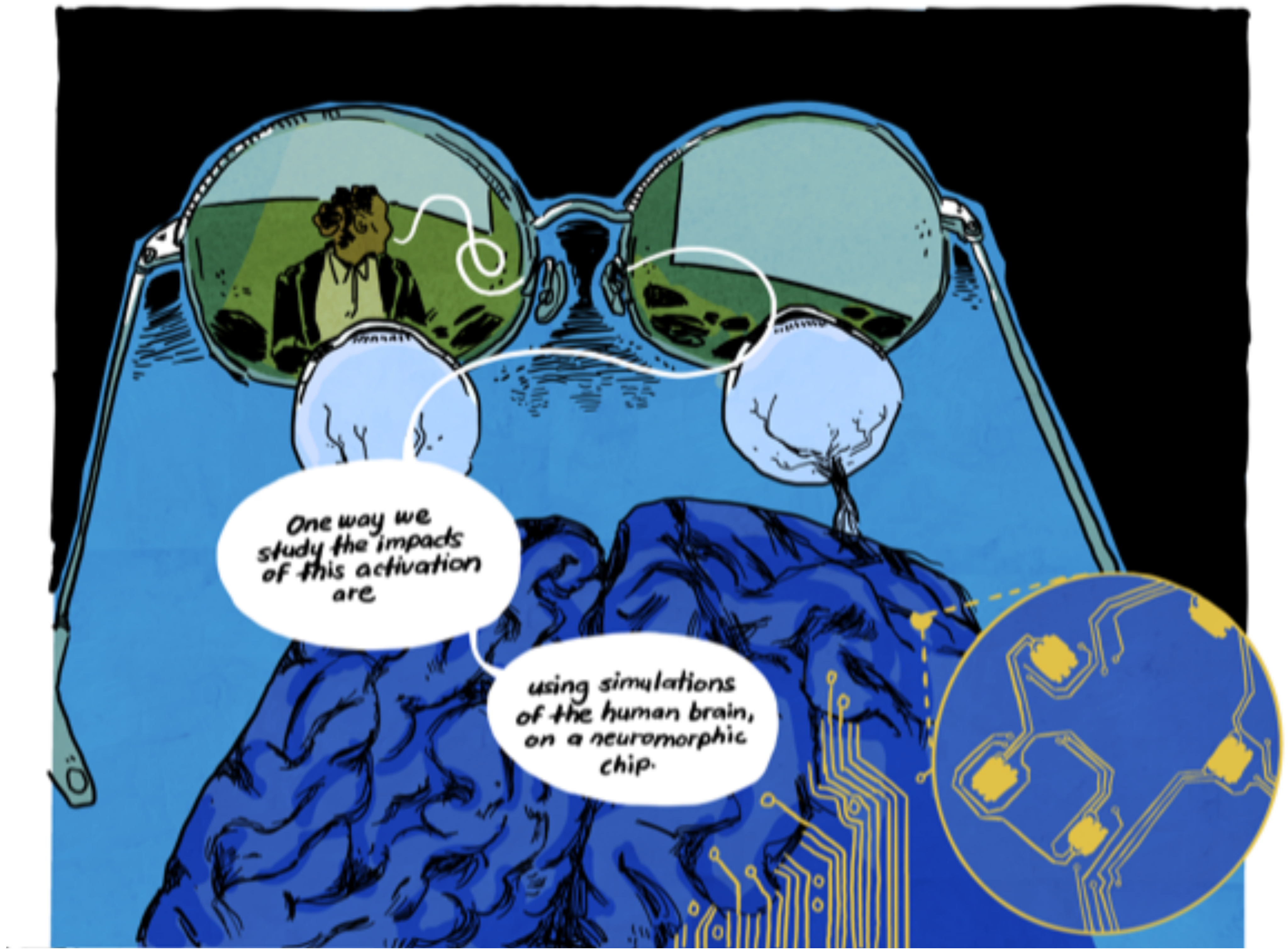
# Dynamical systems for behavioral and neural data


Luca Mazzucato

Depts of Biology, Mathematics & Physics
Institute of Neuroscience
University of Oregon


CSHA Computational and Cognitive Neuroscience Summer School




Artwork by Audra McNamee

# Dynamical systems for behavioral and neural data

- TODAY: hidden Markov models, neural data & RNNs:
  - Slides
    - Spontaneous movements - HMM with Autoregressive observations
    - Naturalistic foraging - HMM with Linear Model emissions
    - HMM best practices
  - Live coding session
    - Mystery dataset
    - Fit Linear Model, Mixture of Linear Model
    - Sample and fit data w/ LM-HMM
    - Model selection for # of hidden states

- Neural data
    - Behavior is metastable (Datta) -> Metastable neural activity (Zach)?
        - spike data + HMM tutorial
        - Poisson HMM basics fit to Zach’s session
        - Calculate dimensionality vs shuffled
- RNN
    - Attractor networks
    - HMM fit to attractors

# Cortical computations via metastable activity

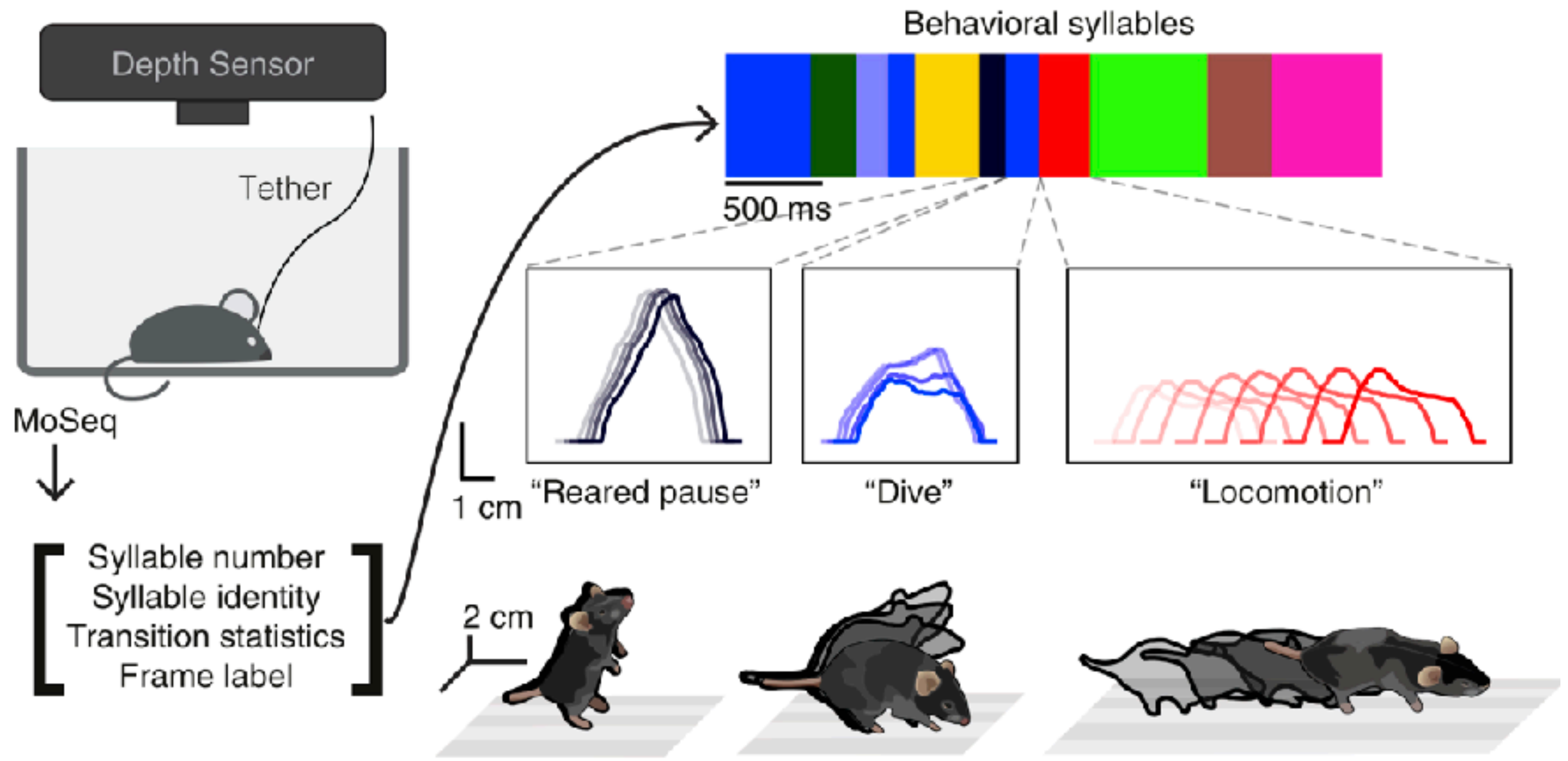

Luca Mazzucato

Depts. of Mathematics, Biology & Physics

Institute of Neuroscience

University of Oregon

# Complex temporal structure in behavior




Variability sources:

1 - many syllables / actions

2 - many words / sequences

3 - transition times (action duration)

[Wiltschko et al, 2015]
[Markovits et al, 2018]

# Complex temporal structure in behavior

# …originating from neural circuit dynamics?

**Neural mechanisms underlying the temporal organization of naturalistic animal behavior**

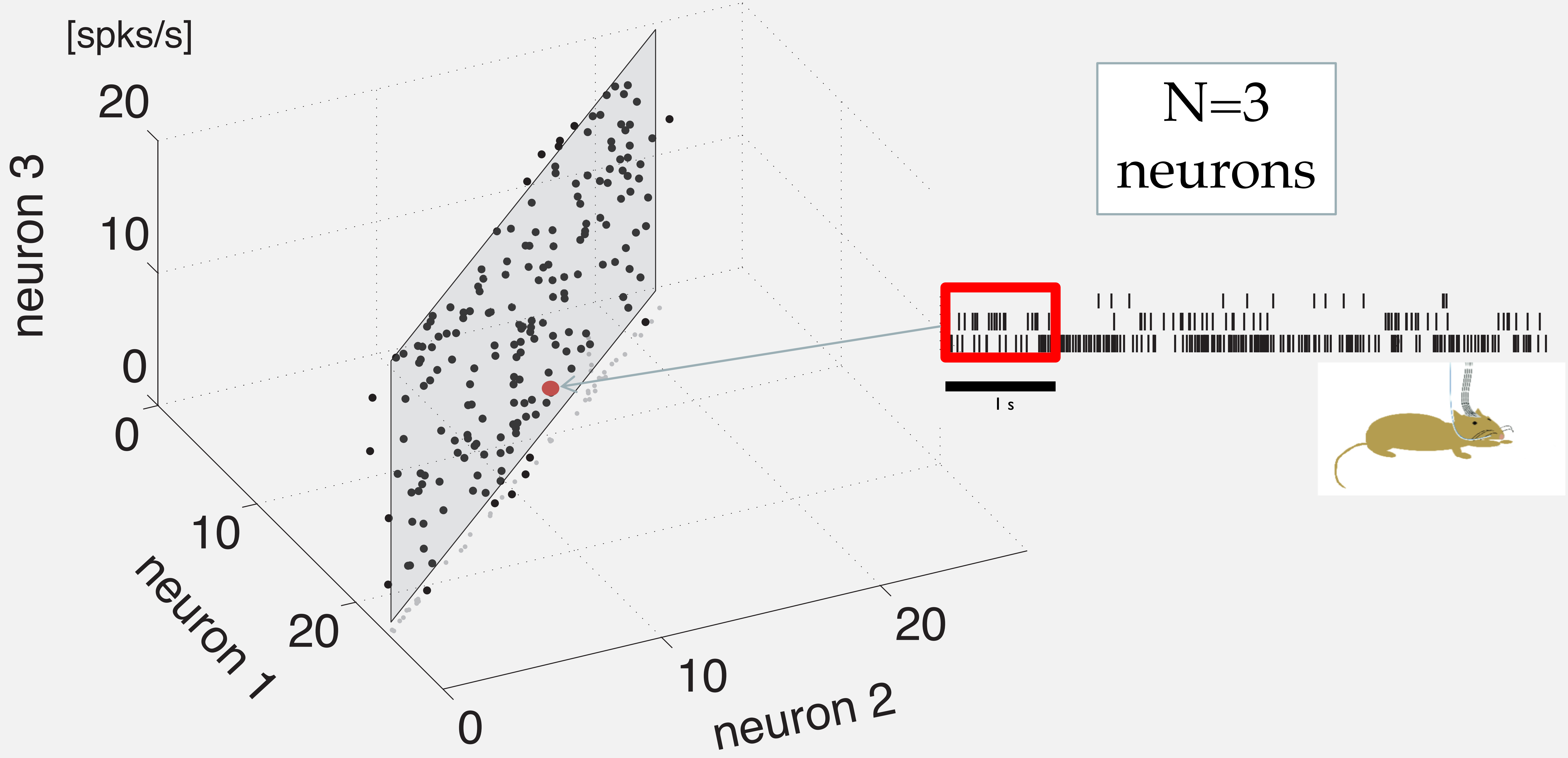
Luca Mazzucato*

eLife

Population activity is a point in N-dimensional space

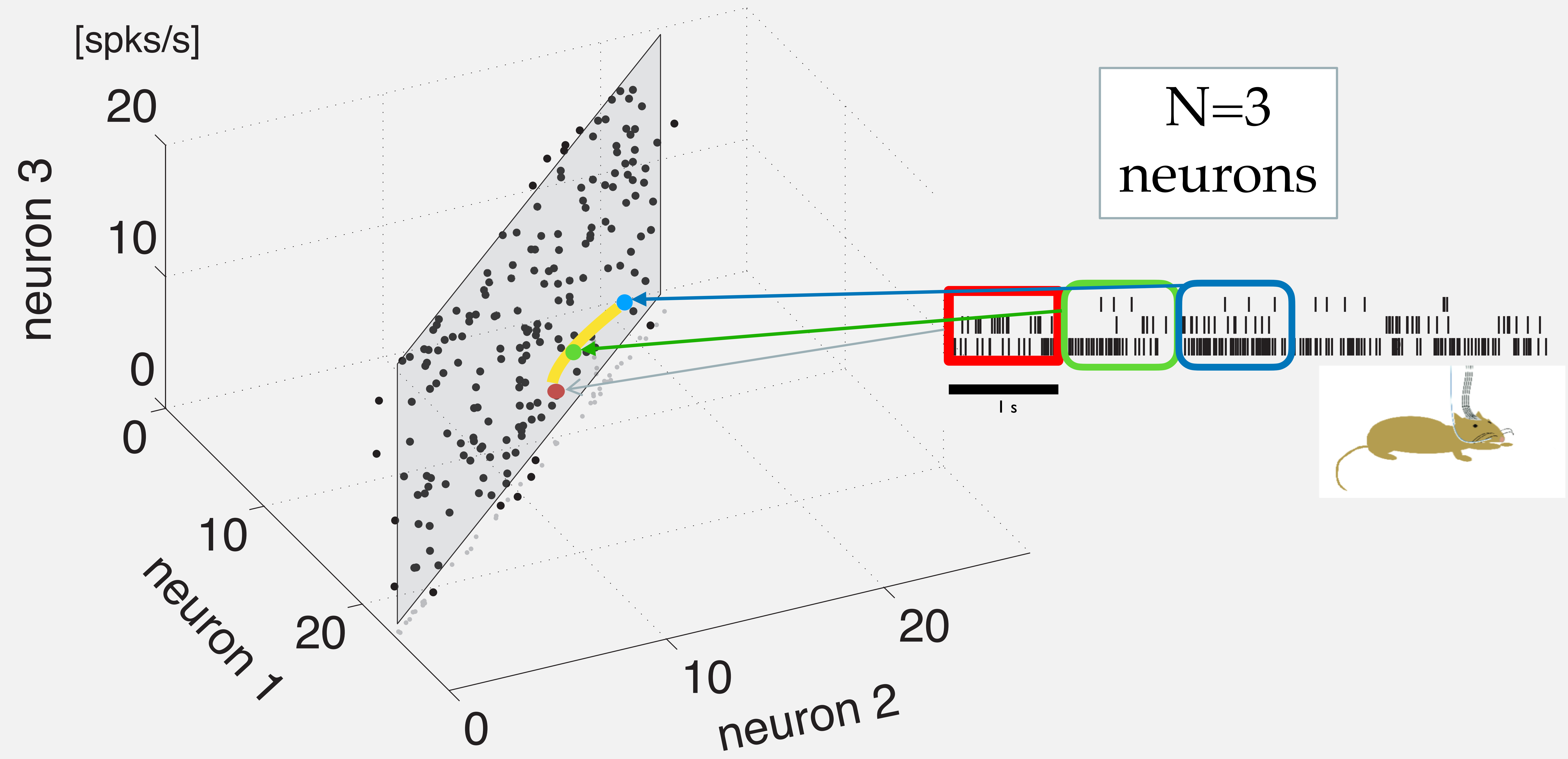
N=3 neurons


$$\overrightarrow{X}(t_1) = [0, 12, 38] \text{ spks/s}$$




N=3
neurons

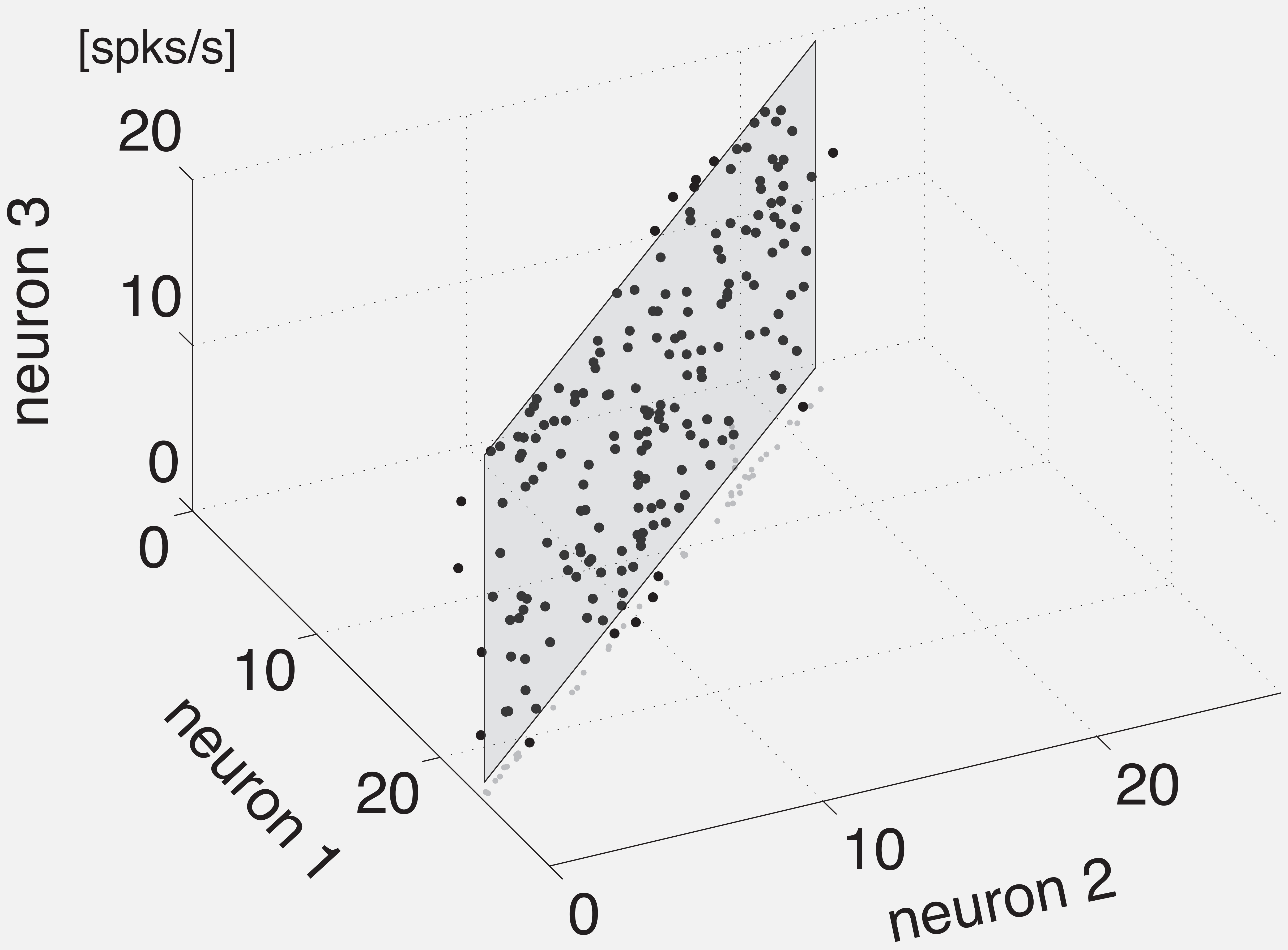
$d$ = smallest # of coord.
axes accounting for
neural pop. activity




N=3
neurons

$d$ = smallest # of coord.
axes accounting for
neural pop. activity

# Neural dimensionality




N=3
neurons

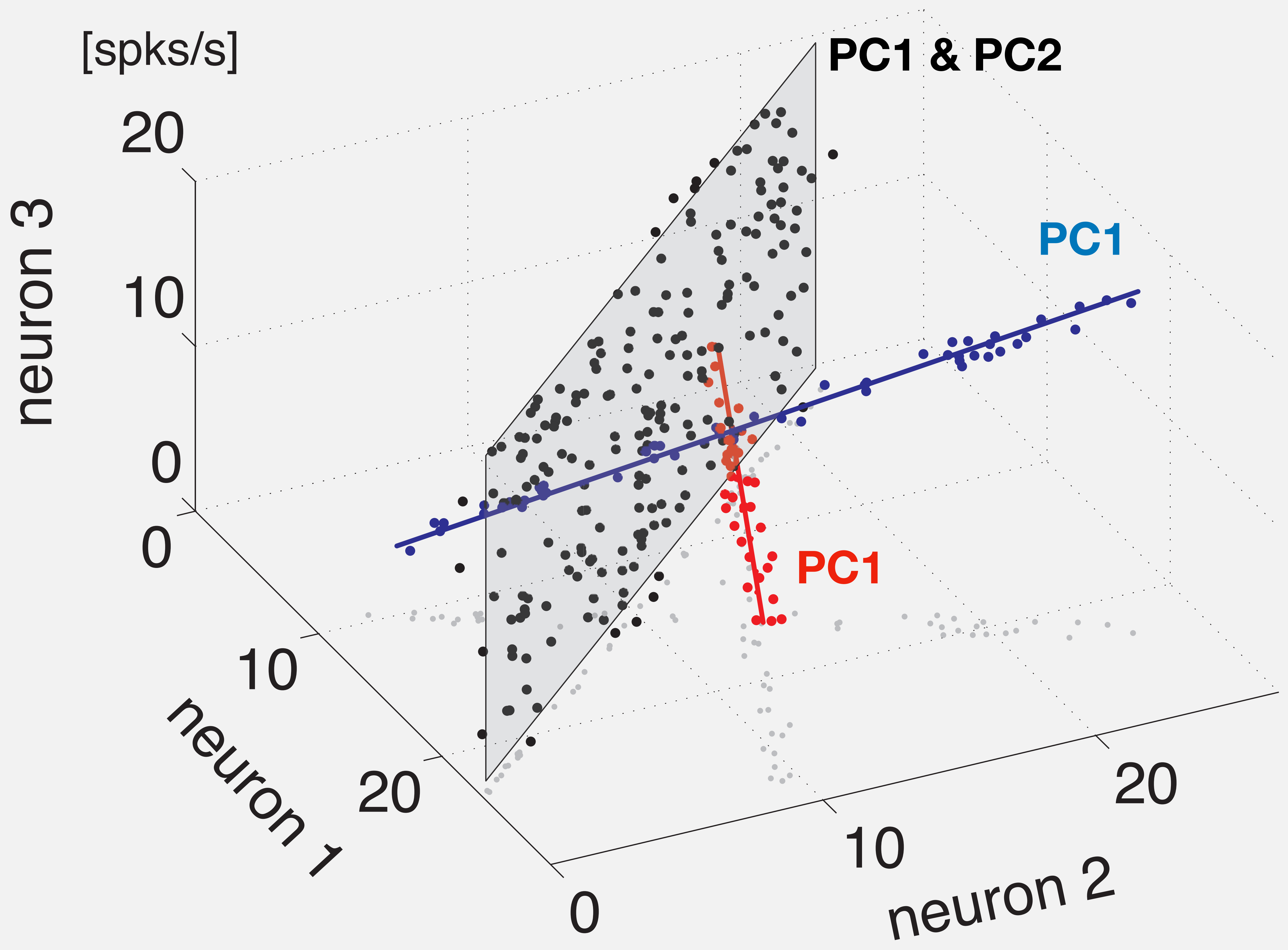
2 coordinate axes
required:

$d = 2$

$d$ = smallest # of coord. axes accounting for neural pop. activity

# Neural dimensionality

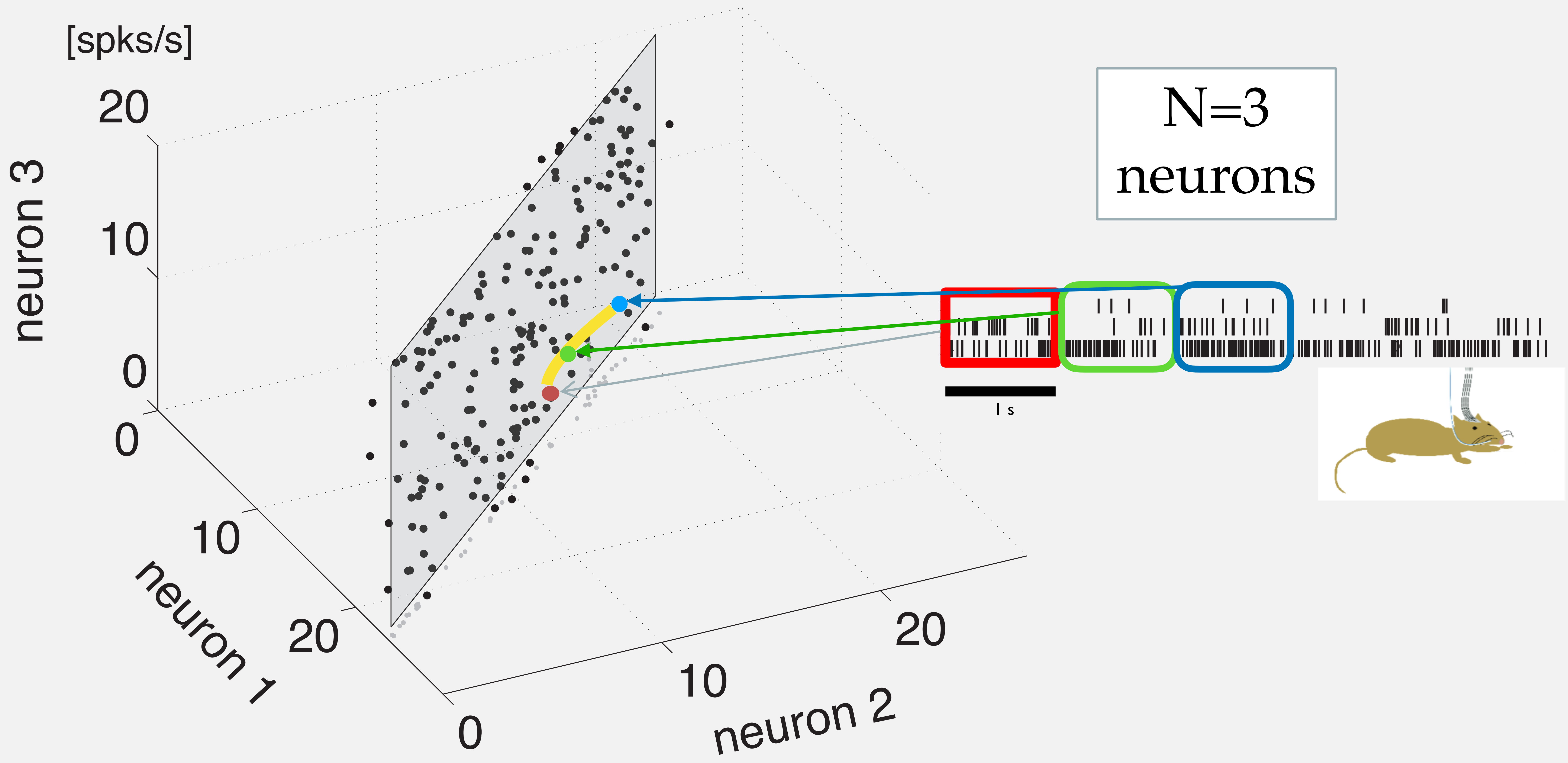
**Principal Component Analysis (PCA)**




N=3
neurons

- ongoing d=2
- stimulus A d=1
- stimulus B d=1

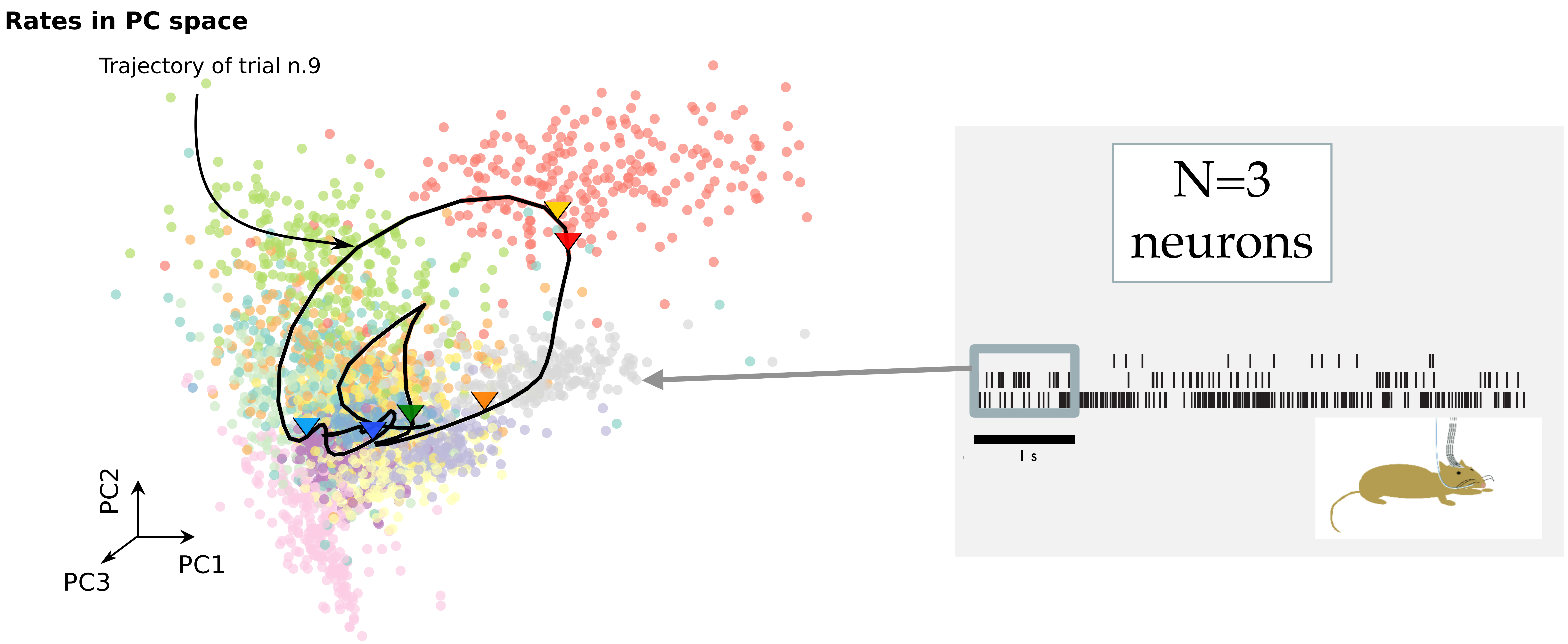
$d$ = smallest # of coord. axes accounting for neural pop. activity




