Inferring low-rank network models from neural activity

Adrian Valente[1], Jonathan Pillow[2, 3], Srdjan Ostojic[1]

1. Group for Neural Theory, LNC2, DEC, École Normale Supérieure, Paris 2. Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

3. Department of Psychology, Princeton University, Princeton, NJ, USA

Abstract

- Trained RNNs can offer good descriptions of collective neural activity but are difficult to interpret.
- Low-rank networks keep similar characteristics yet offer great interpretability.
- We present LINT *(Low-rank Inference from Neural Trajectories)*, a method to infer interpretable connectivity from recordings.
- Our method retrieves low-dimensional task-related subspaces, as well as computational mechanisms at the circuit level.

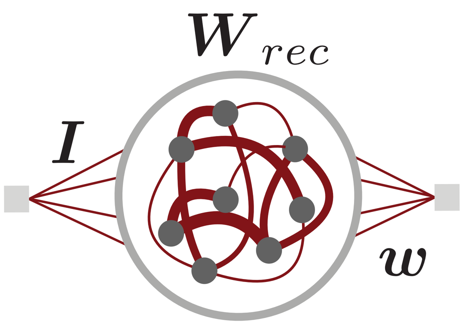
Classes of networks

flexibility

interpretability

Full-rank RNN: $\mathcal{O}(N^2)$ parameters

Trainable $\boldsymbol{W}_{rec}, \boldsymbol{I}, \boldsymbol{w}$



Low-rank RNN: $\mathcal{O}(N)$ parameters

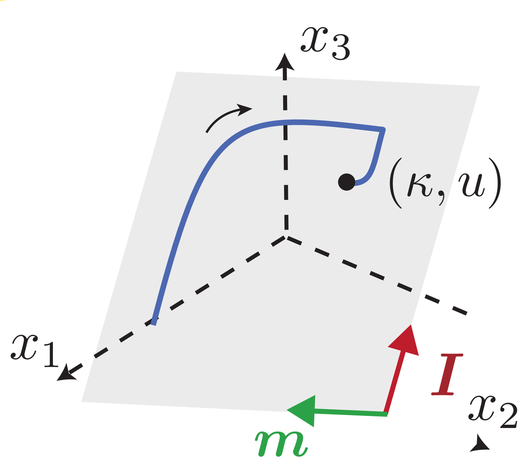
[Mastrogiuseppe & Ostojic 2018]

$$\boldsymbol{W}_{rec} = \sum_{r=1}^{R} \boldsymbol{m}^{(r)} \boldsymbol{n}^{(r)T}$$

Trainable $\boldsymbol{m}^{(r)}, \boldsymbol{n}^{(r)}$

Dynamics on R-dim subspace:

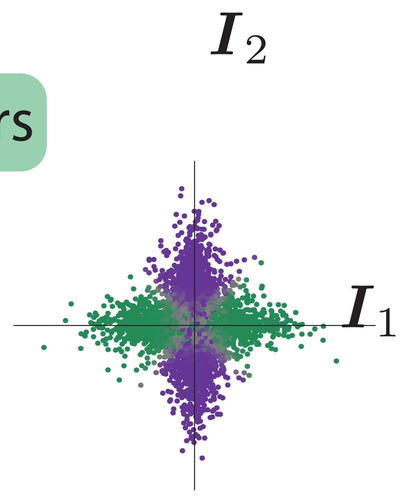
$$\boldsymbol{x}(t) = \sum_{r=1}^{R} \kappa_r(t) \boldsymbol{m}^{(r)} + \sum_{s=1}^{N_{in}} v_s(t) \boldsymbol{I}^{(s)}$$



Multi-population RNN: $\mathcal{O}(1)$ parameters

[Dubreuil, Valente et al. 2022]

$$(m_i^1, n_i^1, I_i, w_i) \sim \sum_{p=1}^{P} \mathcal{N}(\boldsymbol{\mu}_p, \boldsymbol{\Sigma}_p)$$



