Learning to learn with reward-gated local plasticity

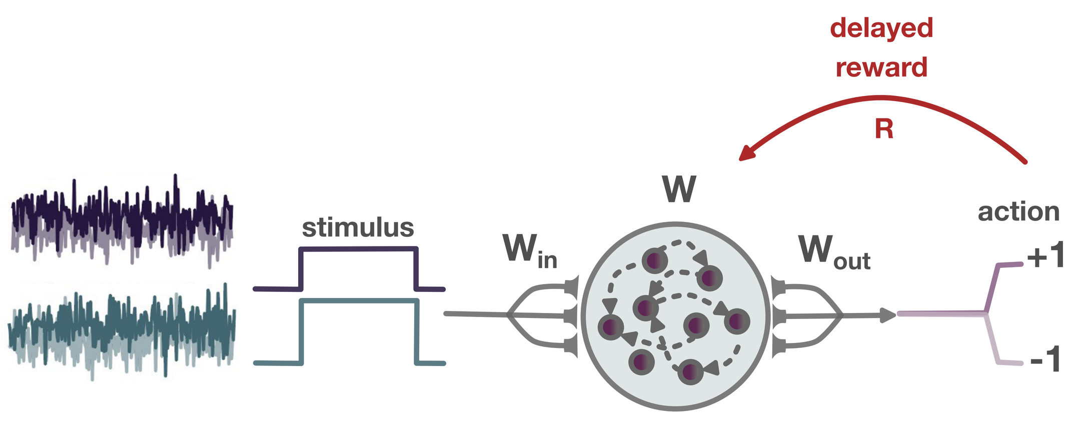
Dimitra Maoutsa
dimitra.maoutsa@gmail.com

Identifying three-factor learning rules that support common neuroscience task aquisition

Which plasticity rules promote task acquisition and which solutions do they learn?

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 that 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.

Discovering plasticity rules



network dynamics

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

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

bio-plausible training

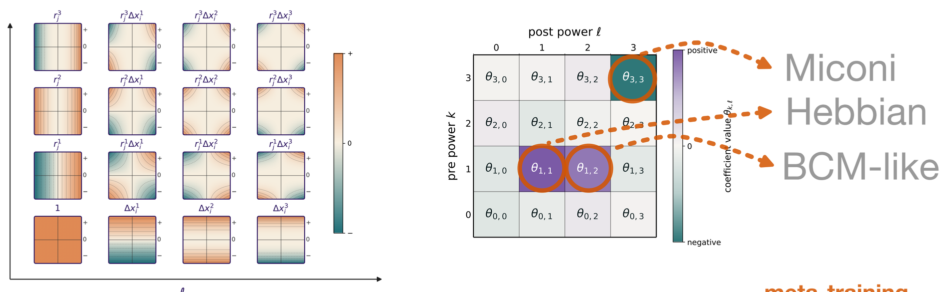
weight update $\Delta w_{ij} = \eta\, e_{ij}^T\,(R - \bar{R})$

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

parametrized rule

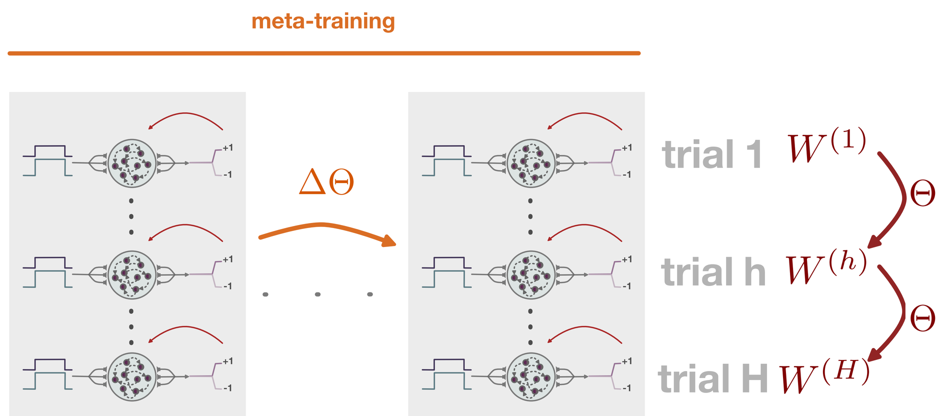


Common rules are a point in this plasticity space

optimization

Maximize objective

$$J(\Theta) = \sum_h \langle \gamma^h R^{(h)} \rangle_S$$



factorize objective

$$\frac{\partial J}{\partial \Theta} = \sum_h \frac{\partial J}{\partial W^{(h)}} \frac{\partial W^{(h)}}{\partial \Theta}$$

behavioral sensitivity — plasticity sensitivity

plasticity sensitivity

Define sensitivity parameters that we propagate throughout learning **Tangent propagation through learning**