N=3
neurons

$d$ = smallest # of coord. axes accounting for neural pop. activity

# Population activity



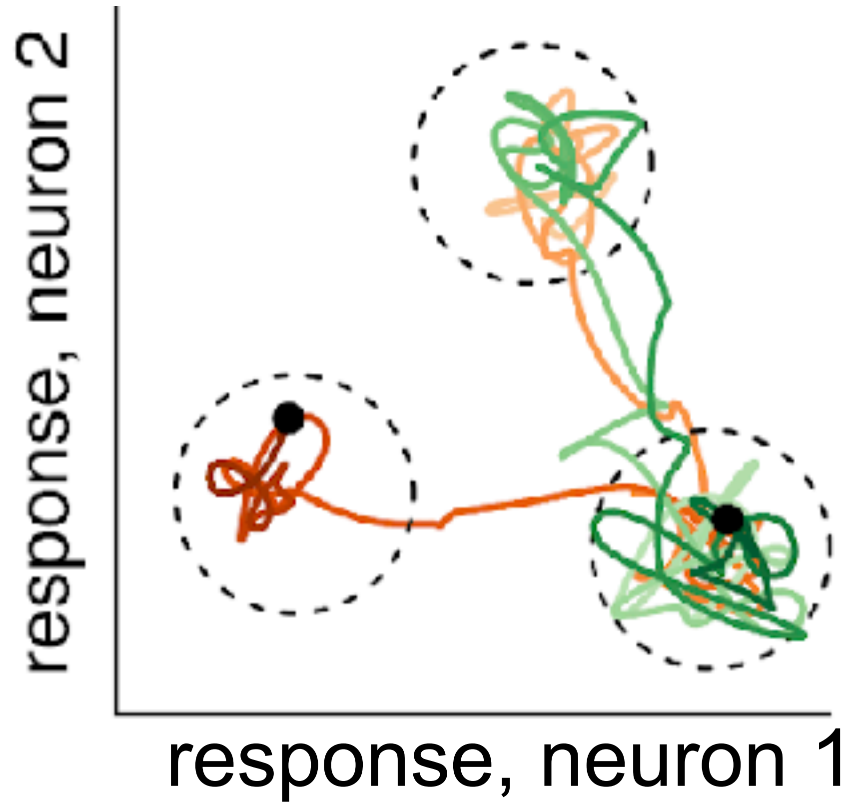
# Temporal dynamics in neural circuits

Dynamical repertoires:

Nonlinear

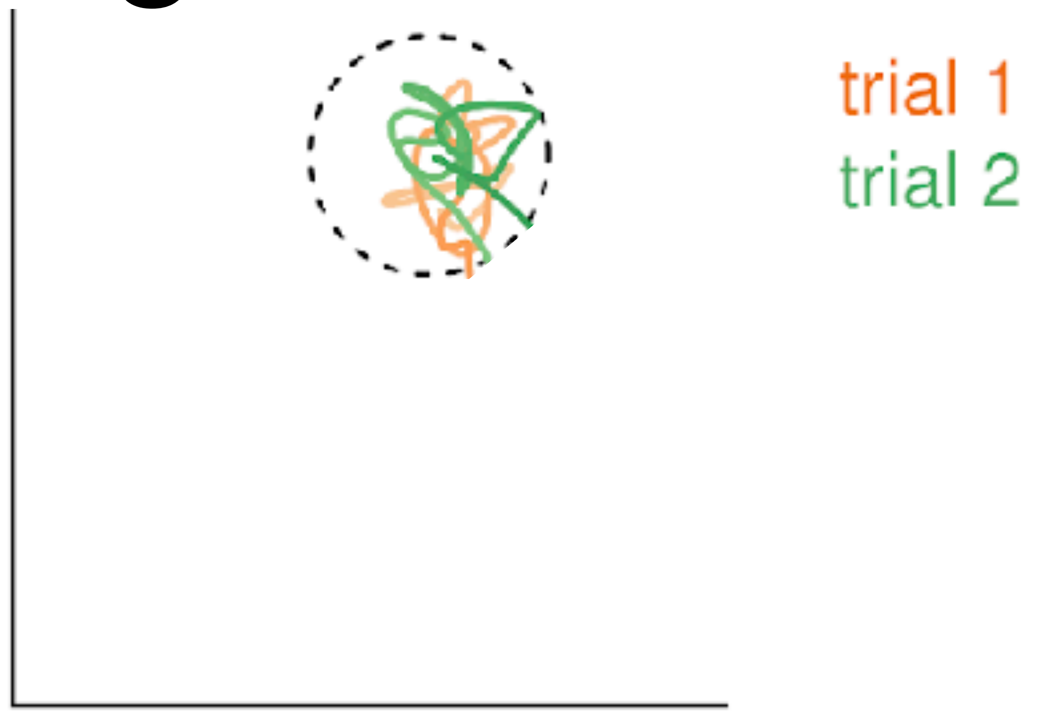
Linear

Multi-attractor

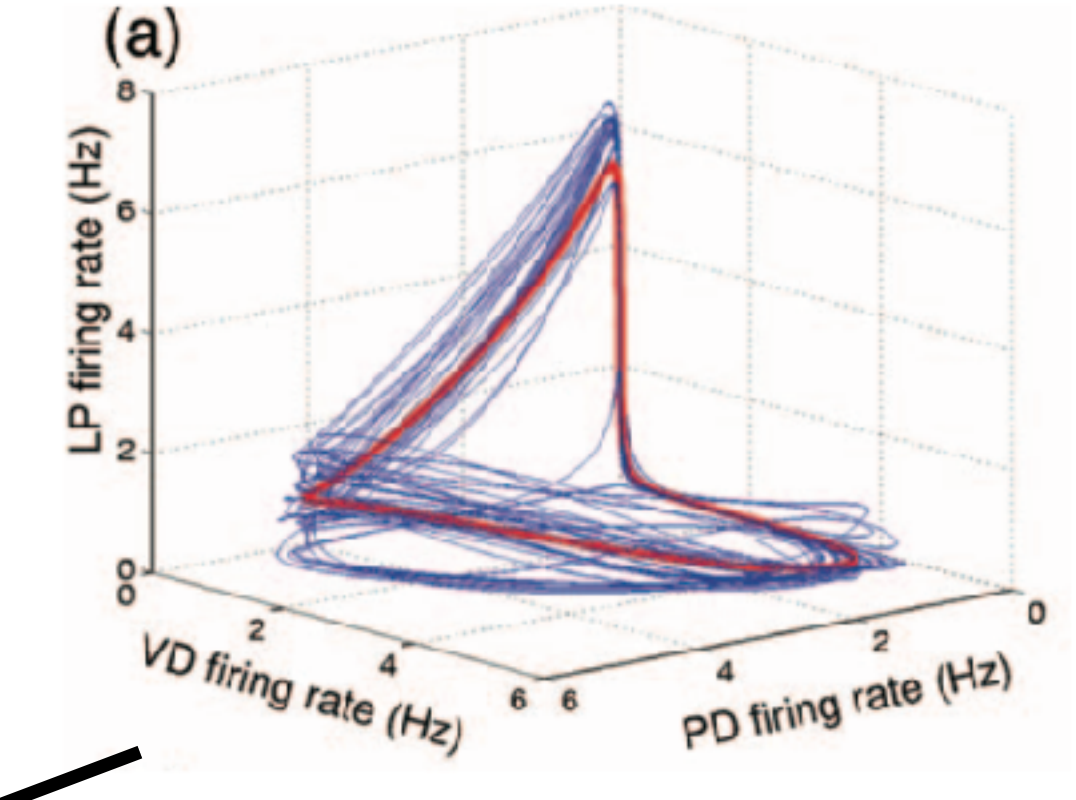



Chaos

Single-attractor




Oscillation

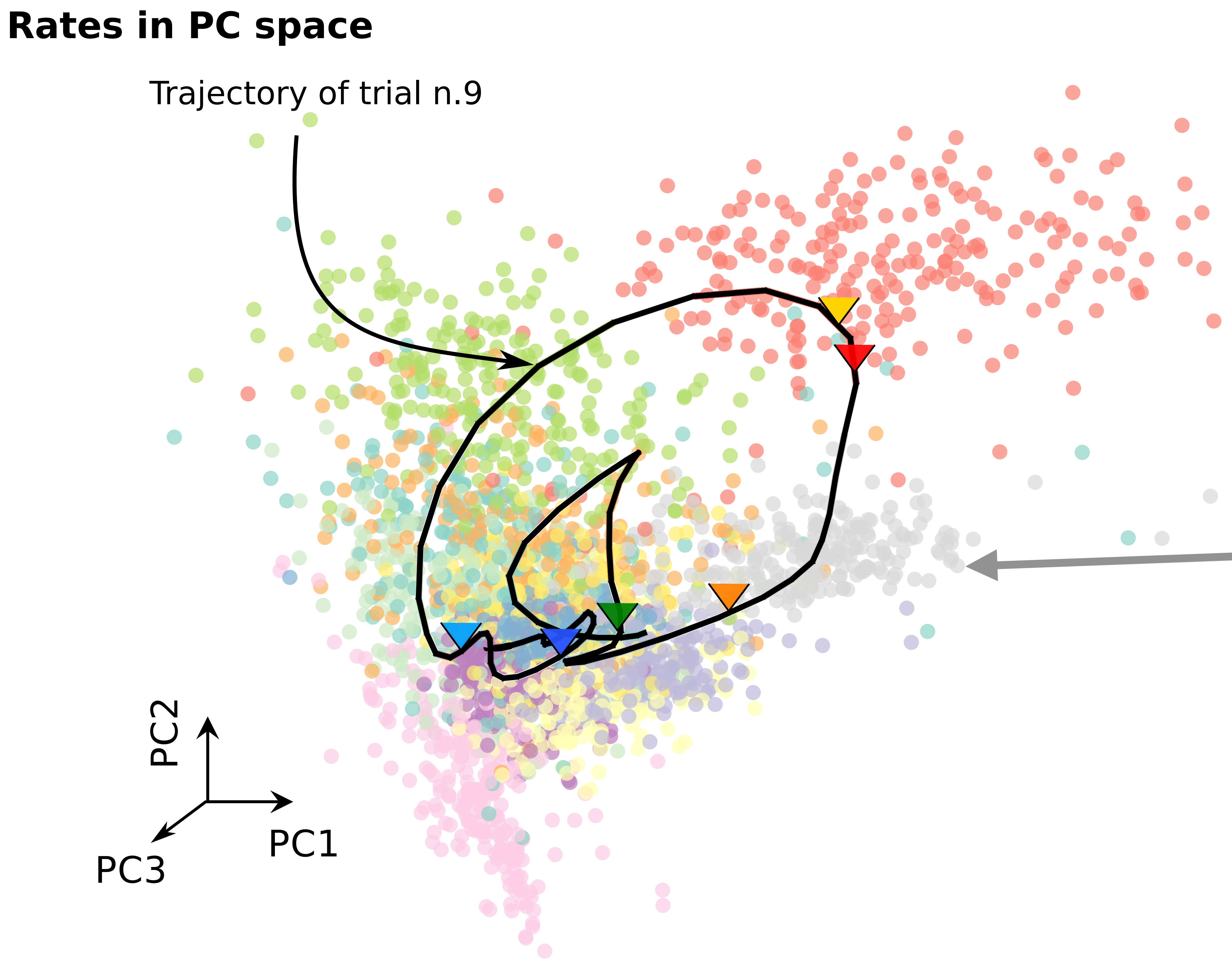


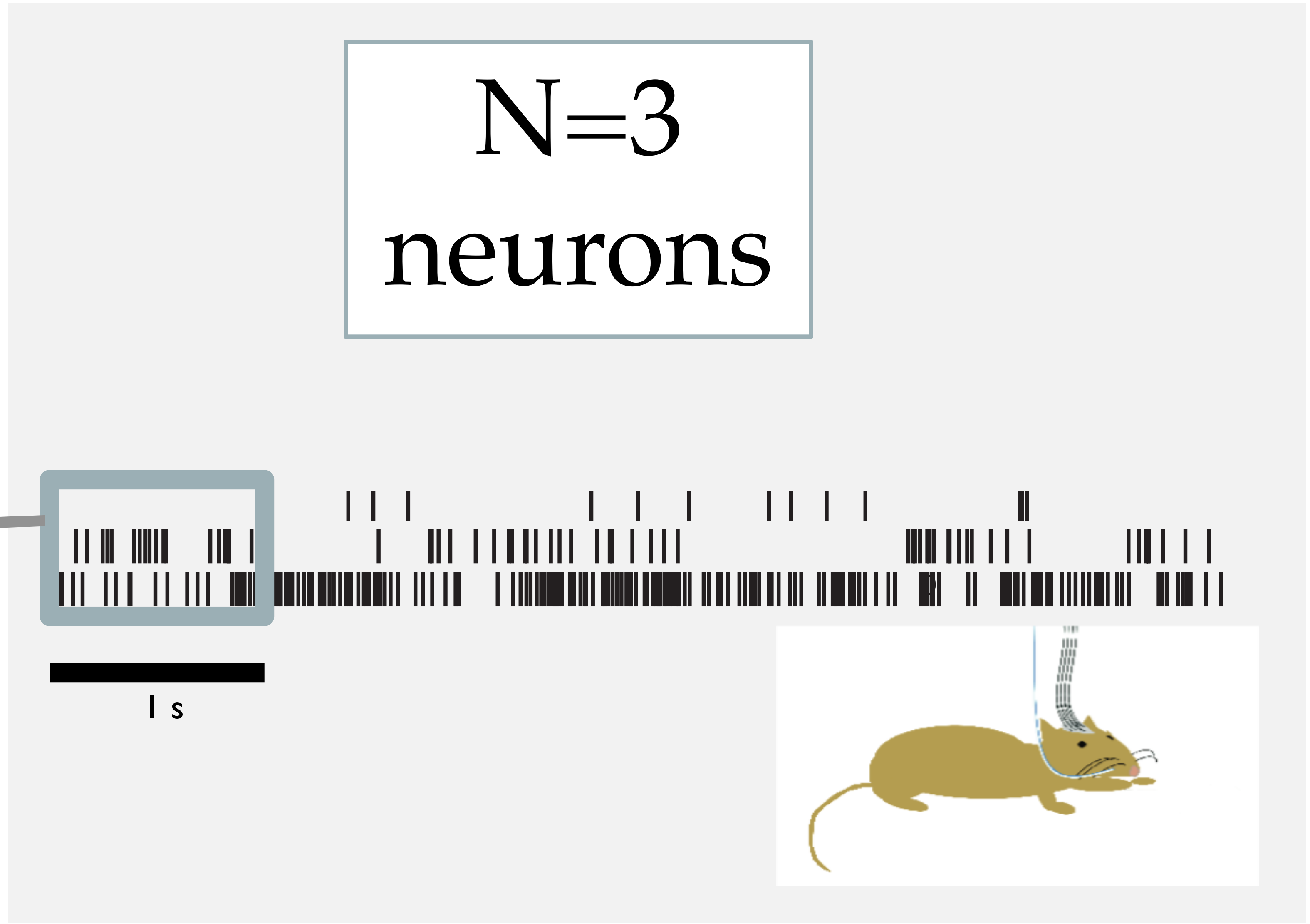

Jumping

Varying smoothly

[Rabinovich et al, 2006]
[Hennequin et al, 2018]

# Population activity



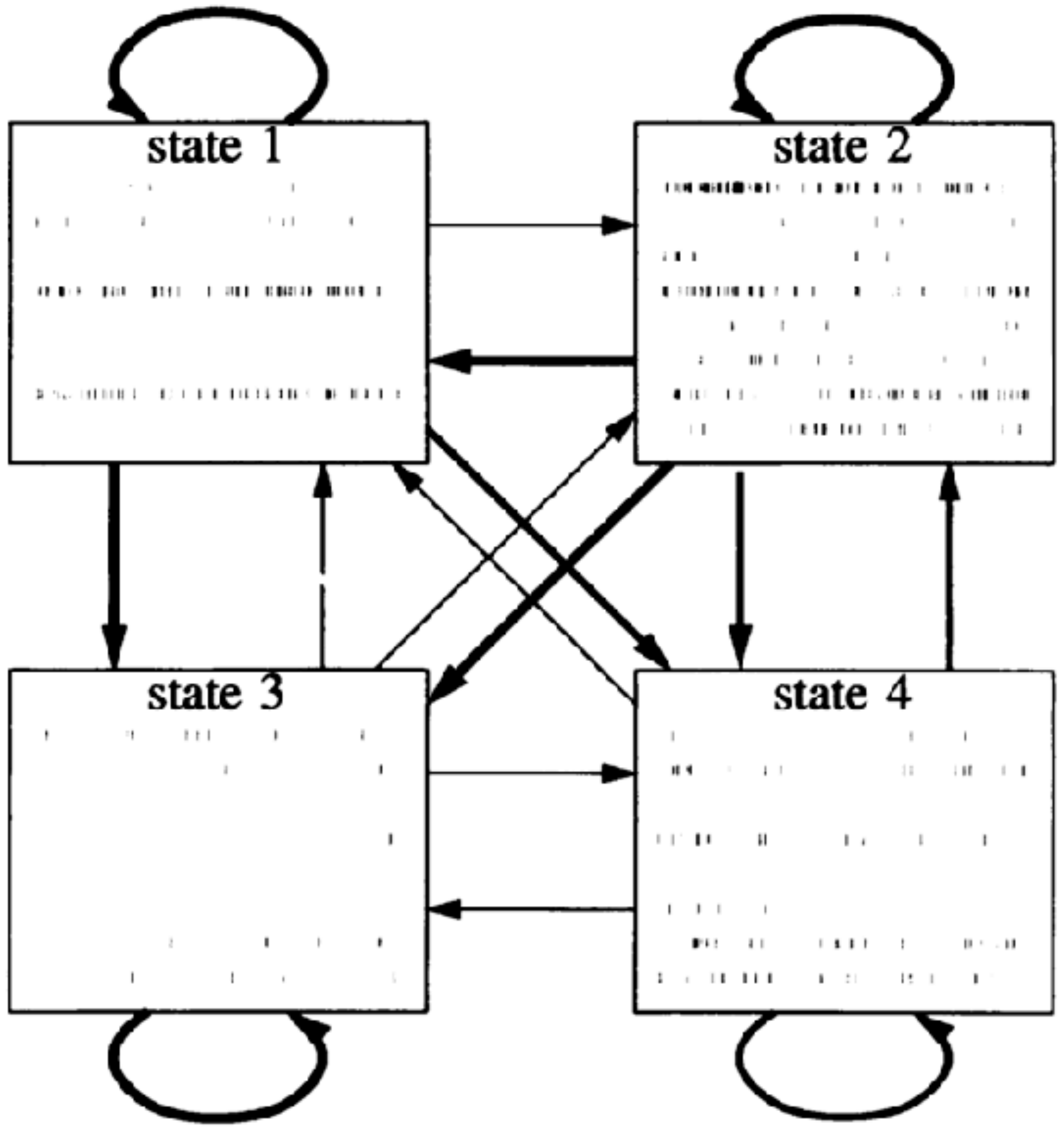
# Metastable activity

*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 8616–8620, September 1995
Neurobiology

## Cortical activity flips among quasi-stationary states

**(single-unit activity/behaving monkeys/hidden Markov models)**

Moshe Abeles*†, Hagai Bergman*, Itay Gat‡, Isaac Meilijson§, Eyal Seidemann§, Naftali Tishby‡, and Eilon Vaadia*



# Goals

**Part I - Theory**

- **Metastable activity in cortical circuits**
- **Hidden Markov Models fit to metastable activity**
- **Modeling metastable activity with attractor networks**

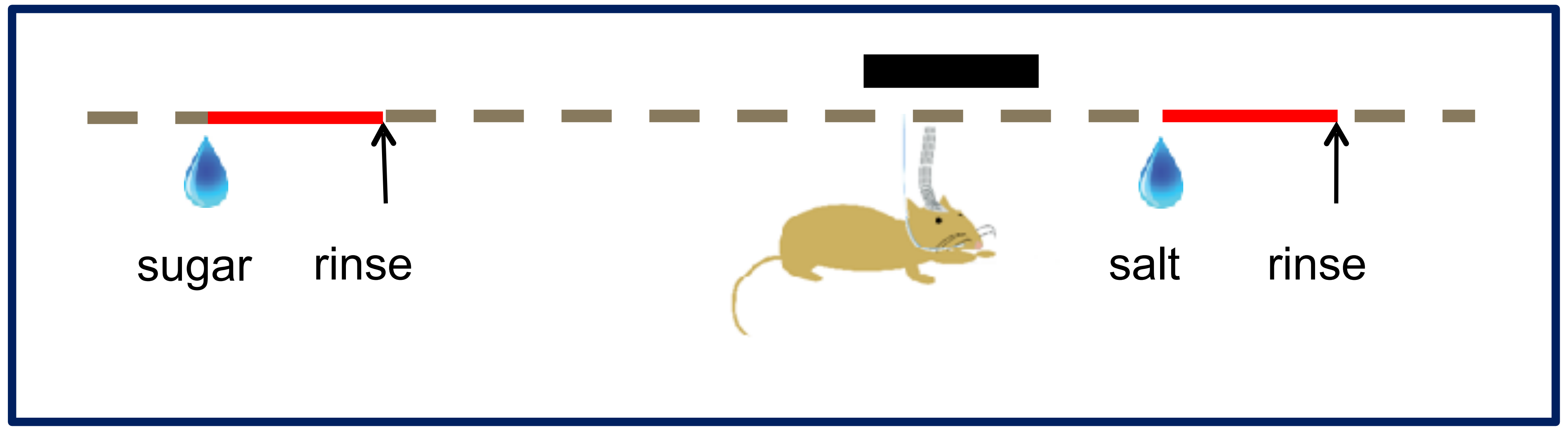
**Part II - Exercise**

- **HMM: model selection with gaussian emissions**
- **Fit Poisson-HMM to spike trains**
- **Challenge: A model of arousal in attractor networks**

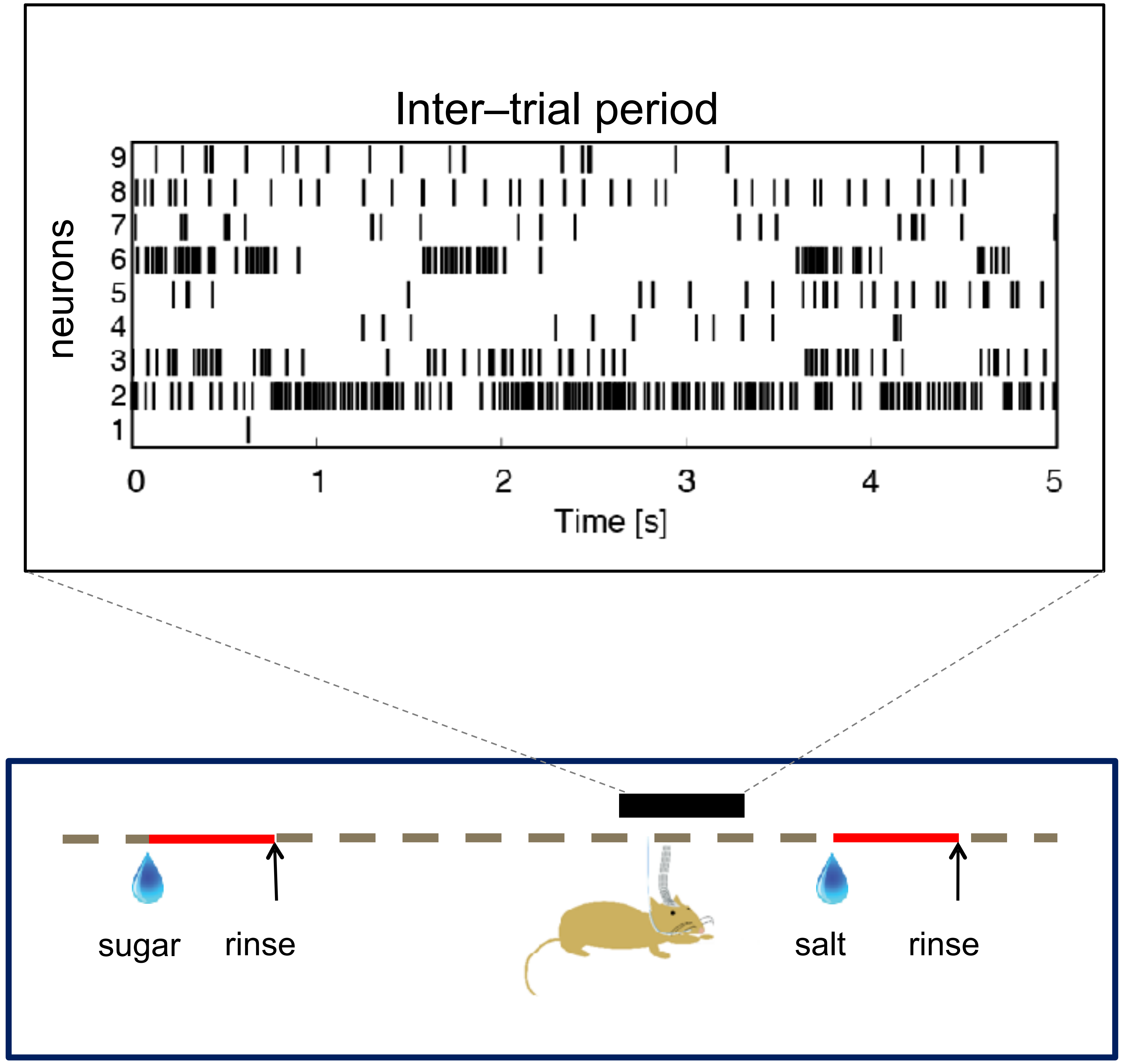
Fit Hidden Markov Models
to neural spike trains…