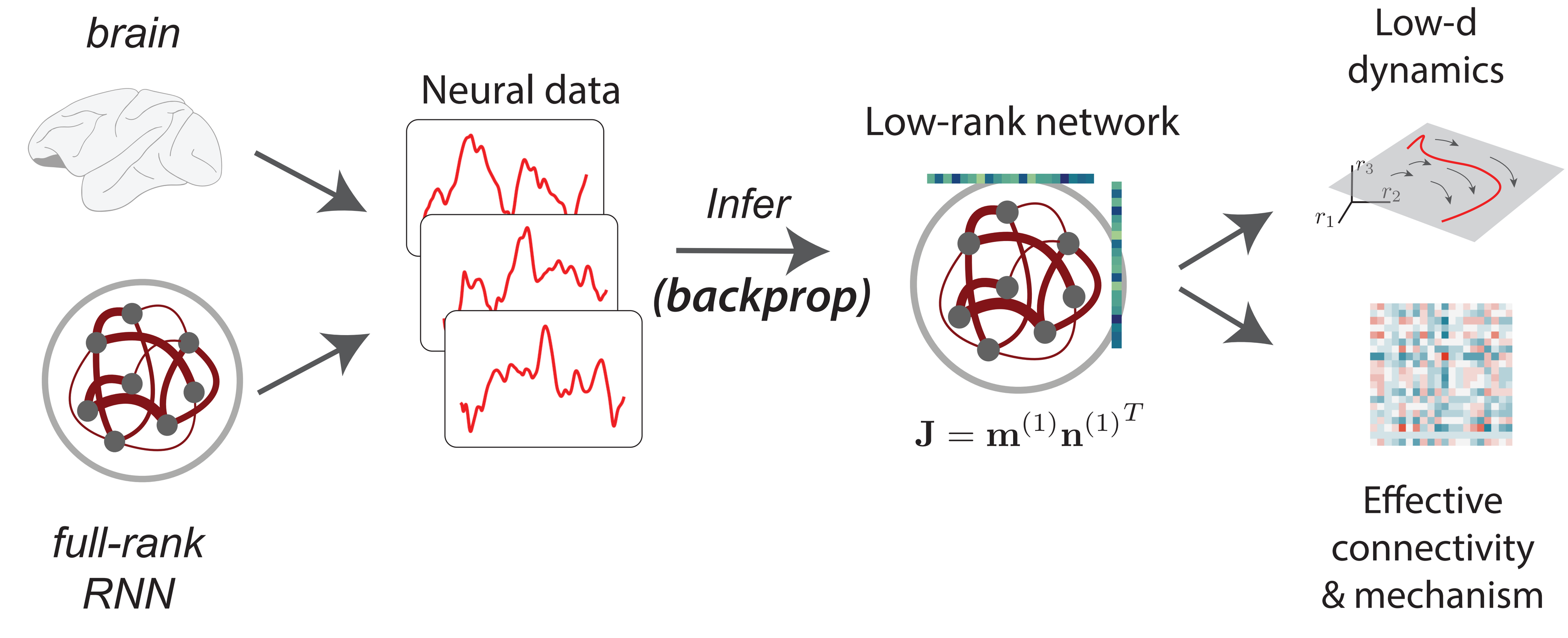
Trainable $(\boldsymbol{\mu}_p, \boldsymbol{\Sigma}_p)$

Analytically tractable dynamics:

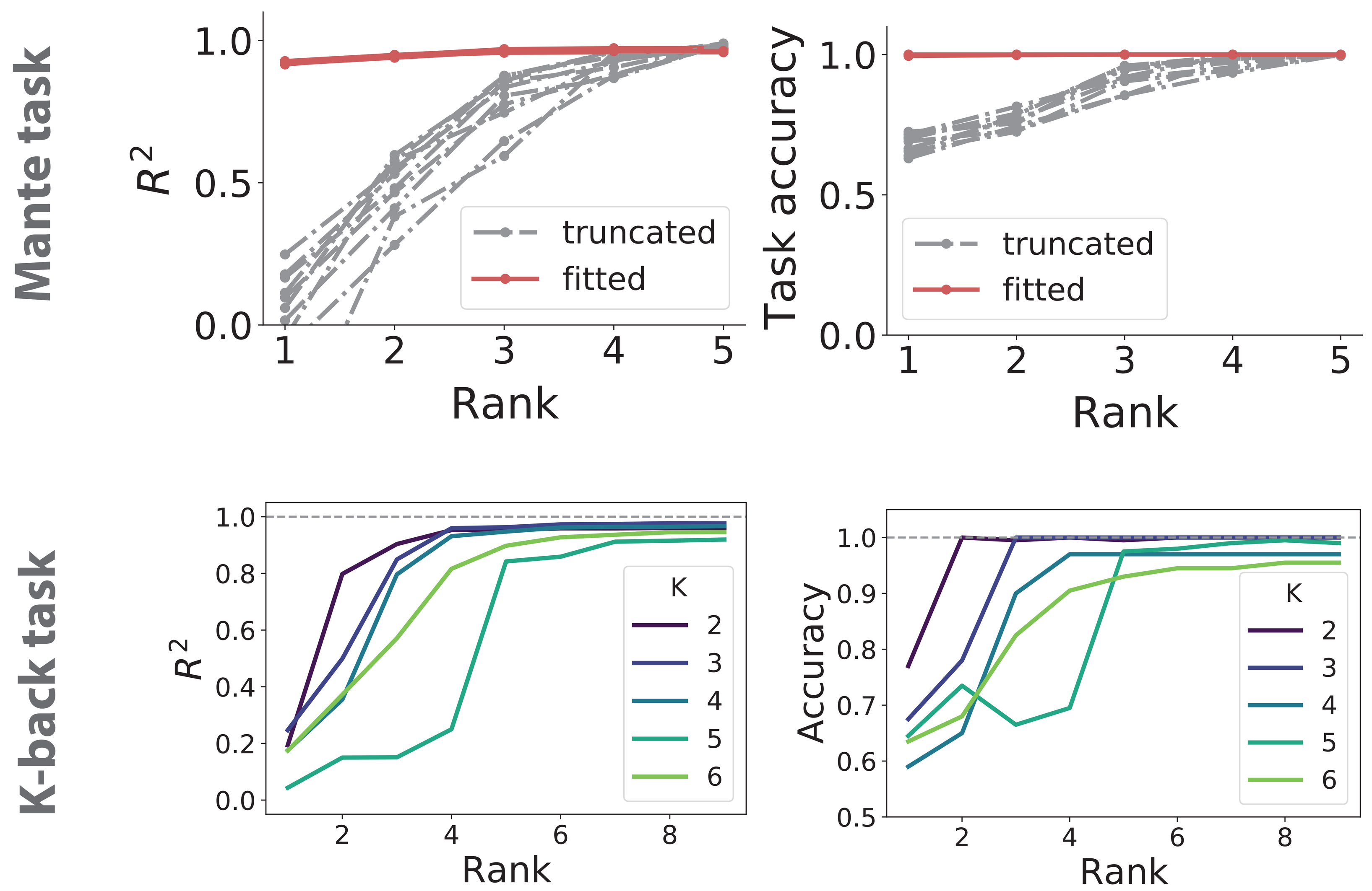
$$\dot{\kappa}_1 = -\kappa_1 + \tilde{\sigma}_{n_1 m_1} \kappa_1 + \tilde{\sigma}_{n_1 m_2} \kappa_2 + \tilde{\sigma}_{n_1 W_{in}} u(t)$$
$$\dot{\kappa}_2 = -\kappa_2 + \tilde{\sigma}_{n_2 m_1} \kappa_1 + \tilde{\sigma}_{n_2 m_2} \kappa_2 + \tilde{\sigma}_{n_2 W_{in}} u(t)$$

with functional connectivities: $\tilde{\sigma}_{ab} = \sum_{p=1}^{P} \sigma_{ab}^{p} \langle \phi' \rangle_p$

