Learning to learn dynamical associations with reward-gated local plasticity

Dimitra Maoutsa
dimitra.maoutsa@gmail.com

ICLR International Conference On Learning Representations
Workshop on New Frontiers in Associative Memories

Introduction

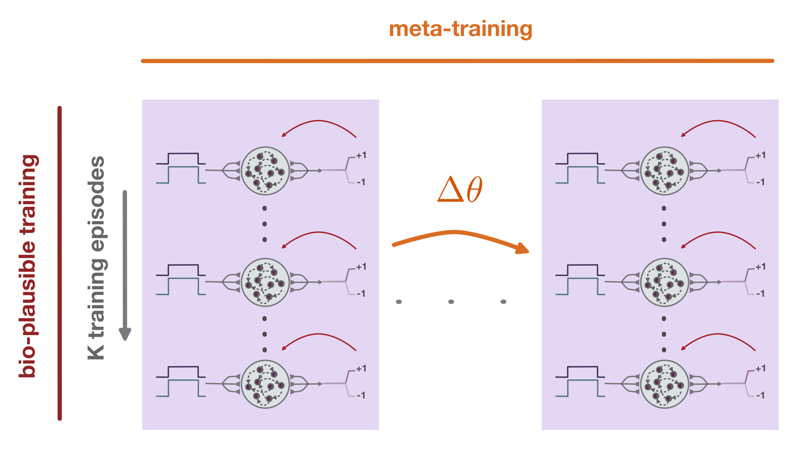
Task-trained recurrent networks optimized with standard methods (BPTT, FORCE) often converge onto similar solutions, although multiple alternatives could perform equally well. Yet, this **"simplicity bias"** may be a property of the learning rule rather than the task itself. We therefore ask here: how does **the plasticity rule variant shape the resulting solution space**, and **do different biologically plausible rules converge to different solution classes**? We first introduce a framework that discovers **local three-factor plasticity rules** to identify rules that **reliably support task acquisition**. We then probe the dynamical landscape and **emerging representations to characterize systemtic learning rule-solution biases.**