…and investigate metastable activity

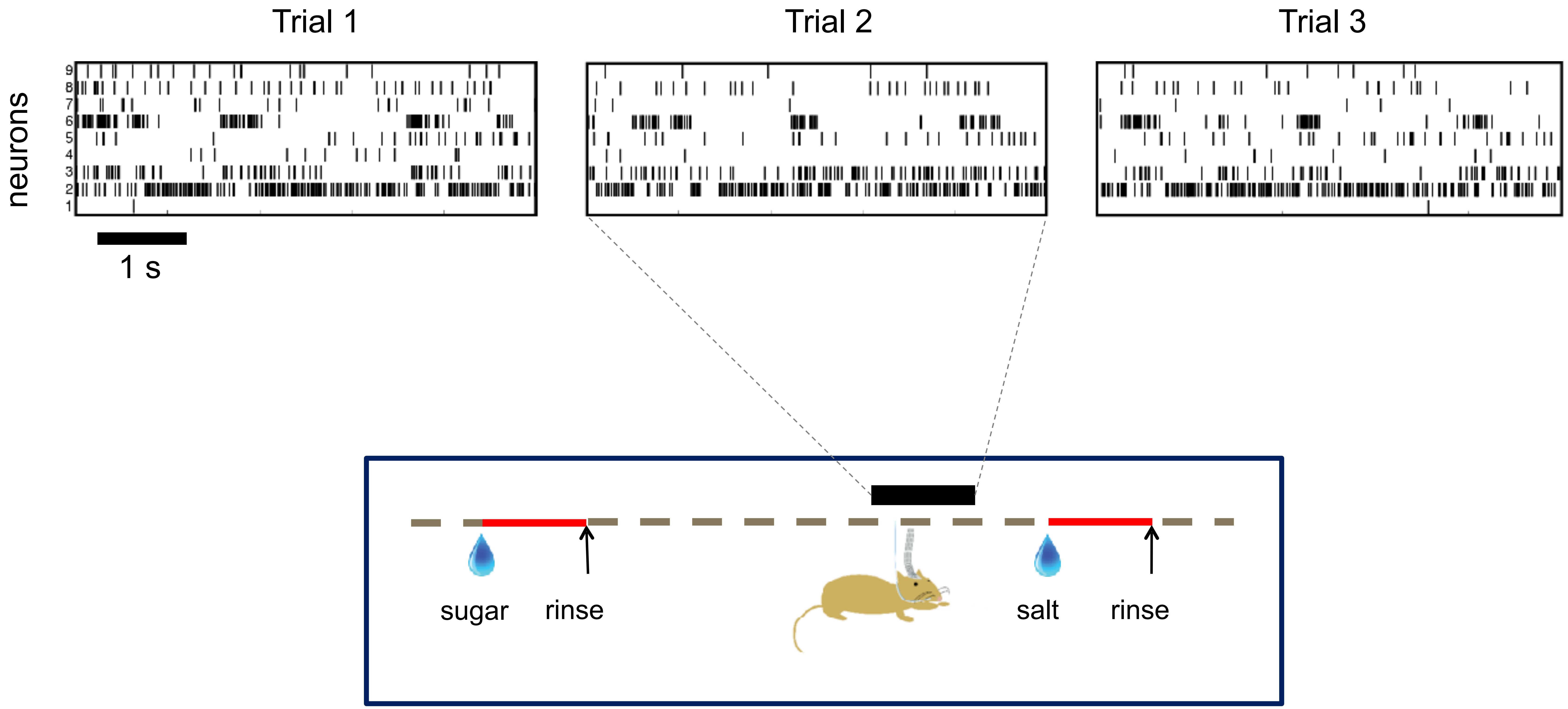
# HMM for neural spike trains



# HMM for neural spike trains

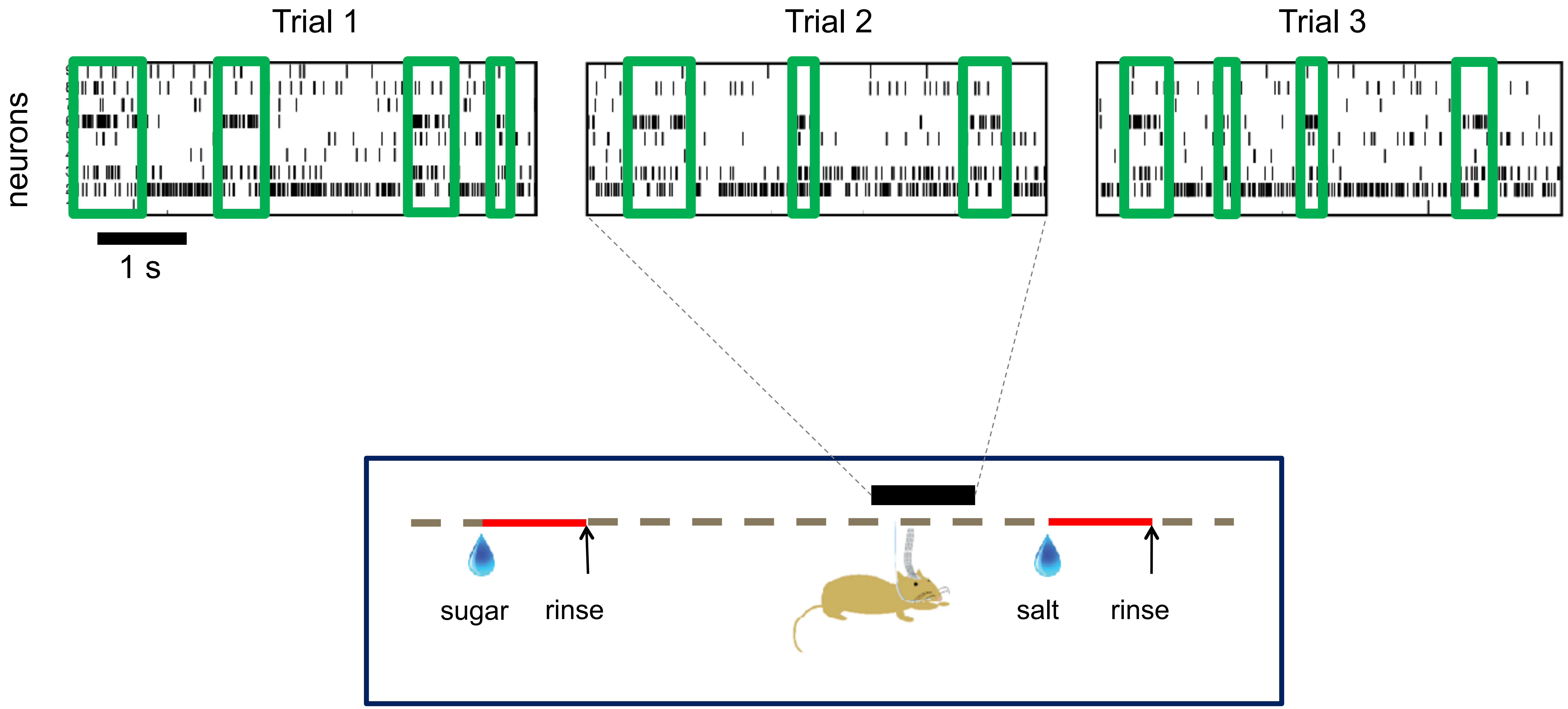


# HMM for neural spike trains



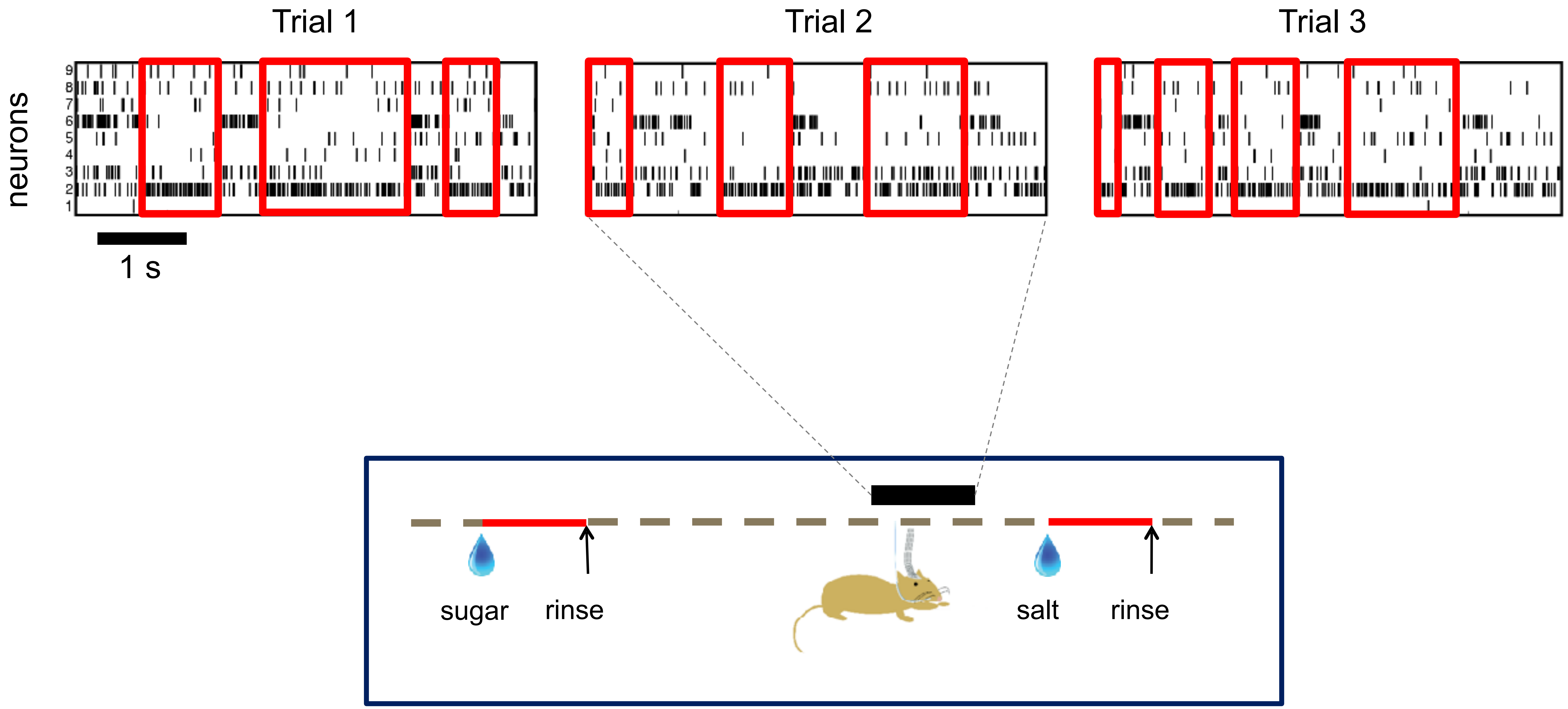

[Mazzucato et al., 2015, 2016]

# HMM for neural spike trains




[Mazzucato et al., 2015, 2016]

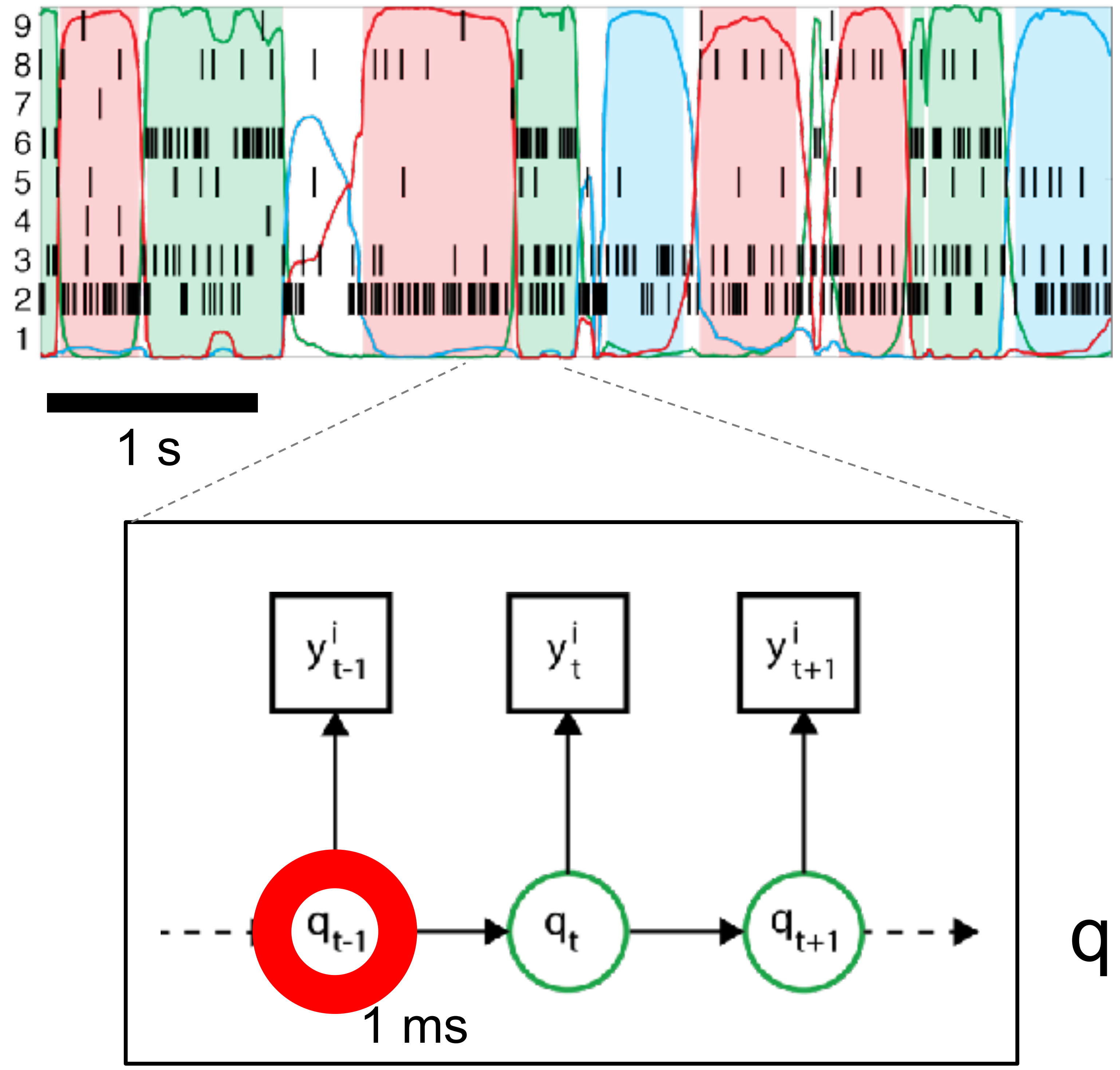
# HMM for neural spike trains




[Mazzucato et al., 2015, 2016]

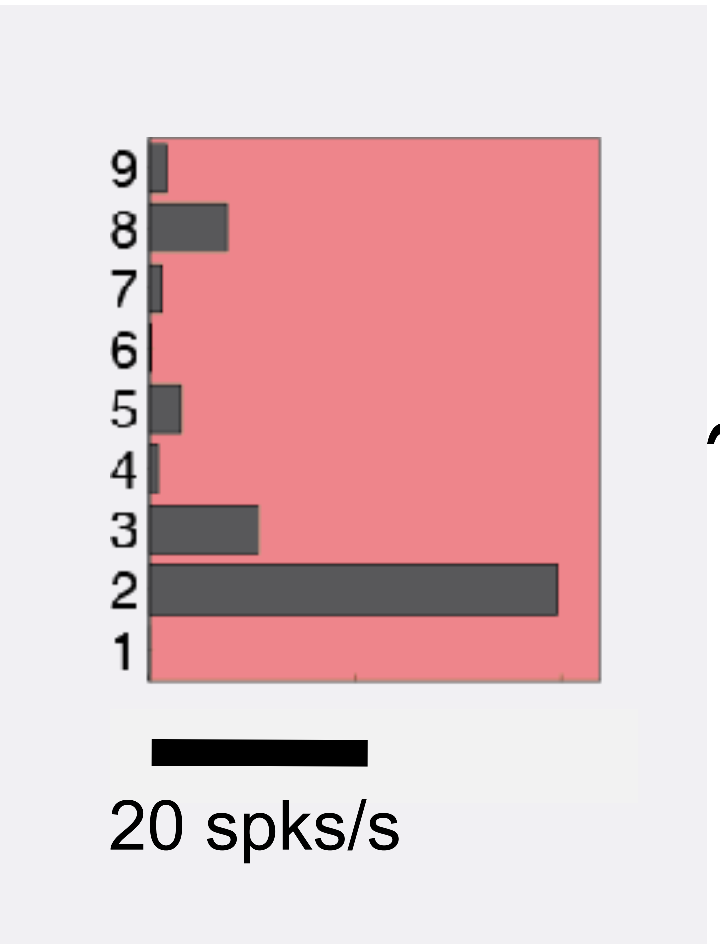
# HMM for neural spike trains

- Hidden Markov Model fit




q = hidden state

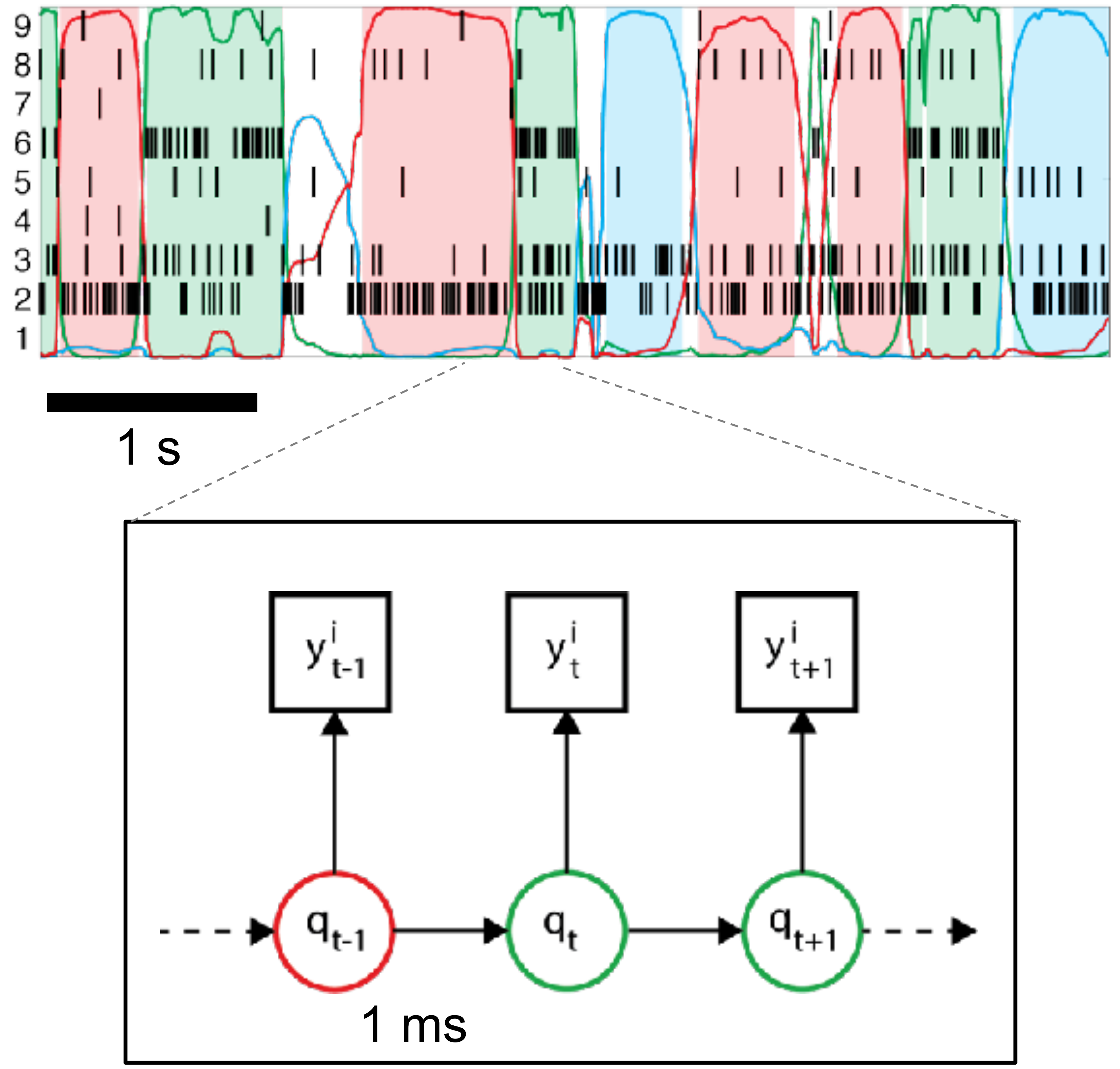
Firing rate vector $r_i$

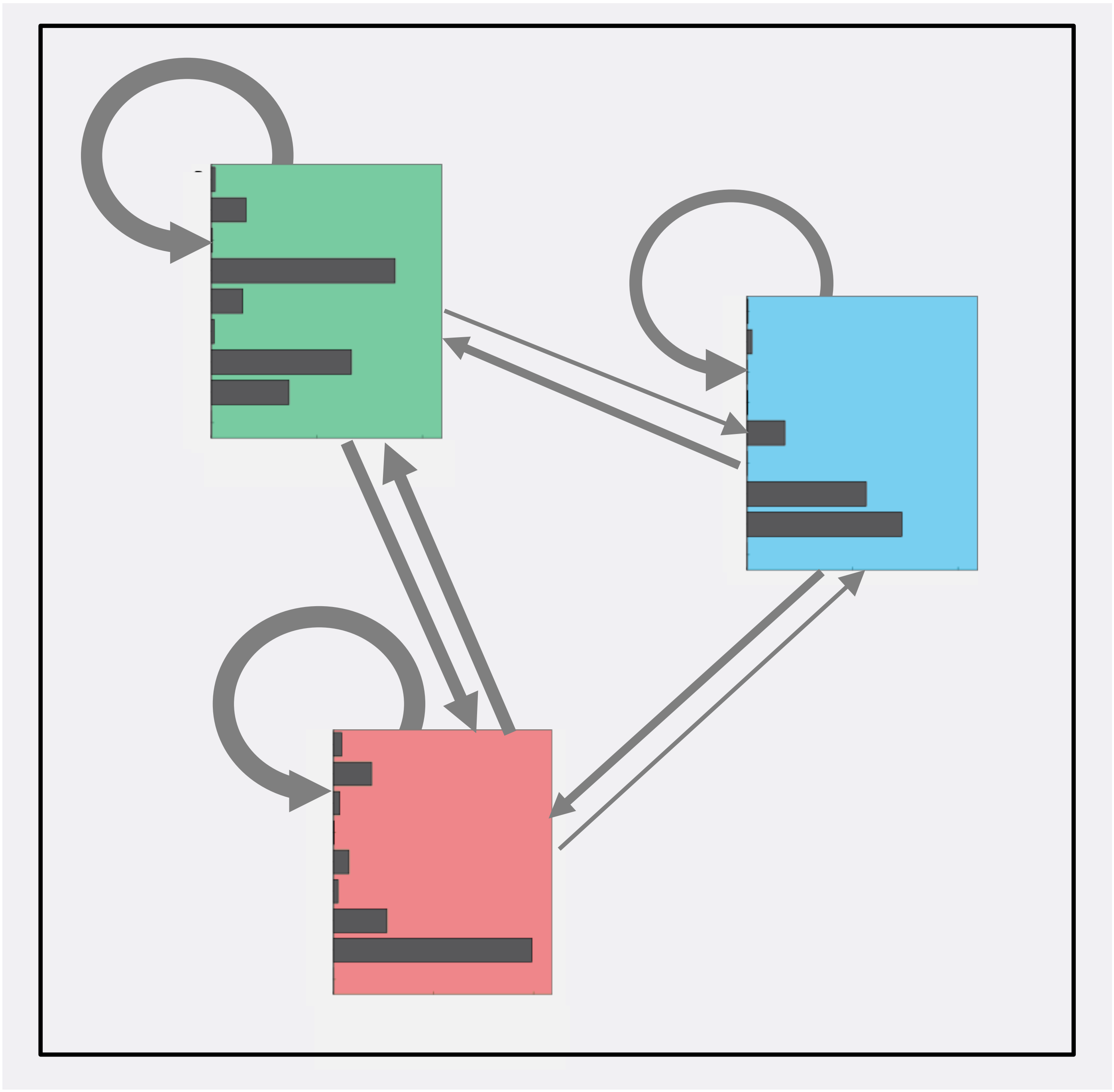



$$\sim y^i_t = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

# HMM for neural spike trains

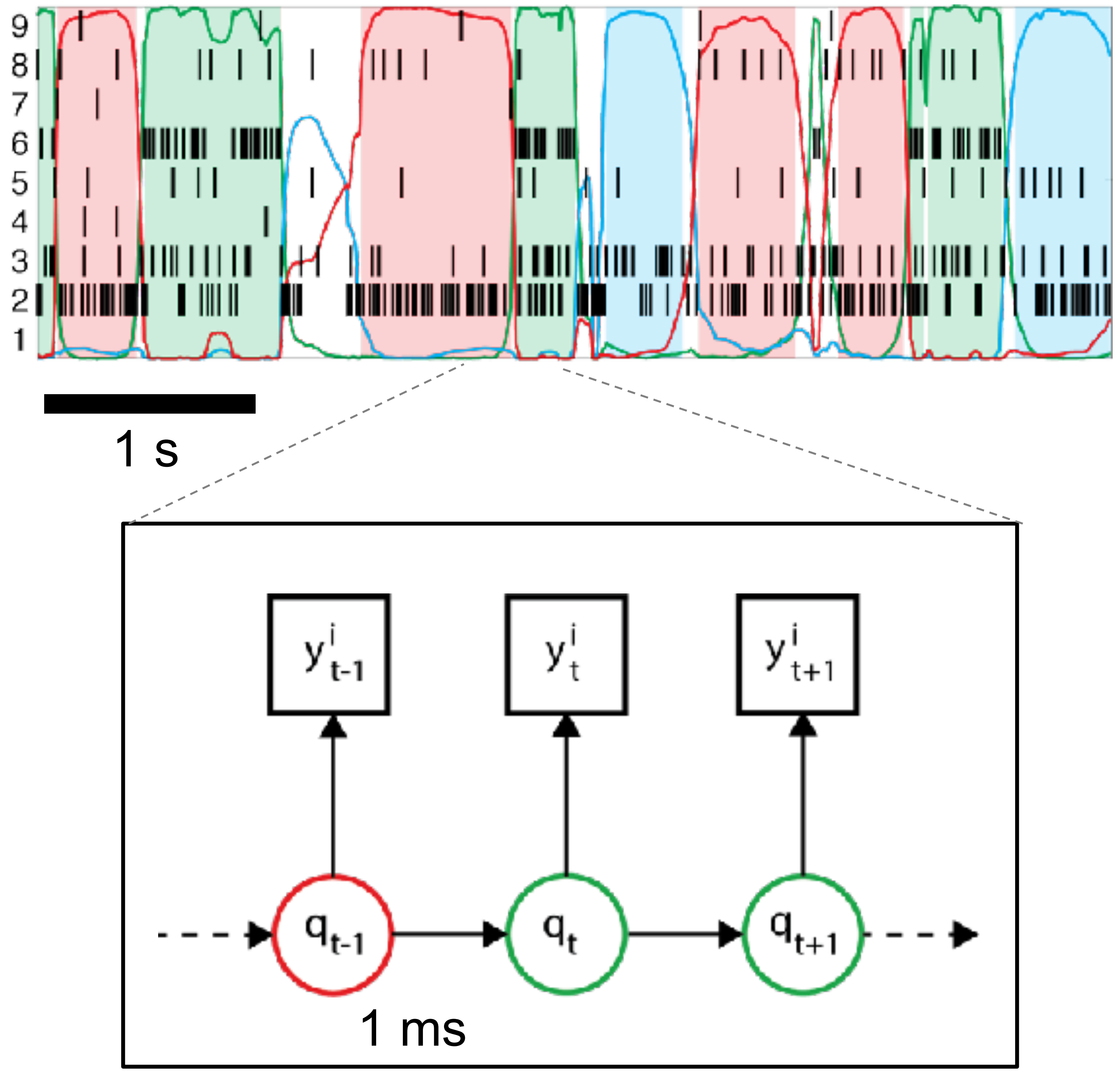
- Hidden Markov Model fit



# HMM for neural spike trains

- Hidden Markov Model fit




- Binning spikes $\Delta t$=1-5 ms
- Tradeoff: speed vs. accuracy:

Poisson $p(y_i(t) = k | S_t = m) = \frac{(\nu_i(m)dt)^k e^{-\nu_i(m)dt}}{k!},$

Bernoulli $p(y_i(t) = 1 | S_t = m) = 1 - e^{\nu_i(m)dt}.$

- “Poisson” HMM → “Bernoulli” HMM

# HMM for neural spike trains

## Model selection

How do we choose the number of hidden states?