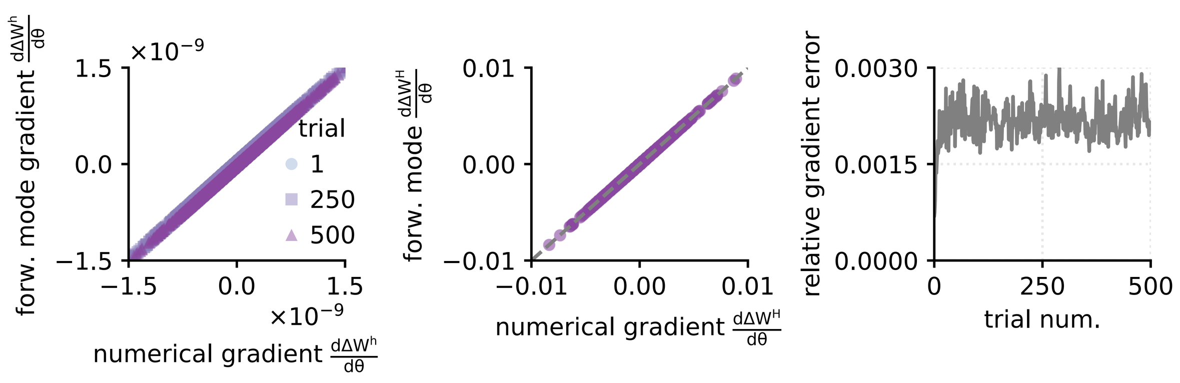
state tangent $\chi_{k,\ell}^{t+1} = \chi_{k,\ell}^t + \alpha\left(-\chi_{k,\ell}^t + \mathbf{W}^{(h)}\,(\text{diag}(\phi'(\mathbf{x}^t))\cdot\chi_{k,\ell}^t) + \mathbf{U}_{k,\ell}^{(h)}\mathbf{r}^t\right)$

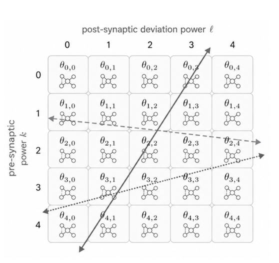
trace tangent $\psi_{k,\ell}^{t+1} = \alpha_x\,\psi_{k,\ell}^t + (1-\alpha_x)\,\chi_{k,\ell}^{t+1},$

eligibility tangent

$$\mathbf{z}_{k,\ell}^{t+1} = \mathbf{z}_{k,\ell}^t + \text{d}t\,(\boldsymbol{\Delta}\mathbf{x}^t)^\ell \otimes (\mathbf{r}^t)^k + \text{d}t\sum_{\kappa,\lambda}\Big[\theta_{\kappa,\lambda}\,\lambda\,(\boldsymbol{\Delta}\mathbf{x}^t)^{\lambda-1}(\psi_{k,\ell}^t - \chi_{k,\ell}^{t+1}) \otimes (\mathbf{r}^t)^\kappa + \theta_{\kappa,\lambda}\,(\boldsymbol{\Delta}\mathbf{x}^t)^\lambda \otimes \kappa\,(\mathbf{r}^t)^{\kappa-1}\,(\text{diag}(\phi'(\mathbf{x}^t))\cdot\chi_{k,\ell}^t)\Big]$$