Meta-learning plasticity rules

optimization

Minimize loss

$$\mathcal{L}(\Theta) = -\langle \sum_h R^{(h)} \rangle$$



reinforcement learning

Policy $\pi_\Theta(\boldsymbol{\Delta W}|\mathcal{S})$ plasticity rule

Action $\boldsymbol{\Delta W}$

Reward $\sum_{h'=h+1}^{H-1} R^{h'}$

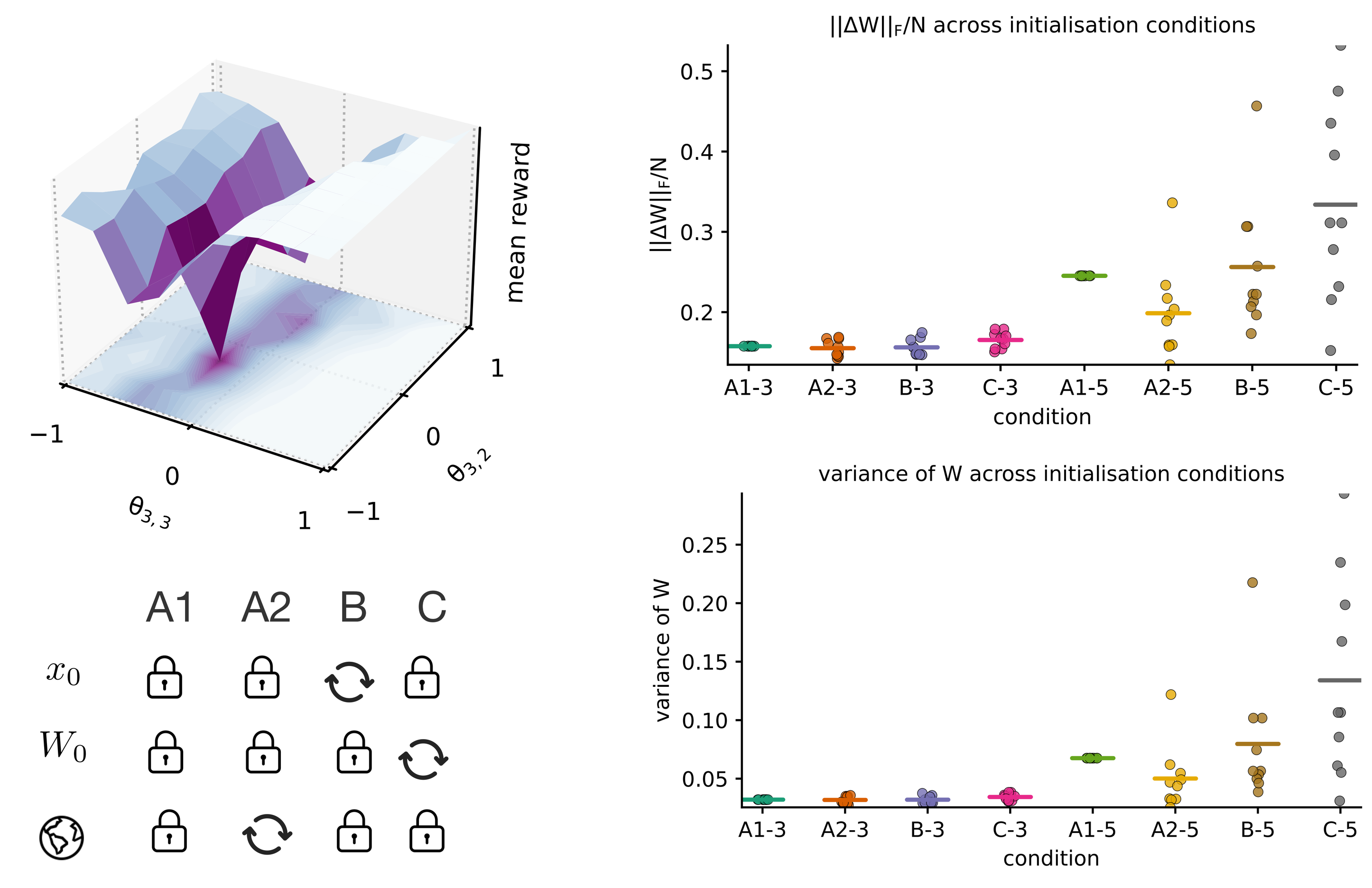
REINFORCE estimator

$$\nabla_\Theta \langle R \rangle \approx \langle (R-\bar{R}) \nabla_\Theta \log \pi_\Theta(\boldsymbol{\Delta W} \mid \mathcal{S}) \rangle$$

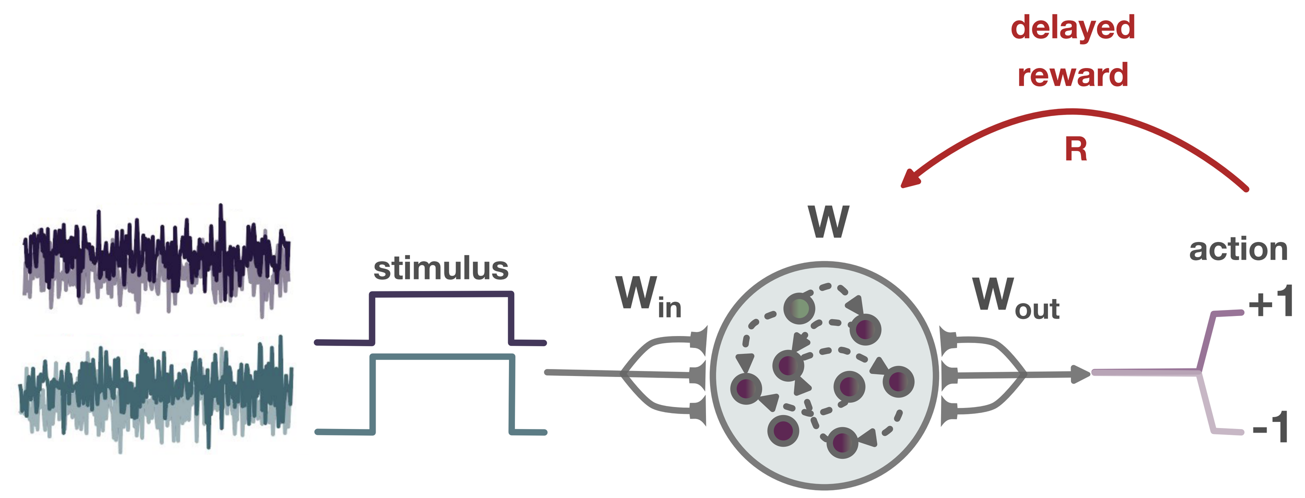
$$\frac{\partial}{\partial \theta_{k\ell}} \left\langle \sum_{h=1}^{H} R^{(h)} \right\rangle = \left\langle \sum_{h=1}^{H-1} \left(\sum_{h'=h}^{H} \delta R^{(h'+1)} \right) (\partial_{\theta_{k\ell}} \mu_{ij}^{(h)}) \frac{1}{\sigma_w} \xi_{ij} \right\rangle$$