1. Information criterion (AIC, BIC)
2. Cross-validation

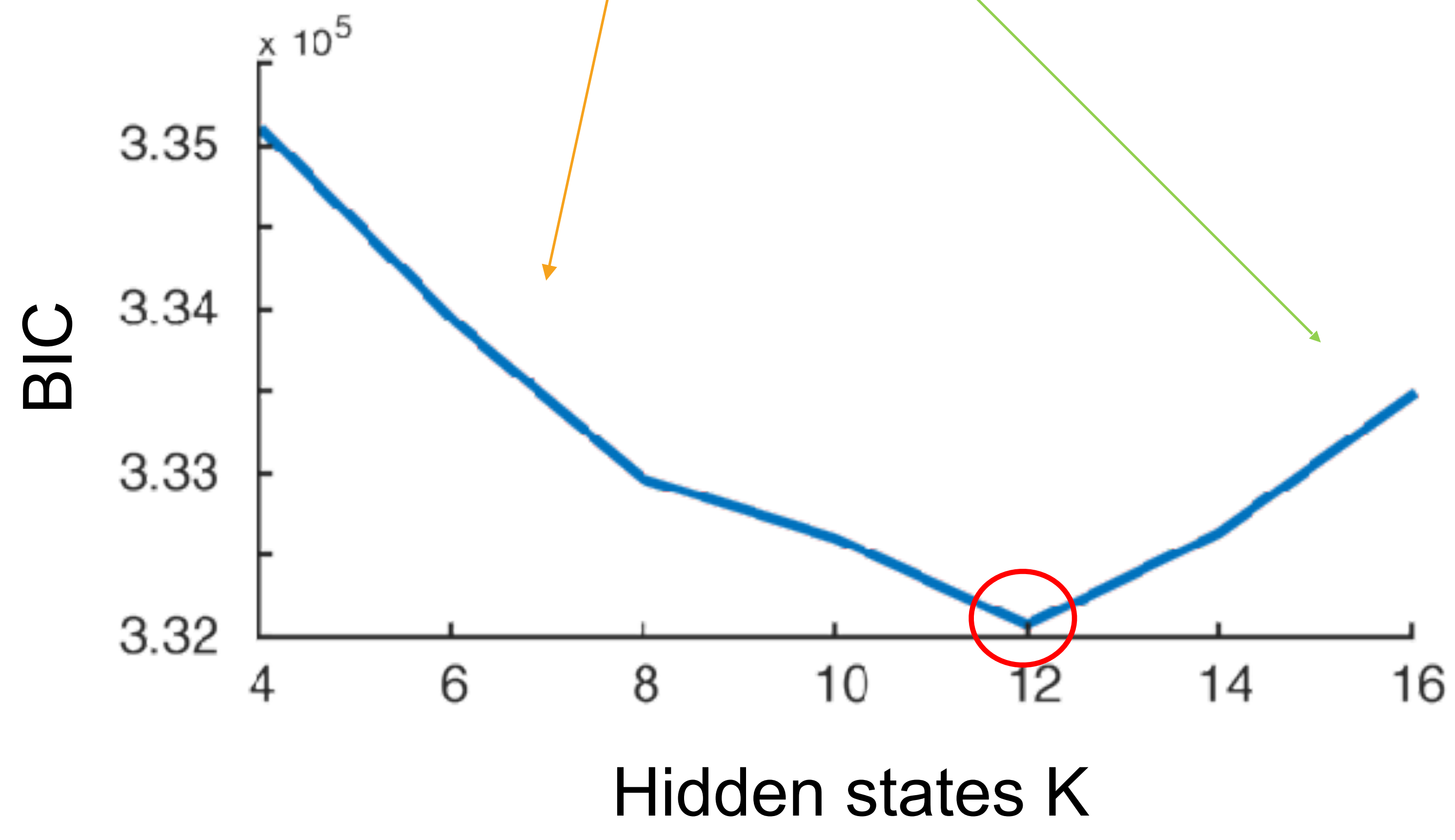
# HMM for neural spike trains

- Information criterion:
  # of HMM parameters: $p(K)=K(K-1)+KN$

K=states
N=neurons

$$\text{BIC} = -2\,\boxed{\text{LL(K)}} + \boxed{\text{p(K)}}\,\log(\text{T})$$

T=# of datapoints



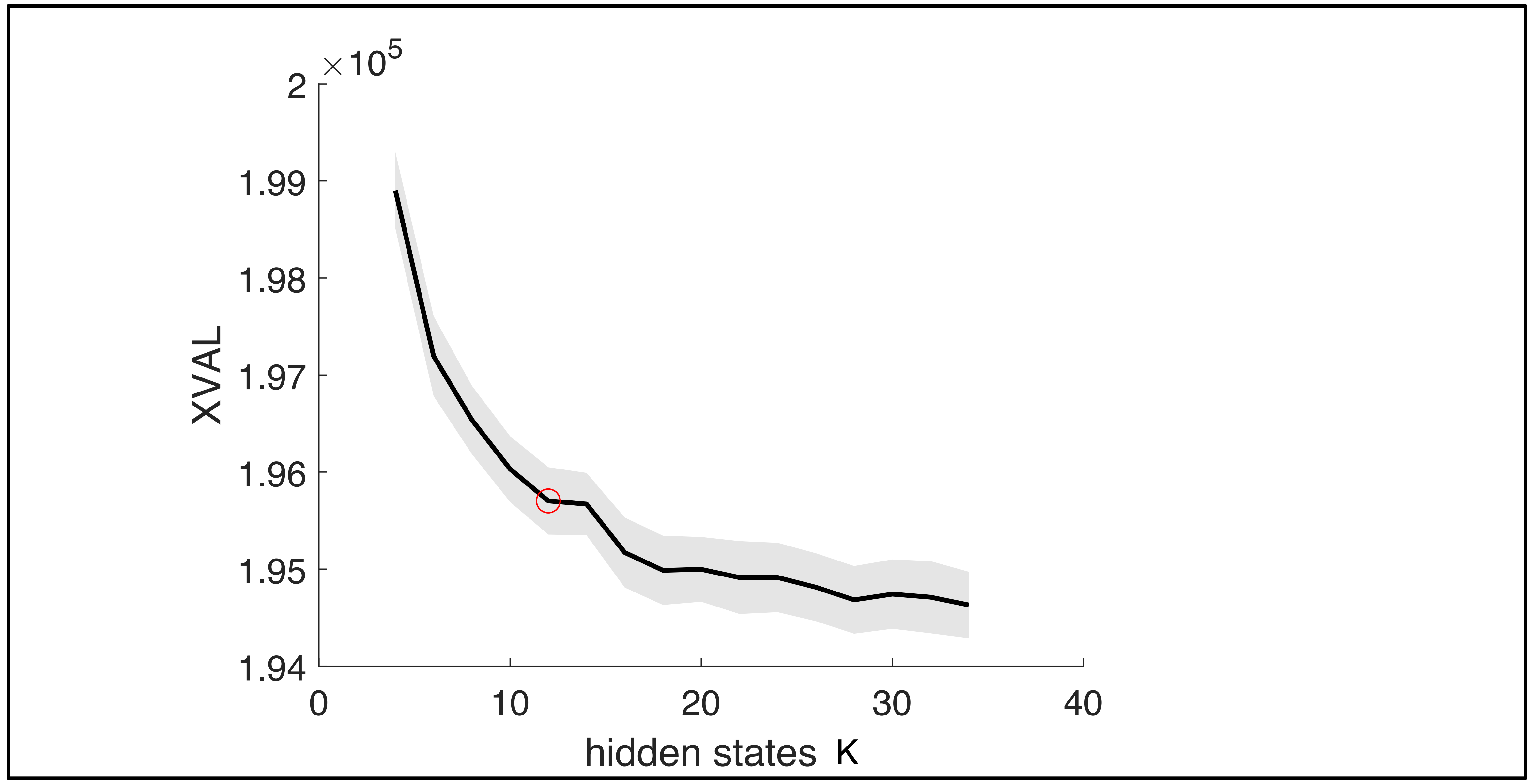

[Recanatesi et al., 2020]
[Sadacca et al., 2016]
[Engel et al., 2016]

# HMM for neural spike trains

- Cross-validation:

1. Randomly split trials in training set $S_1$ and test set $S_2$
2. Fit HMM with $K$ states on $S_1$
3. Estimate log-likelihood of test set $L(S_2|K)$

Select point with largest curvature (2nd deriv.)



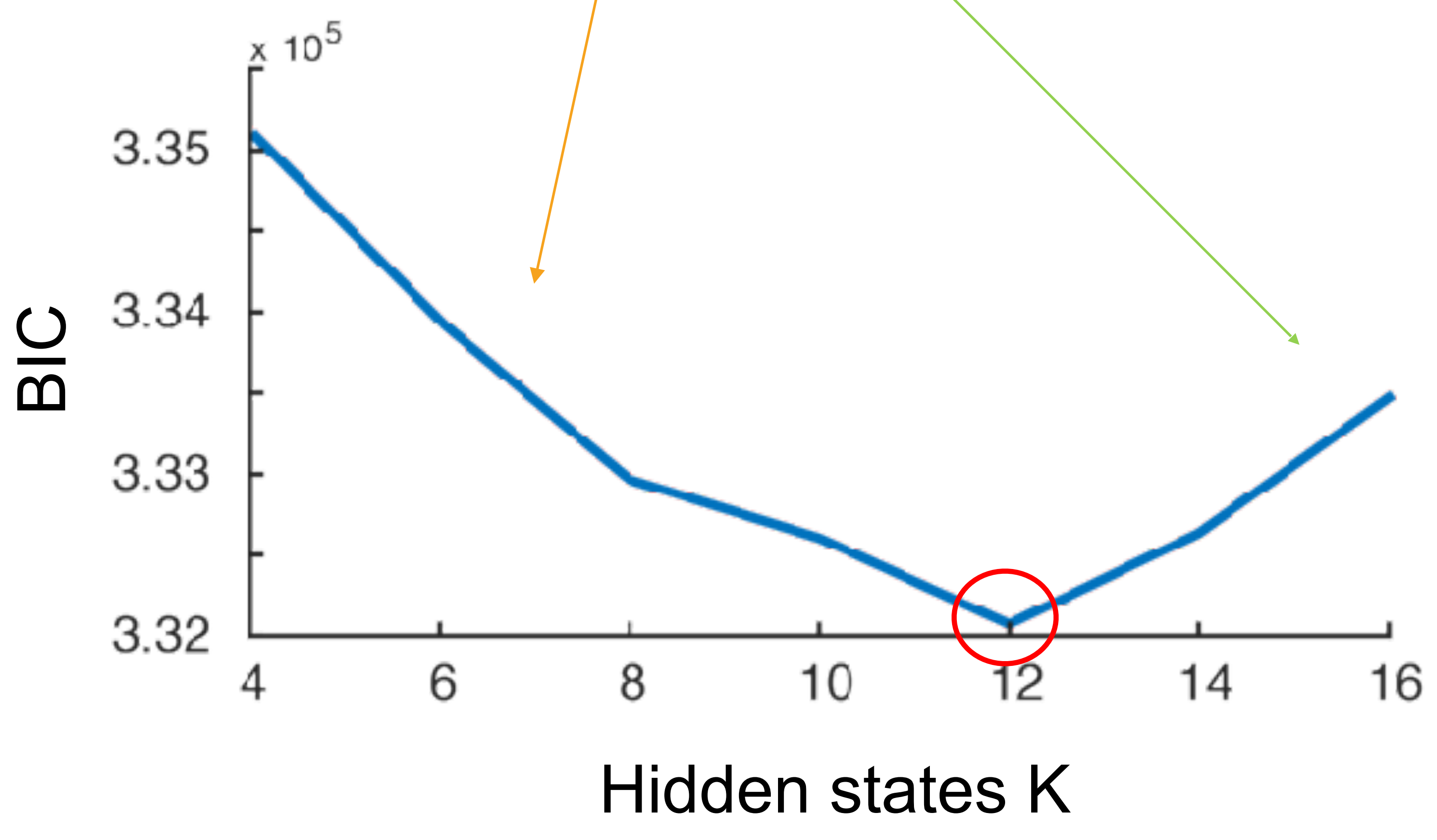

[Recanatesi et al., 2020]
[Sadacca et al., 2016]
[Engel et al., 2016]

# HMM for neural spike trains

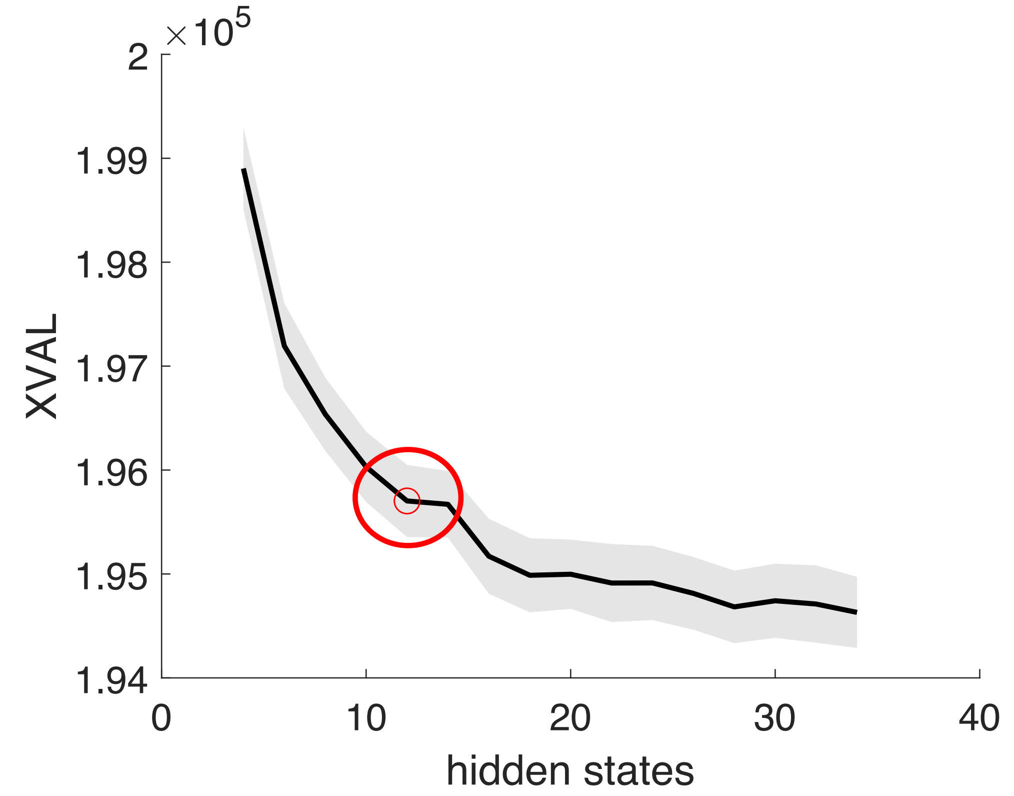
How do we choose the number of hidden states?

1. Information criterion (AIC, BIC)
2. Cross-validation

$$\mathrm{BIC} = -2\,\boxed{\mathrm{LL}(K)} + \boxed{p(K)}\,\log(T)$$




XVAL

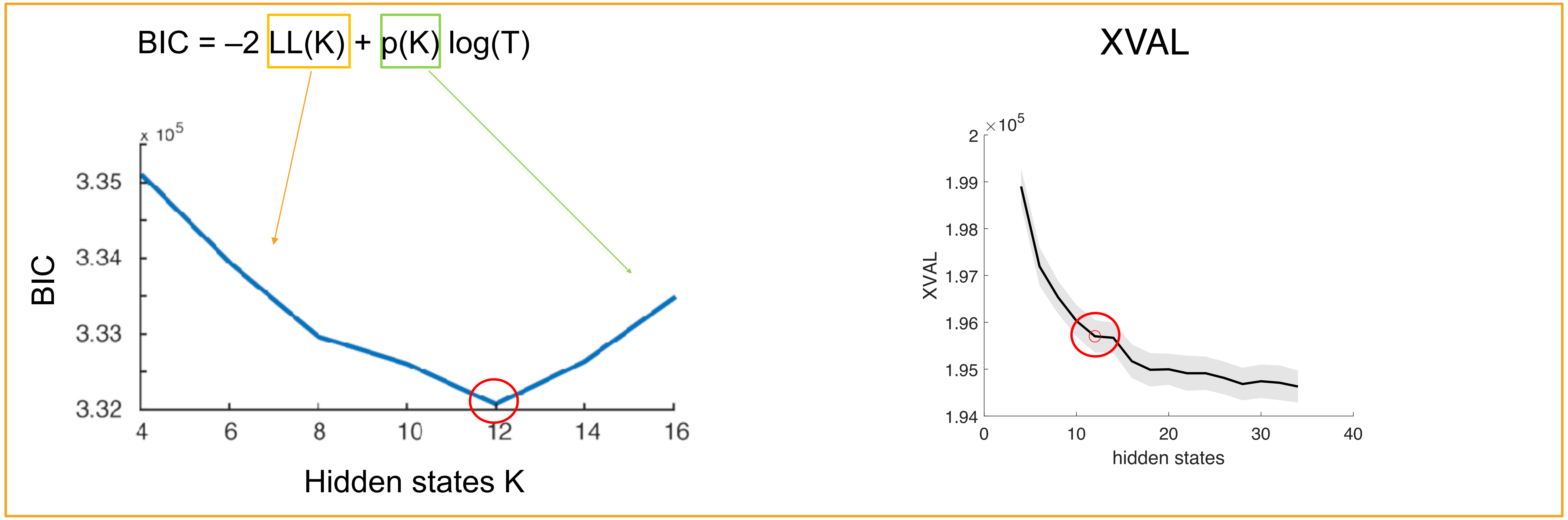



[Recanatesi et al., 2020]
[Sadacca et al., 2016]
[Engel et al., 2016]

# HMM for neural spike trains

## Local minima

- Transition and emission matrices optimized via Expectation-Maximization
- Not convex —> many local minima
- Run several times with random initial conditions, pick best run.




[Recanatesi et al., 2020]
[Sadacca et al., 2016]
[Engel et al., 2016]

# HMM for neural spike trains

In practice:

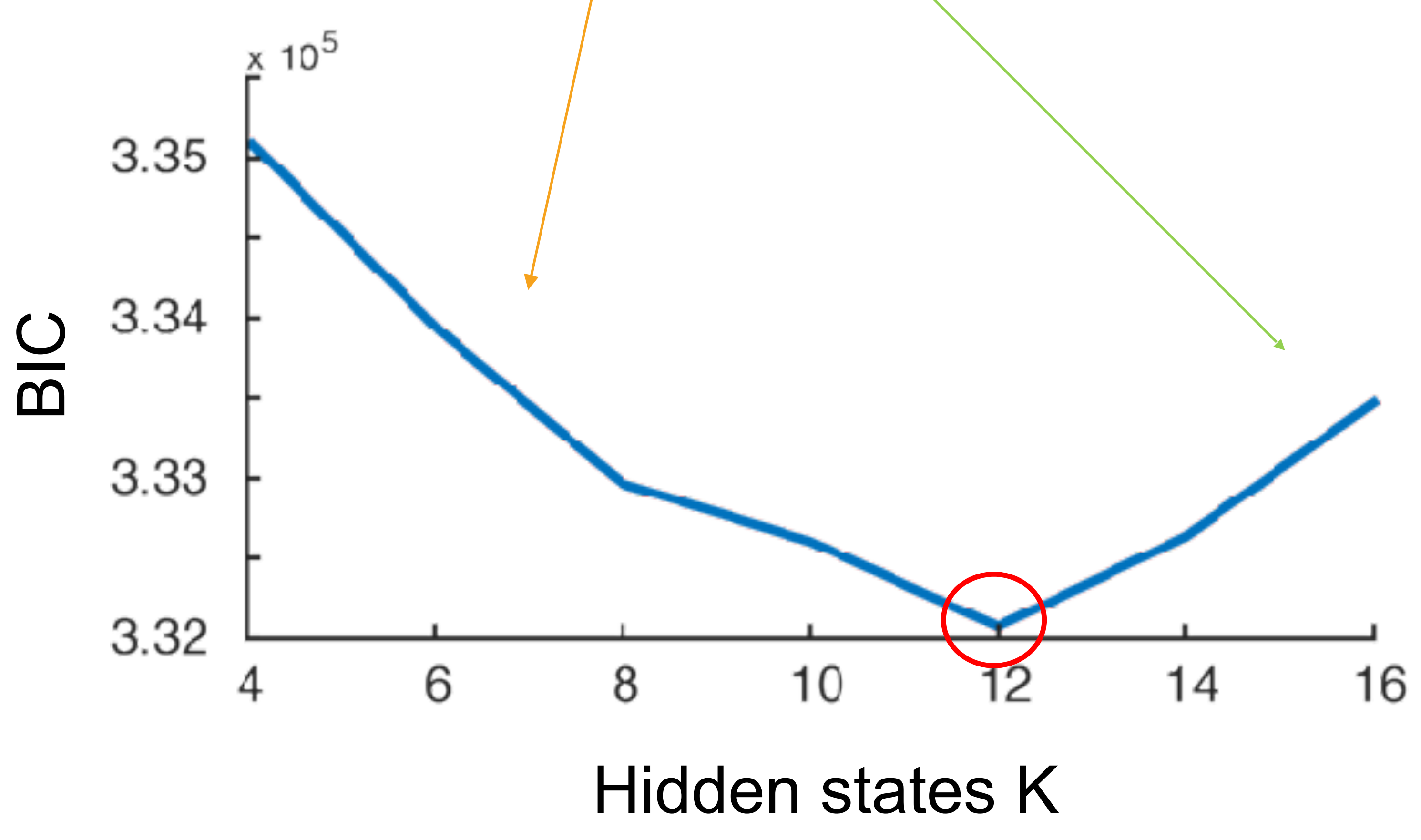
• Most “extra states” have very low posterior probabilities *p*($S_t$=*m*|*data*) in all trials

→ post-hoc state admission criterion:

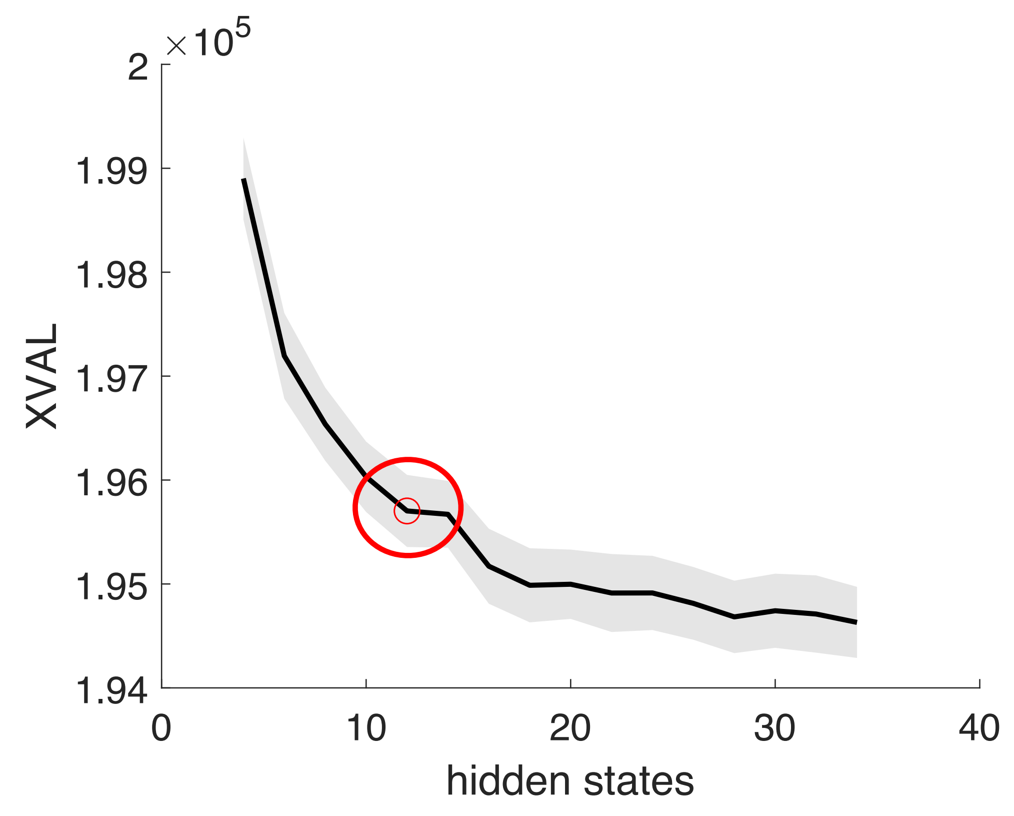
A state is detected at time *t* if it ***p*($S_t$=*m*|*data*)>80% for at least one 50 ms bins**

→ ***only keep “admissible states” for the rest of the analyses***

BIC = –2 $\boxed{\text{LL(K)}}$ + $\boxed{\text{p(K)}}$ log(T)

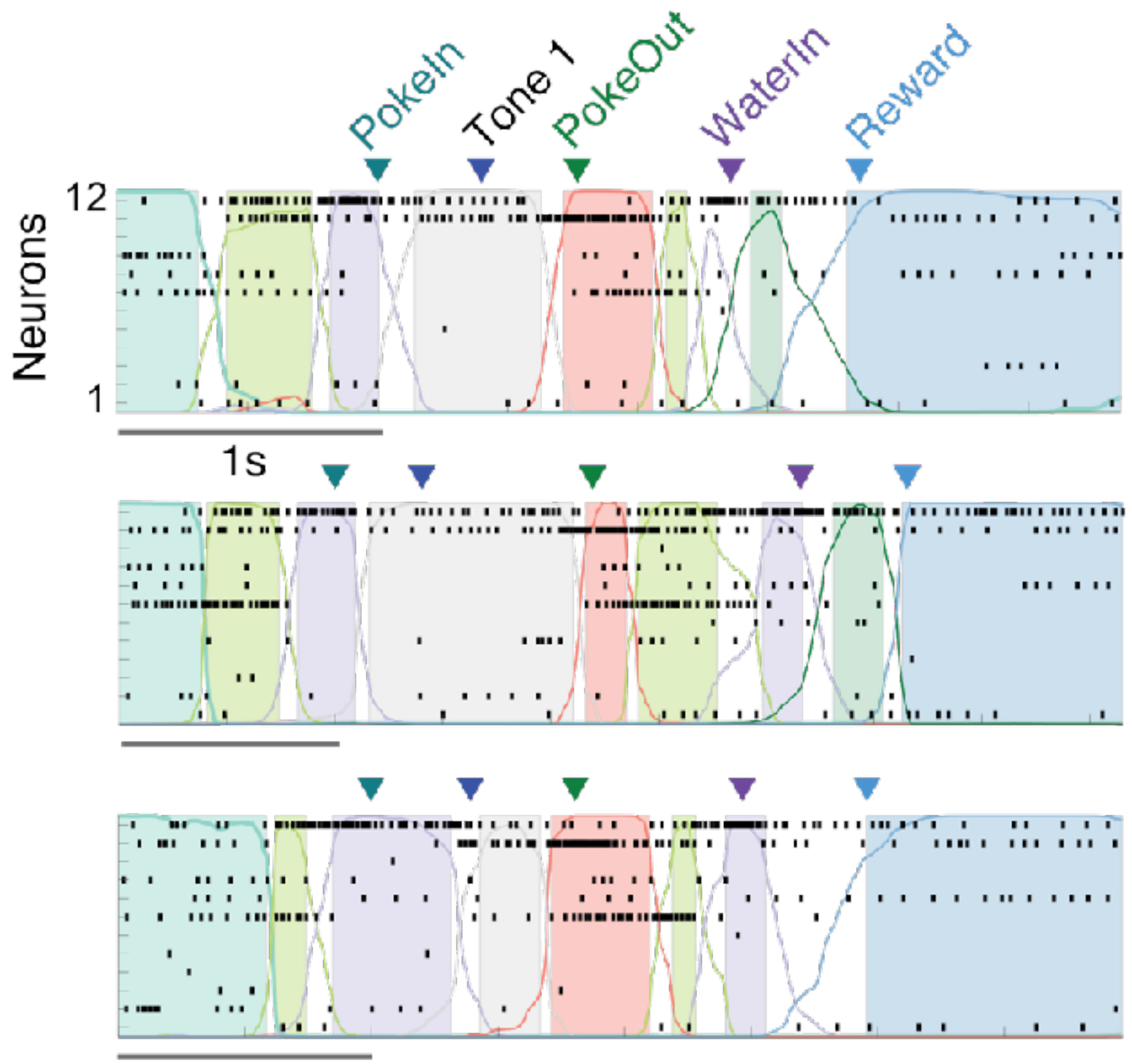


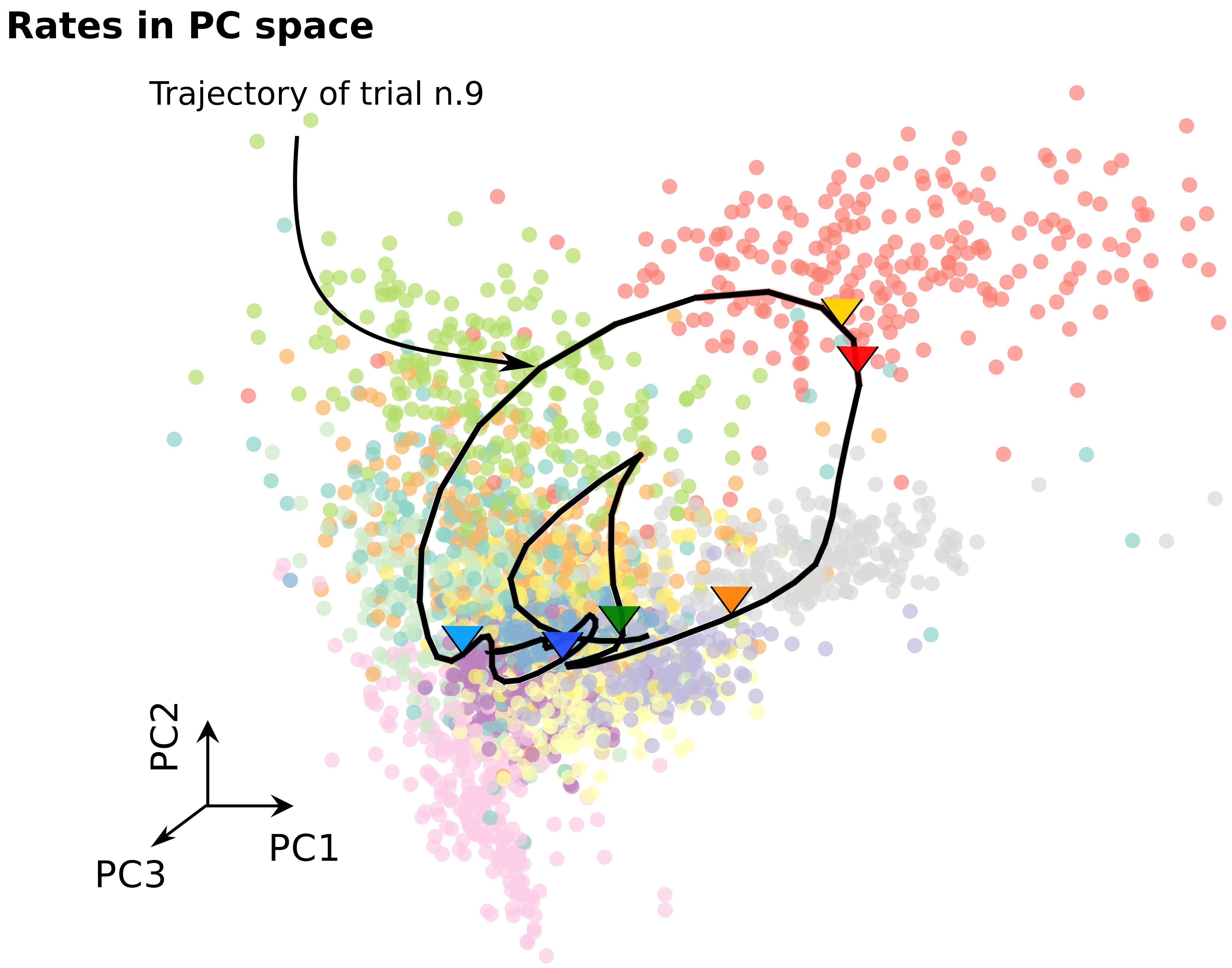

XVAL




[Recanatesi et al., 2020]
[Sadacca et al., 2016]
[Engel et al., 2016]