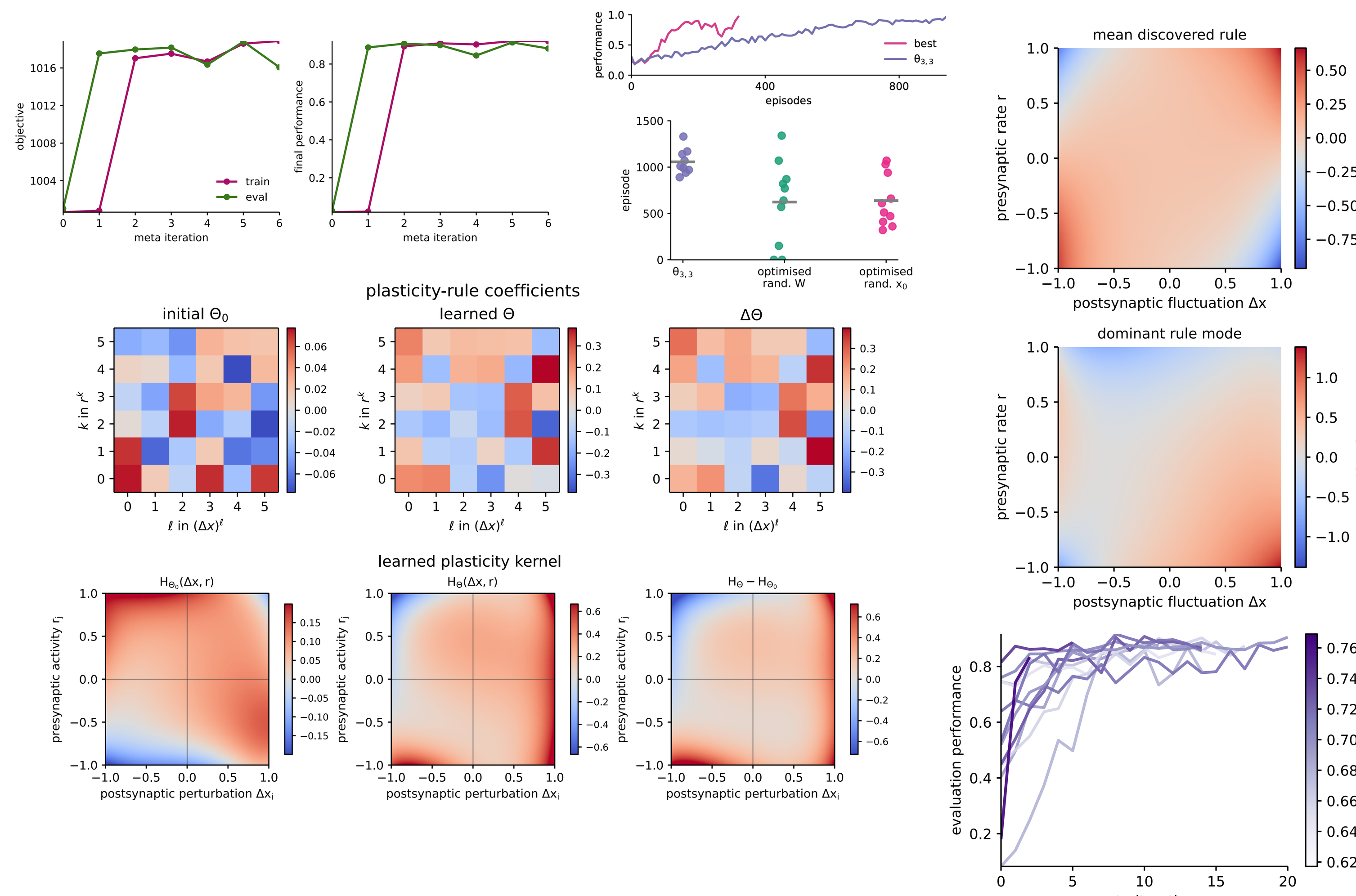
weight tangent $\mathbf{U}_{k,\ell}^{(h+1)} = \mathbf{U}_{k,\ell}^{(h)} + \frac{\partial \boldsymbol{\mu}^{(h)}}{\partial \theta_{k,\ell}}$





