Using artificial neural networks to uncover relationships between network structures and computations

Maturation and Plasticity in Biological and Artificial Neural Networks
Cargese, October 2024

Alexis Dubreuil
CNRS, Bordeaux University

BORDEAUX IMN neurocampus

cnrs

université de BORDEAUX

Outline

➡ Introduction: cognitive modeling

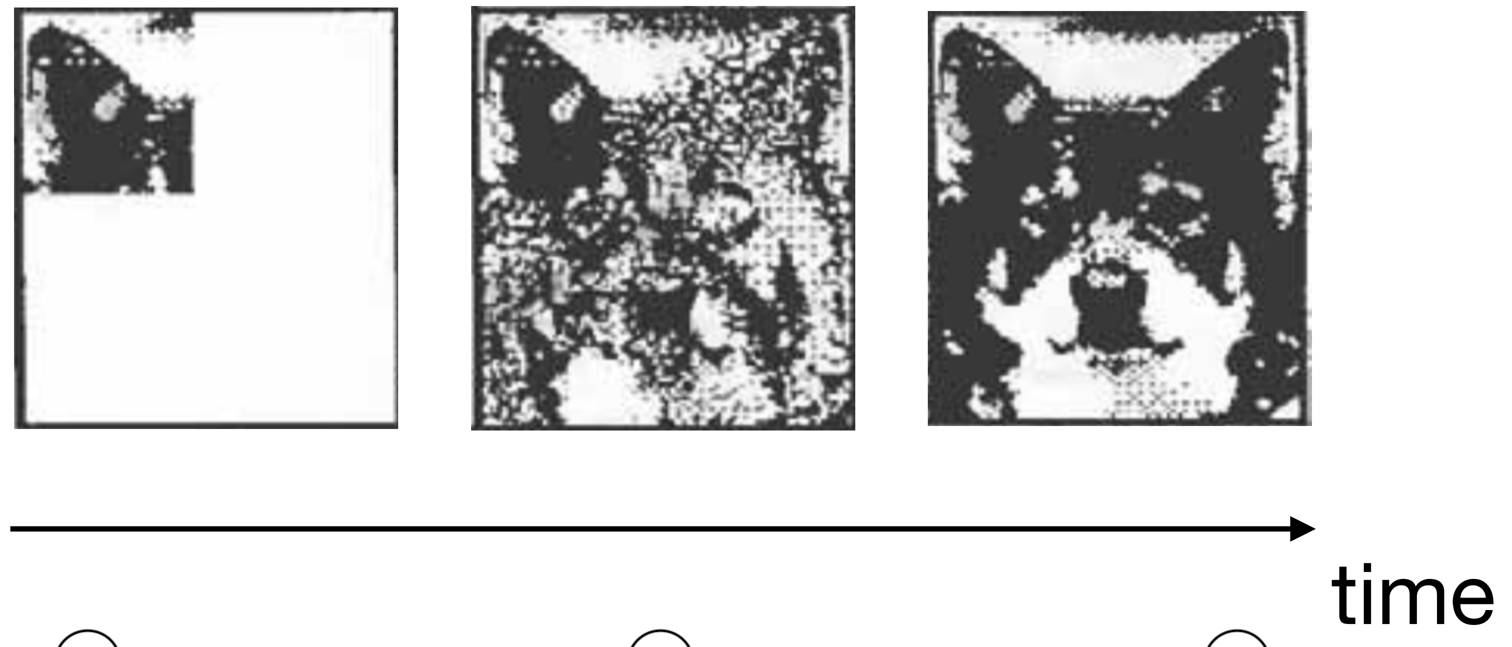
Theory-based reverse-engineering of artificial neural networks

On the computational role of population structure

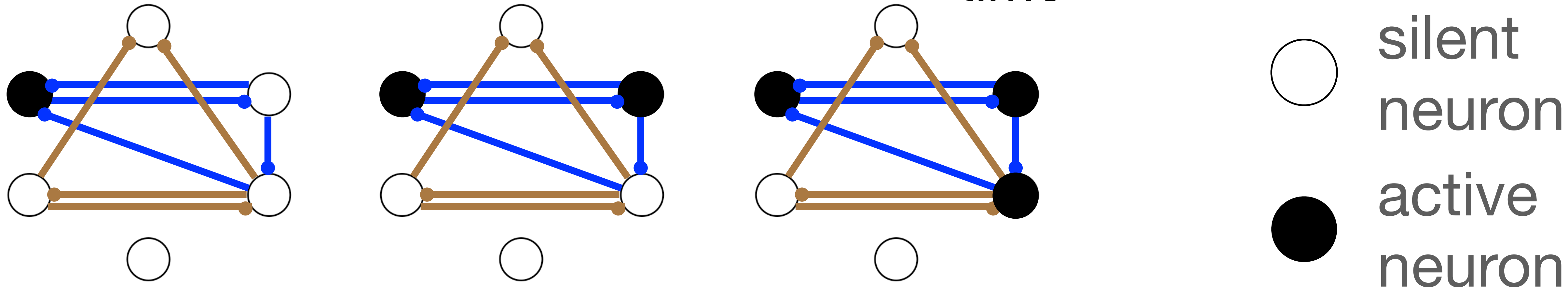
Modeling temporally structured behaviors

Neural implementation of working memory:

Associativity of working memory:



Neural implementation:
positive feed-back loops



$\kappa_A = 0.25$ $\kappa_A = 0.5$ $\kappa_A = 1$

Formalisation:

synaptic structure $$W_{rec,ij} = r_i^{\mu=1} r_j^{\mu=1} + r_i^{\mu=2} r_j^{\mu=2} + \ldots$$

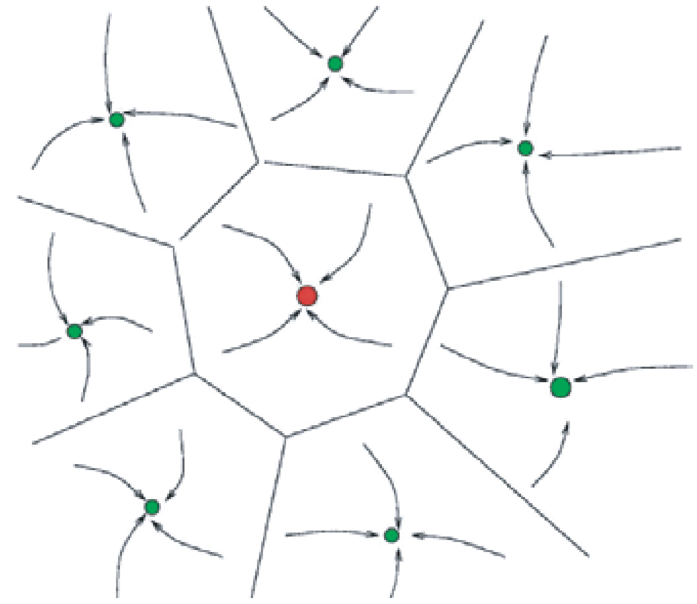
electrical activity $$\frac{d\kappa_A}{dt} = F(\kappa_A(t))$$

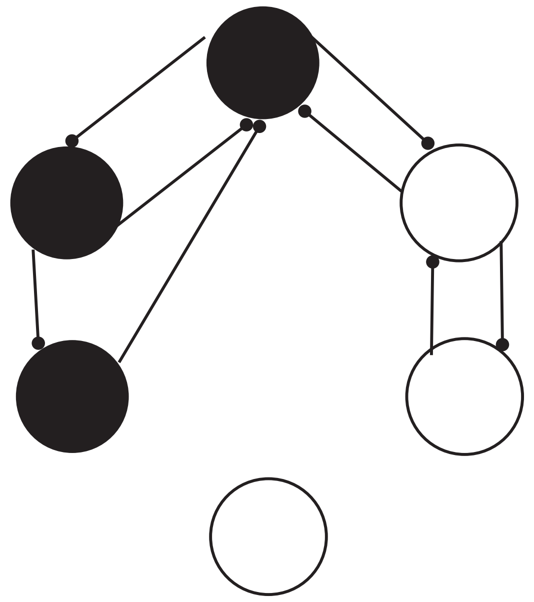
Hopfield, 1982

Derrida, Gardner, Zippelius,1987

Neural implementation of working memory:

Associativity of working memory





cognitive properties

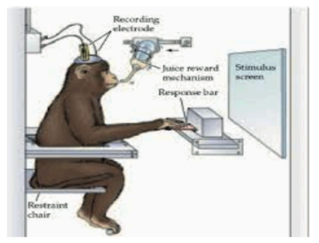
Trajectories of network configurations

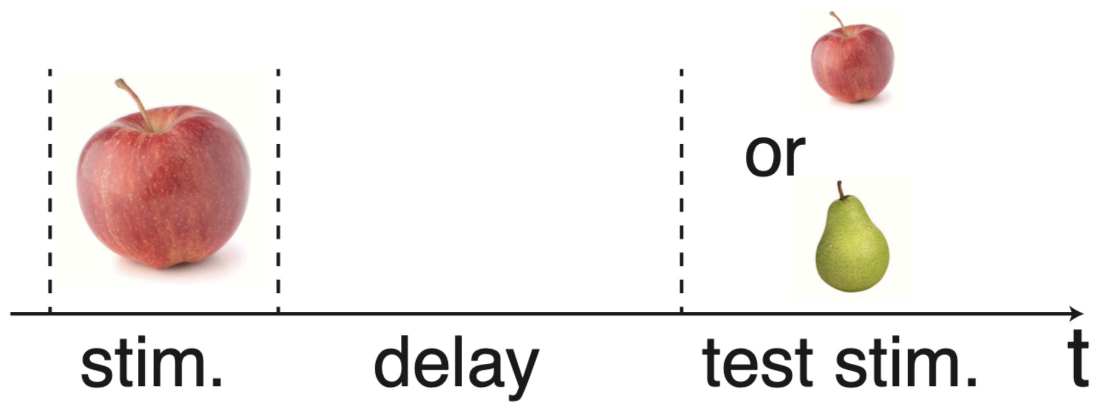
$$\frac{d\kappa_A}{dt} = F(\kappa_A(t))$$

Synaptic structure

Relationship with properties of brain circuits

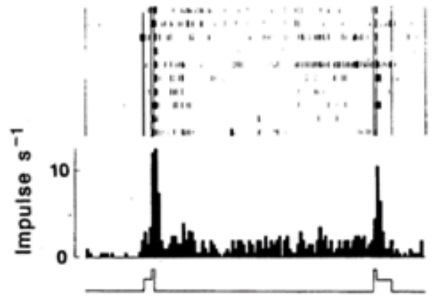
➡ Physiological correlates

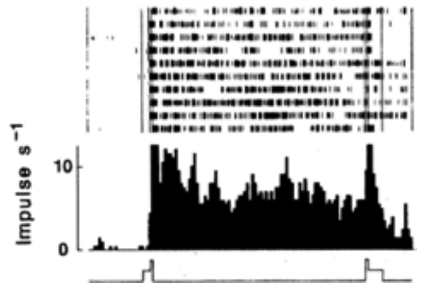




example neuron 1

example neuron 2

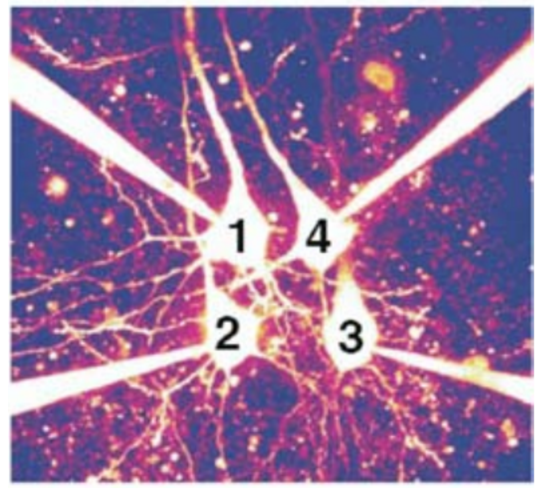




Miyashita, 1988

➡ Anatomical correlates

PFC local connectivity



Proba. of connections = 0.1

Proba. of reciprocal connections = 0.04 > 0.01

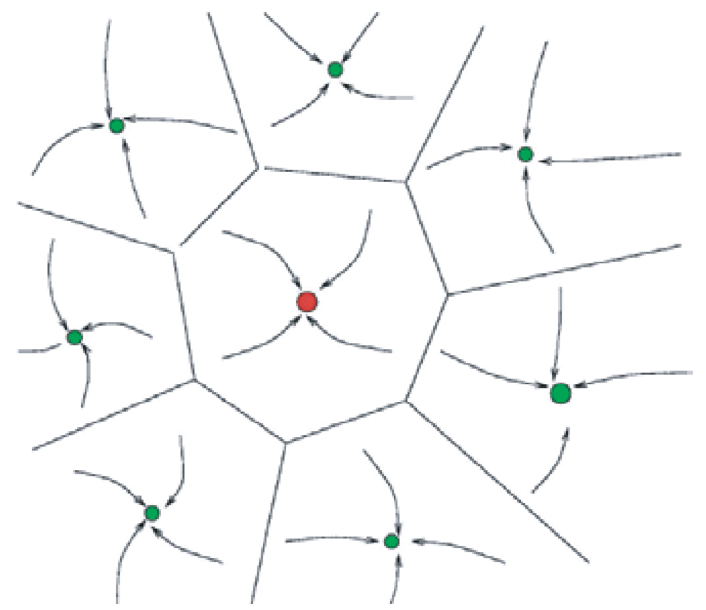
Wang et al, 2006

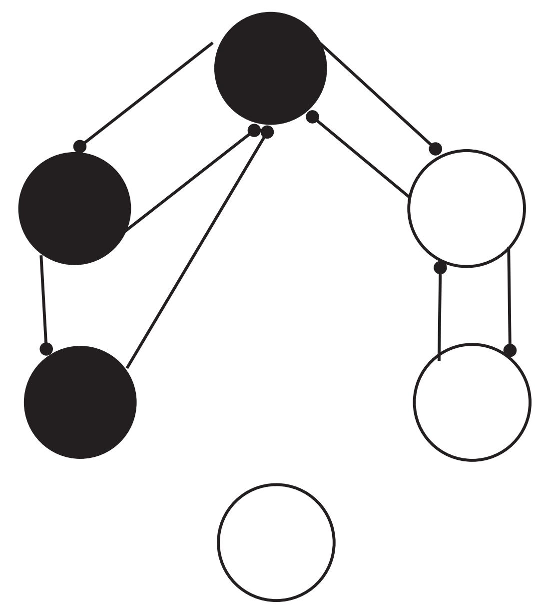
Matches statistics of ANN optimizing memory storage

Brunel 2016

Neural implementation of working memory:

Associativity of working memory





cognitive properties

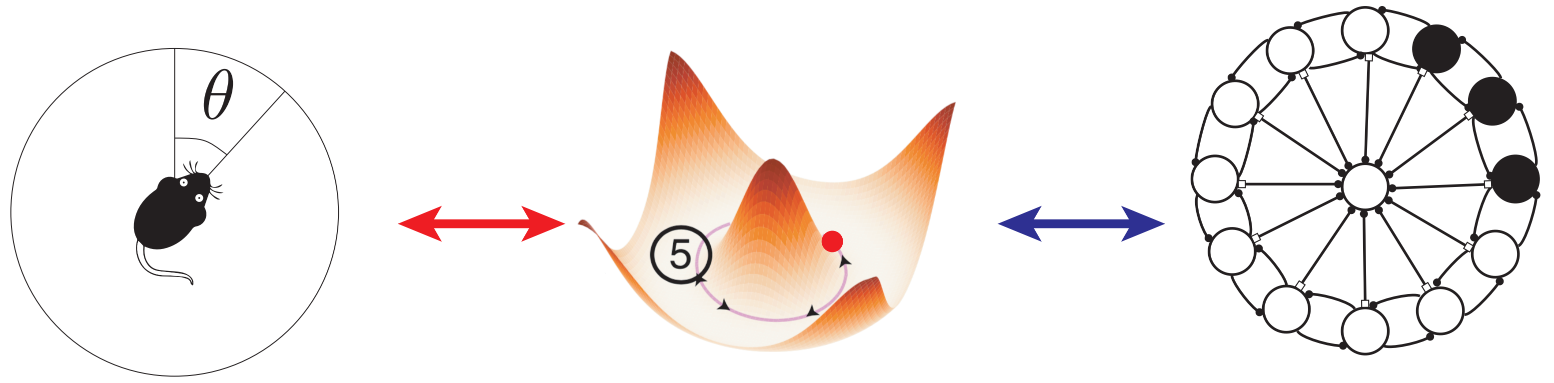
Trajectories of network configurations

$$\frac{d\kappa_A}{dt} = F(\kappa_A(t))$$

Synaptic structure

Modeling of other cognitive phenomena

- Sense of direction and ring model



Amari, 1979 ; Skaggs 1995

Kim, Rouault et al., 2017

- Sense of location and place cells or grid cells

Tsodyks & Sejnowski 1994

Burak & Fiete, 2009

- Decision making

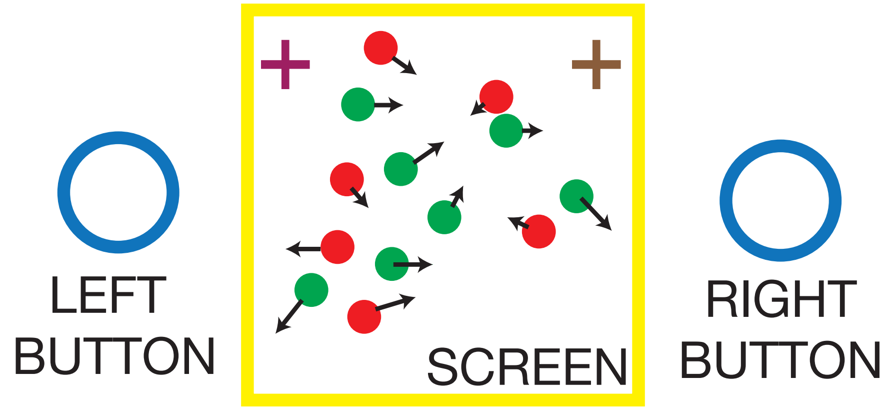
Wang, 2008

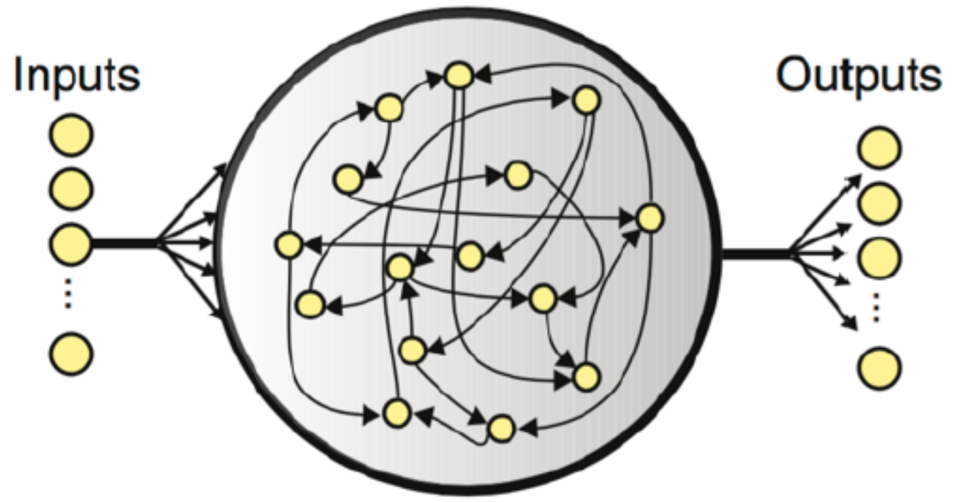
Modeling cognitive processes with artificial neural networks

Train artificial neural networks to perform neuroscience tasks

moving left or right ?

more green (left) or red (right) ?