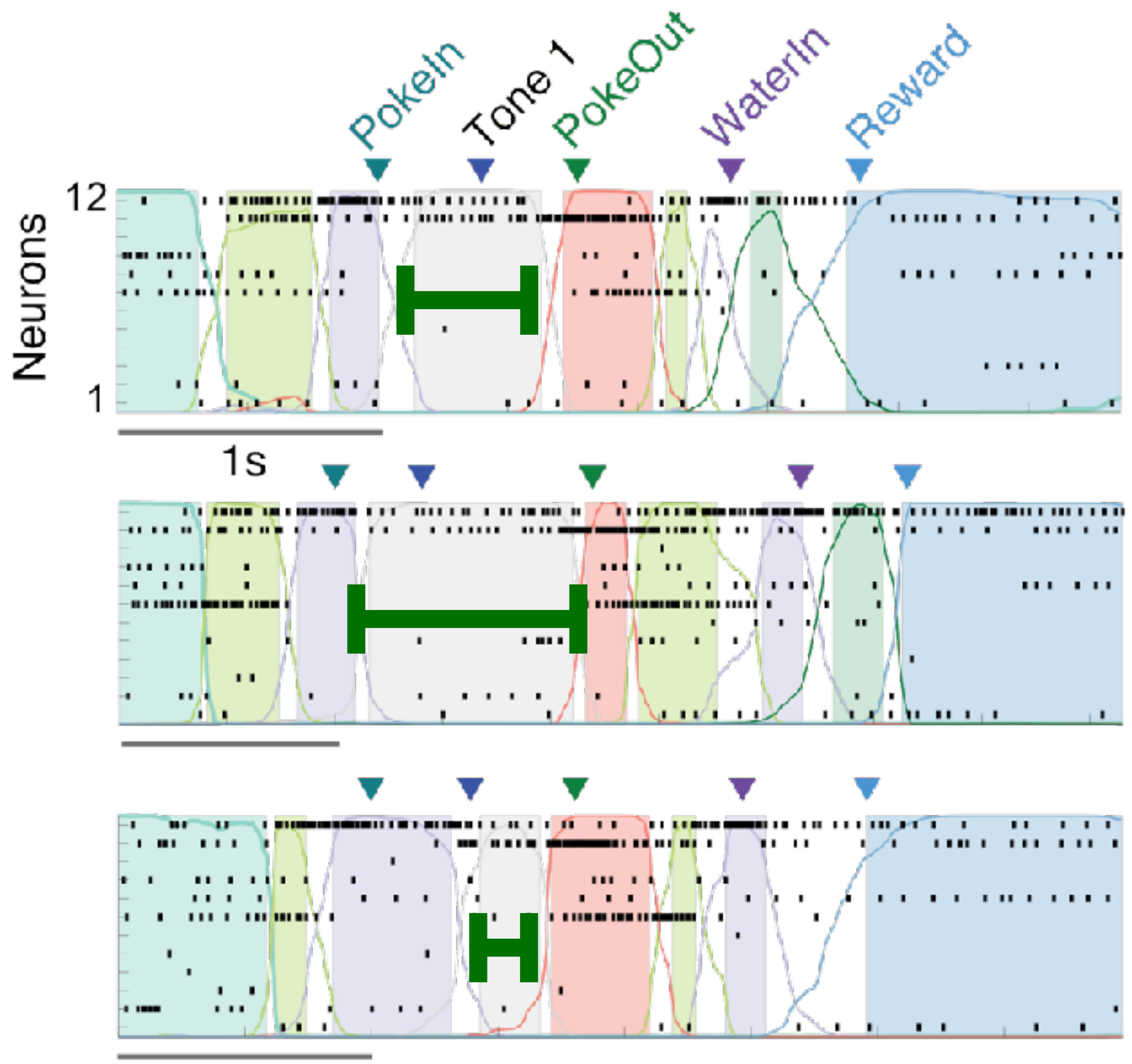
# Secondary motor cortex - sequences of attractors



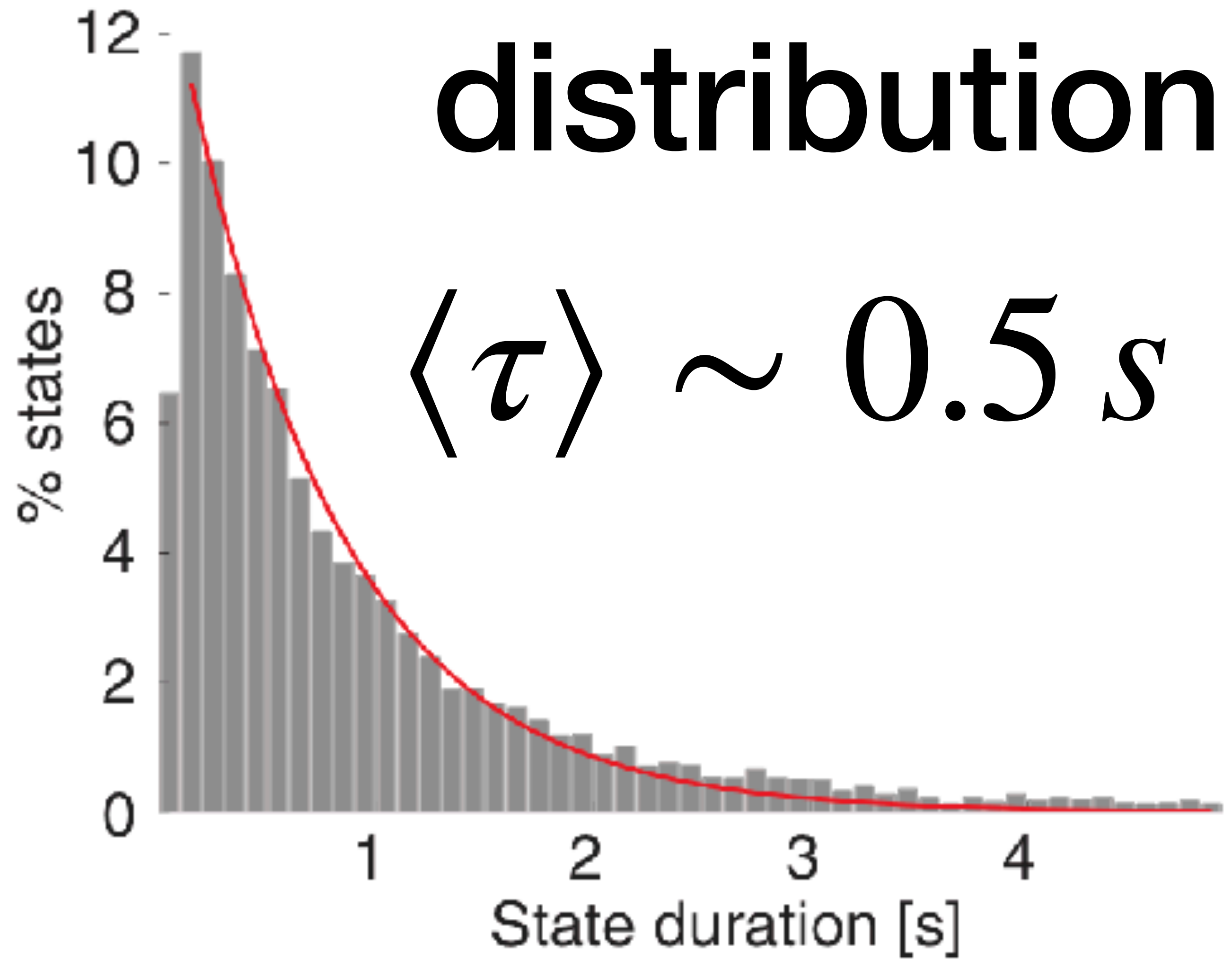

[Recanatesi et al., 2020]

# Secondary motor cortex - sequences of attractors




Exponential dwell time distribution

$$\langle \tau \rangle \sim 0.5\, s$$




[Recanatesi et al., 2020]

# Goals

**Part I - Theory**

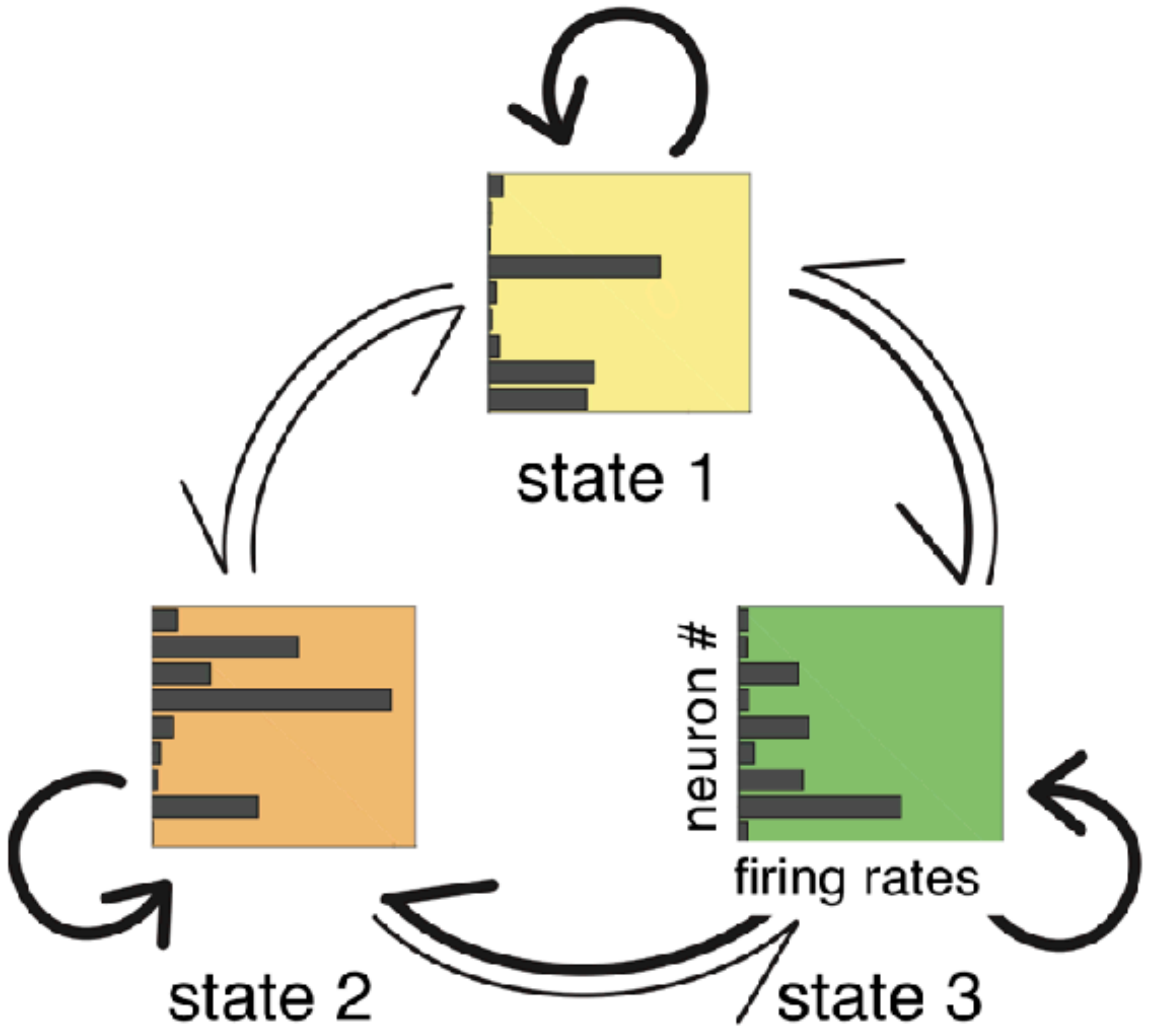
- Metastable activity in cortical circuits
- Hidden Markov Models fit to metastable activity
- Modeling metastable activity with attractor networks

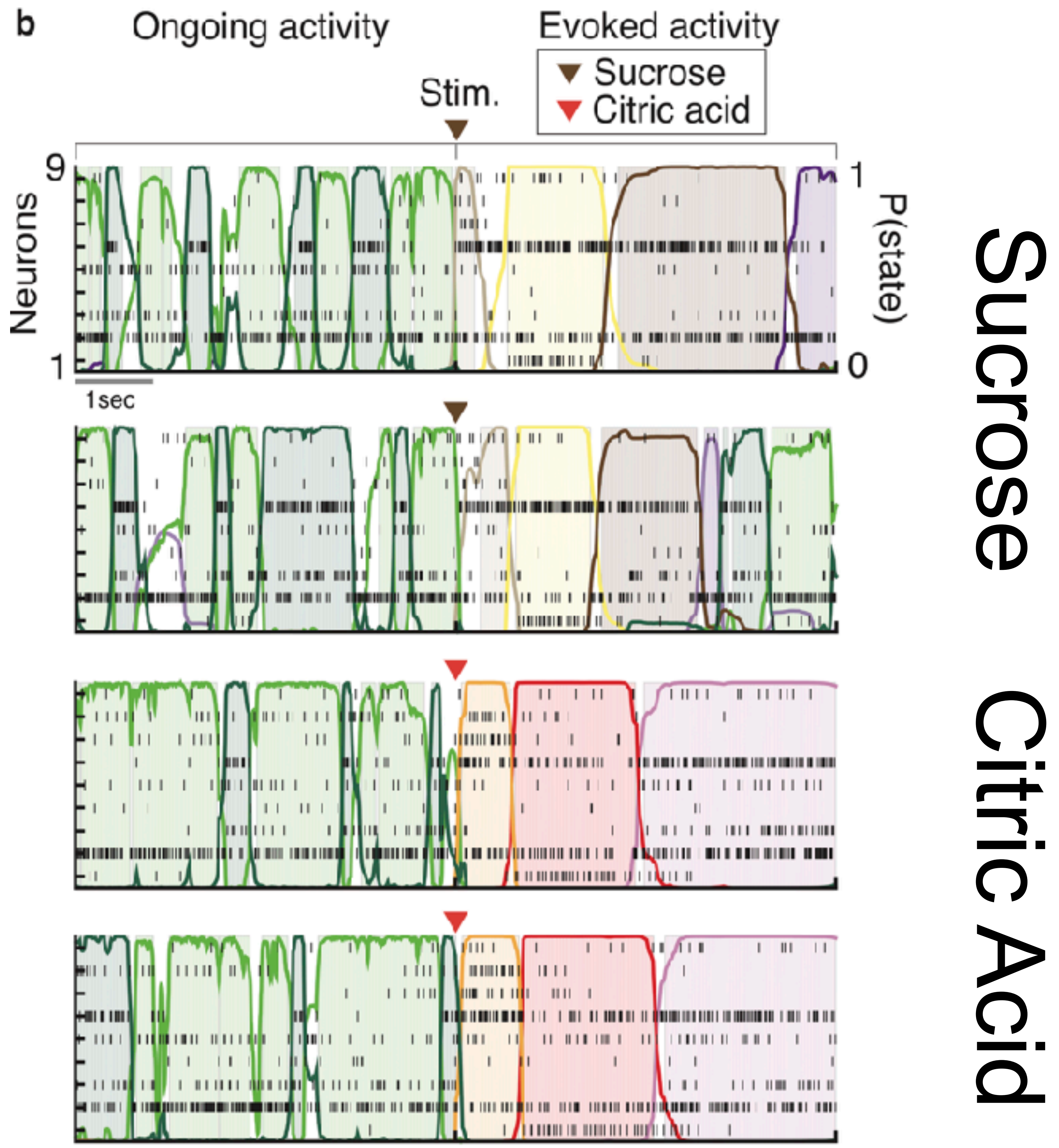
**Part II - Exercise**

- HMM: model selection with gaussian emissions
- Fit Poisson-HMM to spike trains
- Challenge: A model of arousal in attractor networks

# Metastable activity: sensory processing

**Poisson-Hidden Markov Model**

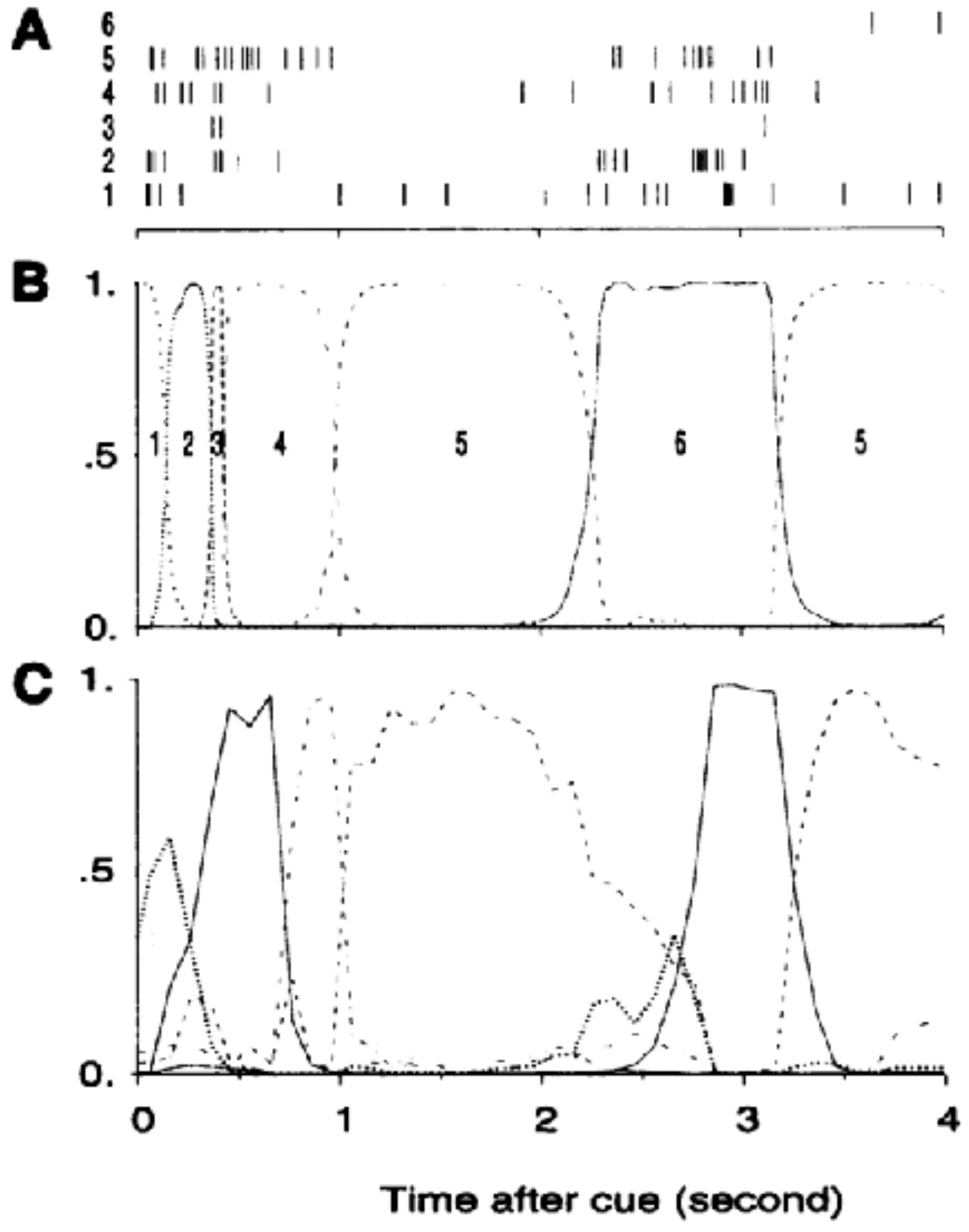



[Mazzucato et al., 2015]

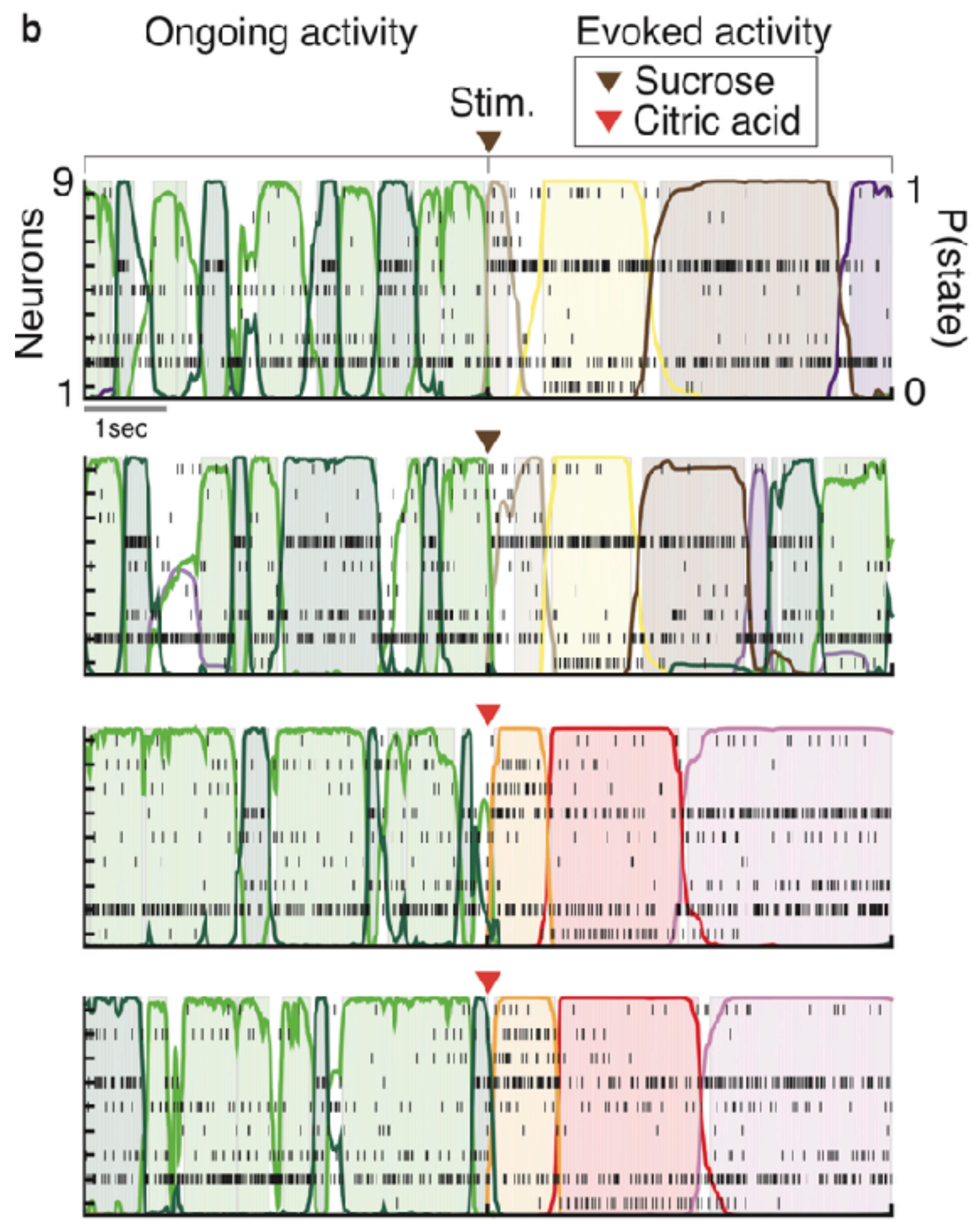
# Metastable activity: sensory processing

## Visual (in PFC)



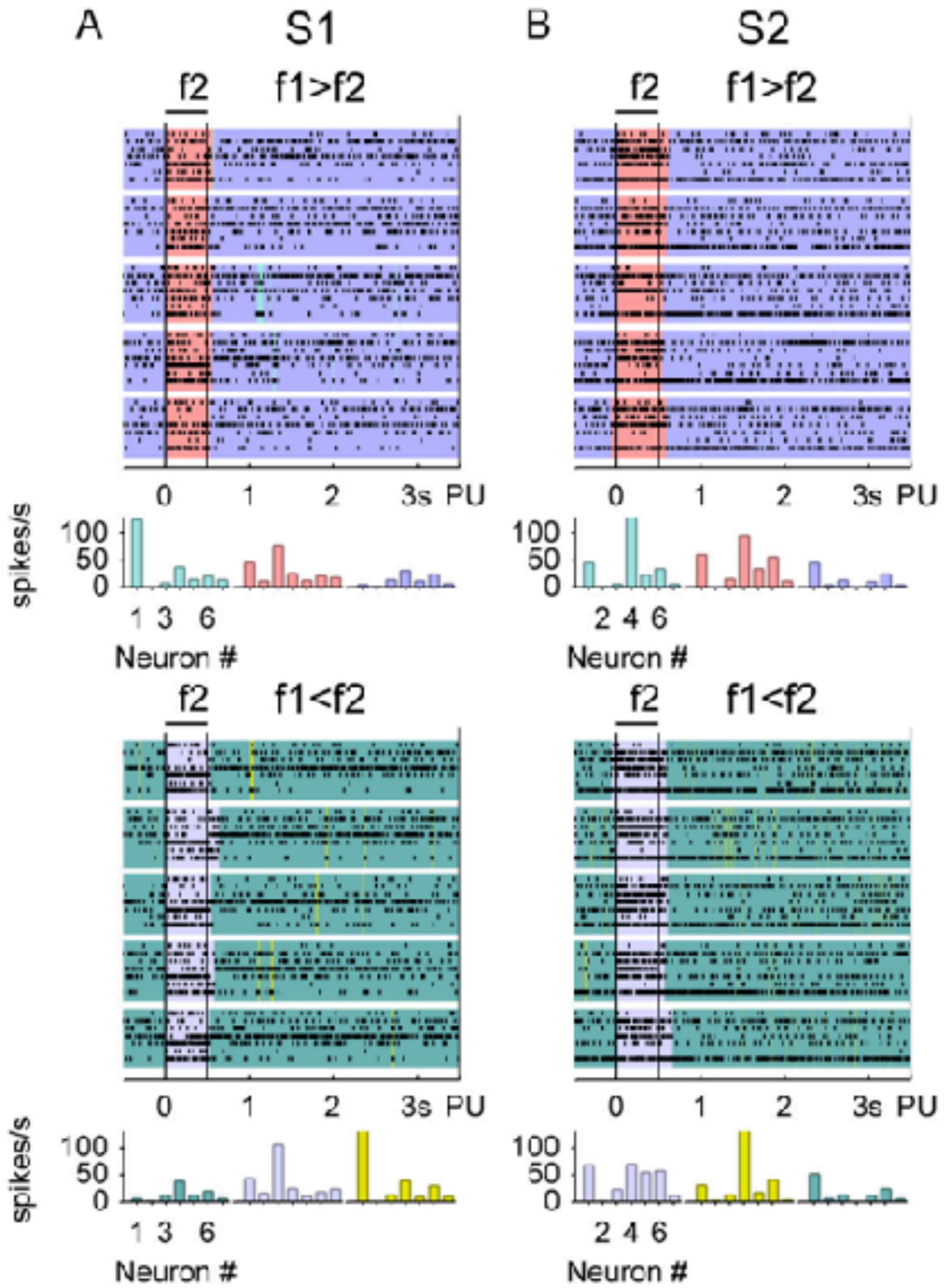

[Abeles et al, 1995]

## Gustatory




[Jones et al, 2007]
[Mazzucato et al, 2015]

