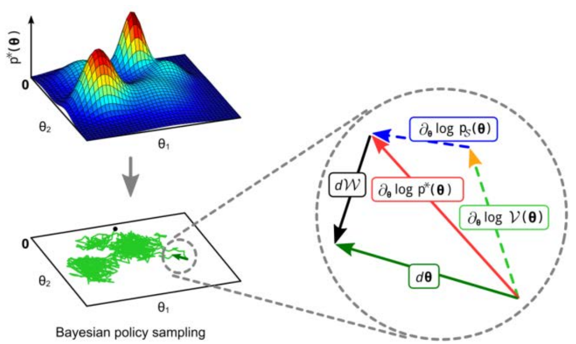
3. On the abstract level of reinforcement learning theory our model proposes to replace policy gradient by continuous **Bayesian policy sampling**

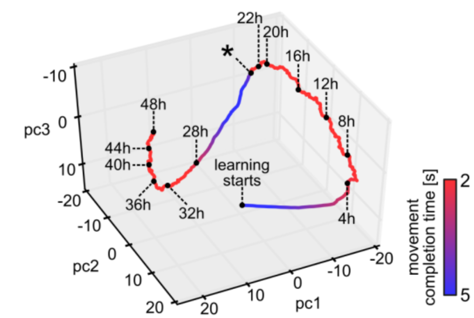
4. This continuous sampling aspect provides **automatic compensation for changes in the network or task**

expected reward



posterior





Summary

- Experimental data suggest a **significant difference** between the organization of neural network plasticity **in the brain** on one hand, and current models for network plasticity in neuromorphic systems and ANNs on the other hand

- Experimentally observed **continuously ongoing stochastic network reconfiguration in the brain** supports **exploration for self-organization and reinforcement learning**

- We propose that this model provides a step towards **lifelong autonomous learning capability** of neuromorphic systems