Approach: fit low-rank RNNs to neural recordings

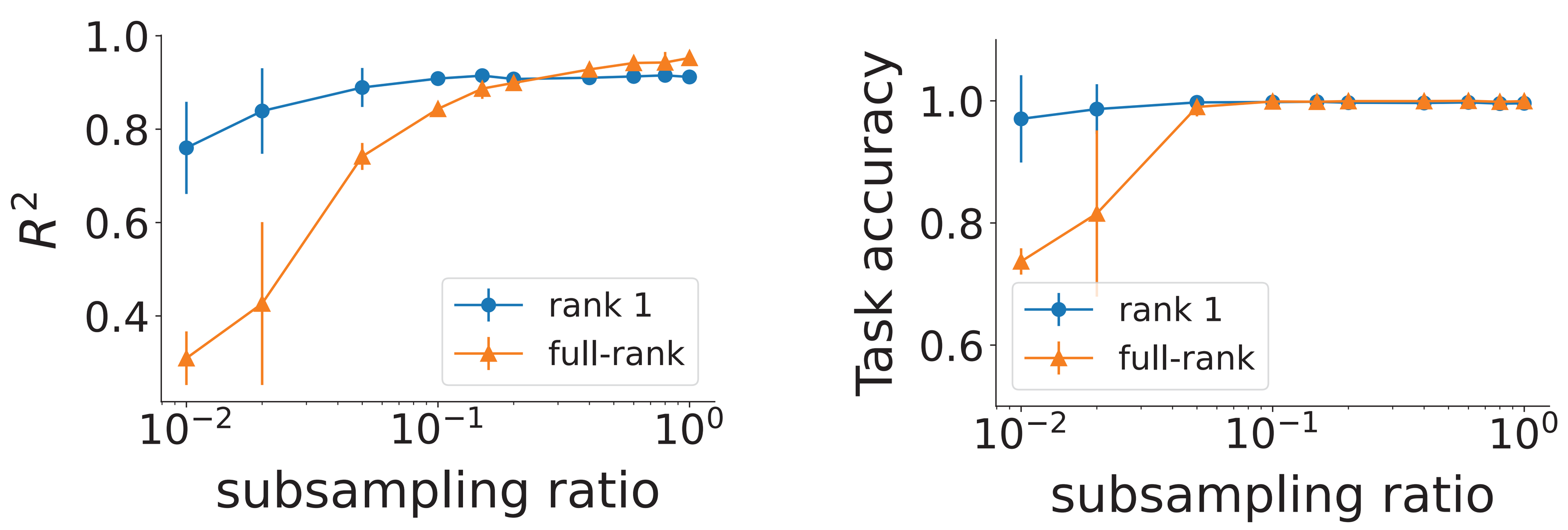


Reproducing full-rank nets with low-rank ones



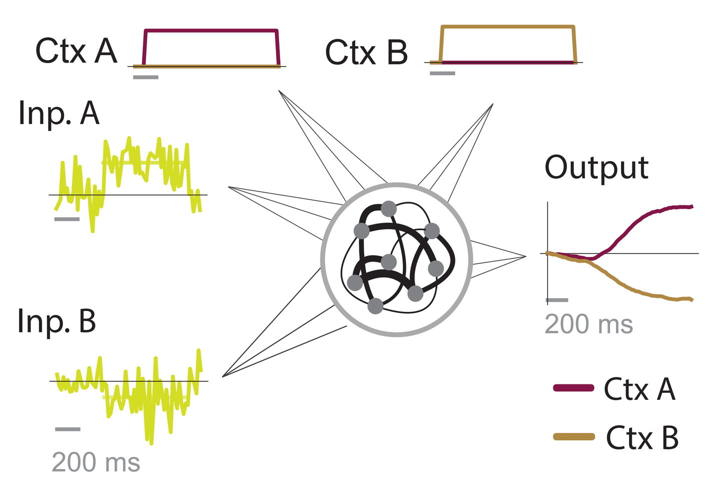
Neuron subsampling

Results for networks fitted to subsampled trajectories of a full-rank, 1000 neurons network