## Somatosensory

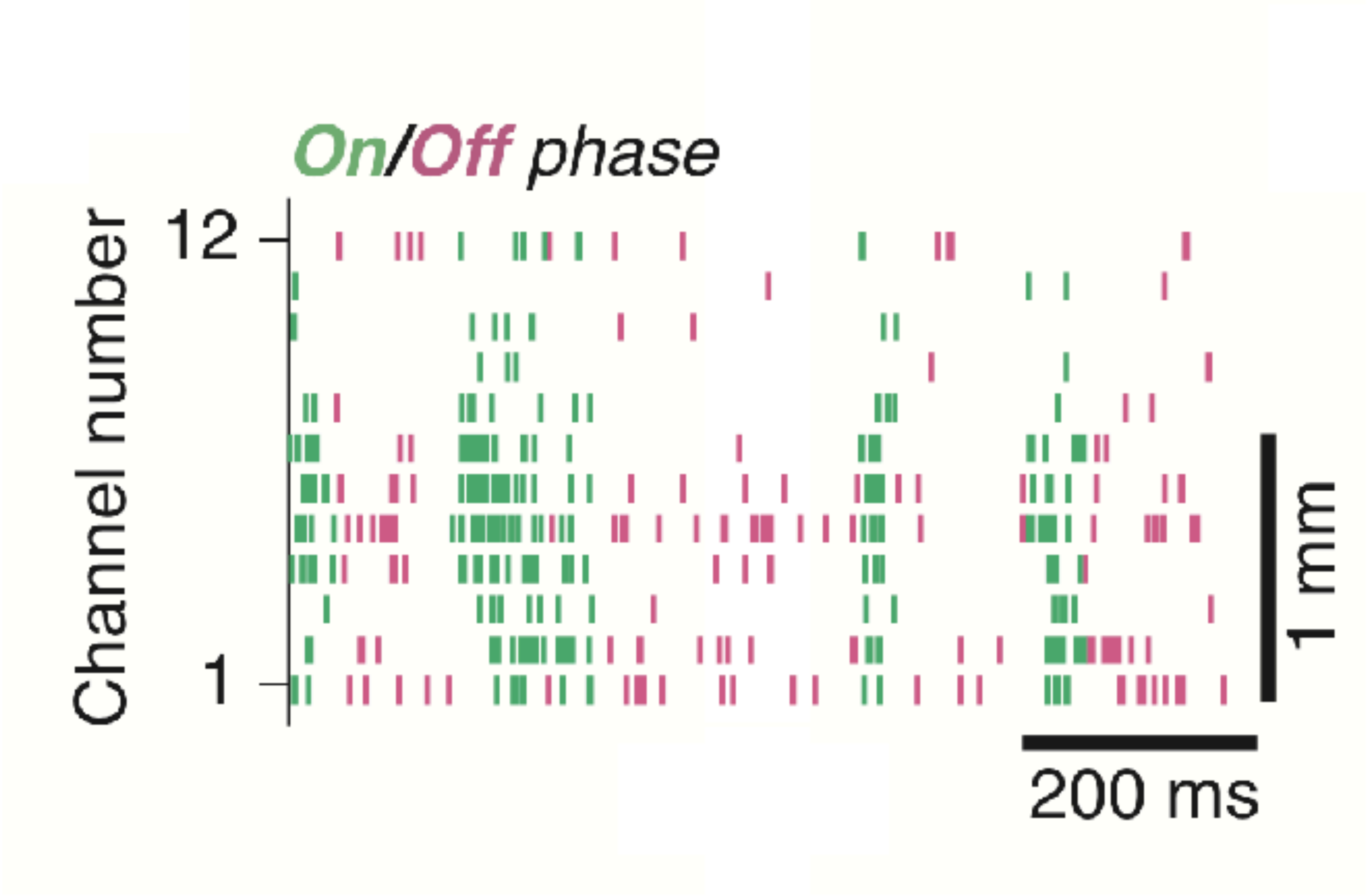


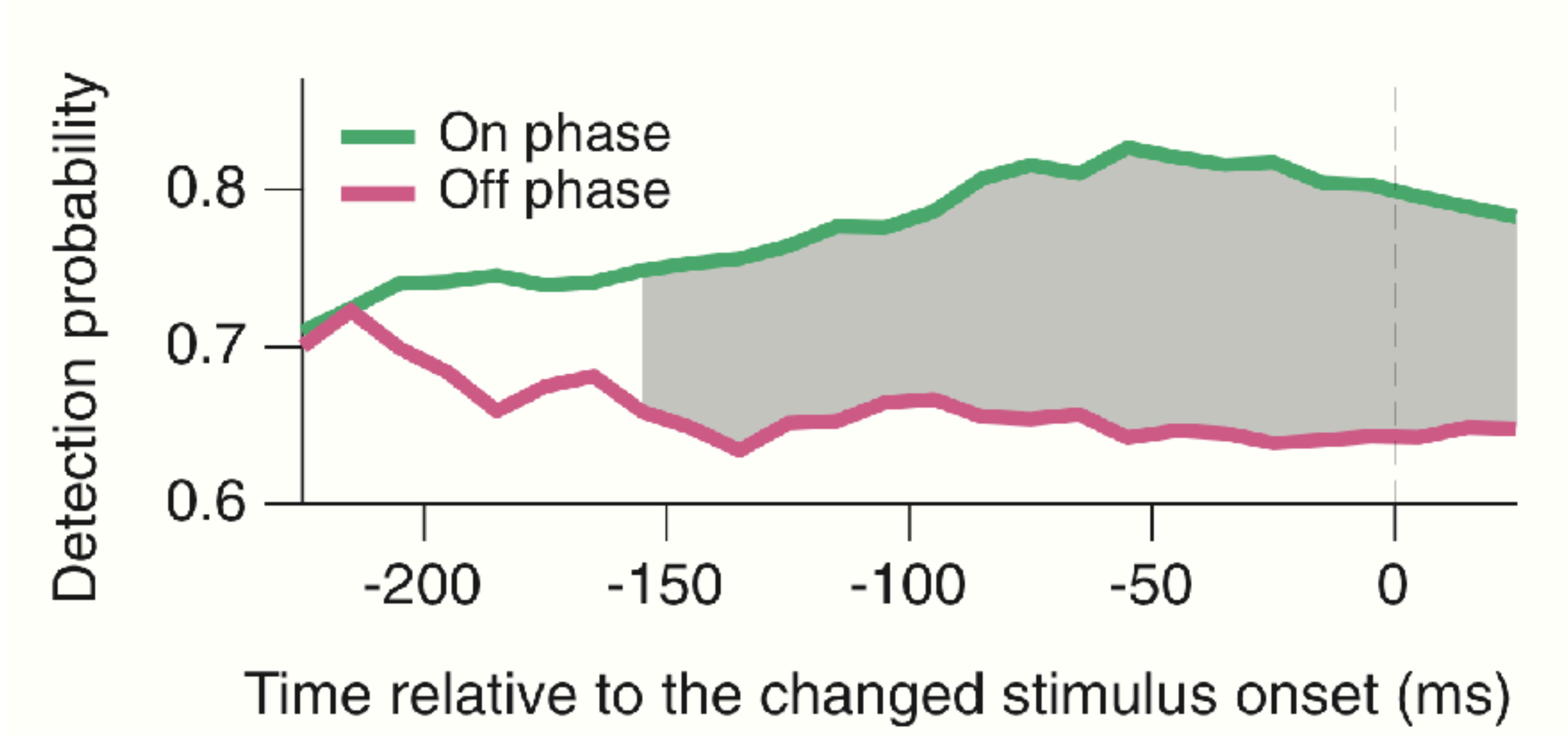

[Ponce-Alvarez et al, 2012]

# Metastable activity: cognition

Selective attention task in V4 (monkey)

*Better task performance during On*



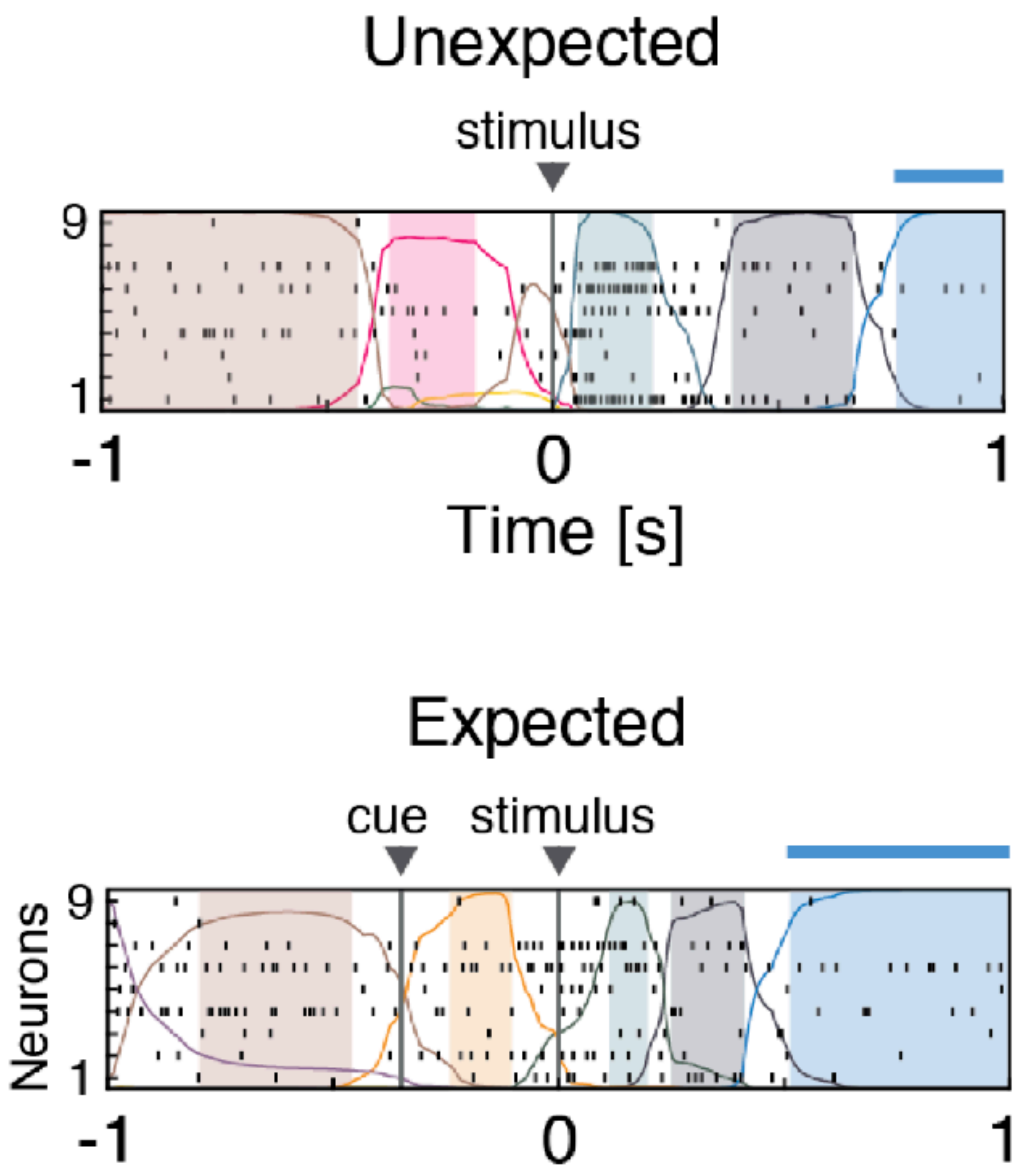

[Engel et al, 2016]

# Metastable activity: cognition

Expectation of taste

in gustatory cortex

**Expected** stimuli
are recognized **faster**
than unexpected ones




[Mazzucato et al, 2019]

# Goals

**Part I - Theory**

- **Metastable activity in cortical circuits**
- **Hidden Markov Models fit to metastable activity**
- **Modeling metastable activity with attractor networks**

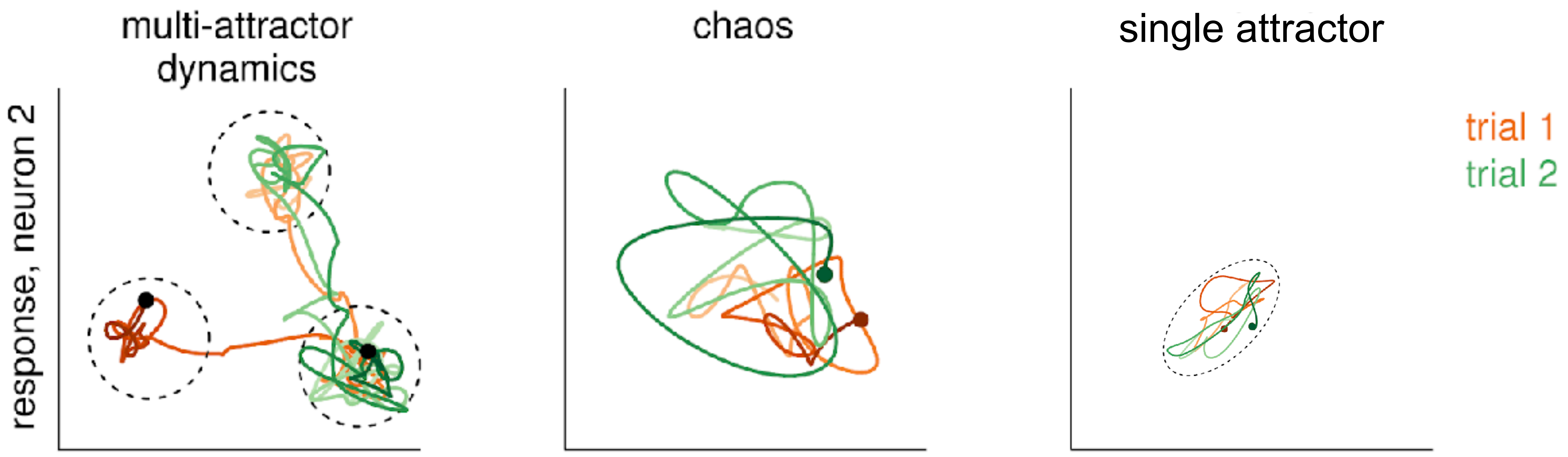
**Part II - Exercise**

- **HMM: model selection with gaussian emissions**
- **Fit Poisson-HMM to spike trains**
- **Challenge: A model of arousal in attractor networks**

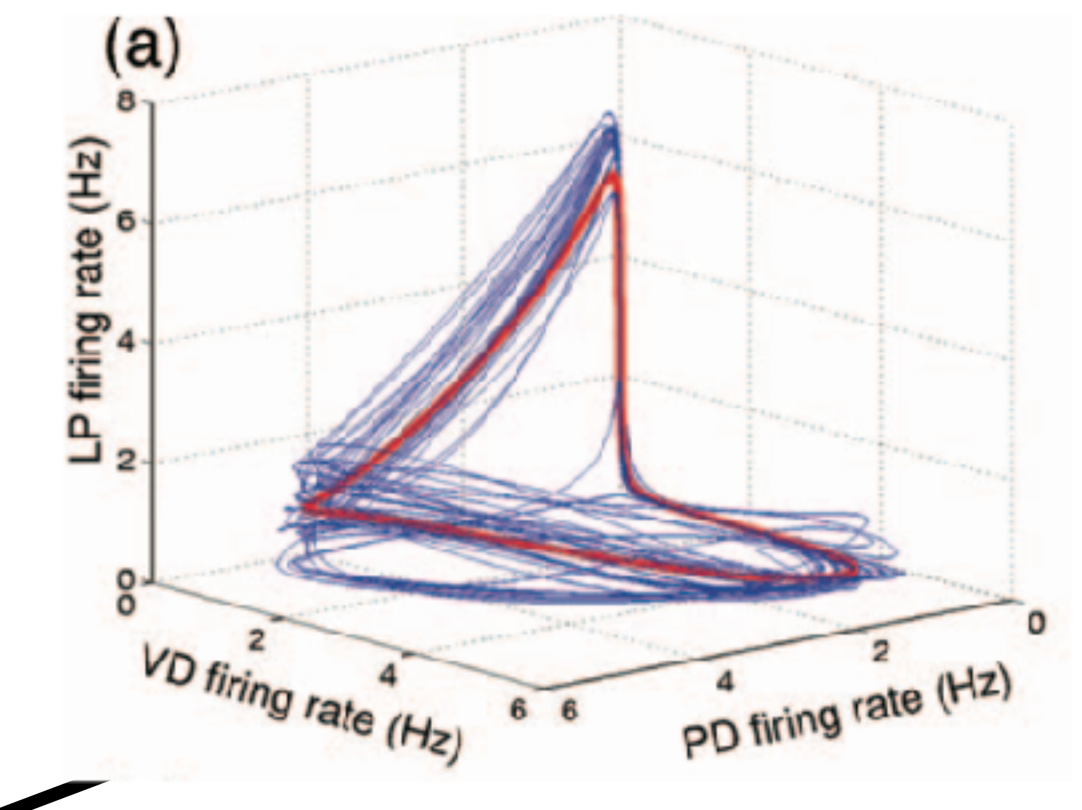
# Modeling metastable neural activity with attractor networks

# Temporal dynamics in neural circuits

Dynamical repertoires:




oscillations




Jumping

Varying smoothly

[Rabinovich et al, 2006]
[Hennequin et al, 2018]

# Clustered networks of leaky-integrate-and-fire neurons

# LIF neuron

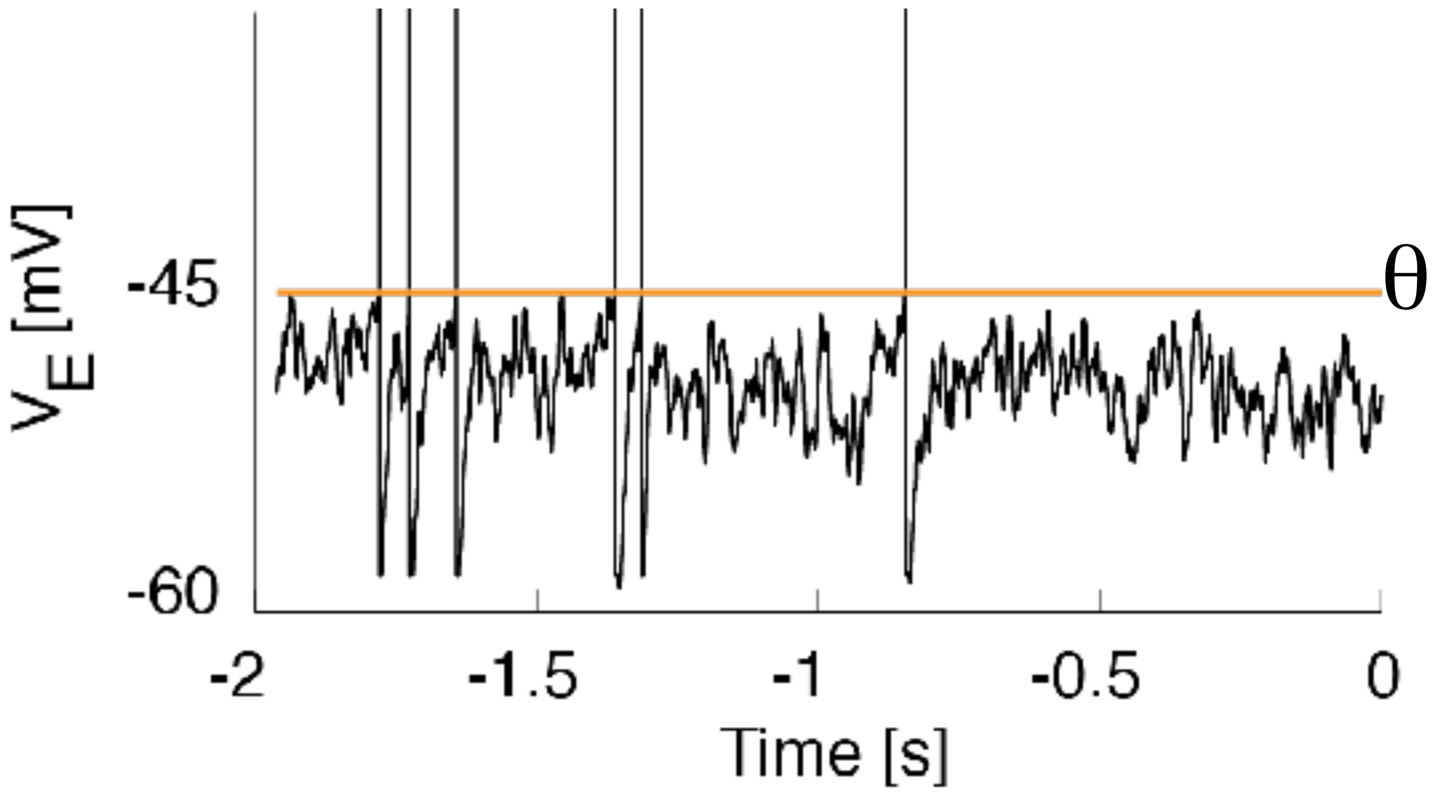
PSC: $$\tau_s \frac{dI_i}{dt} = -I_i + I_{ext}(t) + \sum_j^{pre} J_{ij} \sum_n^{spikes} \delta(t - t_j^{(n)})$$

Potential: $$\tau_m \frac{dV_i}{dt} = V_L - V_i + I_i(t)$$

b.c.: $V(t^*) = \theta$

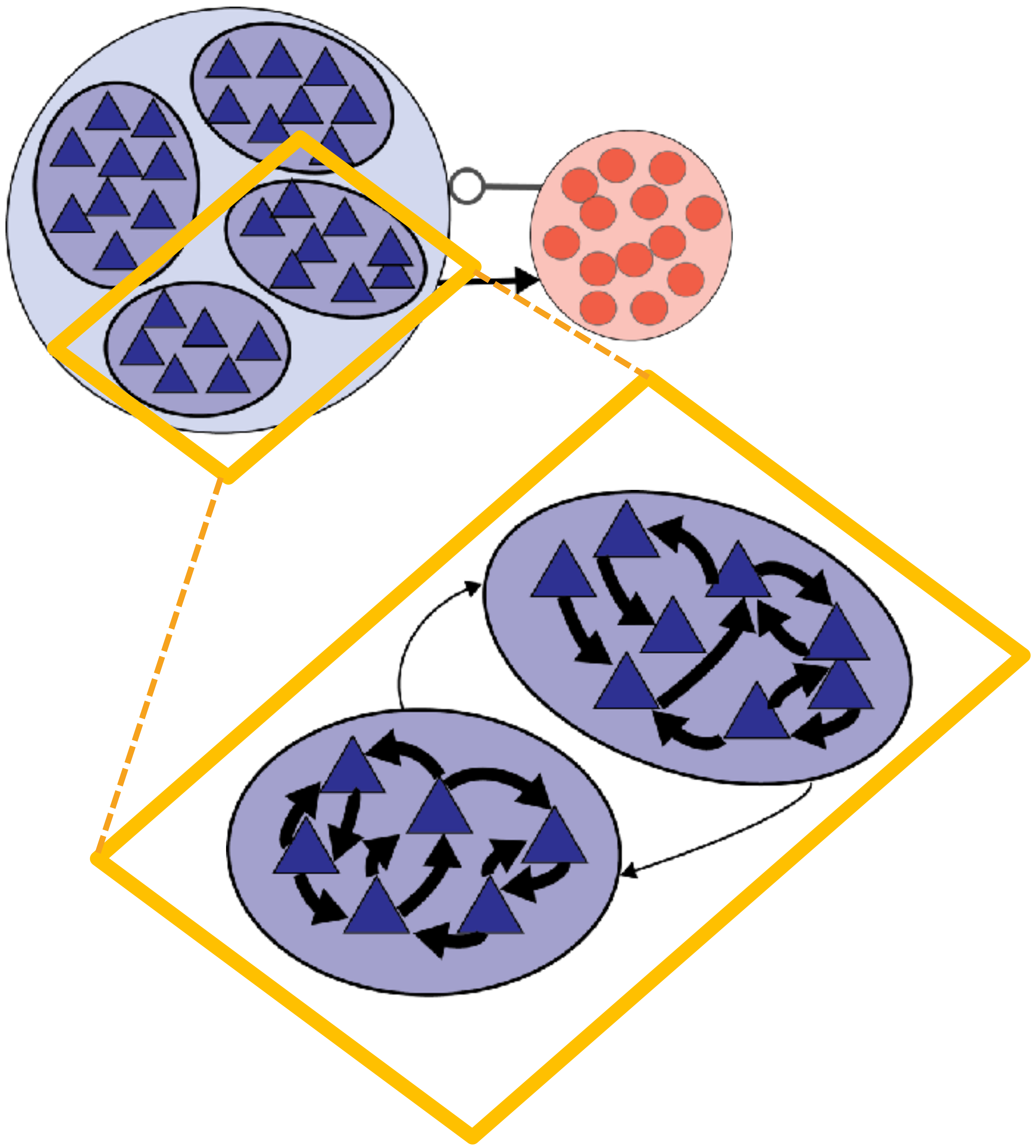
→ spike

→ reset

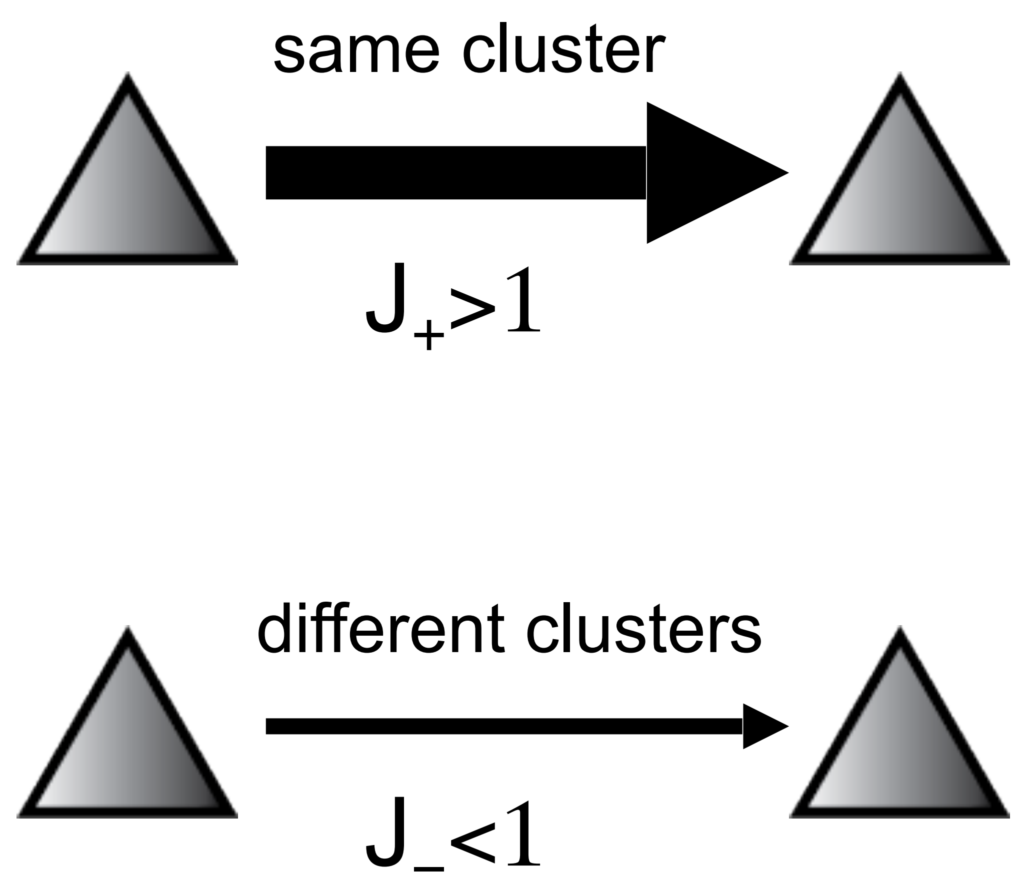



[van Vreeswijk & Sompolinsky, 1996] [Haider et al., 2006]
[Amit & Brunel, 1997] [Renart et al., 2007]