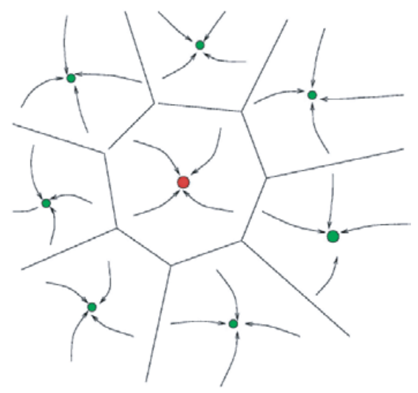


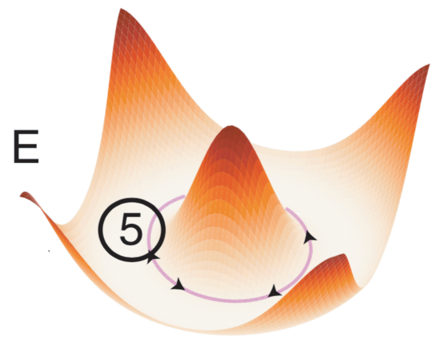


Sussillo, 2014

Barak, 2017

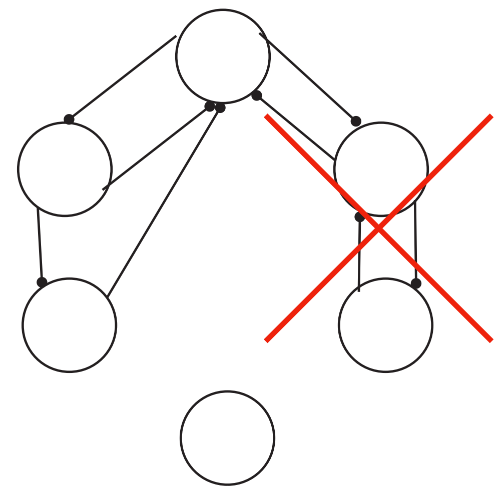
Reverse-engineering trained networks:





$$\frac{d\kappa_A}{dt} = F(\kappa_A(t))$$

Sussillo & Barak, 2013



Yang et al, 2019

➡ No unified framework to link macroscopic dynamics and network’s structure

Outline

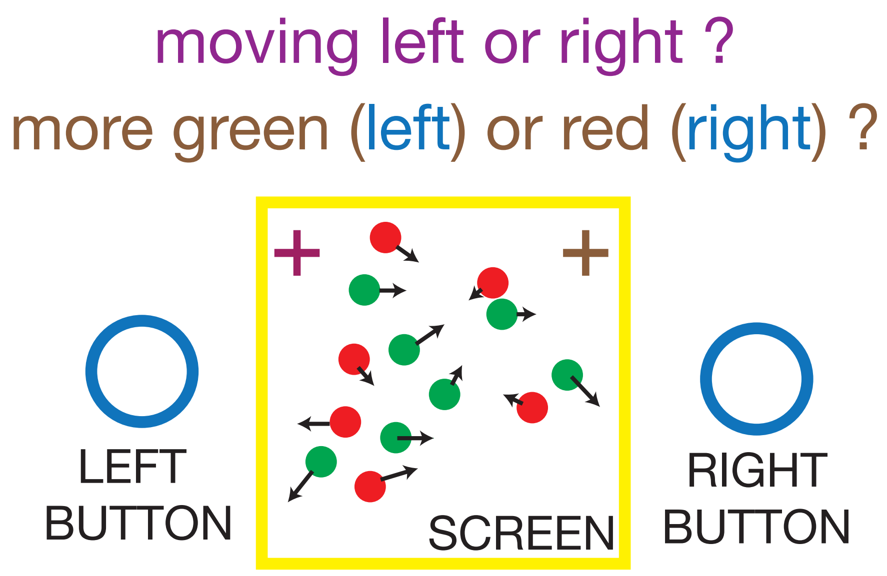
Introduction: cognitive modeling

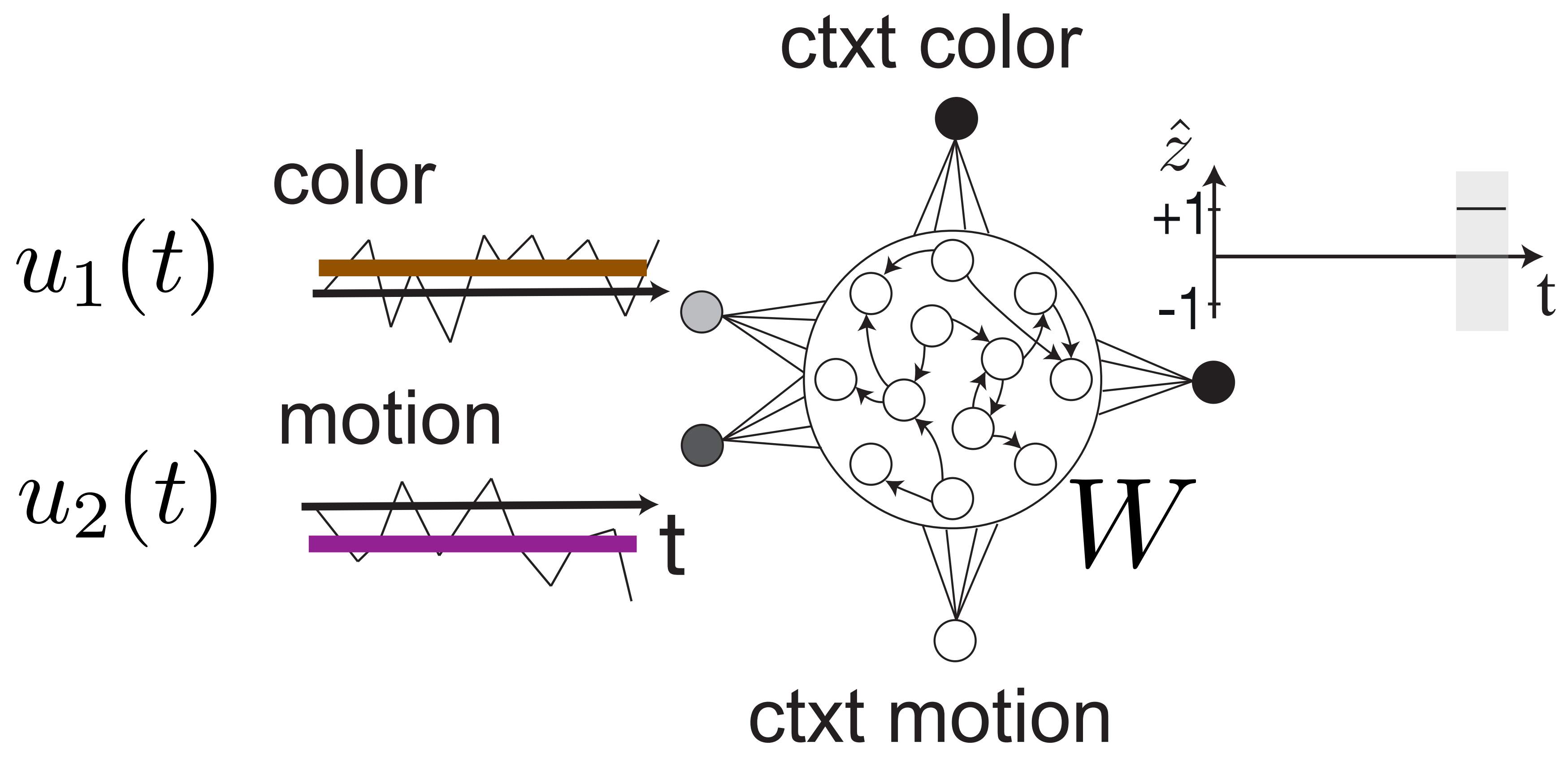
➡ Theory-based reverse-engineering of artificial neural networks

On the computational role of population structure

Modeling temporally structured behaviors

Training artificial neural networks on neuroscience tasks





➡ Recurrent neural networks (amorphous connectivity structure): N neurons

➡ Learning algorithms engineer artificial neural networks:

- training dataset + cost-function

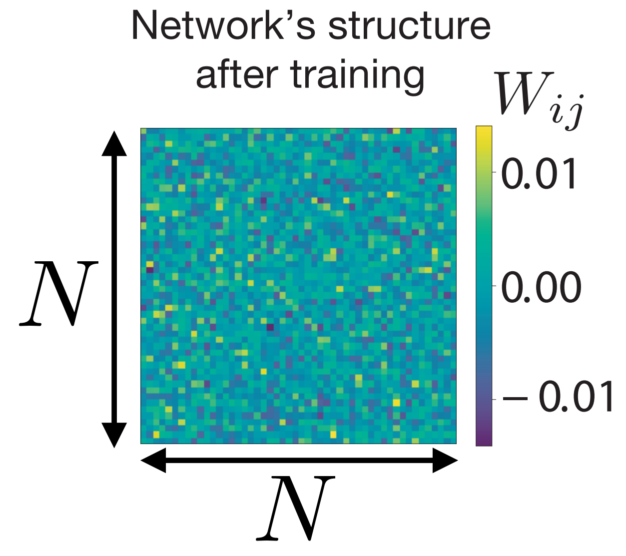
$\{\{$  $,+1\};\{$  $,\ 0\};...\} = \{\{i^{\mu=1},\hat{z}^{\mu=1}\};\{i^{\mu=2},\hat{z}^{\mu=2}\};...\}$

$$\mathcal{L}_W = \sum_{\mu} (z^{\mu}_W - \hat{z}^{\mu})^2$$

- optimize cost-function over network parameters

compute $\nabla \mathcal{L}_W$; update W ; repeat…

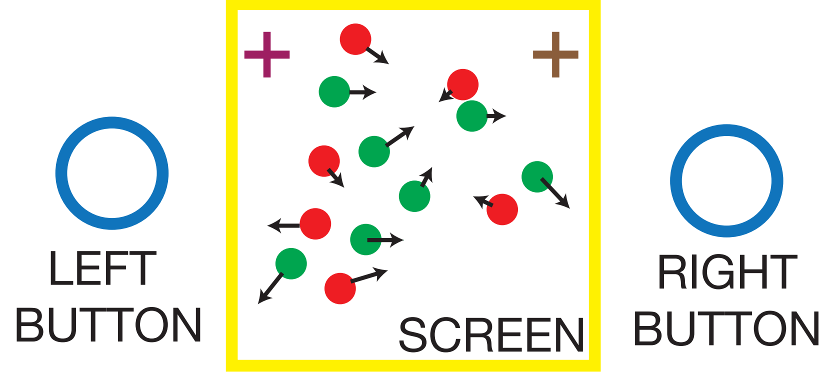
➡ After training:



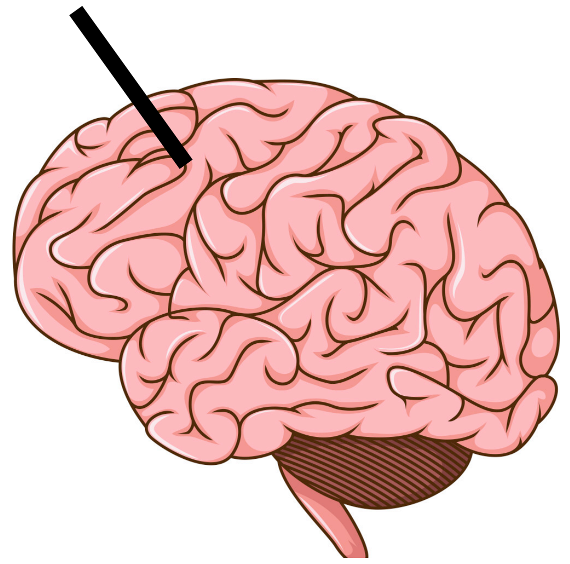
$$N^2 = 10^4/10^6$$

Reverse-engineering brain networks

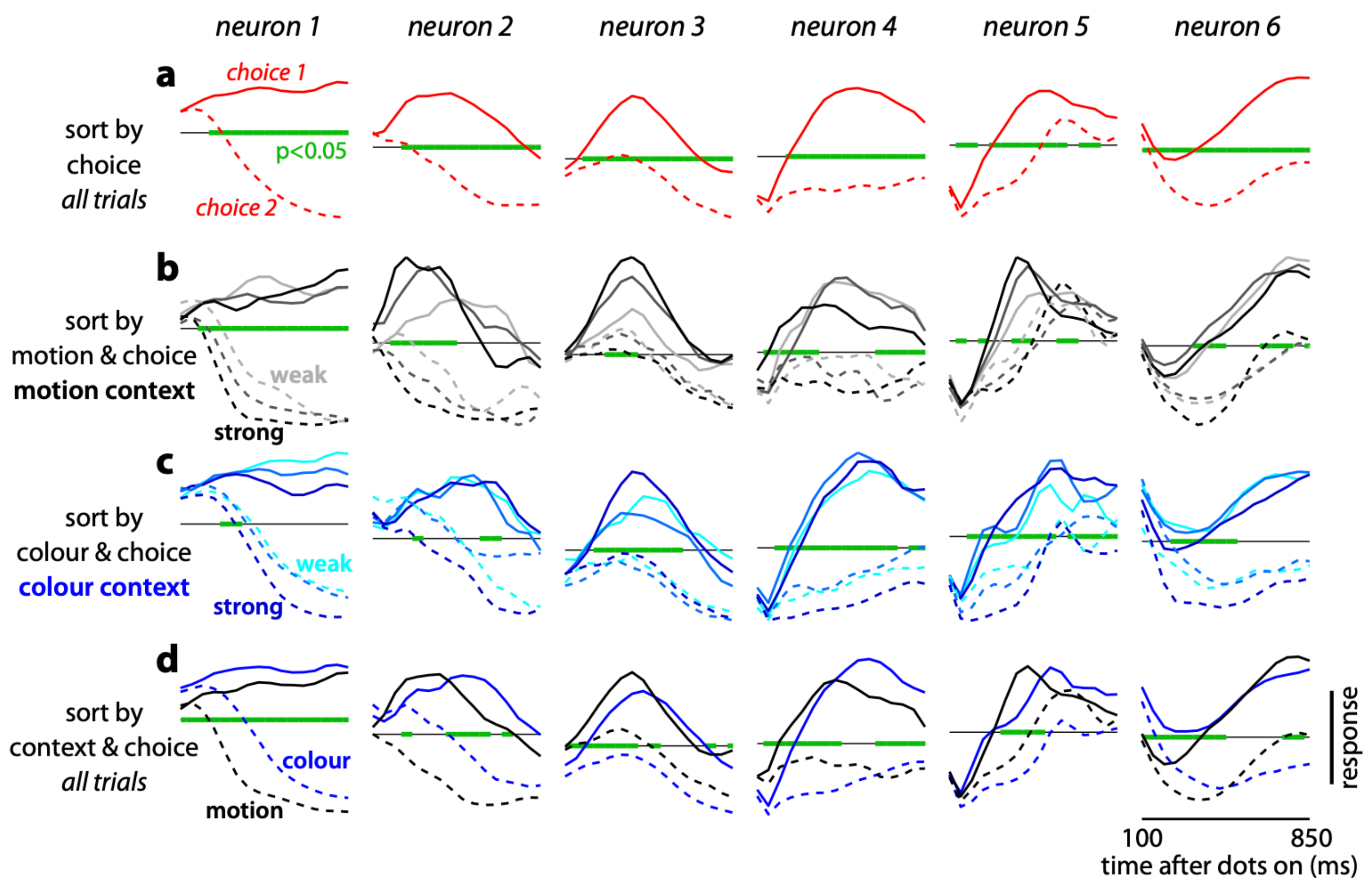
moving left or right ?
more green (left) or red (right) ?