explicit gradient too noisy and too computationally demanding for backprop

Discovering three-factor rules



Network and plasticity dynamics



network dynamics

$$\frac{d\mathbf{x}^t}{dt} = -\mathbf{x}^t + W\phi(\mathbf{x}^t) + W_{\text{in}}\mathbf{u}^t$$

$$\mathbf{r}^t = \phi(\mathbf{x}^t) \doteq \tanh(\mathbf{x}^t)$$

bio-plausible training

weight update $\Delta w_{ij} = \eta\, e_{ij}^T (R - \bar{R}) + \xi_{ij}$

j i

eligibility trace $\dot{e}_{ij}^t = \sum_{k,l} \theta_{k,l} (r_j^t)^k (\bar{x}_i^t - x_i^t)^l$

pre syn. rate — post syn. fluctuations

Subspace-sensitivities provide accurate gradient information

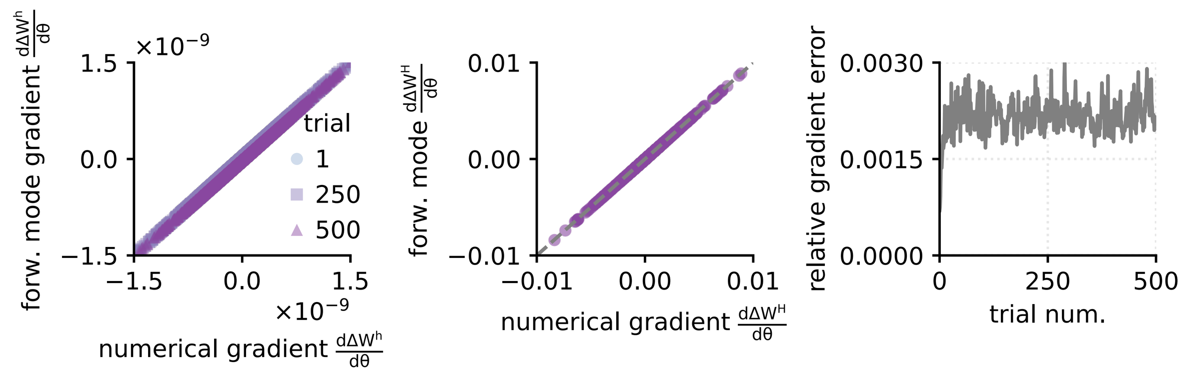
We define sensitivity parameters that we propagate throughout learning **Tangent propagation through learning**

state tangent
$$\boldsymbol{\chi}_{k,\ell}^{t+1} = \boldsymbol{\chi}_{k,\ell}^{t} + \alpha\Big(-\boldsymbol{\chi}_{k,\ell}^{t} + \mathbf{W}^{(h)} \left(\mathrm{diag}(\phi'(\mathbf{x}^t)) \cdot \boldsymbol{\chi}_{k,\ell}^{t}\right) + \mathbf{U}_{k,\ell}^{(h)}\, \mathbf{r}^t \Big)$$

trace tangent
$$\boldsymbol{\psi}_{k,\ell}^{t+1} = \alpha_x\, \boldsymbol{\psi}_{k,\ell}^{t} + (1-\alpha_x)\, \boldsymbol{\chi}_{k,\ell}^{t+1},$$

eligibility tangent
$$\mathbf{z}_{k,\ell}^{t+1} = \mathbf{z}_{k,\ell}^{t} + \mathrm{d}t\, (\boldsymbol{\Delta}\mathbf{x}^t)^{\ell} \otimes \left(\mathbf{r}^t\right)^{k} + \mathrm{d}t \sum_{\kappa,\lambda} \Big[\theta_{\kappa,\lambda}\, \lambda\, (\boldsymbol{\Delta}\mathbf{x}^t)^{\lambda-1} (\boldsymbol{\psi}_{k,\ell}^{t} - \boldsymbol{\chi}_{k,\ell}^{t+1}) \otimes \left(\mathbf{r}^t\right)^{\kappa} + \theta_{\kappa,\lambda}\, (\boldsymbol{\Delta}\mathbf{x}^t)^{\lambda} \otimes \kappa\, \left(\mathbf{r}^t\right)^{\kappa-1} \left(\mathrm{diag}(\phi'(\mathbf{x}^t)) \cdot \boldsymbol{\chi}_{k,\ell}^{t}\right) \Big]$$

weight tangent
$$\mathbf{U}_{k,\ell}^{(h+1)} = \mathbf{U}_{k,\ell}^{(h)} + \frac{\partial \boldsymbol{\mu}^{(h)}}{\partial \theta_{k,\ell}}$$



Different plasticity rules result in different representations and learning dynamics