# Clustered network



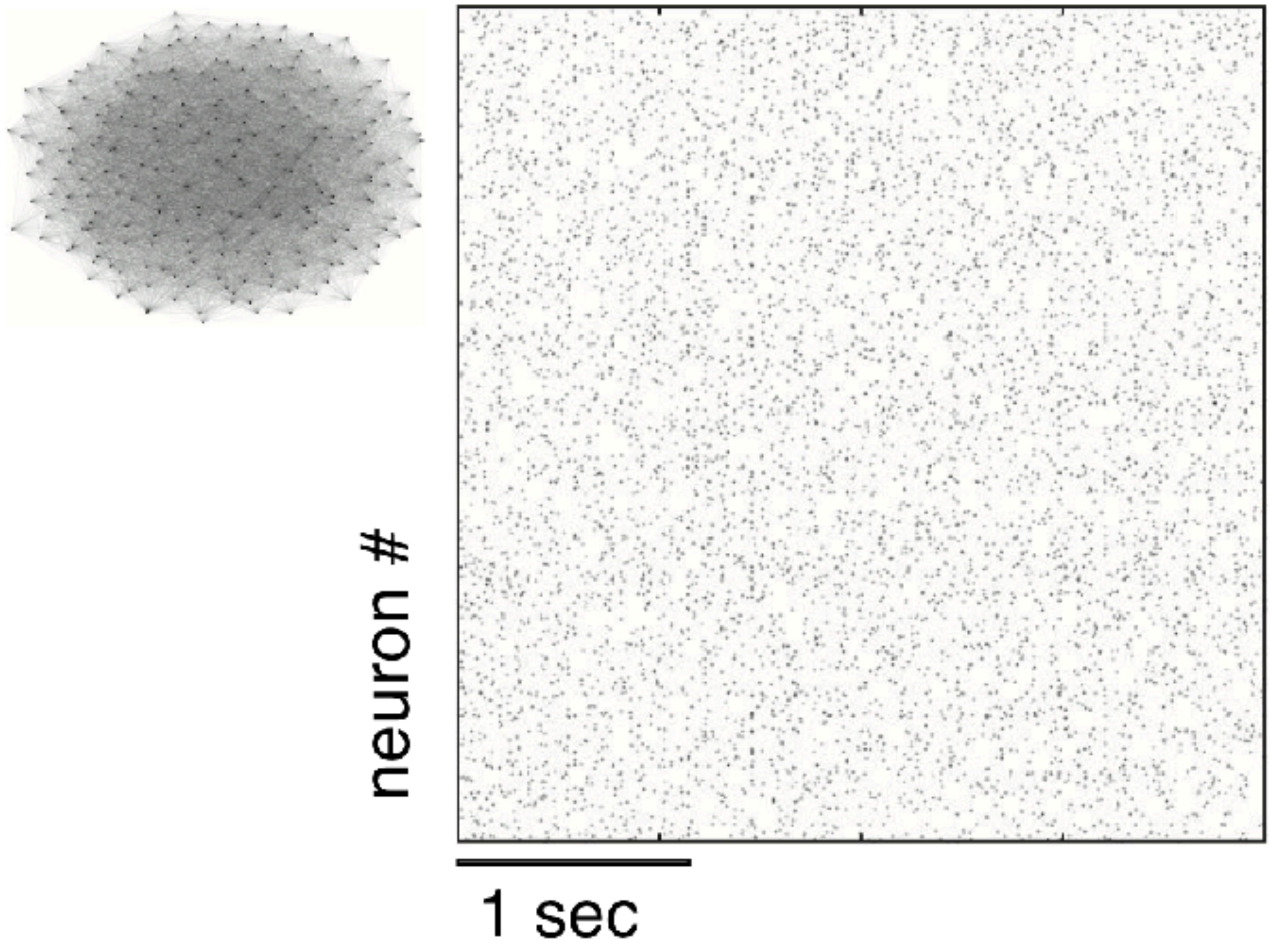
Recurrent E → E synaptic weights:

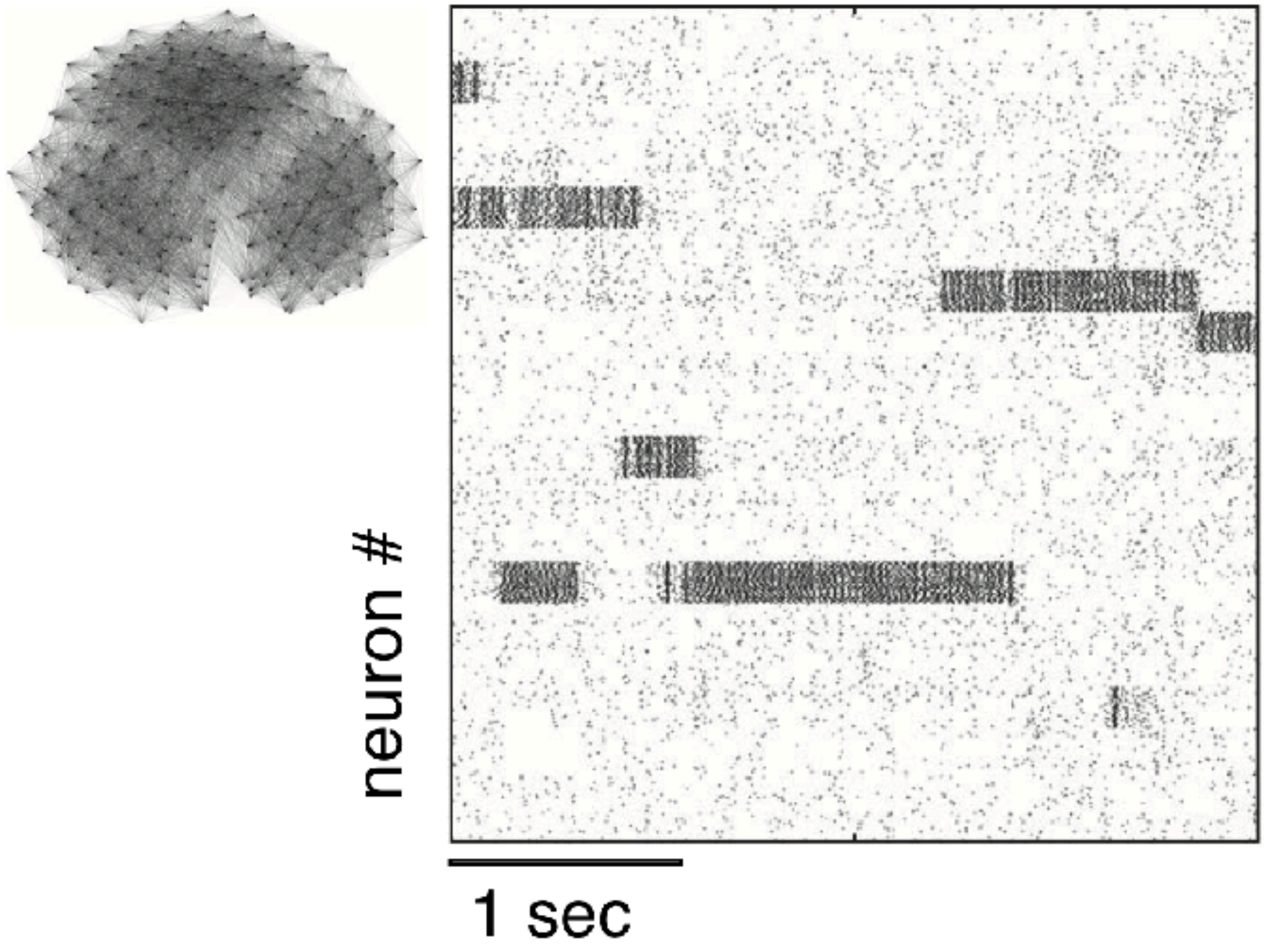




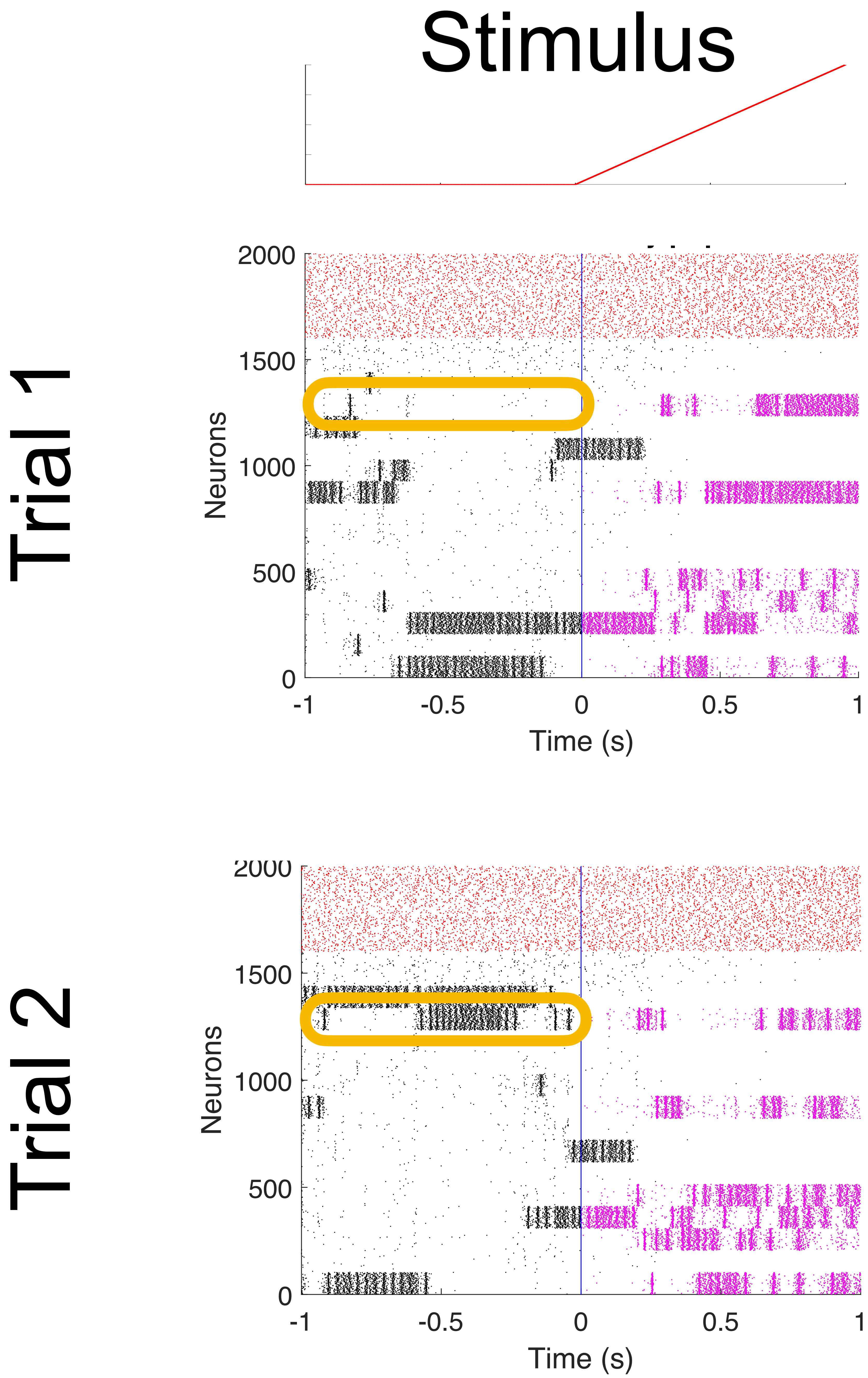
[Amit & Brunel, 1997]
[Deco & Hugues, 2012]
[Litwin–Kumar & Doiron, 2012]
[Mazzucato et al., 2015]

# Clustered network




[Litwin-Kumar & Doiron, 2012]

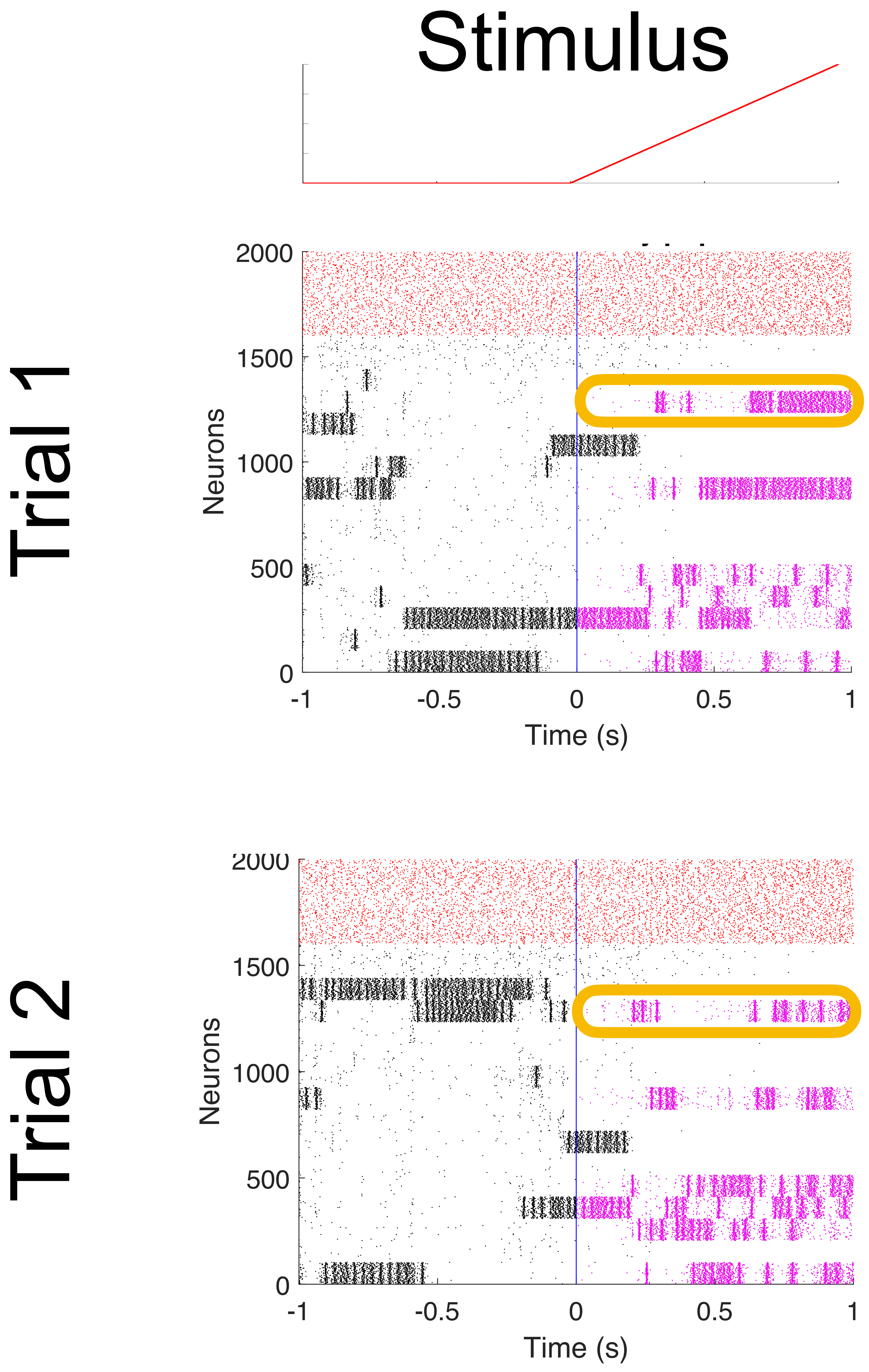



Trial-to-trial variability:

Fano factor = Var(rate) / Mean(rate)

Fano(spont) > Fano(evoked)

[Deco & Hugues, 2012]
[Litwin-Kumar & Doiron, 2012]
[Mazzucato et al, 2015]




Trial-to-trial variability:

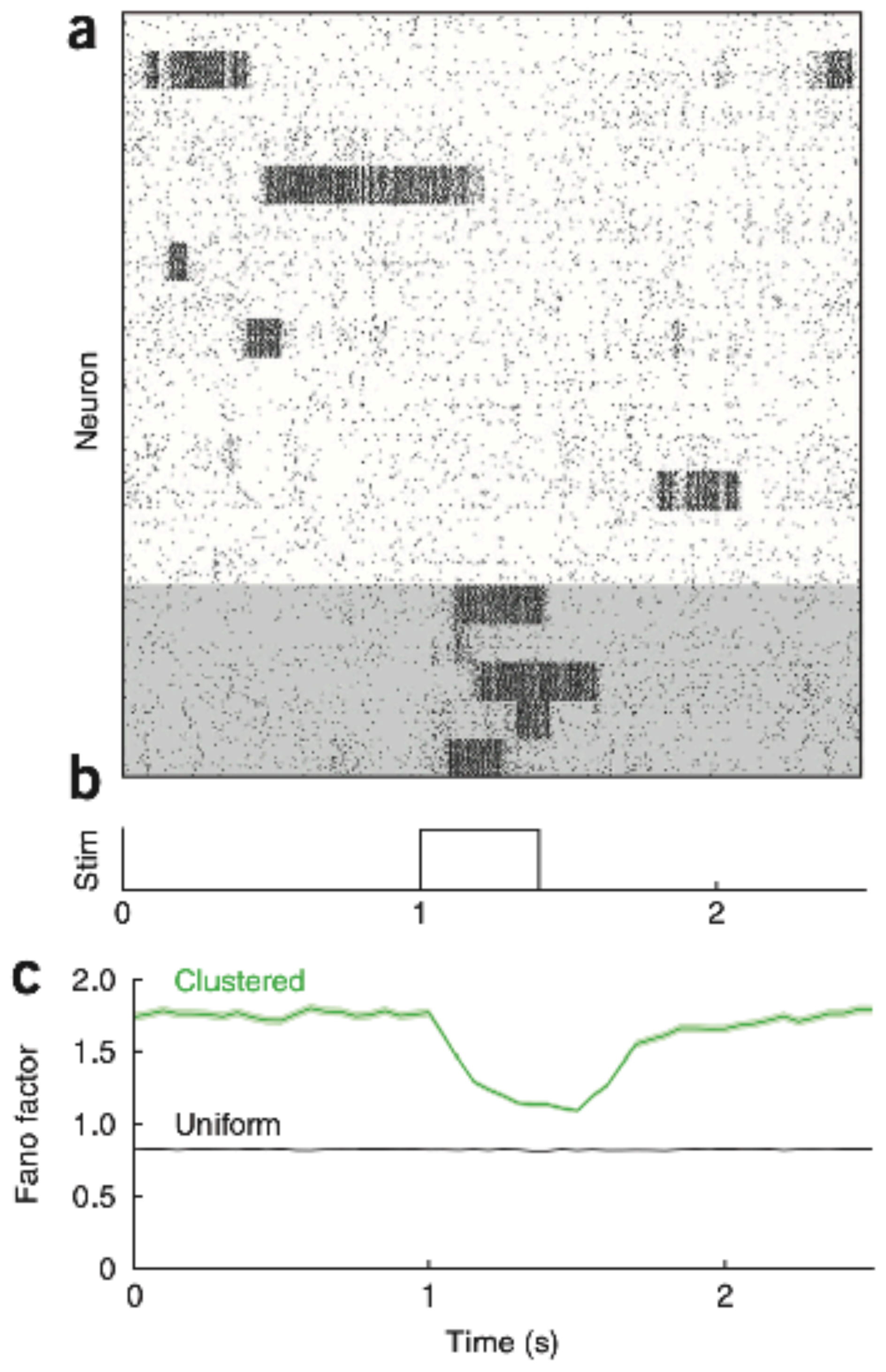
Fano factor = Var(rate) / Mean(rate)

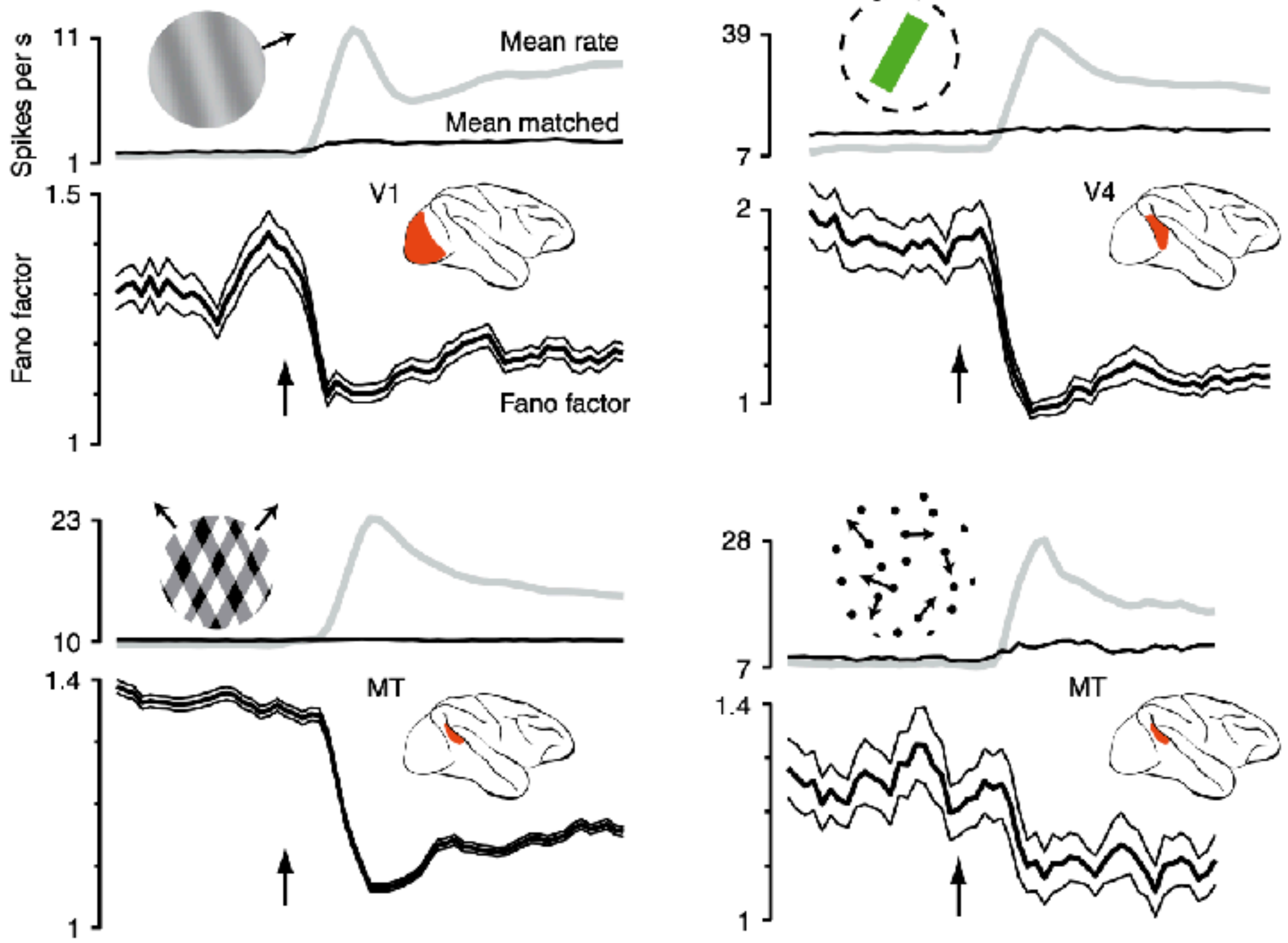
Fano(spont) > Fano(evoked)

[Deco & Hugues, 2012]
[Litwin-Kumar & Doiron, 2012]
[Mazzucato et al, 2015]

MODEL

DATA




Fano(spont) > Fano(evoked)

[Deco & Hugues, 2012]
[Litwin-Kumar & Doiron, 2012]
[Mazzucato et al, 2015]

[Churchland et al, 2010]

# Goals

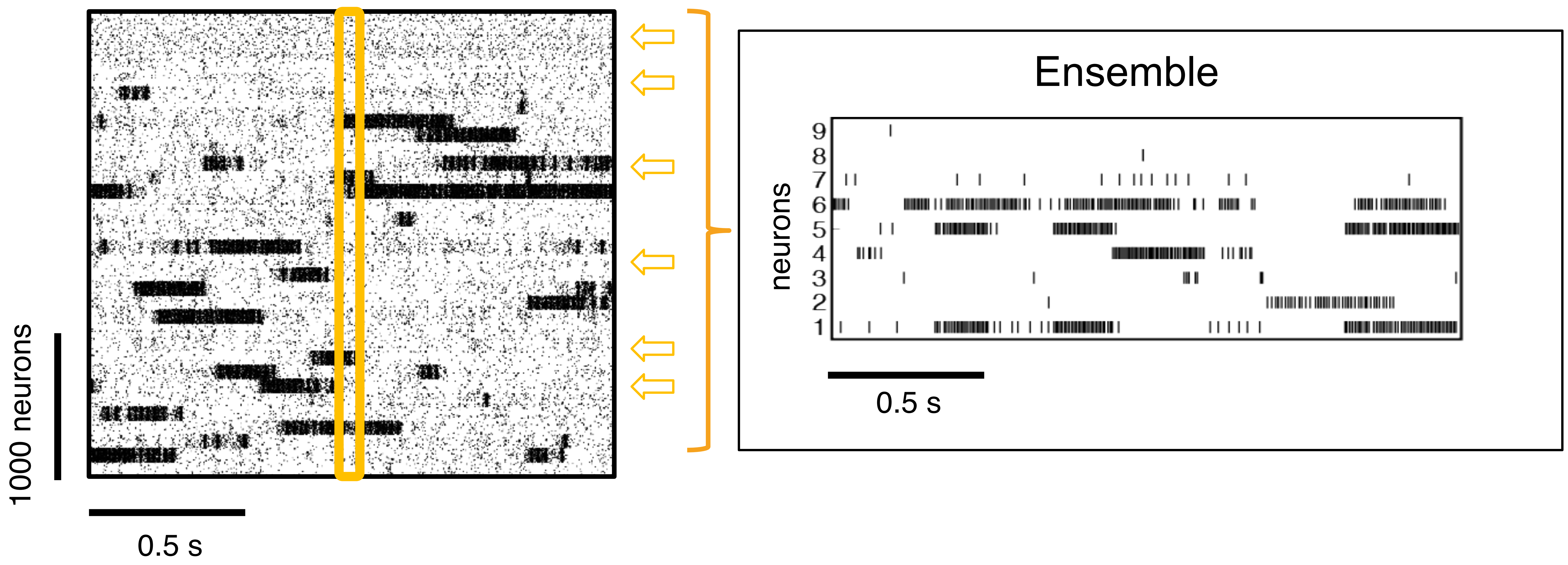
**Part I - Theory**

- **Metastable activity in cortical circuits**
- **Hidden Markov Models fit to metastable activity**
- **Modeling metastable activity with attractor networks**

**Part II - Exercise**

- **HMM: model selection with gaussian emissions**
- **Fit Poisson-HMM to spike trains**
- **Challenge: A model of arousal in attractor networks**

# Clustered network

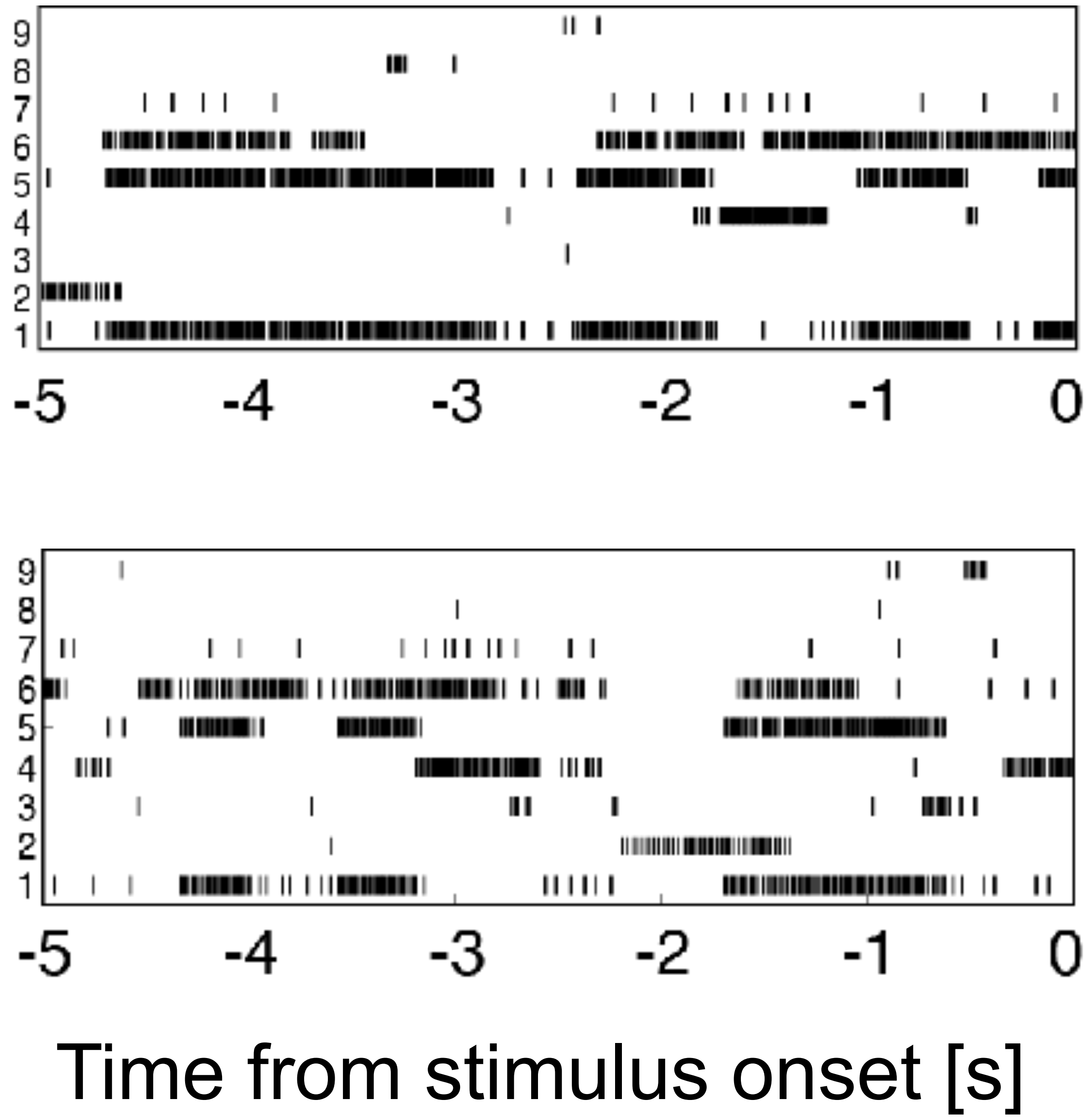



[Mazzucato et al., 2015, 2016]