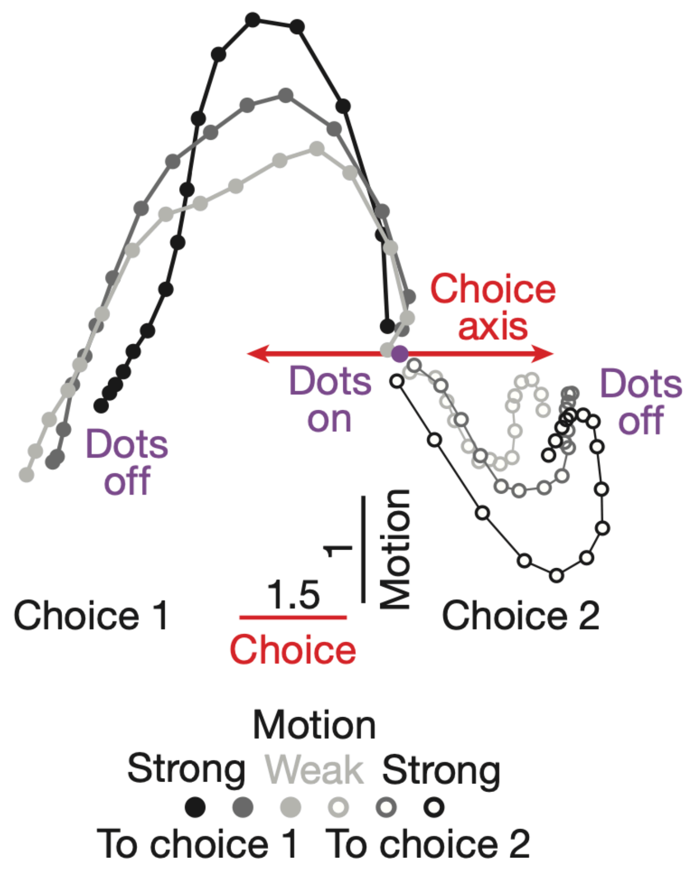
PFC recordings



➡ Tuning of individual neurons



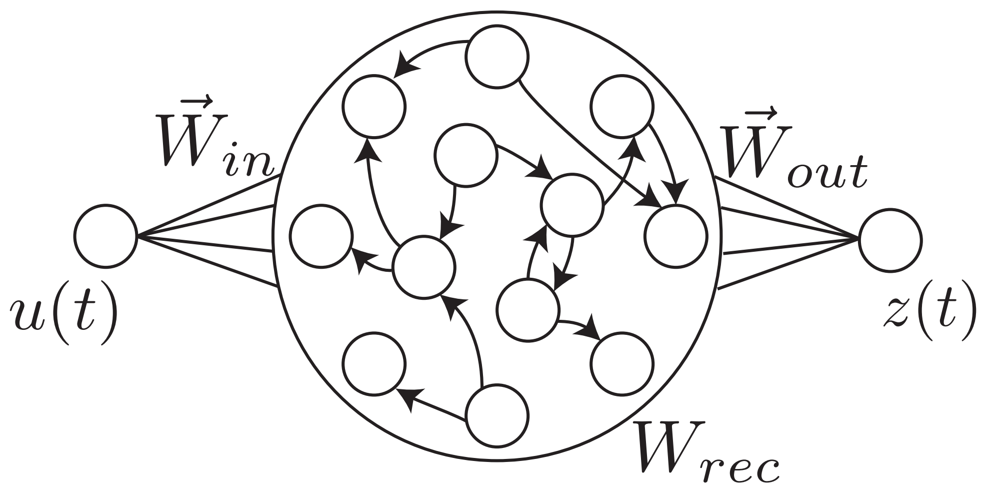
➡ Population level analysis



$$r_{i,t}(k) = \beta_{i,t}(1)\, choice(k) + \beta_{i,t}(2)\, motion(k) + \beta_{i,t}(3)\, color(k) + \beta_{i,t}(4)\, context(k) + \beta_{i,t}(5)$$

Mante, Sussillo, Shenoy, Newsome, 2013

Theory-based reverse-engineering of artificial neural networks

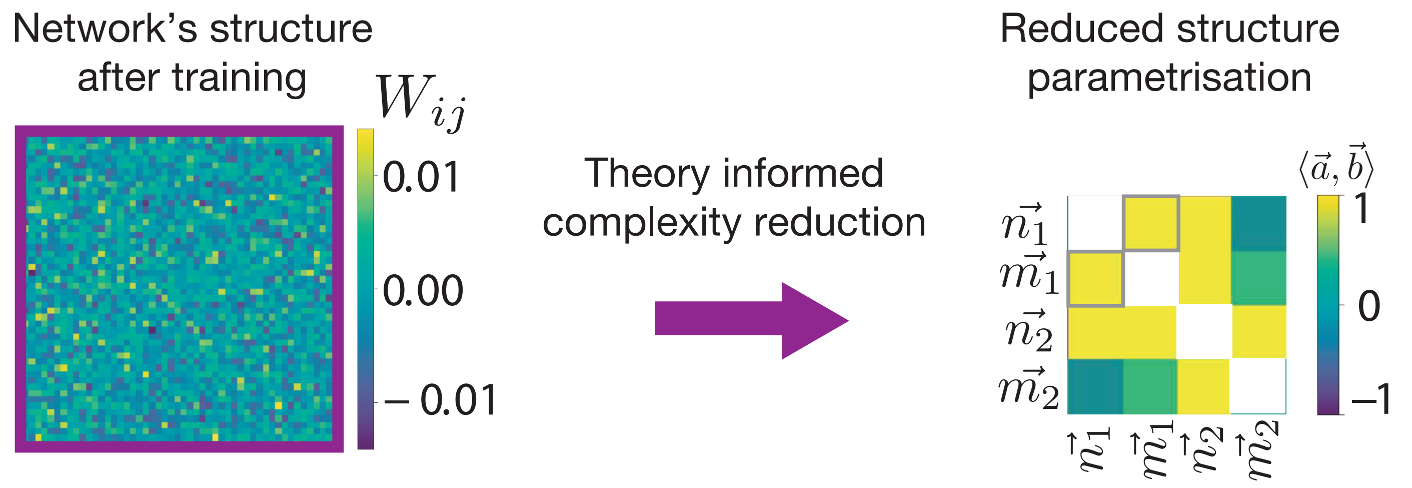


$$\tau \frac{d\vec{x}}{dt} = -\vec{x} + W_{rec}\vec{\Phi}(\vec{x}) + u(t)\vec{W}_{in}$$

$$z(t) = \vec{W}_{out}^T \vec{\Phi}(\vec{x}) \text{ with } \Phi(x) = tanh(x)$$

- Goal: understand how these networks solve a task
- Difficulty: lots of parameters, non-linear systems
- Solution: leverage theoretical results on rate models

Connectivity



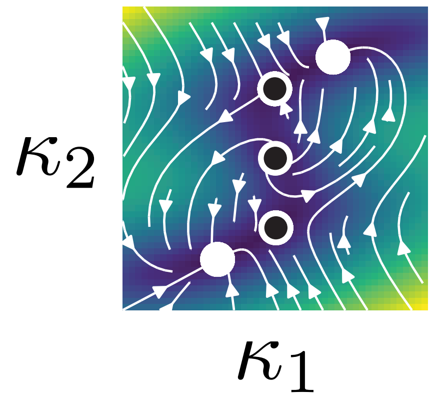
Dynamics

$$\tau \frac{d\vec{x}}{dt} = -\vec{x} + W_{rec}\vec{\Phi}(\vec{x}) + u(t)\vec{W}_{in}$$

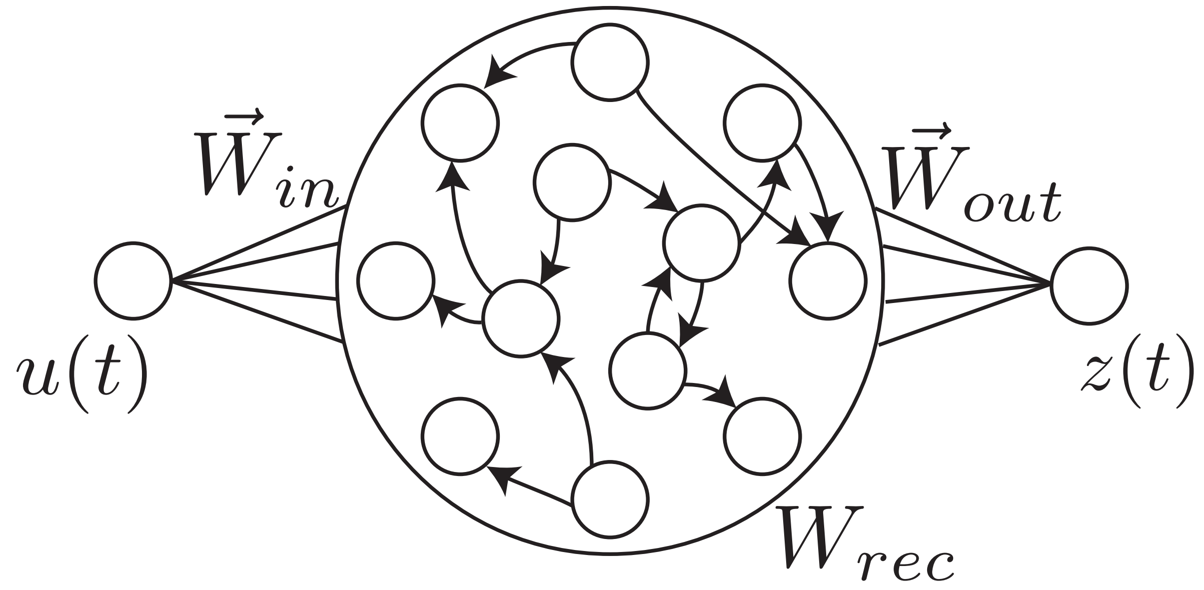
$$\vec{x} \in \mathbb{R}^N \text{ . } N \gg 1$$

$$\frac{d\vec{\kappa}}{dt} = F(\vec{\kappa})$$

$$\vec{\kappa} \in \mathbb{R}^D \text{ . } D = O(1)$$



Low-rank connectivity matrices



$$\tau \frac{d\vec{x}}{dt} = -\vec{x} + W_{rec}\vec{\Phi}(\vec{x}) + u(t)\vec{W}_{in}$$

$$z(t) = \vec{W}_{out}^{T}\vec{\Phi}(\vec{x}) \text{ with } \Phi(x) = tanh(x)$$

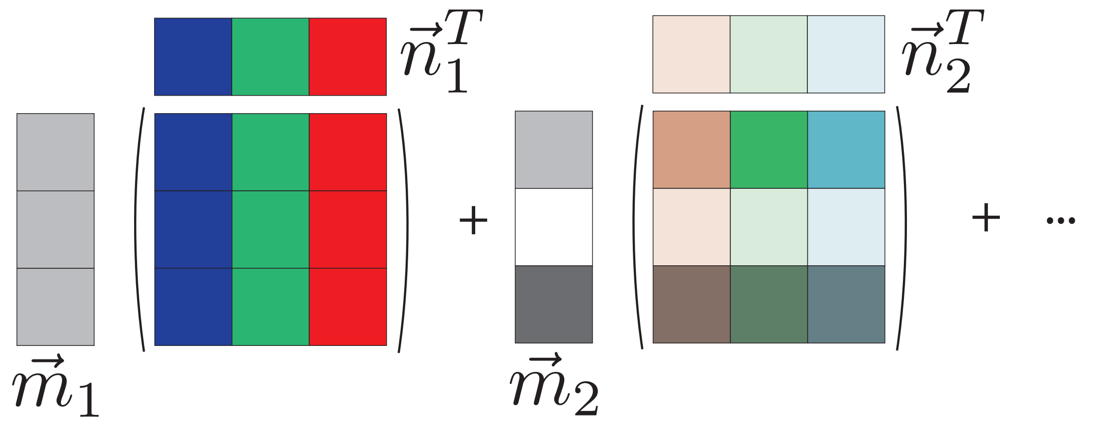
Low-rank recurrent matrix:

$$W_{rec} = \vec{m}_1\vec{n}_1^T + \vec{m}_2\vec{n}_2^T + \ldots$$

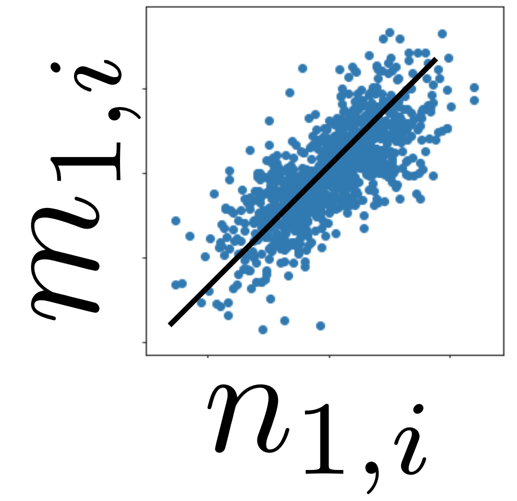
Seung, 1996

Landau & Sompolinksy, 2018

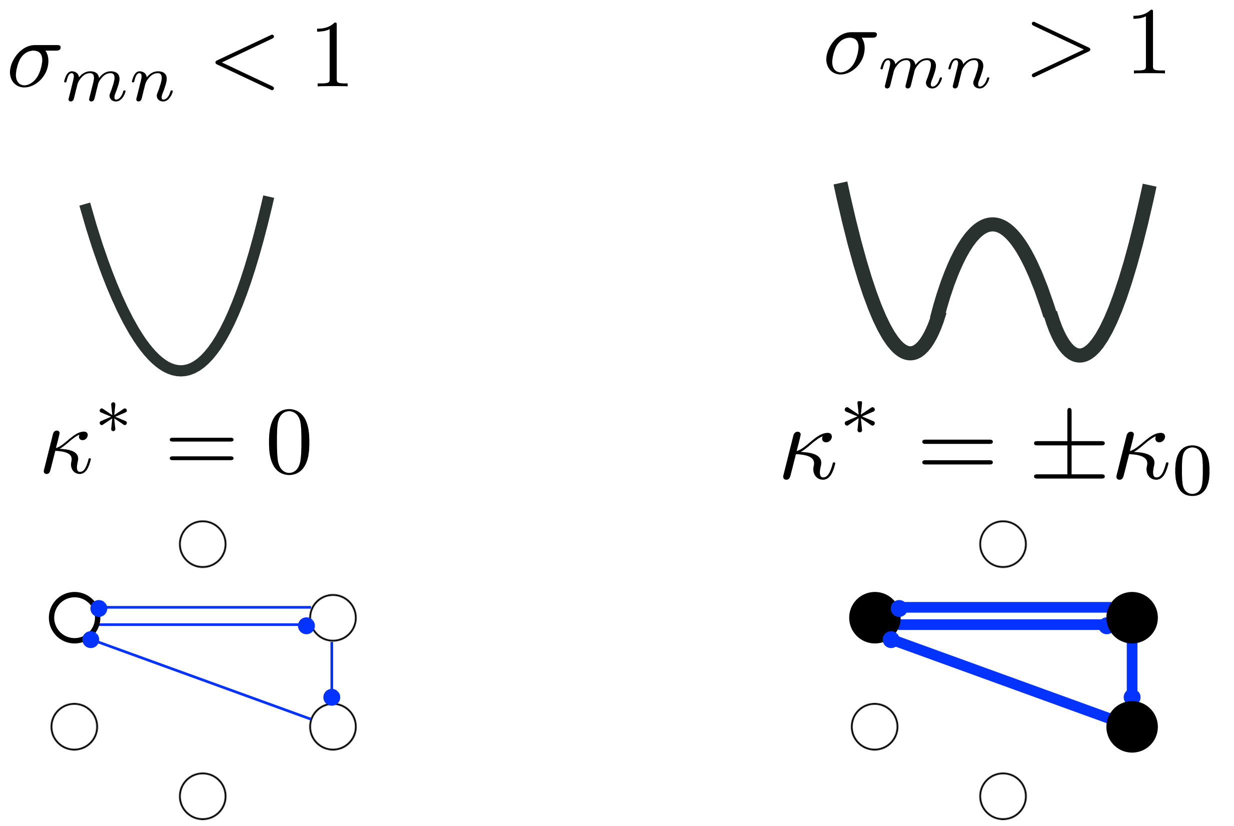
Mastrogiuseppe & Ostojic, 2018



Gaussian connectivity vectors:

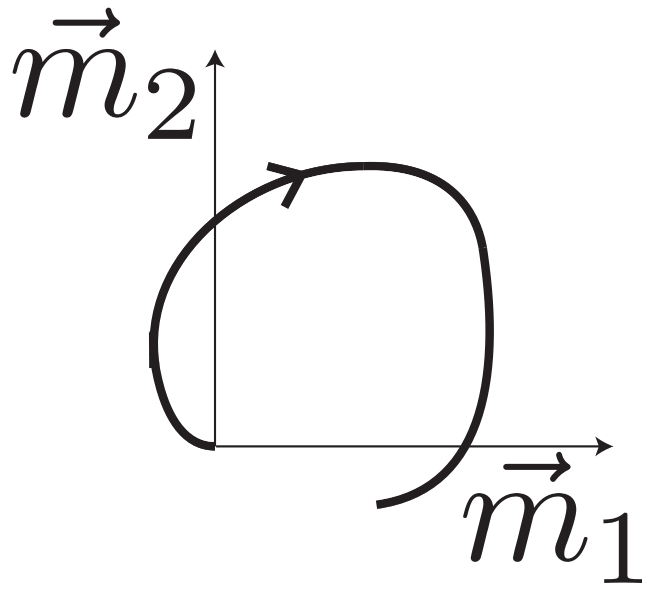


Relate connectivity parameters and dynamical features of RNN.



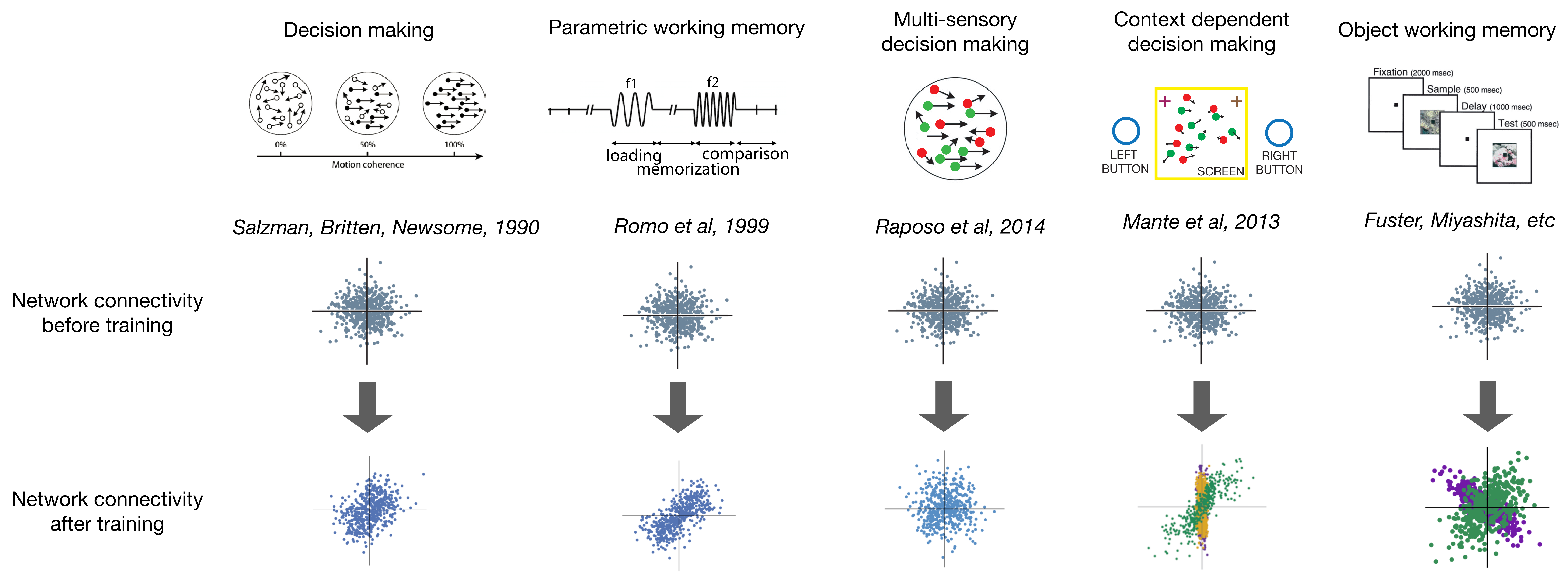
Rank controls dimensionality:

$$W_{rec}\vec{\Phi}(\vec{x}) = \vec{m}_1\underbrace{\vec{n}_1^T\vec{\Phi}(\vec{x})}_{\kappa_1} + \vec{m}_2\underbrace{\vec{n}_2^T\vec{\Phi}(\vec{x})}_{\kappa_2} + \ldots$$