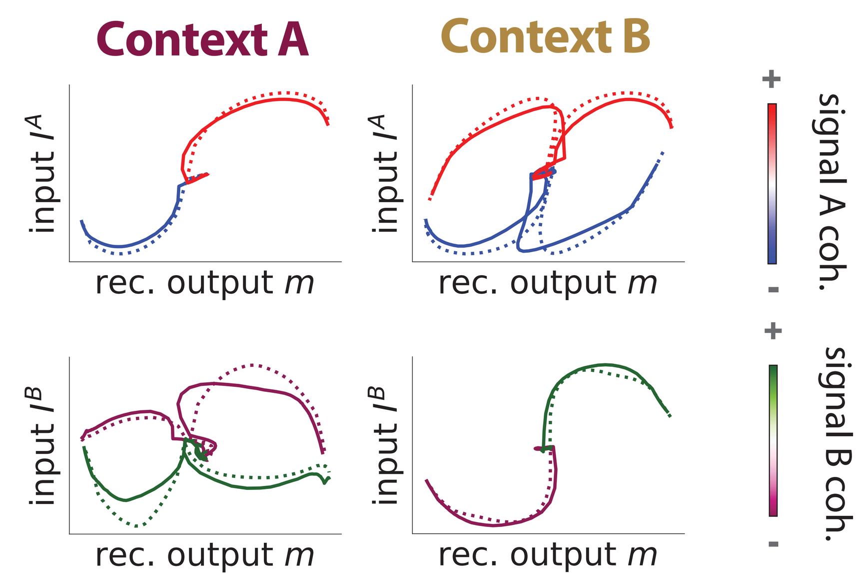


"Opening the black box" of a full-rank network

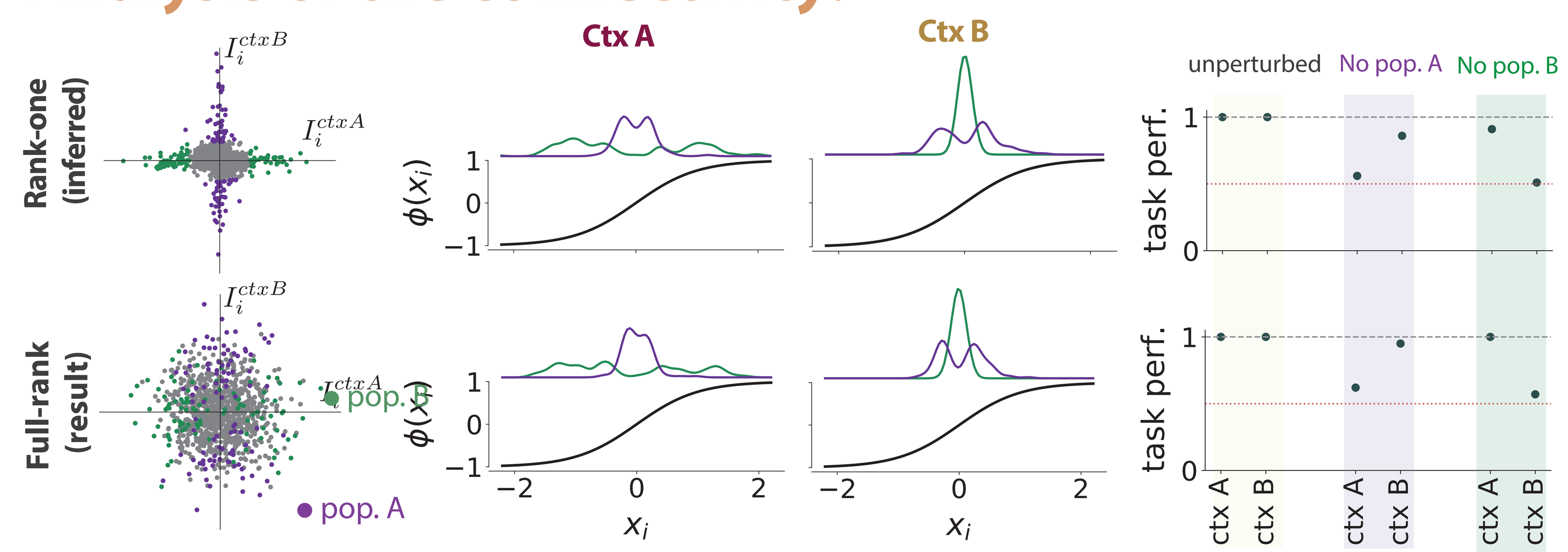
Task:



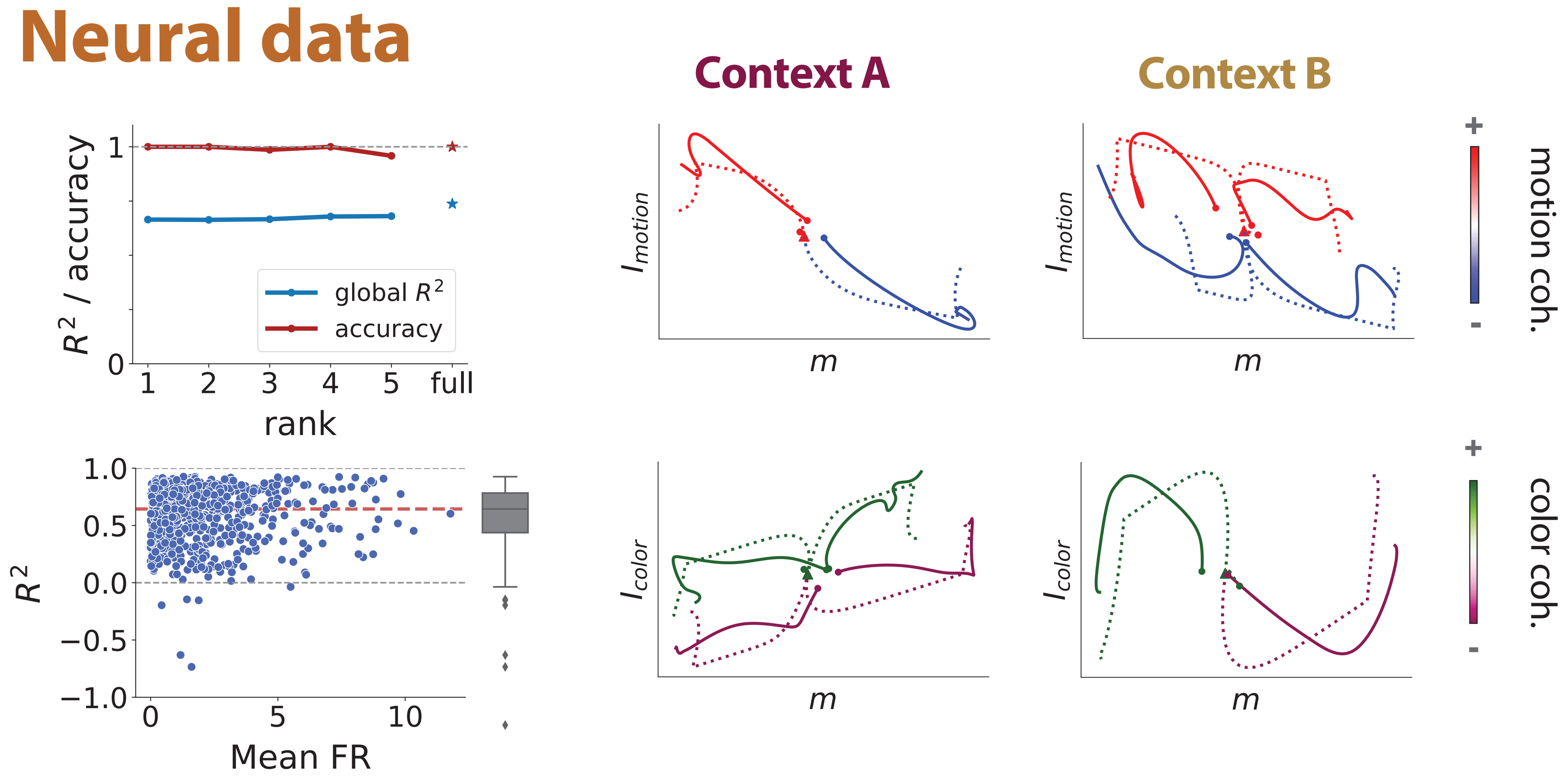
Low-dimensional dynamics:



Analysis of the connectivity:



Neural data



Take-home messages

- Networks with a very **low-rank connectivity** can capture the main aspects of dynamics of unconstrained RNNs and recorded neural data.
- They provide a way to dissect mechanisms in black-box RNNs and make them more **interpretable**.
- They bridge **dimensionality reduction** with the **computational requirements** of the studied task.

Bibliography

Mastrogiuseppe & Ostojic, Linking Connectivity, Dynamics, and Computations in Low-Rank RNNs, Neuron, 2018

Beiran et al., Shaping Dynamics With Multiple Populations in Low-Rank Recurrent Neural Networks, Neural Computation, 2021

Dubreuil, Valente et al., The role of population structure in computations through neural dynamics, Nat Neuro, 2022

Schuessler et al., The interplay between randomness and structure during learning in RNNs, Neurips 2020