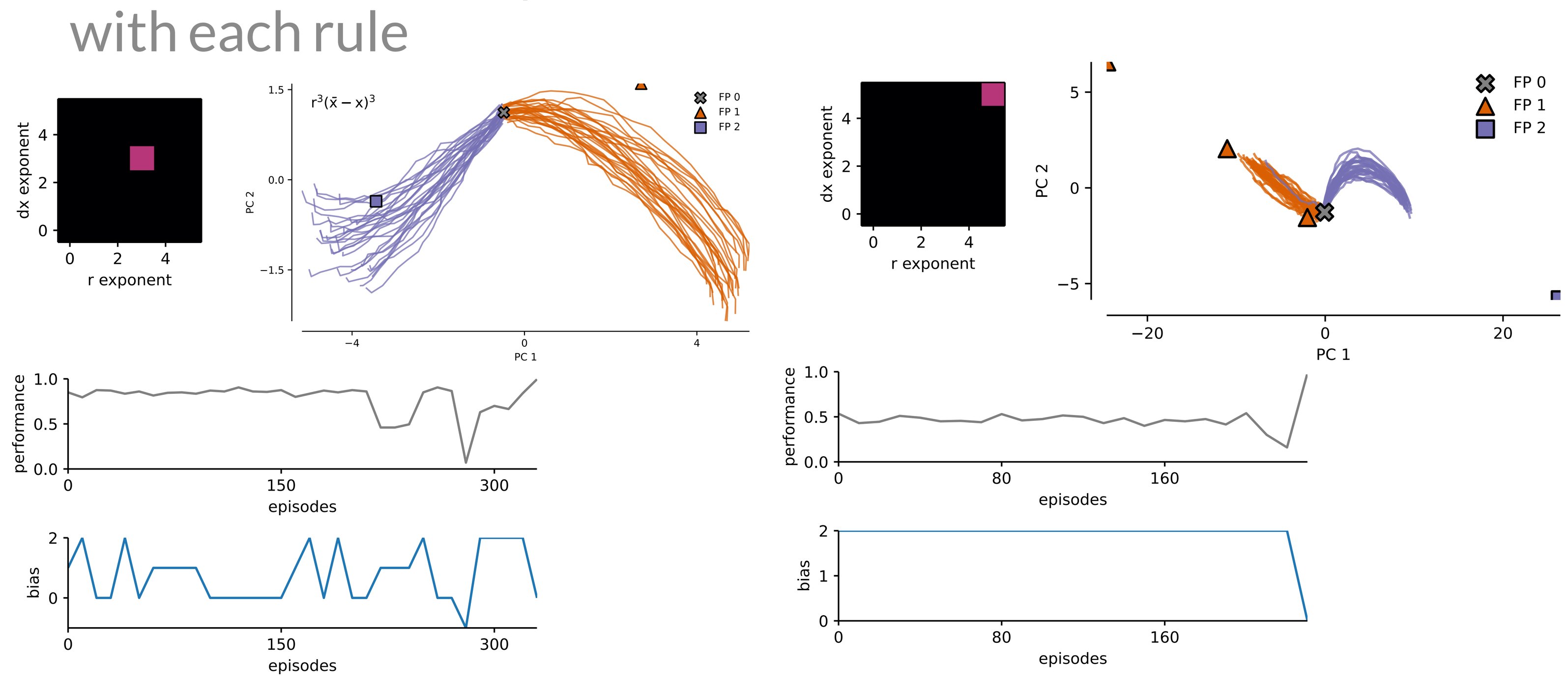
random projections in the plasticity parameter space reduce the computational cost of maintaining an extra 4 state system for every parameter θ $\tilde{\chi}_{(\omega)} = \sum_{k,\ell} u_{k,\ell}^{(\omega)} \chi_{k,\ell}$

Meta-optimization discovers task-learning rules



Characterizing identified rules and solutions

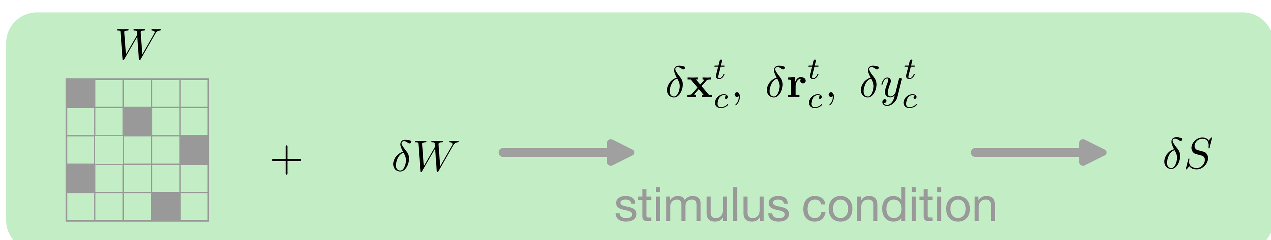
Characterize the dynamical structure of networks trained with each rule



Different rules result in different representational geometries, while the underlying dynamical landscape might be qualitatively similar

behavioral sensitivity

We introduce a differentiable scalar behavioral proxy that correlates with task performance, and use its gradient as a tractable proxy for how recurrent weight changes affect the objective



A small recurrent weight perturbation changes the condition-dependent trajectories

$$S(W) = \langle y_A^t - y_B^t \rangle_{t \in \mathcal{T}_{\text{dec}}}$$

proxy for behavioral sensitivity

$$G_S(W) = \frac{\partial S}{\partial W} \approx \frac{\partial J}{\partial W}$$