$\boxed{\text{macroscopic variables}}$

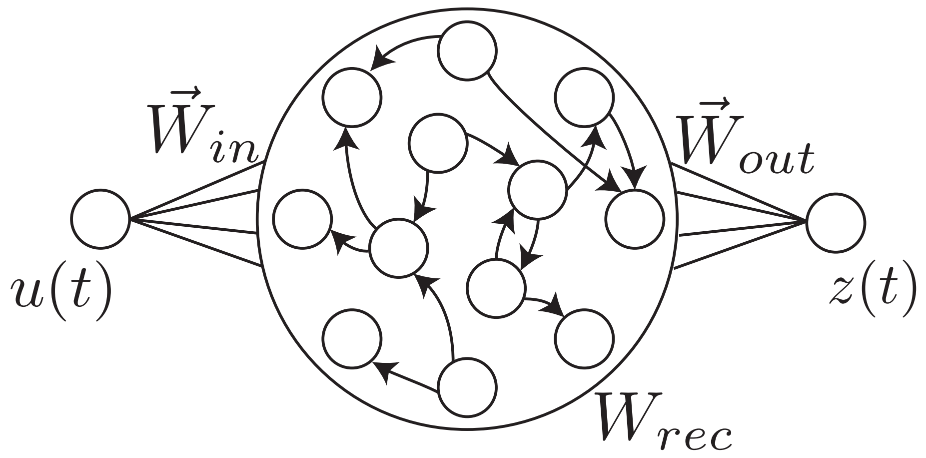
Training low-rank networks to perform various tasks



➡ A priori amorphous networks, for some tasks, segregate into populations

Dubreuil, Valente et al., 2022

Theory: cognitive variables and inputs interact through functional couplings



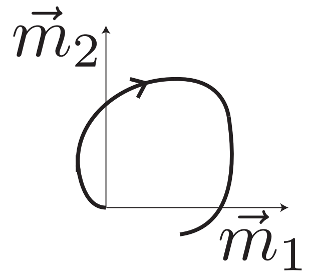
$$\tau \frac{d\vec{x}}{dt} = -\vec{x} + W_{rec}\vec{\Phi}(\vec{x}) + u(t)\vec{W}_{in}$$

$$z(t) = \vec{W}_{out}^T \vec{\Phi}(\vec{x}) \text{ with } \Phi(x) = tanh(x)$$

$$W_{rec} = \vec{m}_1 \vec{n}_1^T + \vec{m}_2 \vec{n}_2^T + \ldots$$

$$\kappa_1 = \langle \vec{x}, \vec{m}_1 \rangle$$
$$\kappa_2 = \langle \vec{x}, \vec{m}_2 \rangle$$

Functional dynamics (e.g. K=2):



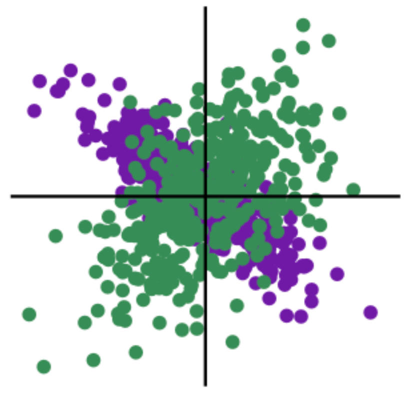
$$\tau \dot{\kappa}_1 = -\kappa_1 + \tilde{\sigma}_{n_1 m_1}\kappa_1 + \tilde{\sigma}_{n_1 m_2}\kappa_2 + \tilde{\sigma}_{n_1 W_{in}} u(t)$$
$$\tau \dot{\kappa}_2 = -\kappa_2 + \tilde{\sigma}_{n_2 m_1}\kappa_1 + \tilde{\sigma}_{n_2 m_2}\kappa_2 + \tilde{\sigma}_{n_2 W_{in}} u(t)$$

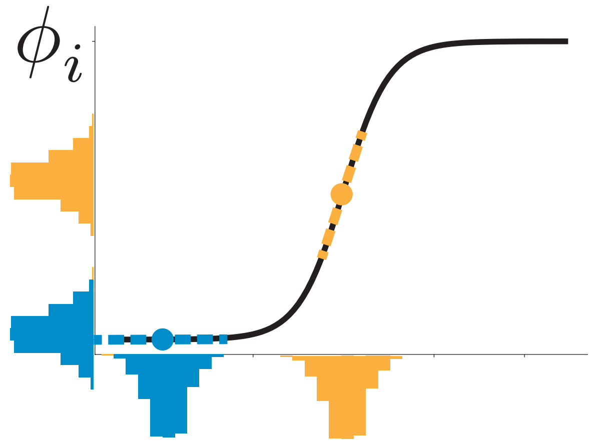
$$|\frac{du}{dt}| \gg \tau$$

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

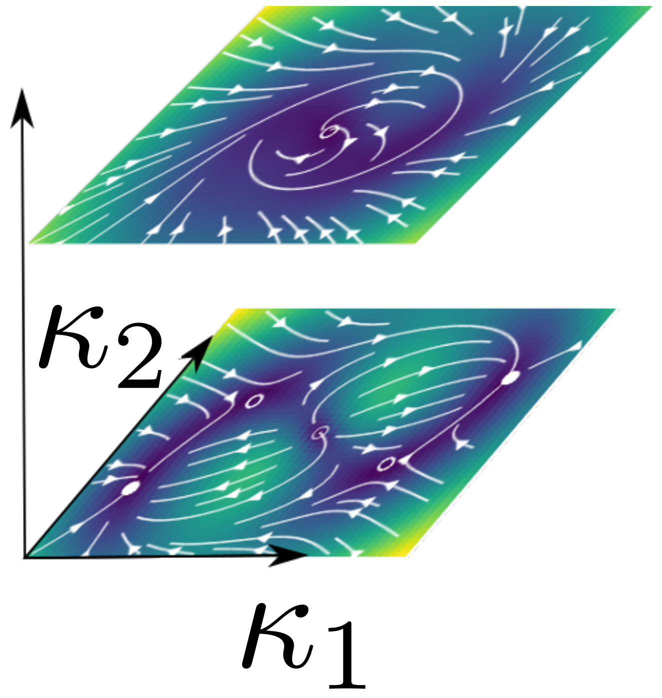
structural

activity dependent

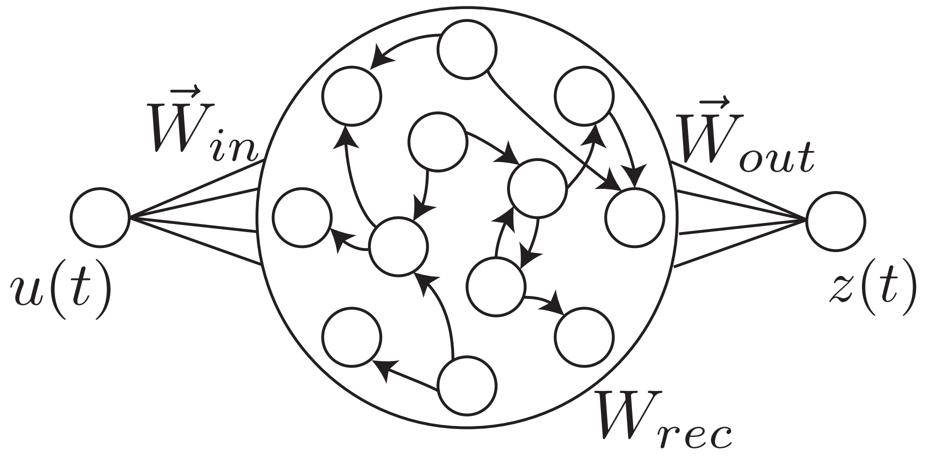




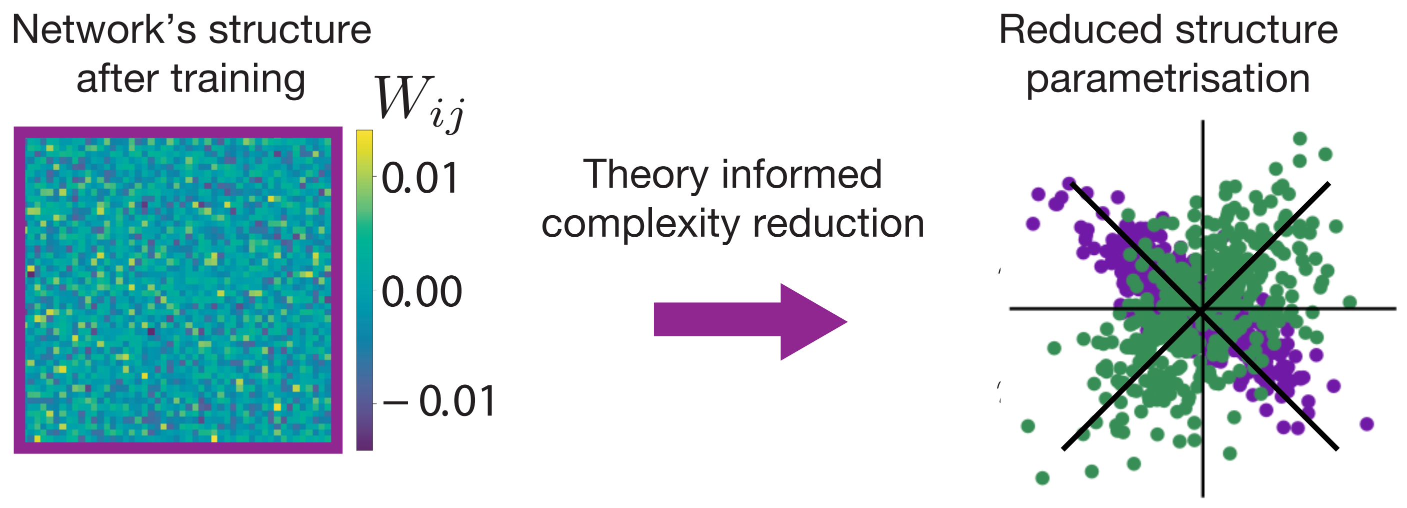
$\langle \Phi' \rangle_{p=1}$



Summary: leveraging theory for reverse-engineering ANN



➡ Picking-up dynamically relevant summary statistics of the connectivity matrix



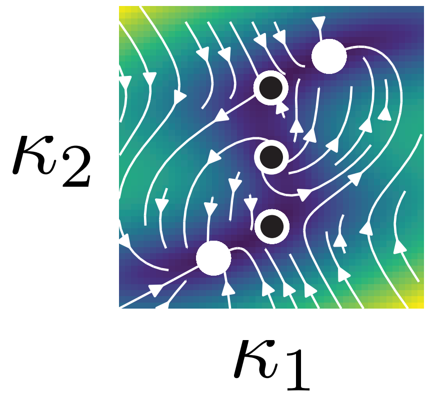
➡ Reduce dimensionality of the dynamics

$$\tau \frac{d\vec{x}}{dt} = -\vec{x} + W_{rec}\vec{\Phi}(\vec{x}) + u(t)\vec{W_{in}}$$

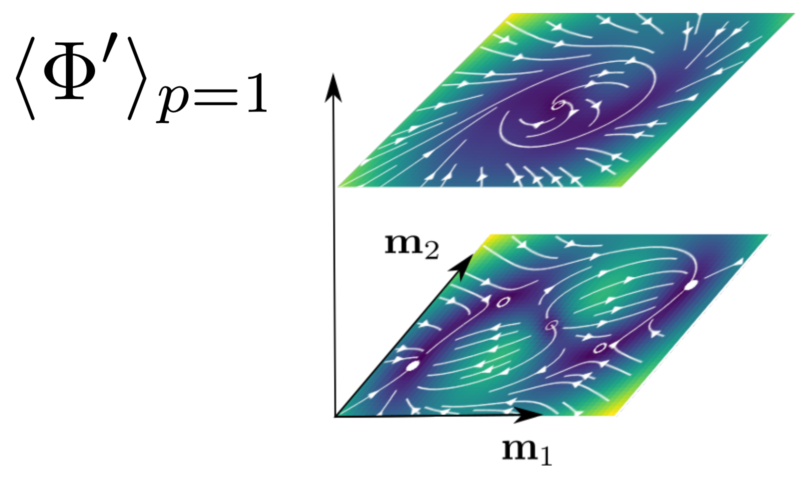
$$\vec{x} \in \mathbb{R}^N \; : \; N \gg 1$$

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

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



➡ Reconfiguration of dynamics through gain modulation of specific populations



Dubreuil, Valente et al., 2022

Outline

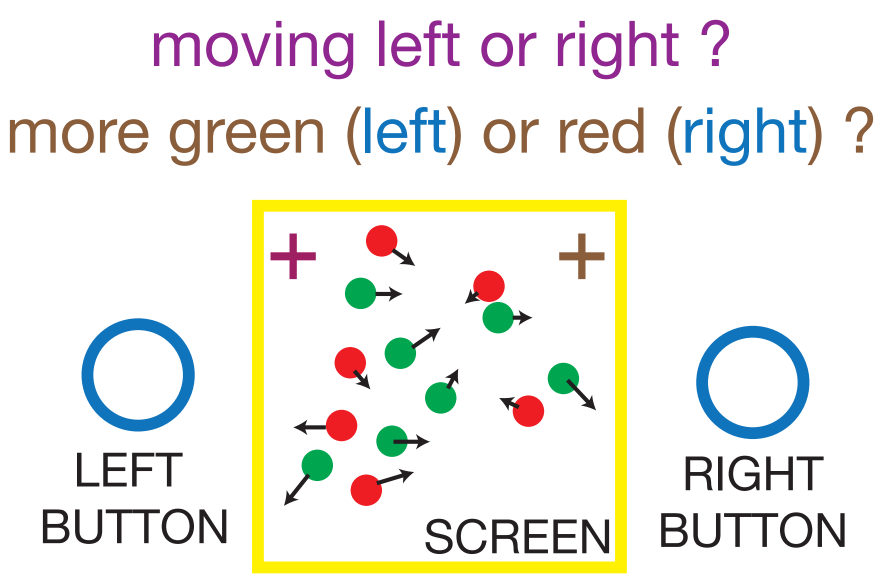
Introduction: cognitive modeling

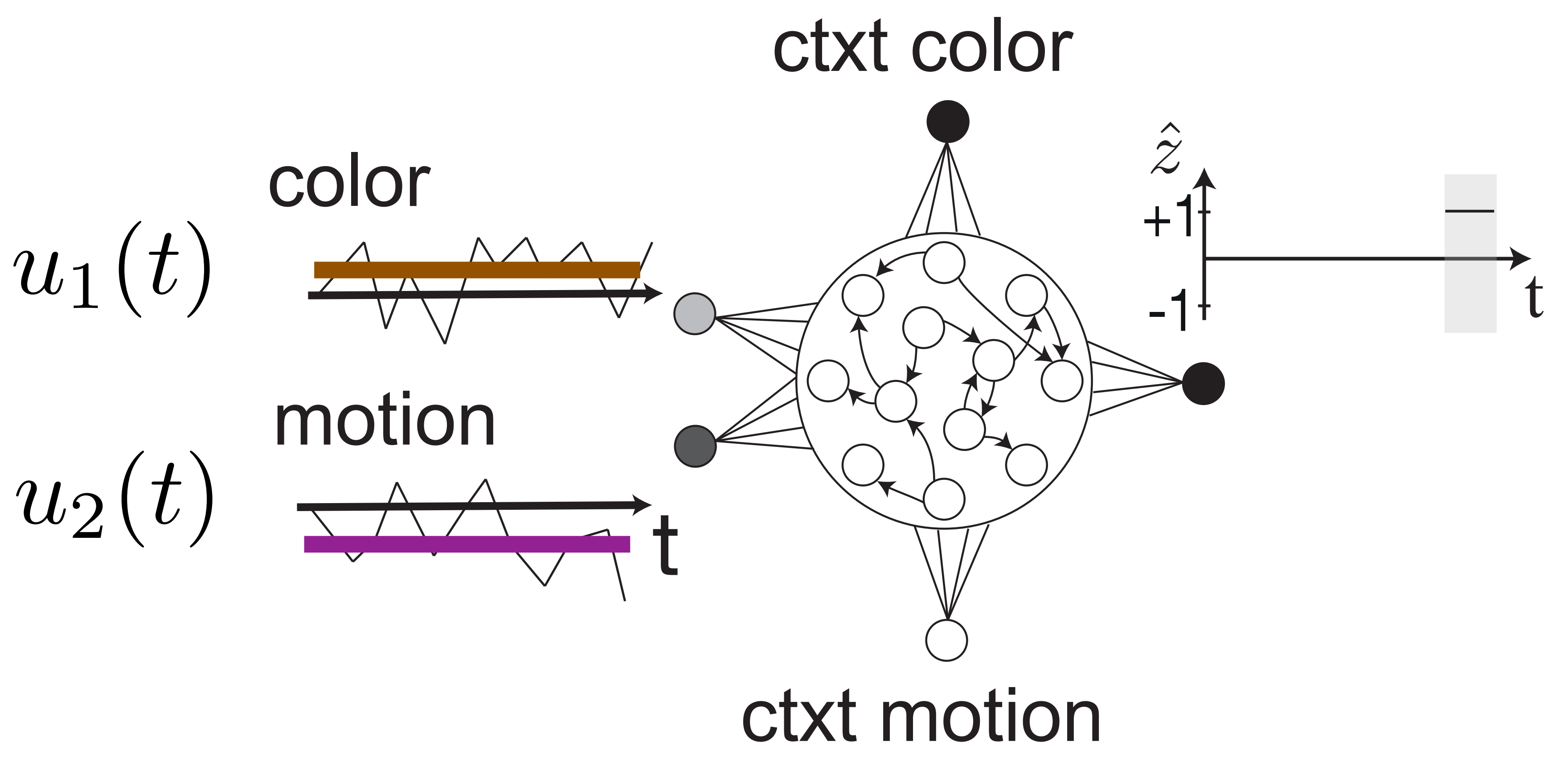
Theory-based reverse-engineering of artificial neural networks

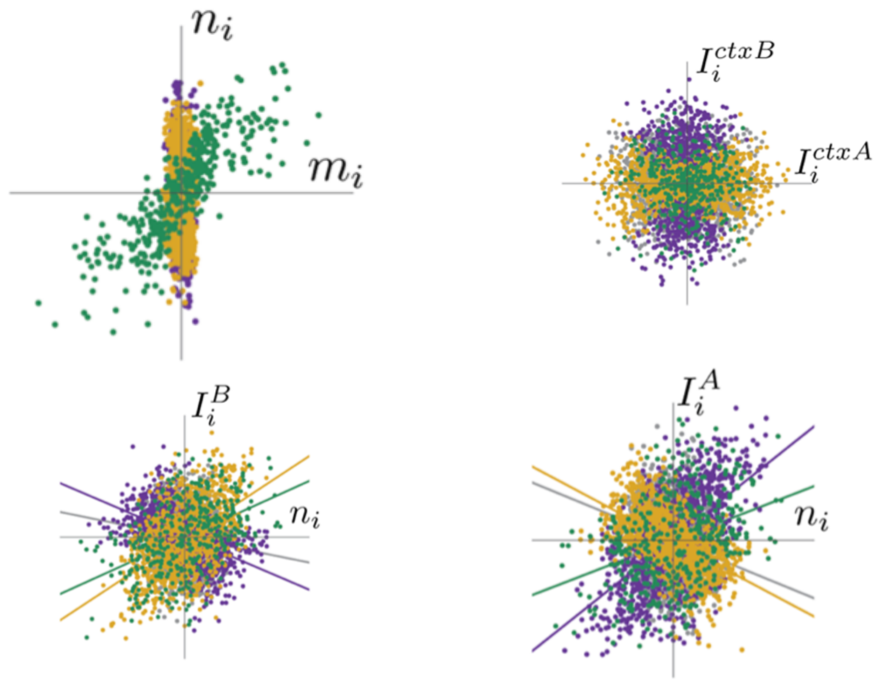
➡ On the computational role of population structure

Modeling temporally structured behaviors

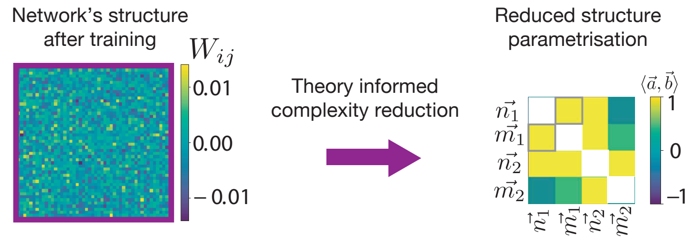
Example: context-dependent decision making







➡ Assess relationship between connectivity vectors (4 input vectors, 2 recurrent vectors, 1 output vector)



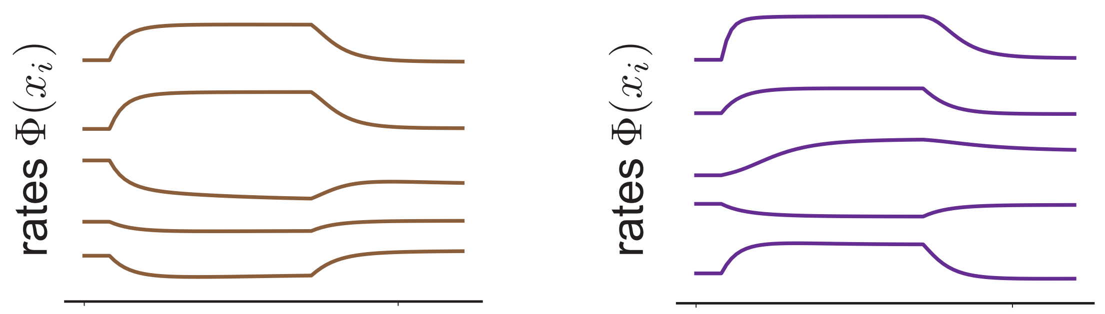
➡ Reduced dynamical system

$$\dot{\kappa} = -\kappa + \tilde{\sigma}_{nm}\kappa + \tilde{\sigma}_{nW_{in1}} u_1(t) + \tilde{\sigma}_{nW_{in2}} u_2(t)$$

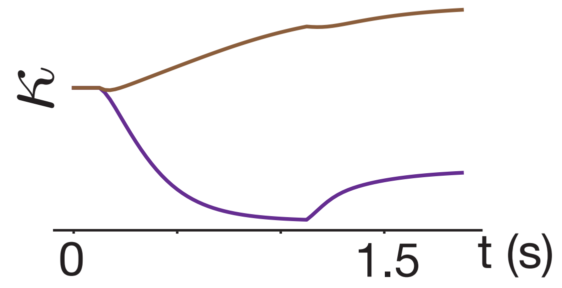
$$\kappa = \langle \vec{x}, \vec{m} \rangle$$

color = +10%

motion = -20%



Cognitive variable temporally integrates the relevant sensory input

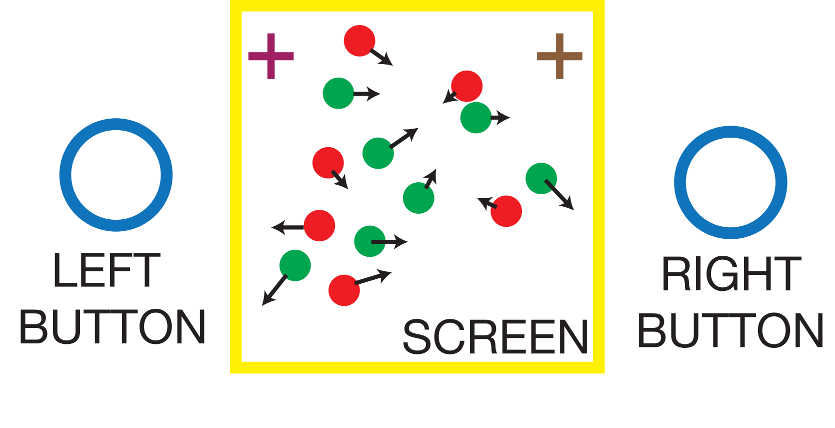


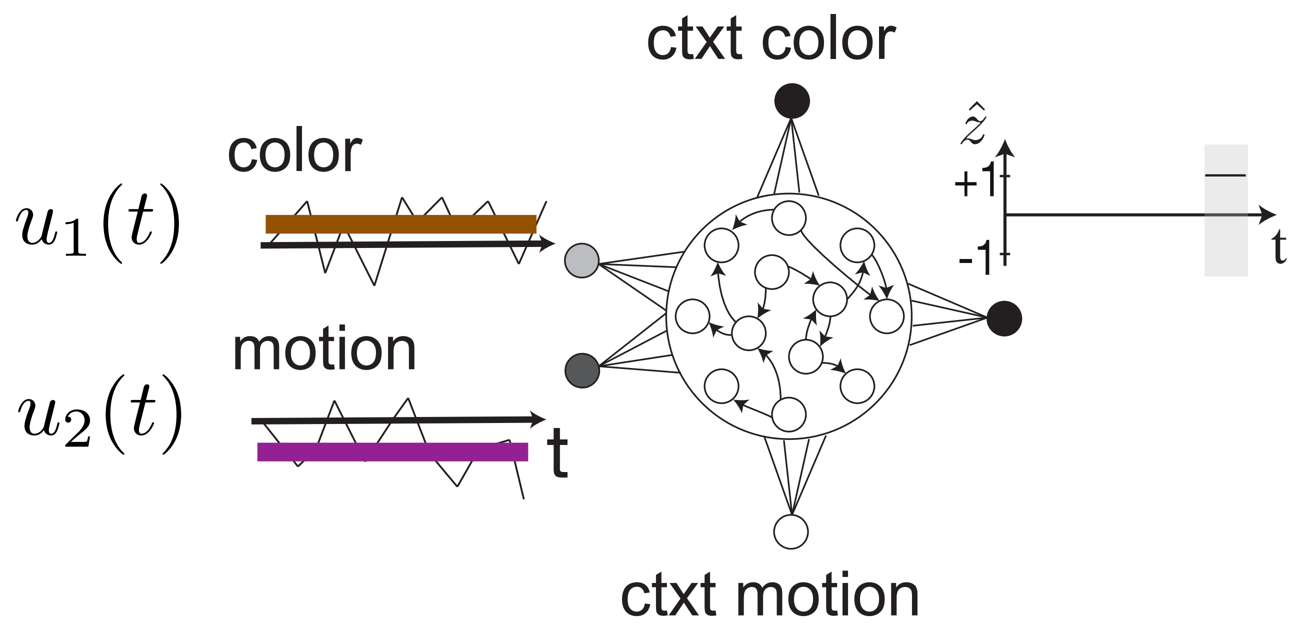
cf recordings in LIP, *Steinemann et al. 2024*

Example: context-dependent decision making

moving left or right ?

more green (left) or red (right) ?



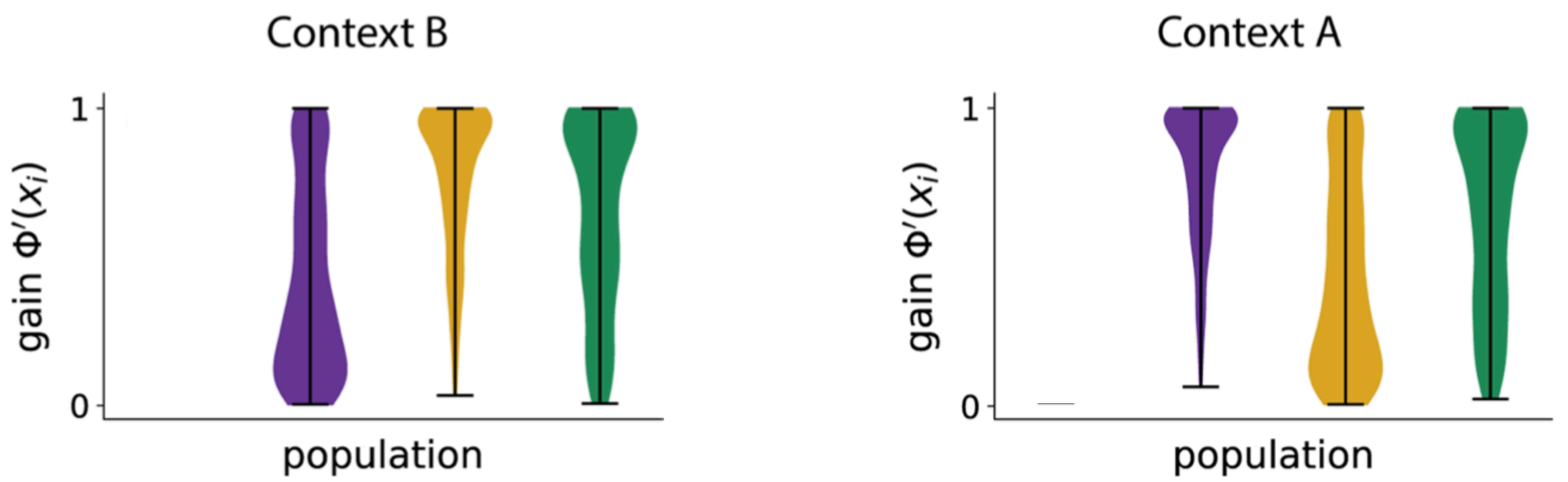


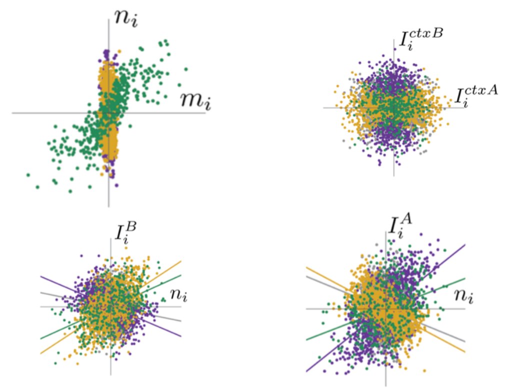
➡ Reduced dynamical system

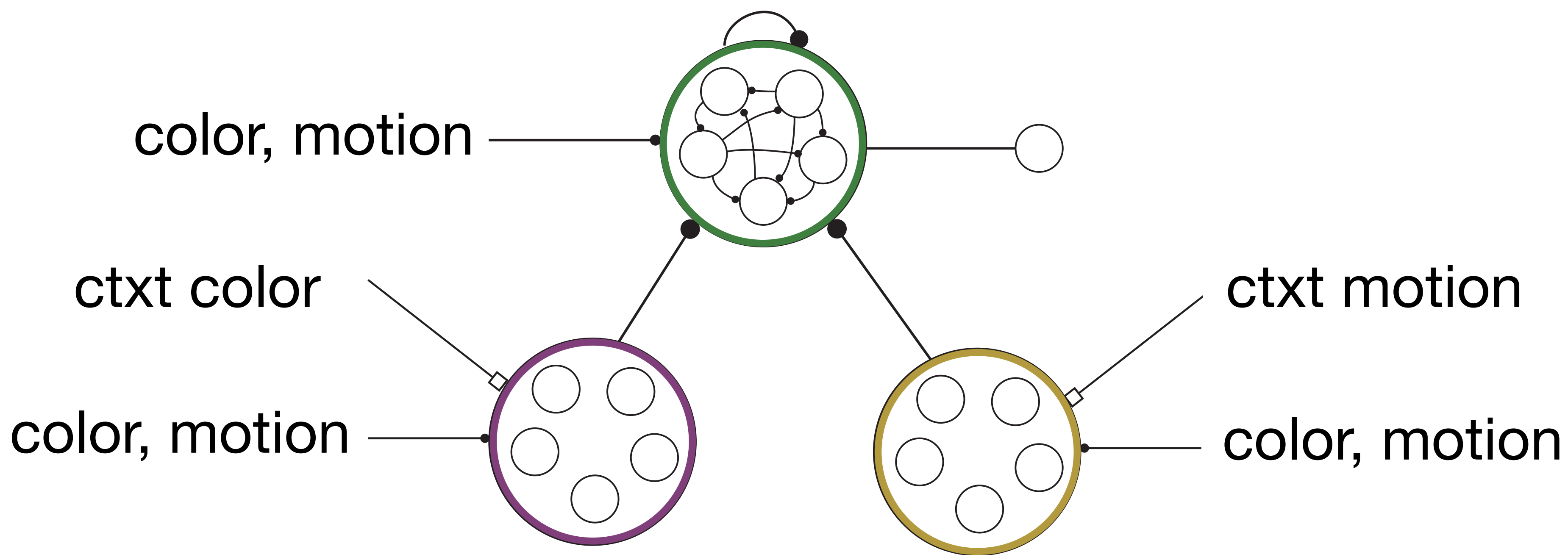
$$\dot{\kappa} = -\kappa + \tilde{\sigma}_{nm}\kappa + \tilde{\sigma}_{nW_{in1}} u_1(t) + \tilde{\sigma}_{nW_{in2}} u_2(t)$$

➡ Functional couplings controlled via gain modulation of specific populations

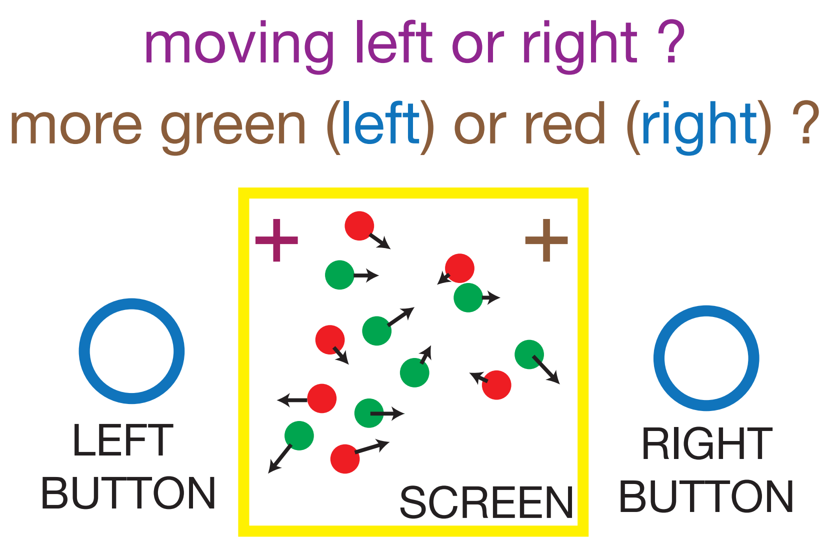
$$\tilde{\sigma}_{ab} = \sum_{p=1}^{P} \sigma^p_{ab} \langle \phi' \rangle_p$$

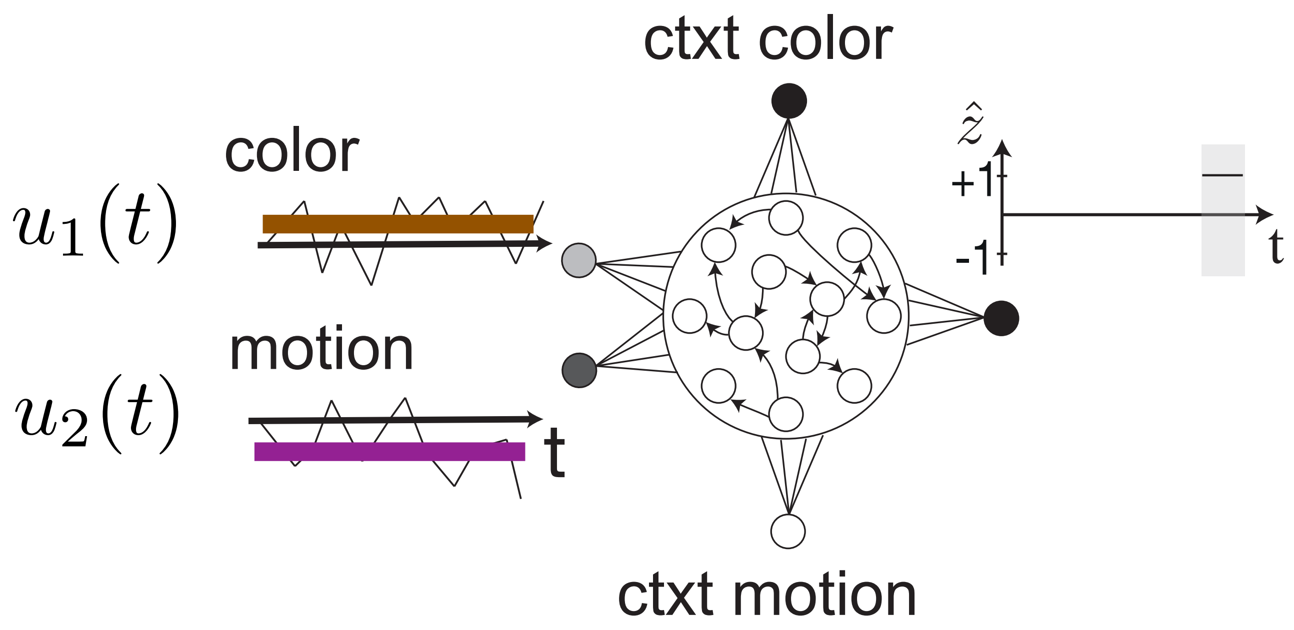


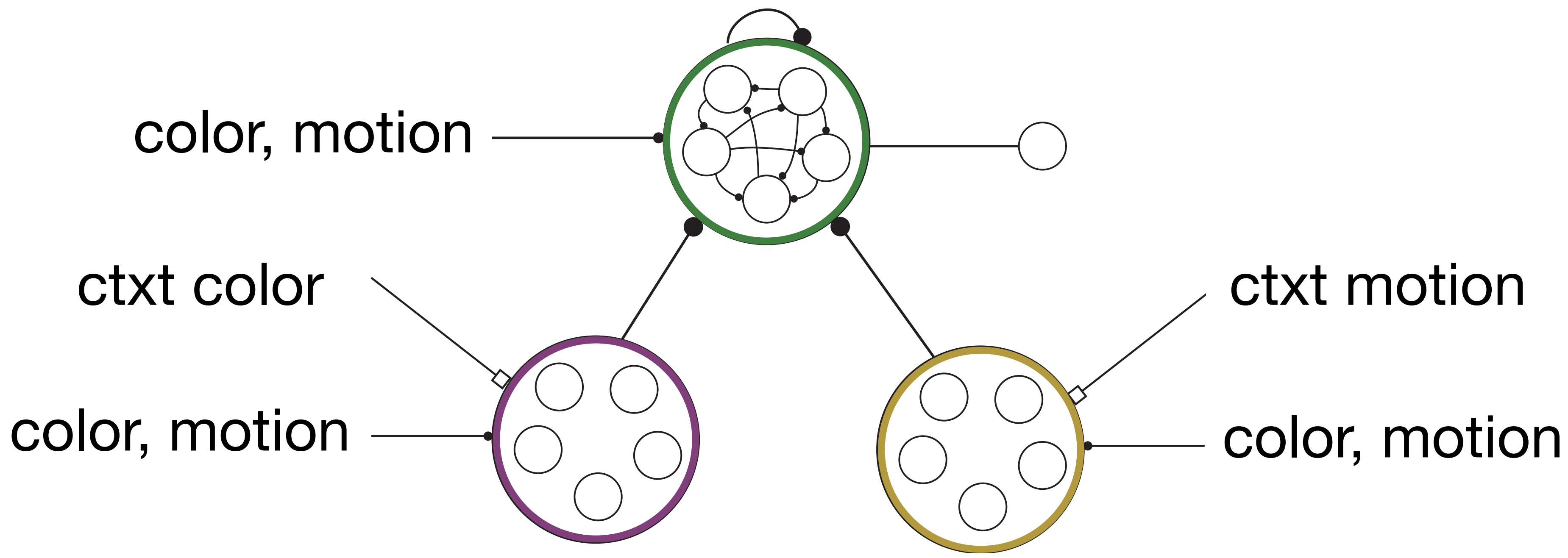




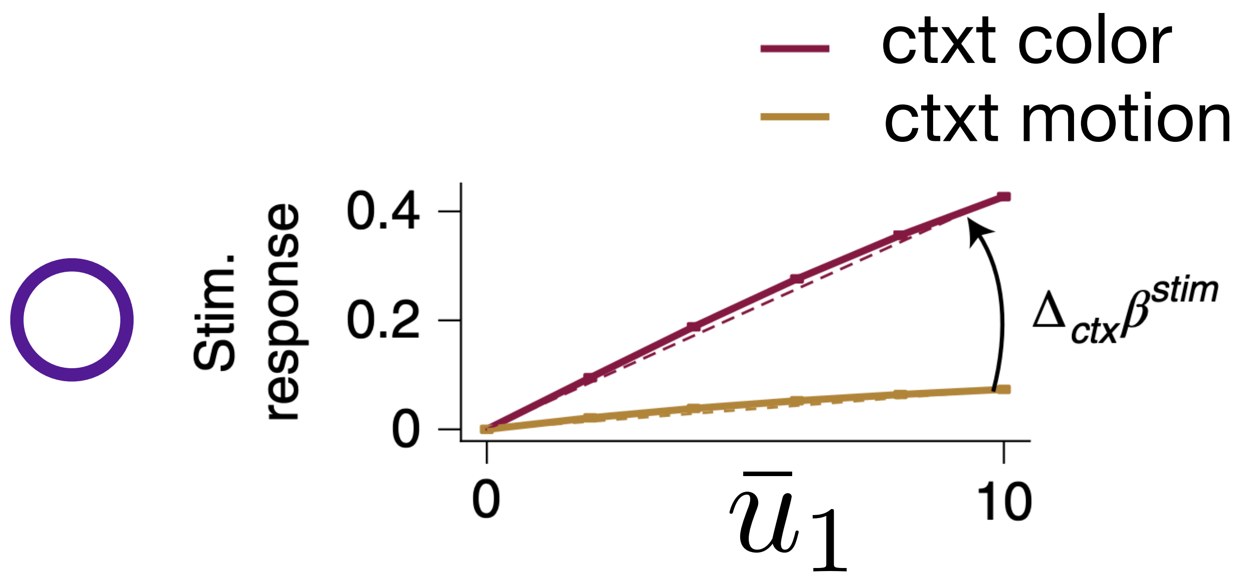
Selectivity profiles of populations: non-linear mixed-selectivity

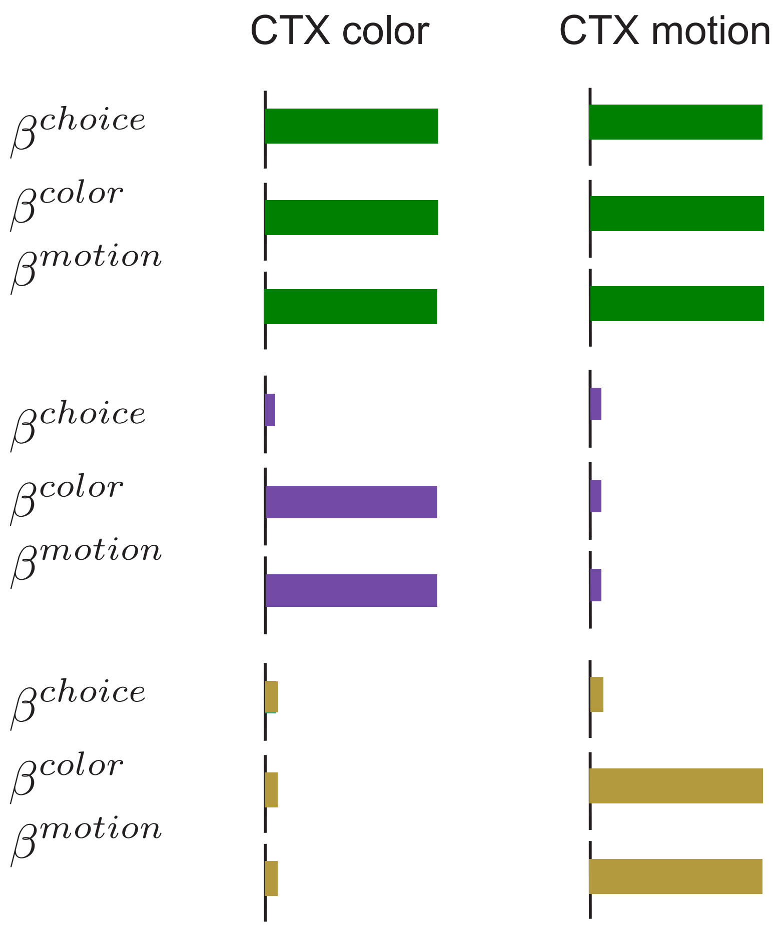






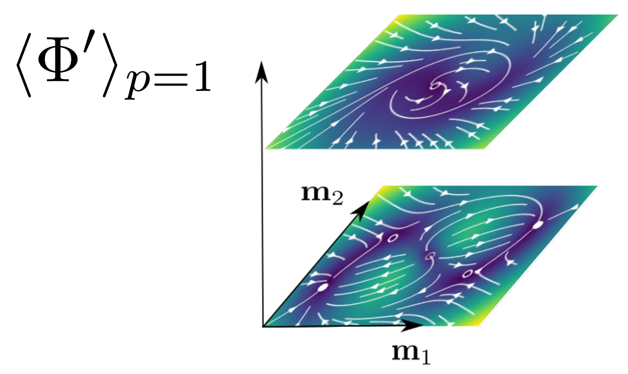
Context-dependent tuning curves





$$r_i = \phi(x_i)$$
$$= \phi\left(u_1(t)W_{in1,i} + u_2(t)W_{in2,i} + c_{col.}W_{c.col.,i}\right)$$

➡ Non-linear mixed selectivity can reflect reconfiguration of network's dynamics

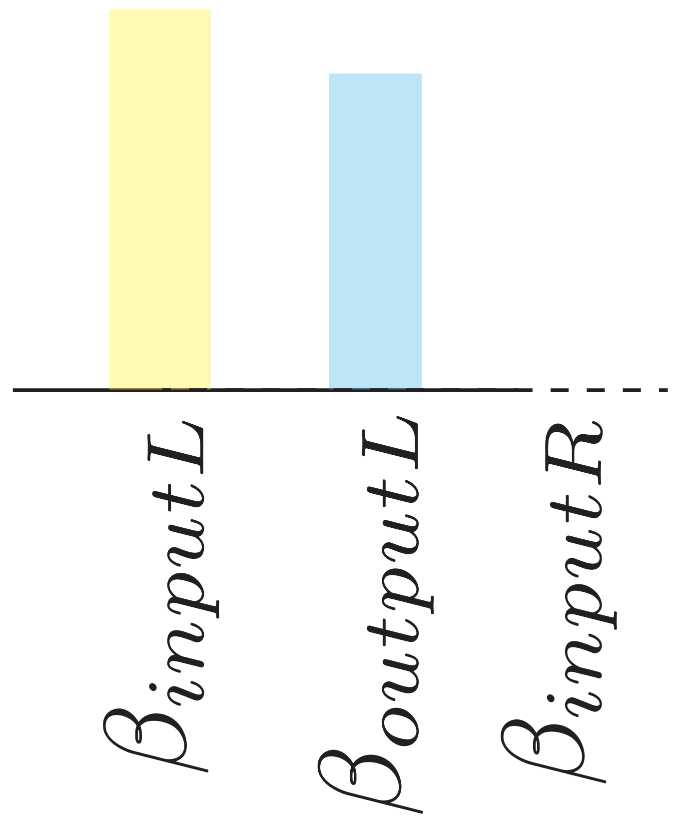


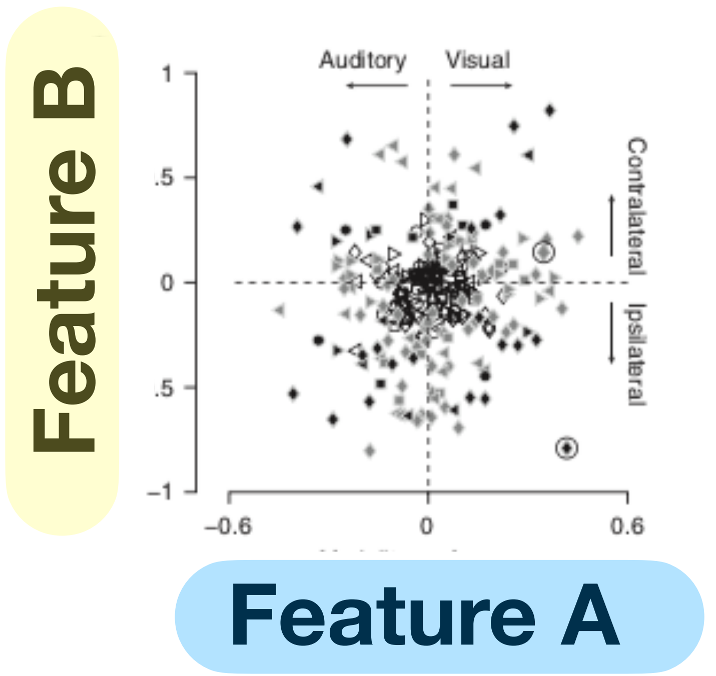
Rigotti et al., 2013

Selectivity profiles of populations: random mixed selectivity

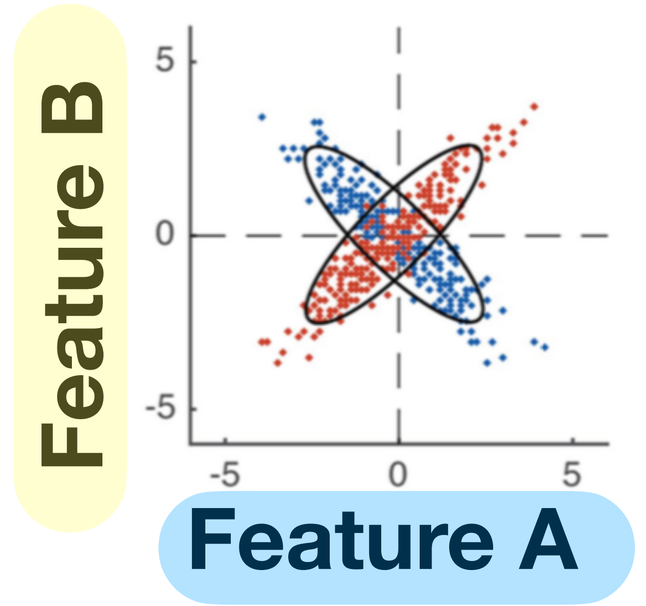
➡ Structure of population-level selectivity profiles

Selectivity profiles



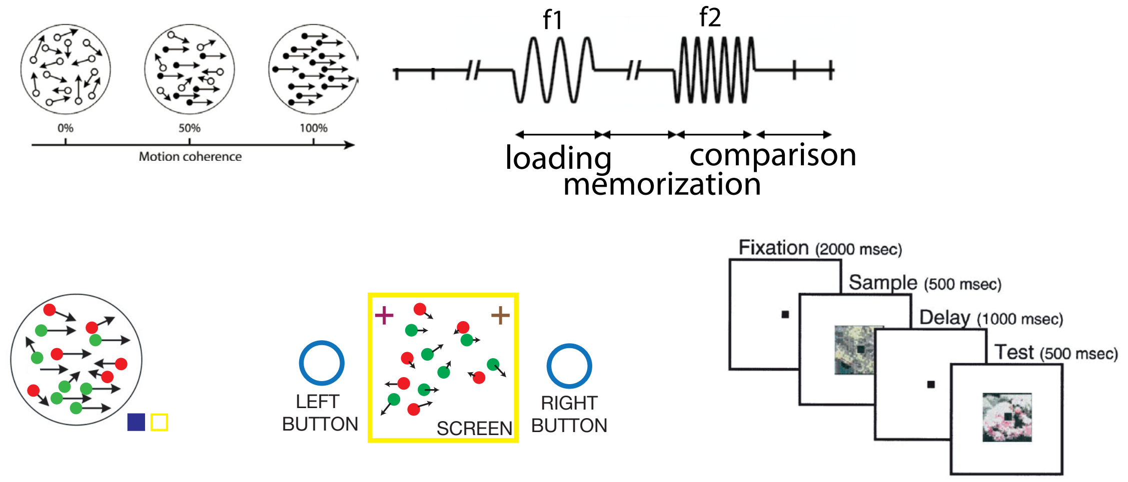


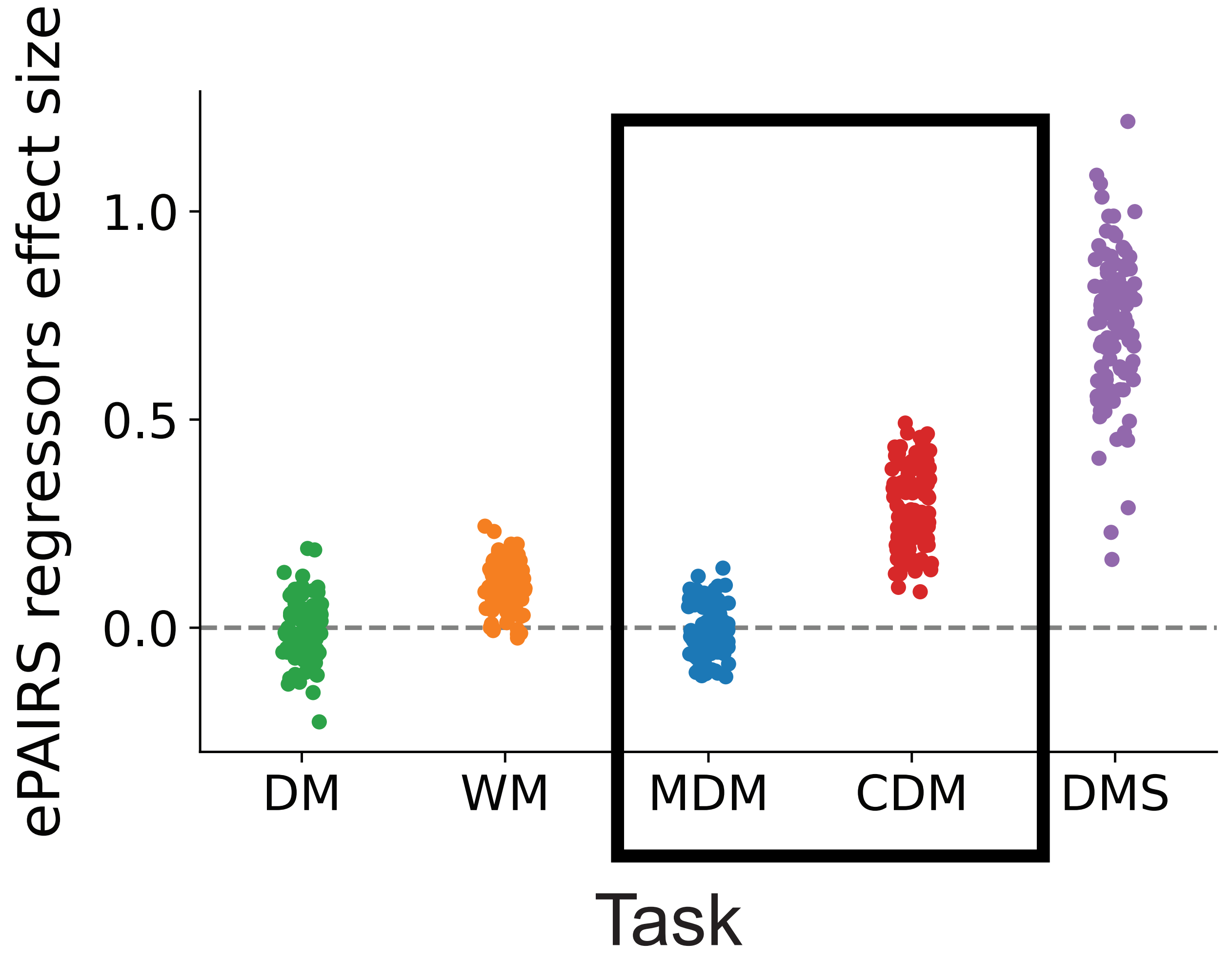
Unstructured selectivity
Raposo et al, 2014



Structured selectivity
Hirokawa et al, 2019

➡ Statistical test for population structure in selectivity

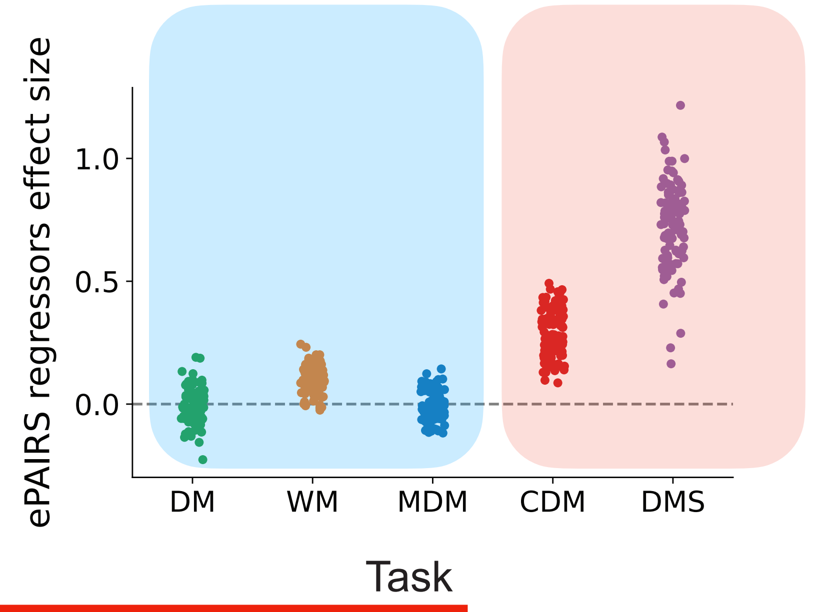




Tasks shape selectivity profiles of populations

Selectivity profiles of populations: random mixed selectivity

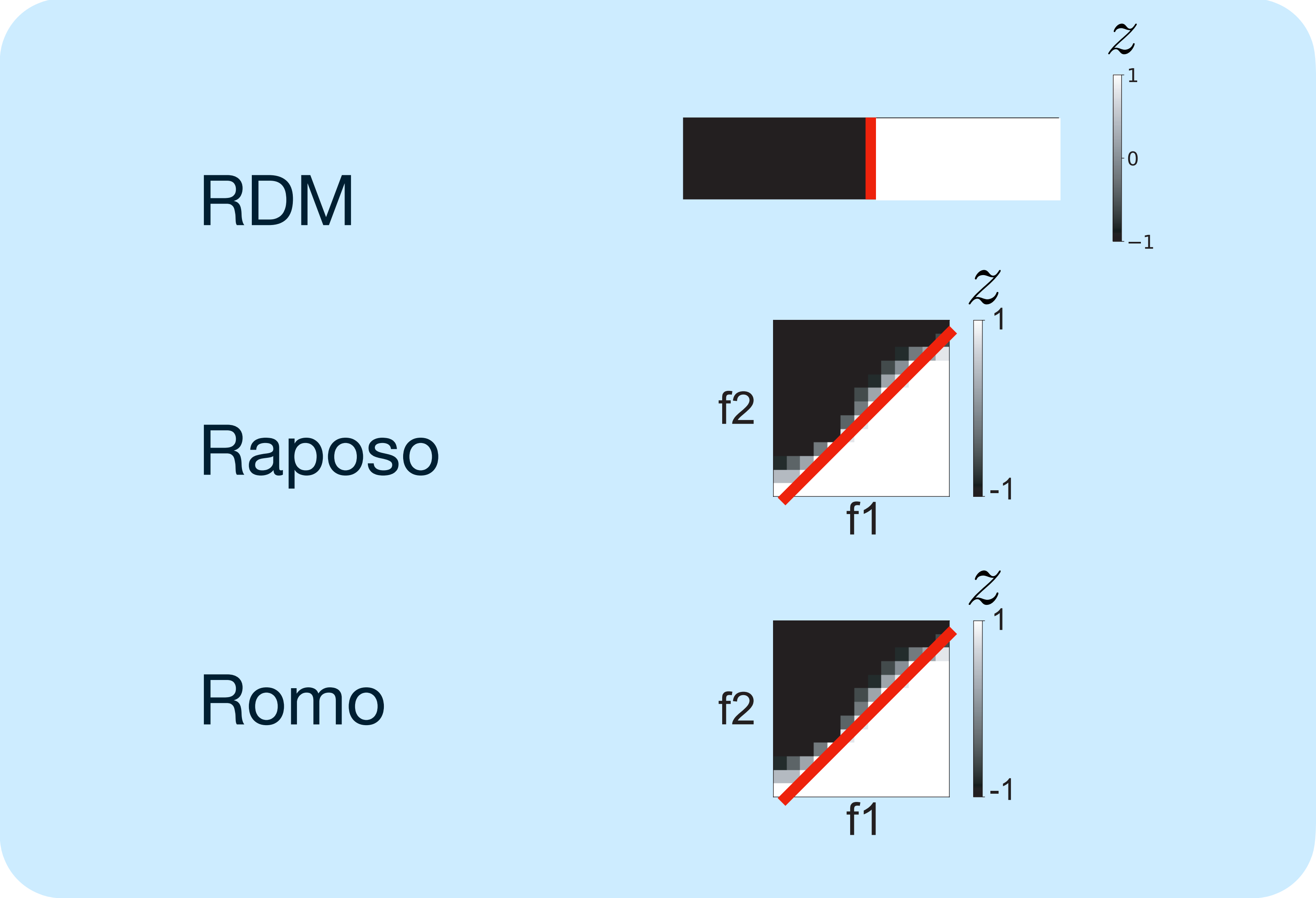
- Statistical test for population structure in selectivity

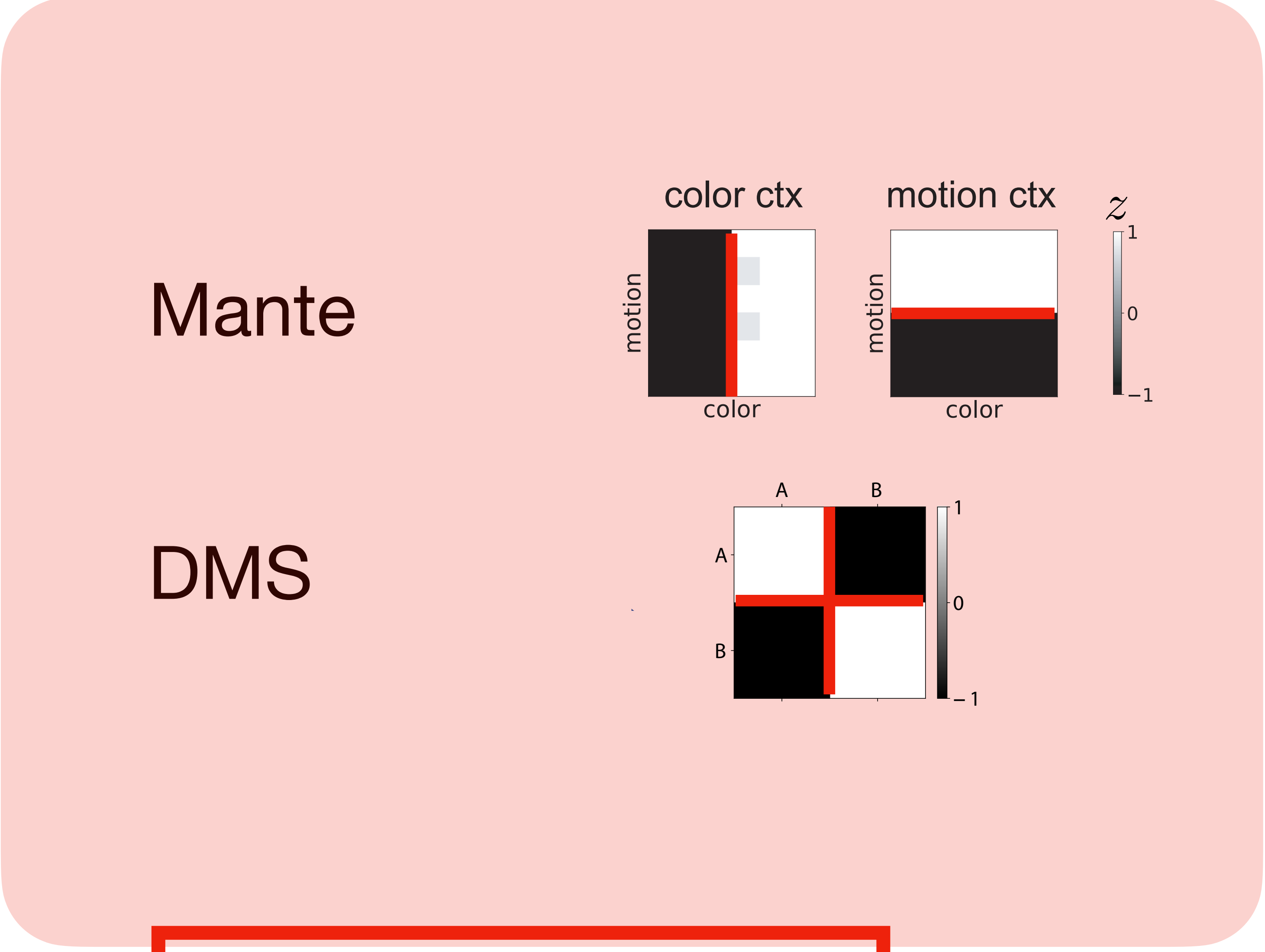


- Training/Reverse-engineering on multiple tasks

stereotyped tasks

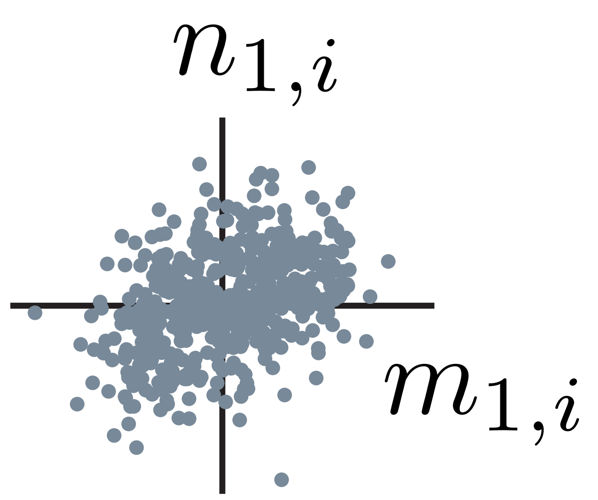
flexible tasks

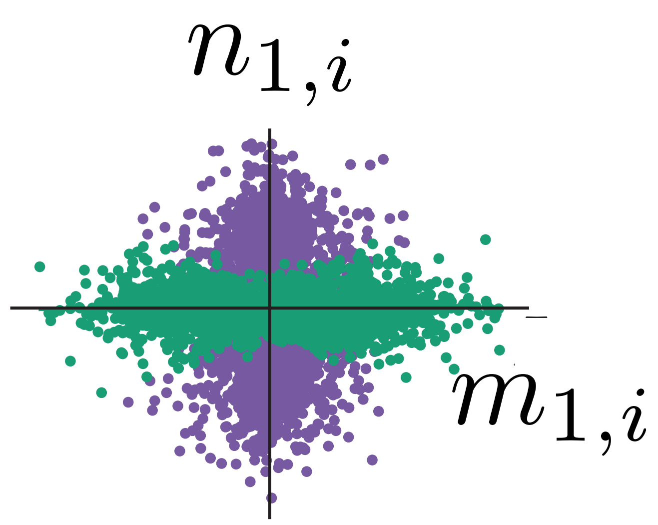




no population structure

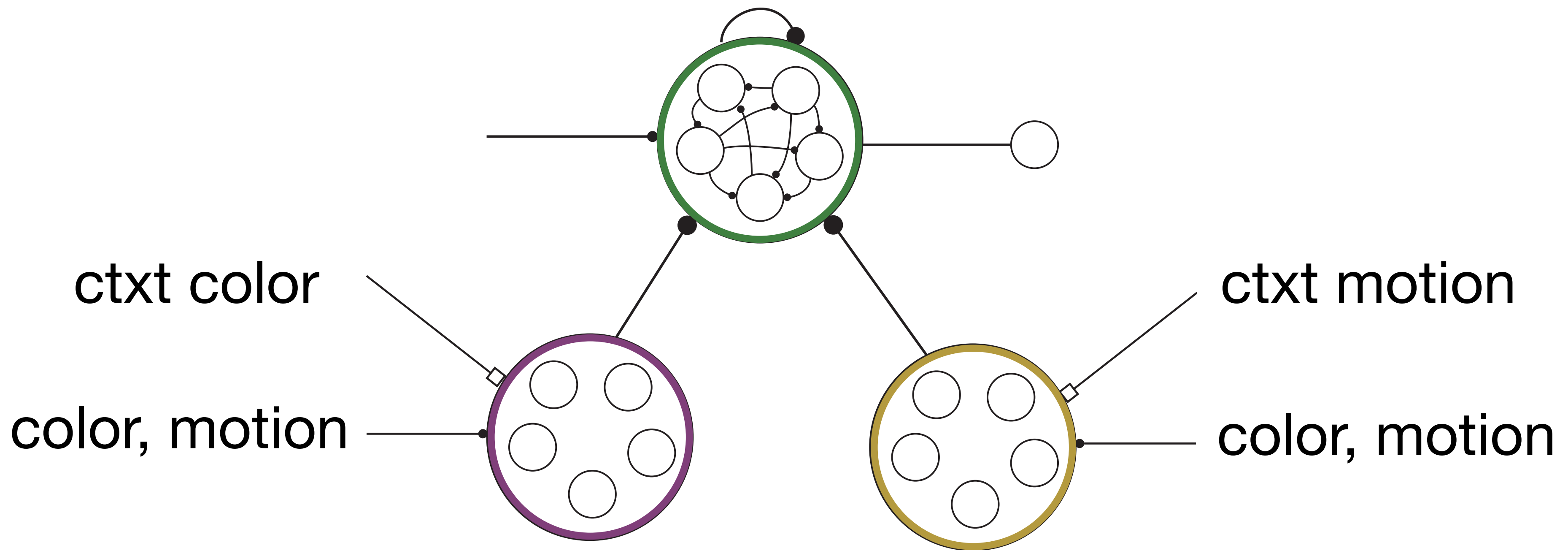
population structure





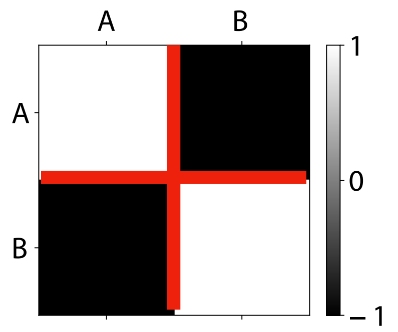
Summary: leveraging theory for reverse-engineering ANN

➡ Reconfiguration of dynamics through gain modulation



$$\dot{\kappa} = -\kappa + \tilde{\sigma}_{nm}\kappa + \tilde{\sigma}_{nW_{in1}}u_1(t) + \tilde{\sigma}_{nW_{in2}}u_2(t)$$

➡ Population structure was task dependent and required for flexible tasks



➡ Population structure can be detected in data

Non-linear mixed selectivity of the purple and gold populations ?

Outline

Introduction: cognitive modeling

Theory-based reverse-engineering of artificial neural networks

➡ On the computational role of population structure

Modeling temporally structured behaviors

Outline

Introduction: cognitive modeling

Theory-based reverse-engineering of artificial neural networks

On the computational role of population structure

➡ Modeling temporally structured behaviors

Long-distance dependencies in behavior

- Natural language processing:

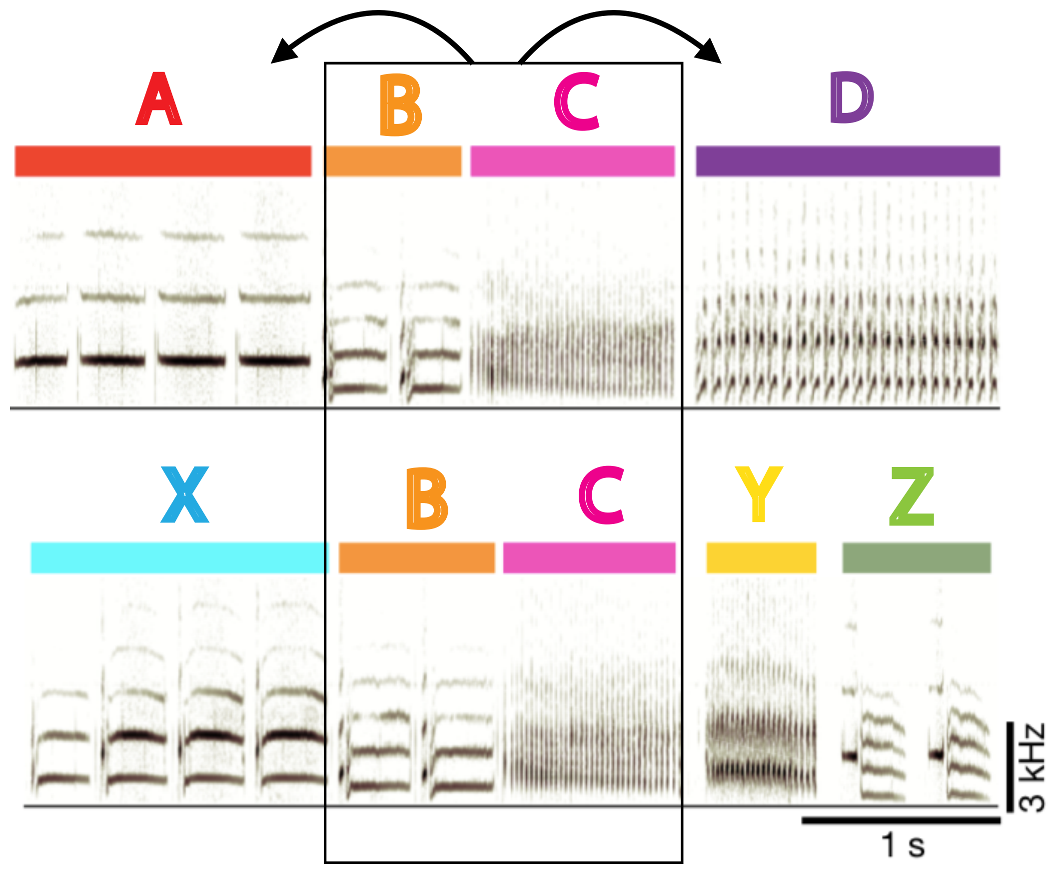
The cat near the trees catches the mouse

The cats near the trees catch the mouse

- Bird songs:

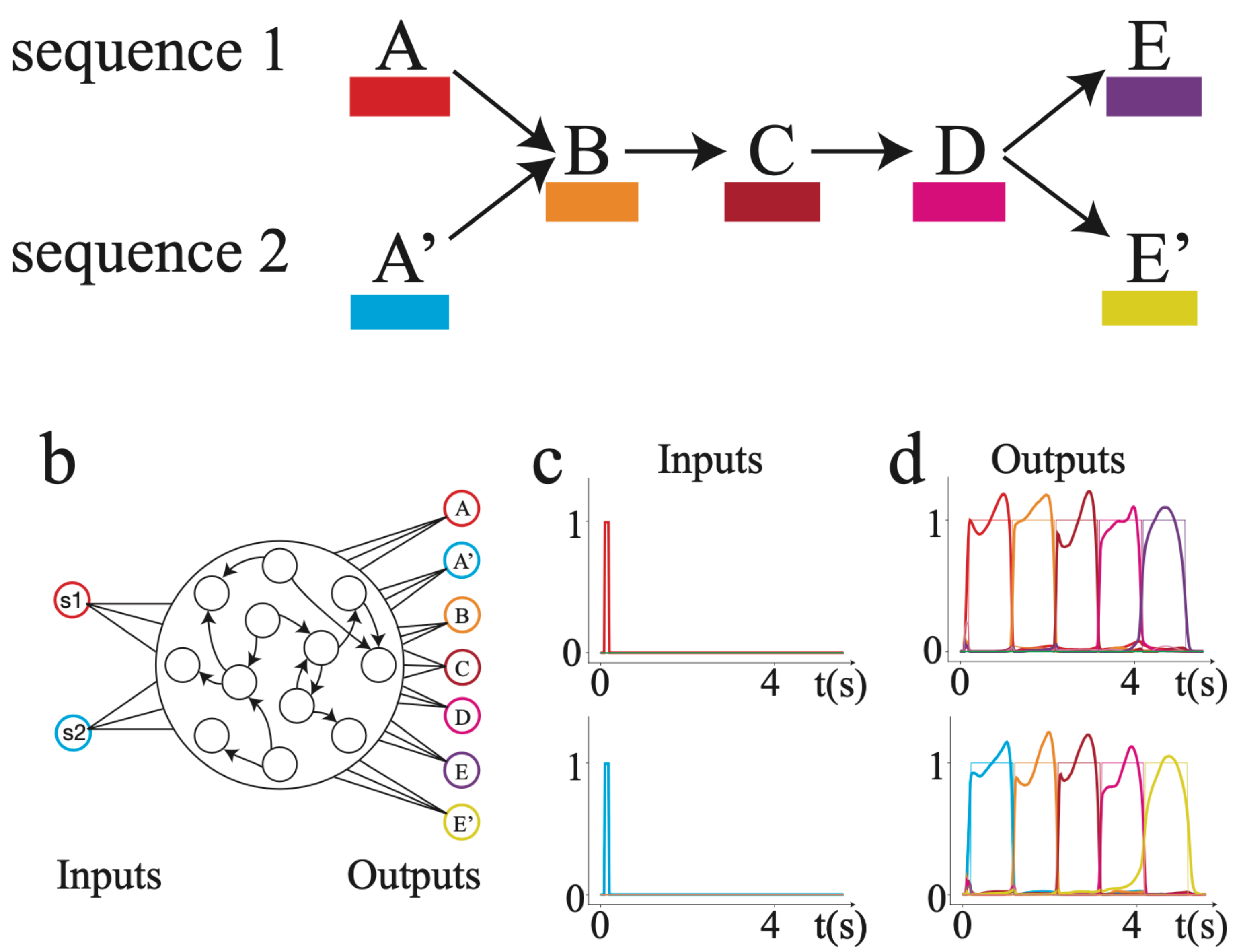
song 1

song 2

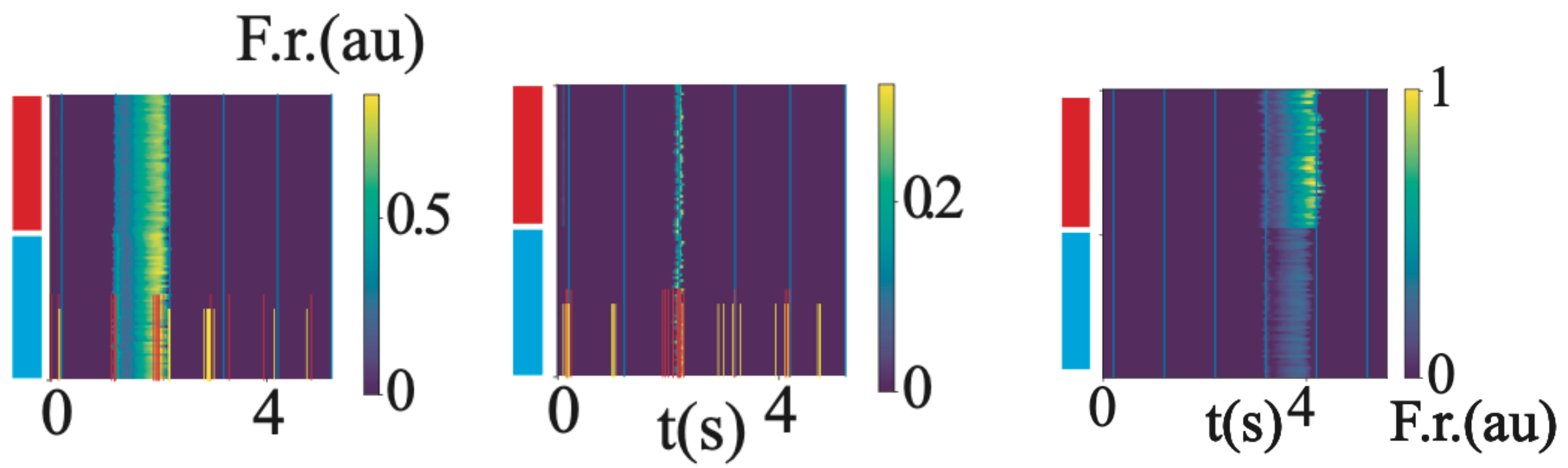


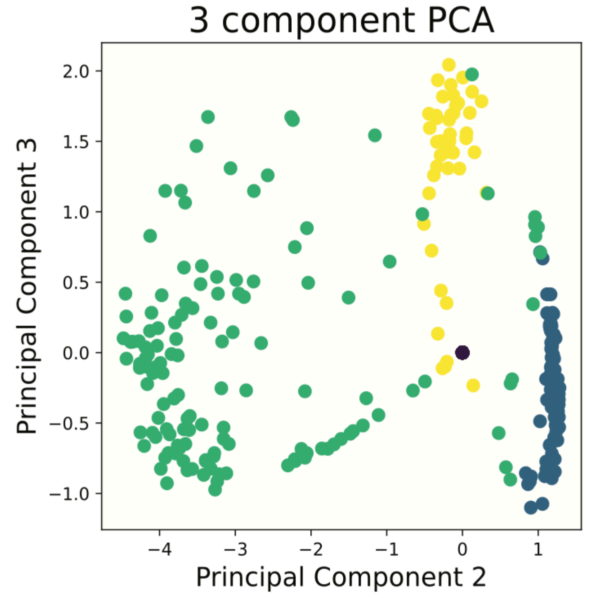
Markovitz et al, 2013

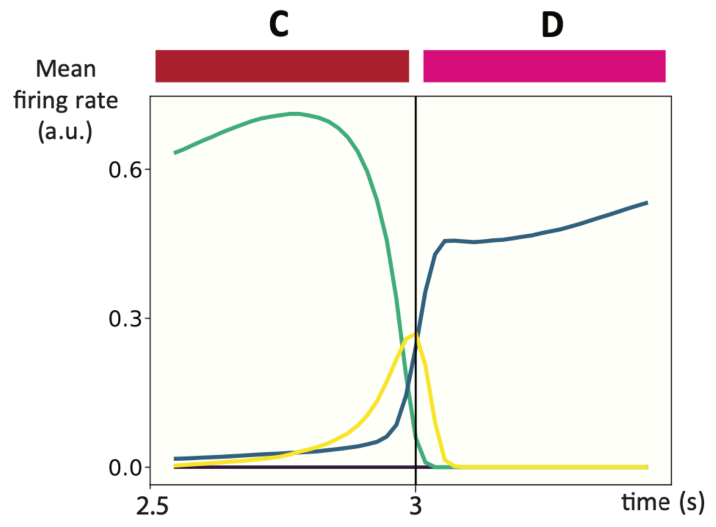
Long-distance dependencies in neural networks



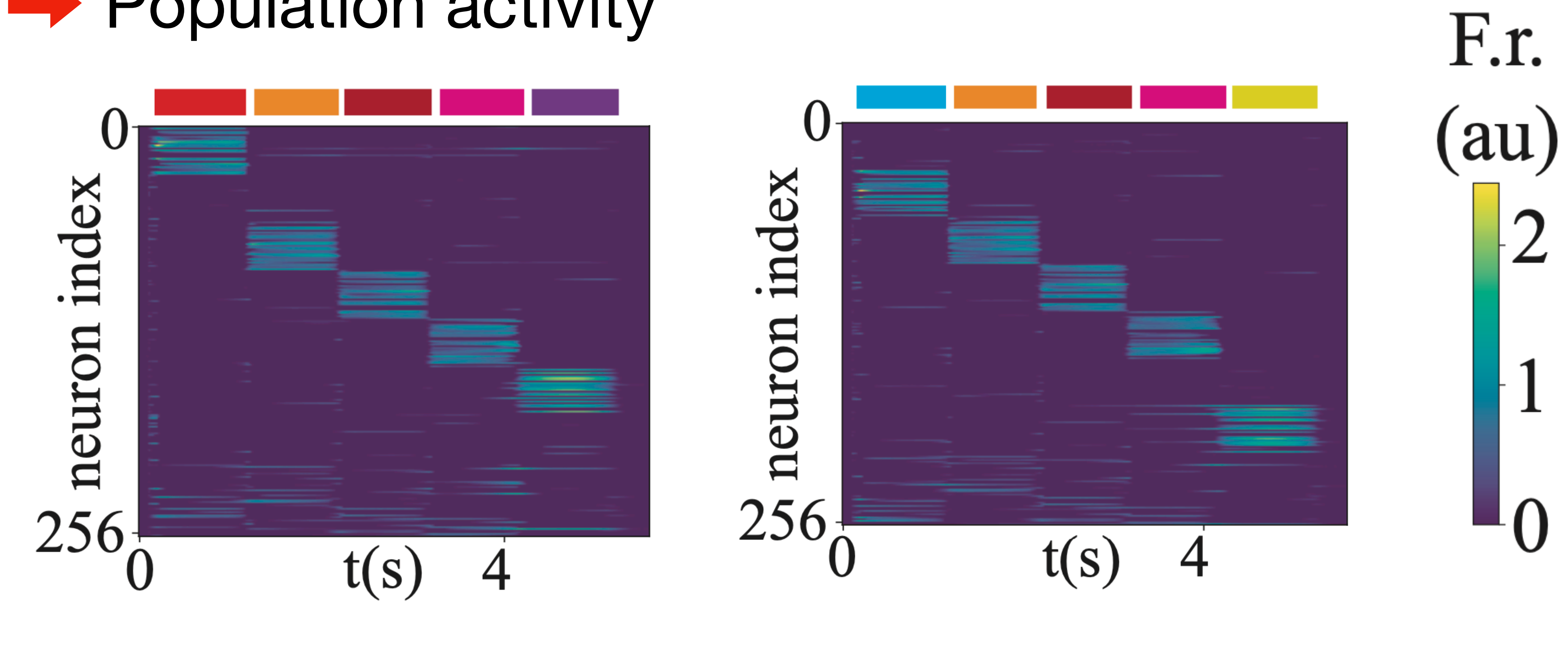
Single neuron activity





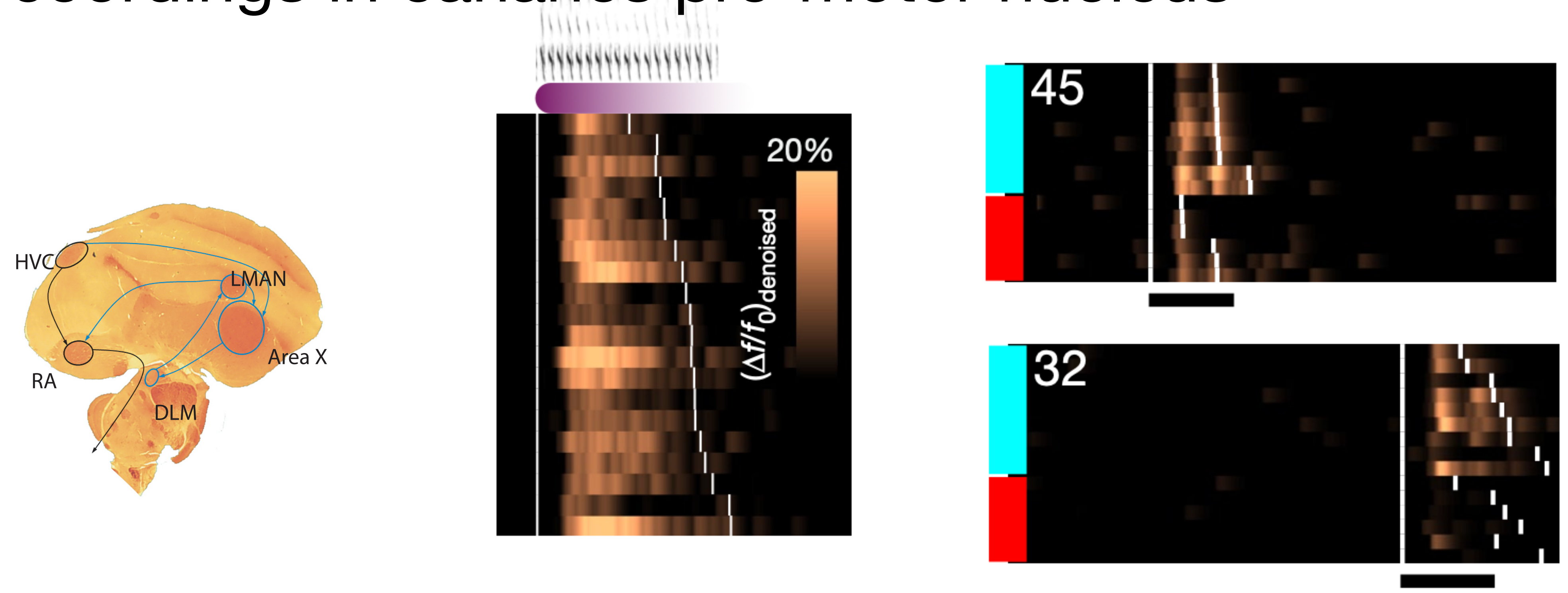


Population activity



Dubreuil & Leblois, 2021

Brain recordings in canaries pre-motor nucleus



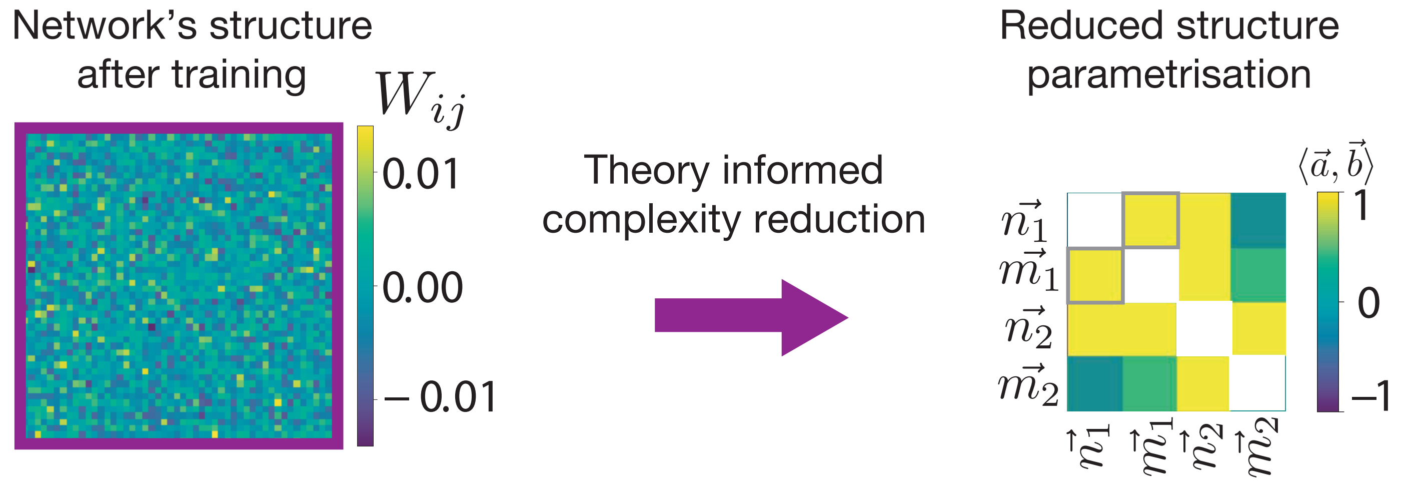
Cohen et al, 2020

Conclusion

- Dynamics of rate networks with low-rank connectivity can be reduced to dynamics on cognitive variables

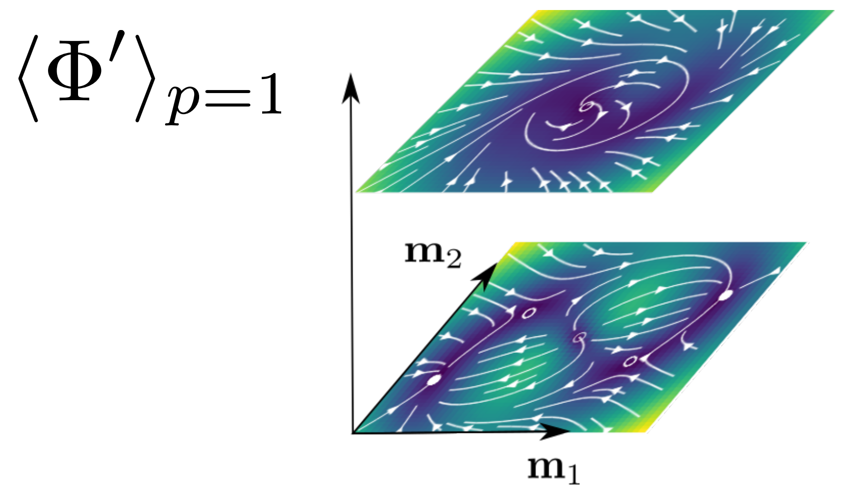
$$\dot{\kappa}_1 = -\kappa_1 + \tilde{\sigma}_{n_1 m_1}\kappa_1 + \tilde{\sigma}_{n_1 m_2}\kappa_2 + \tilde{\sigma}_{n_1 W_{in}} u(t)$$
$$\dot{\kappa}_2 = -\kappa_2 + \tilde{\sigma}_{n_2 m_1}\kappa_1 + \tilde{\sigma}_{n_2 m_2}\kappa_2 + \tilde{\sigma}_{n_2 W_{in}} u(t)$$

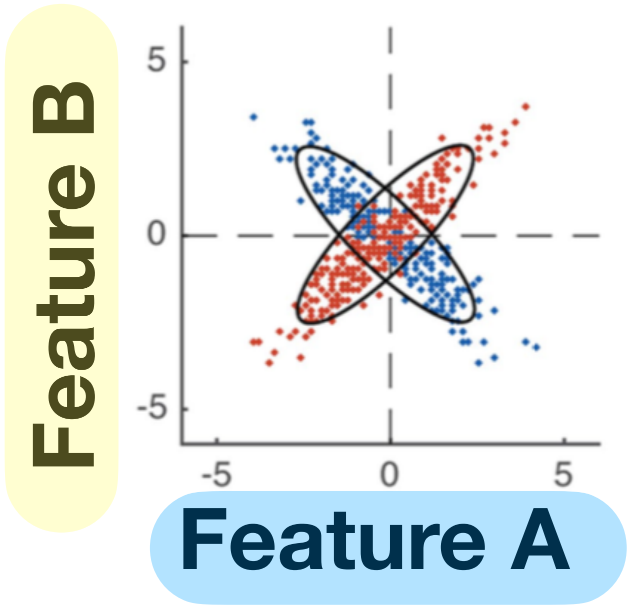
- Use of theory to reverse-engineer artificial neural networks



- Gain modulation of specific populations can reconfigure network's dynamics

  Neural correlates: specific connectivity profiles





Thanks

Group for Neural Theory, ENS, Paris

Adrian Valente
Francesca Mastrogiuseppe
Manuel Beiran
Srdjan Ostojic

Physics Laboratory, ENS, Paris

Rémi Monasson

Institute for Neurodegenerative Diseases, Bordeaux

Arthur Leblois
Vimala Maillard