# Clustered network

DATA

MODEL

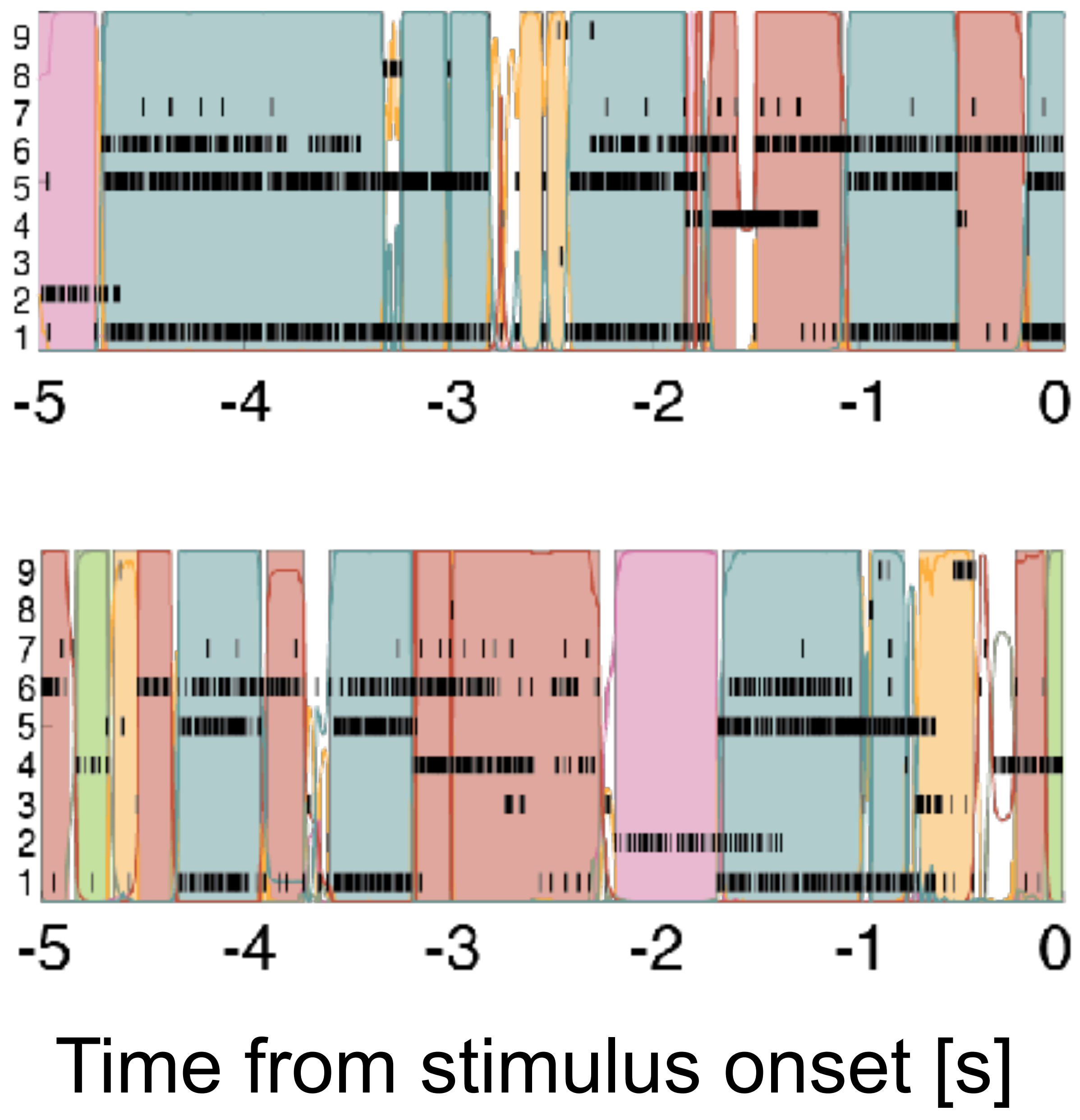



[Mazzucato et al., 2015]

# Clustered network

DATA

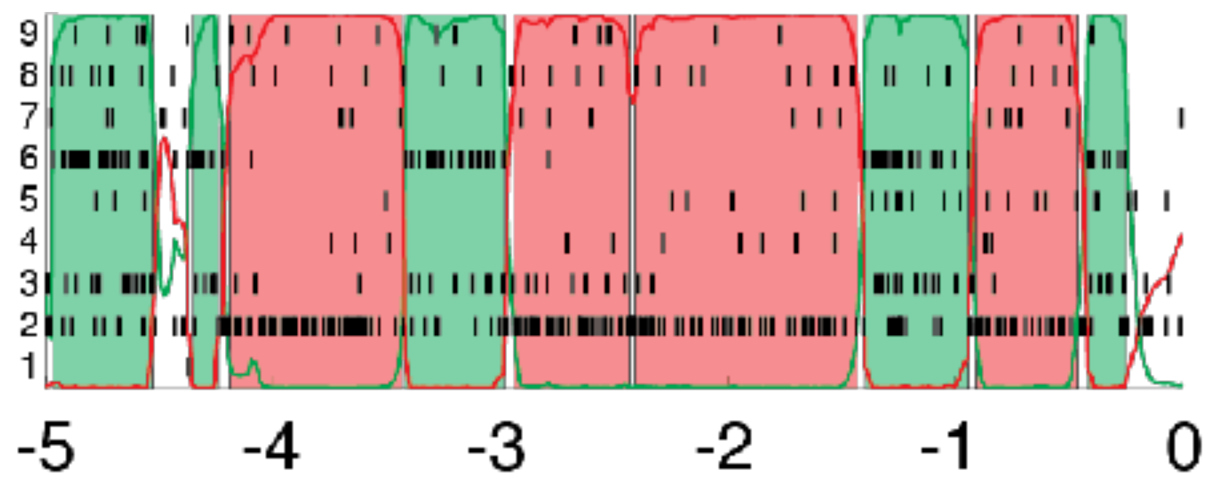
MODEL

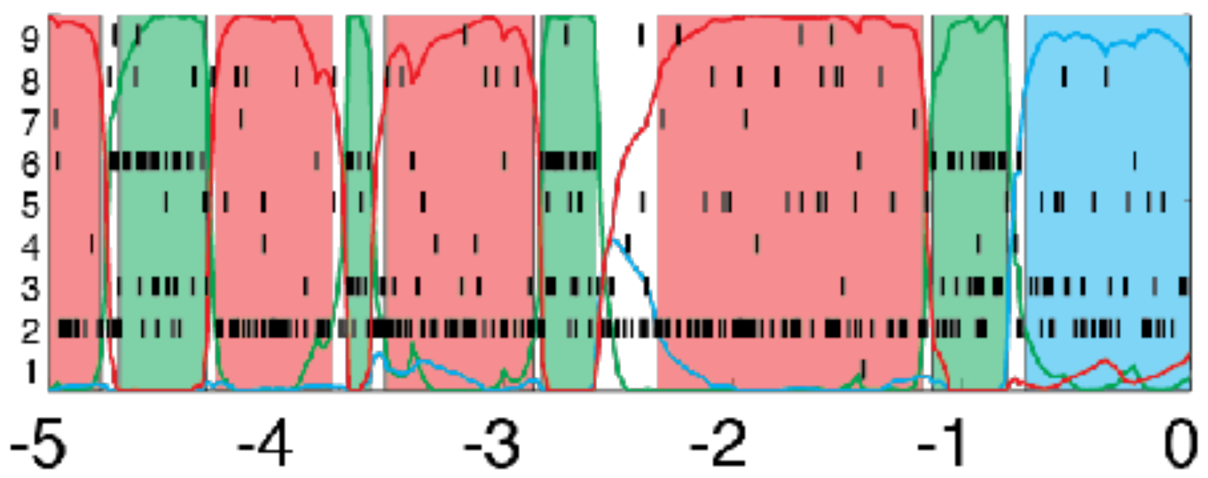



[Mazzucato et al., 2015]

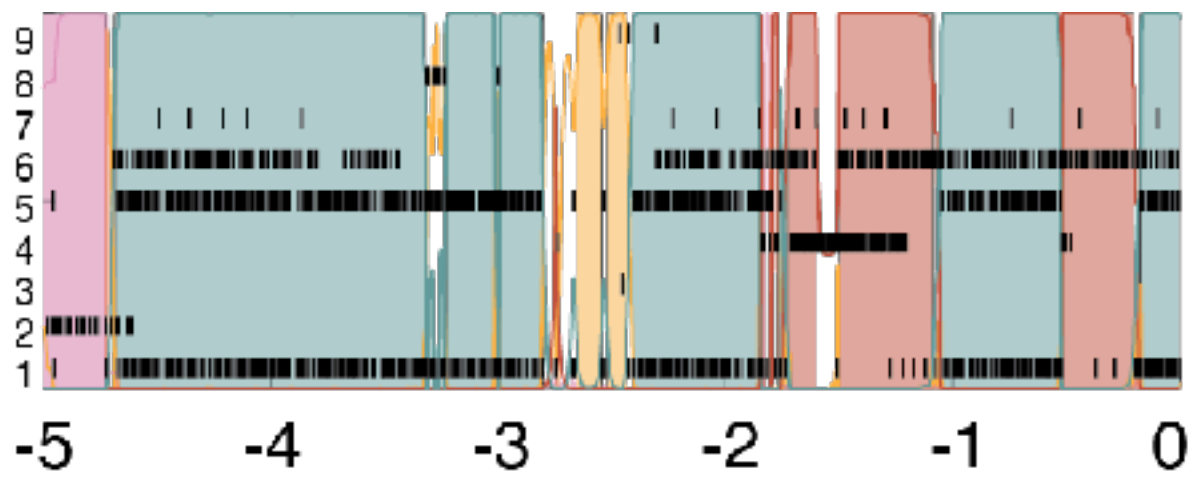
# Clustered network

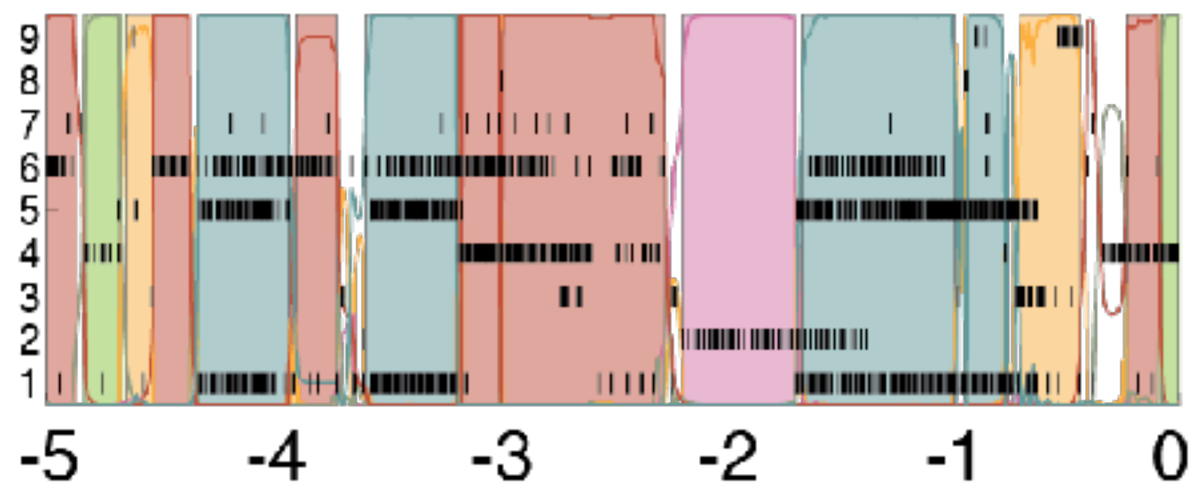
| DATA | MODEL |
|---|---|




Time from stimulus onset [s]




Time from stimulus onset [s]

[Mazzucato et al., 2015]

# Goals

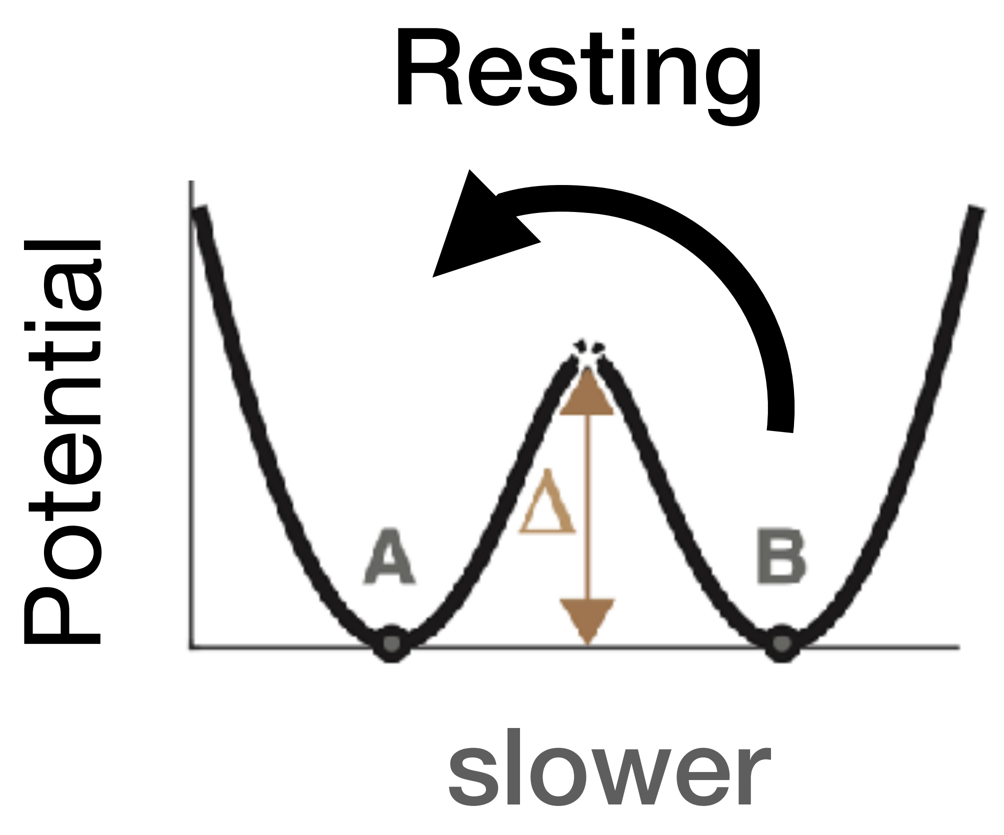
**Part I - Theory**

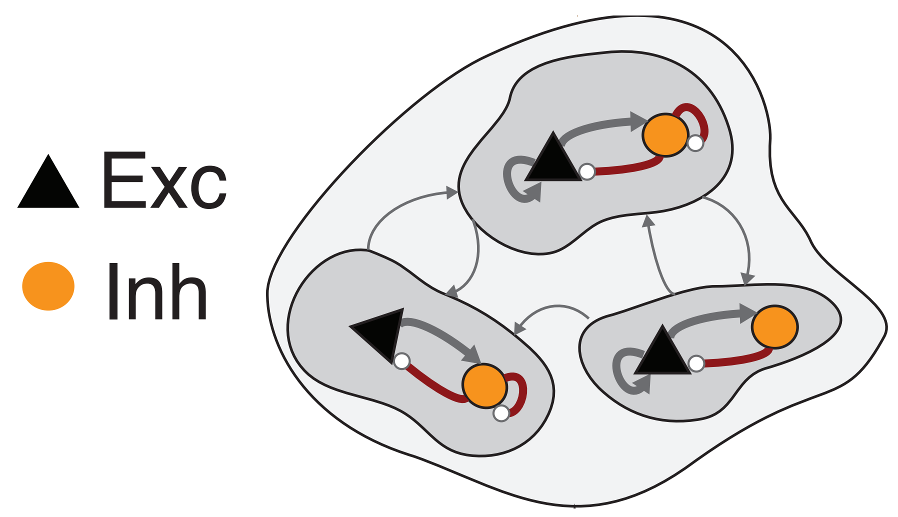
- **Metastable activity in cortical circuits**
- **Hidden Markov Models fit to metastable activity**
- **Modeling metastable activity with attractor networks**

**Part II - Exercise**

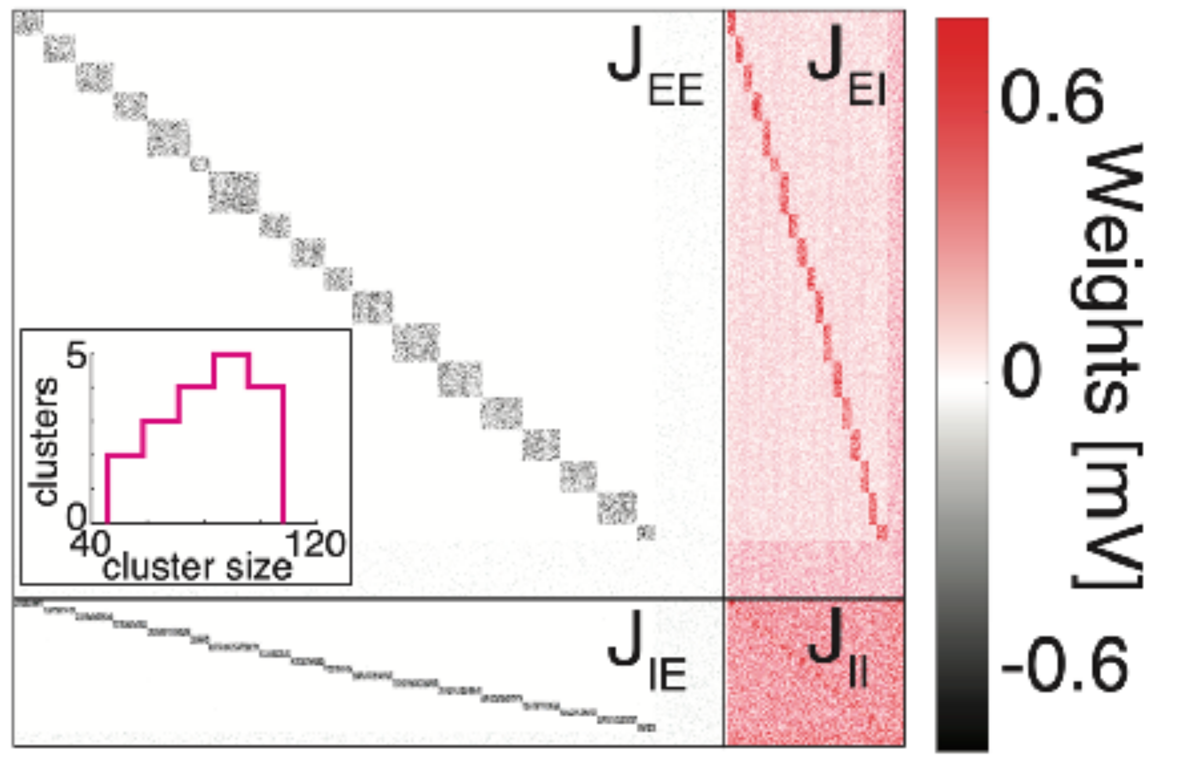
- **HMM: model selection with gaussian emissions**
- **Fit Poisson-HMM to spike trains**
- **Challenge: A model of arousal in attractor networks**

# Ongoing activity

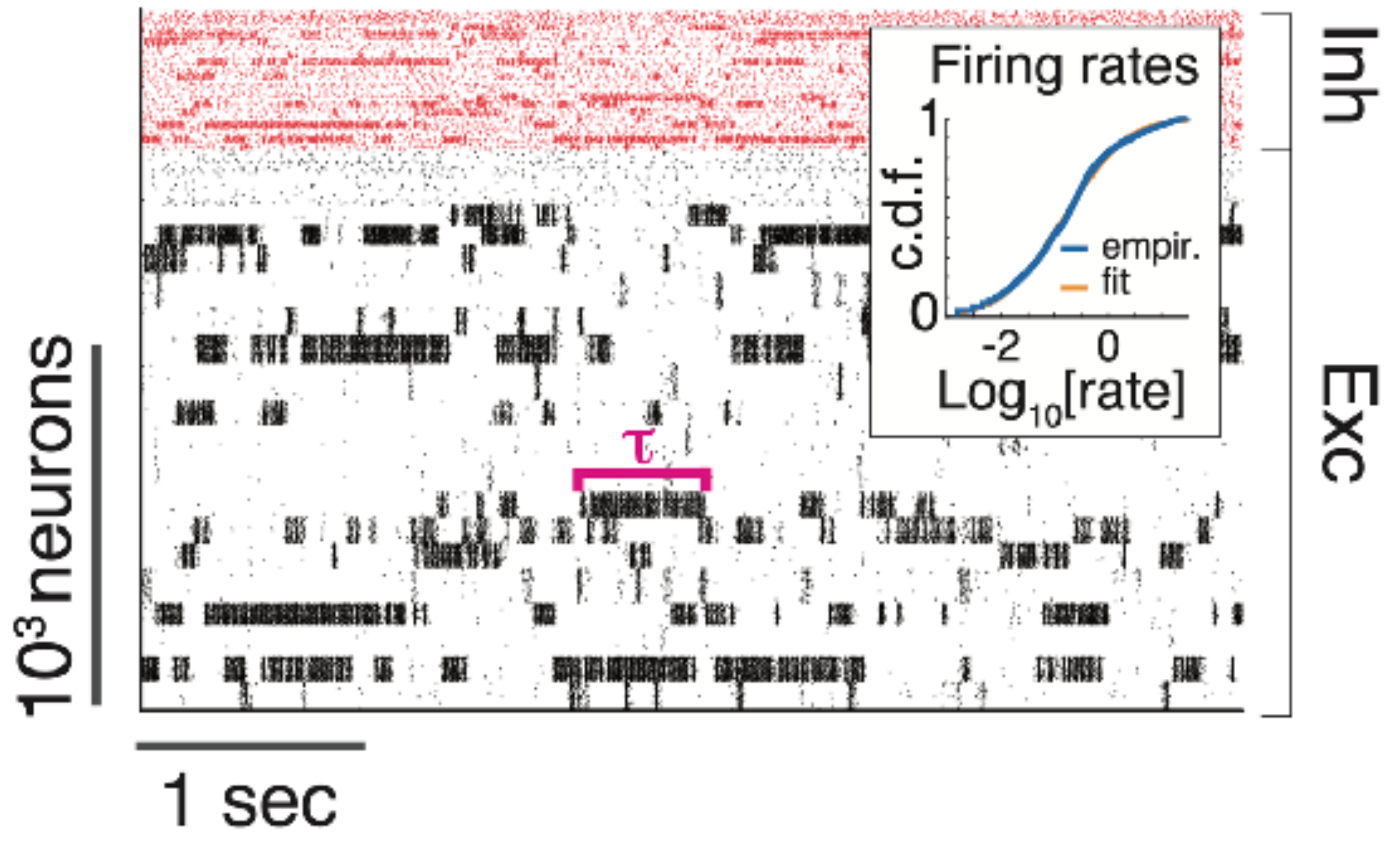


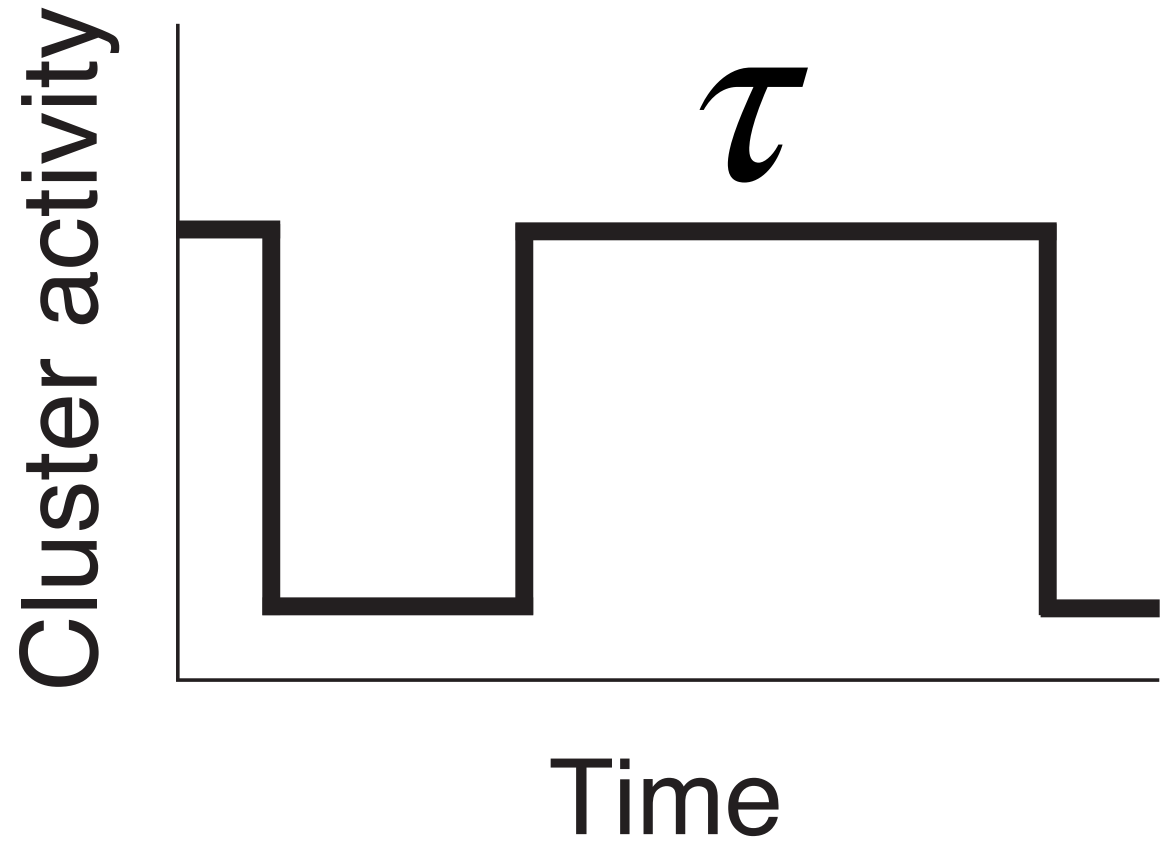




## Synaptic weights




## Ongoing activity

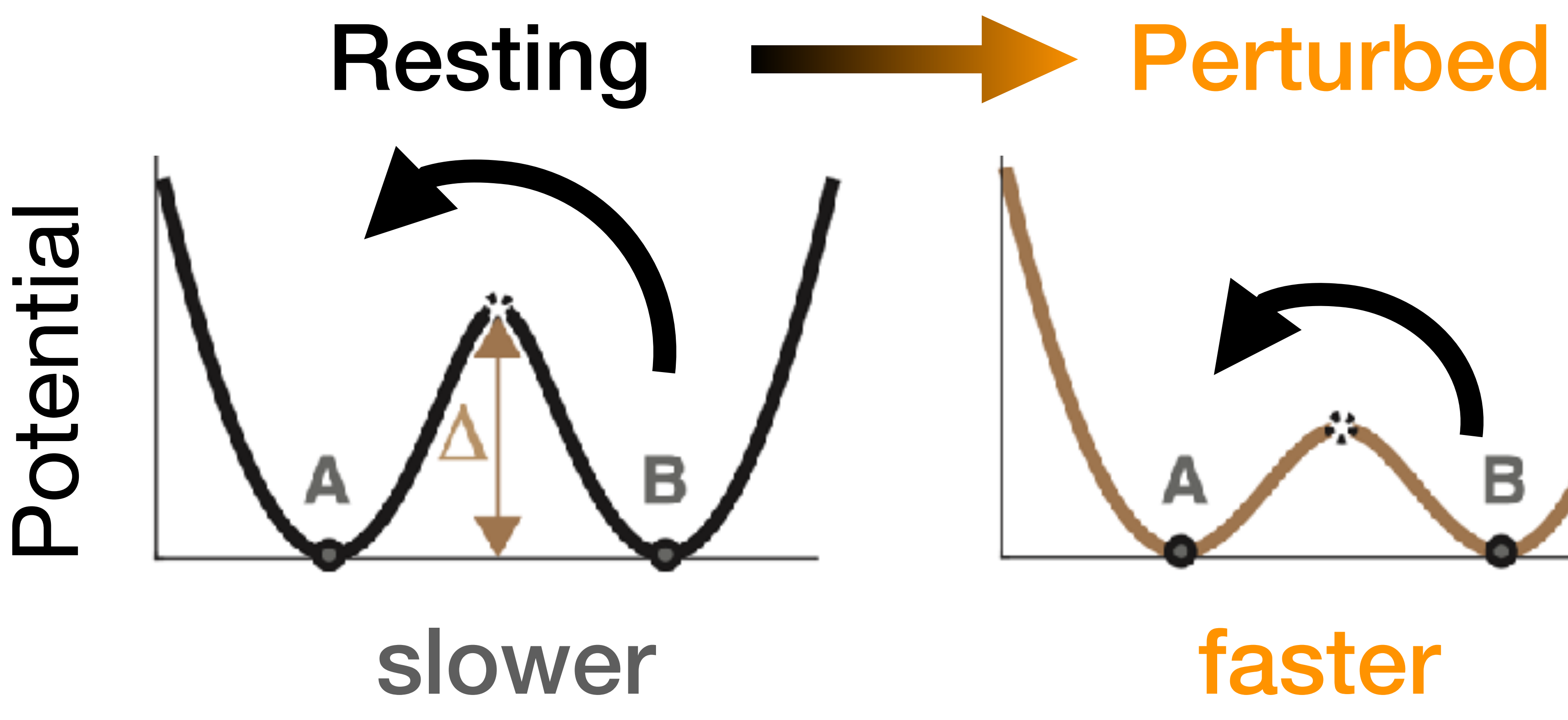


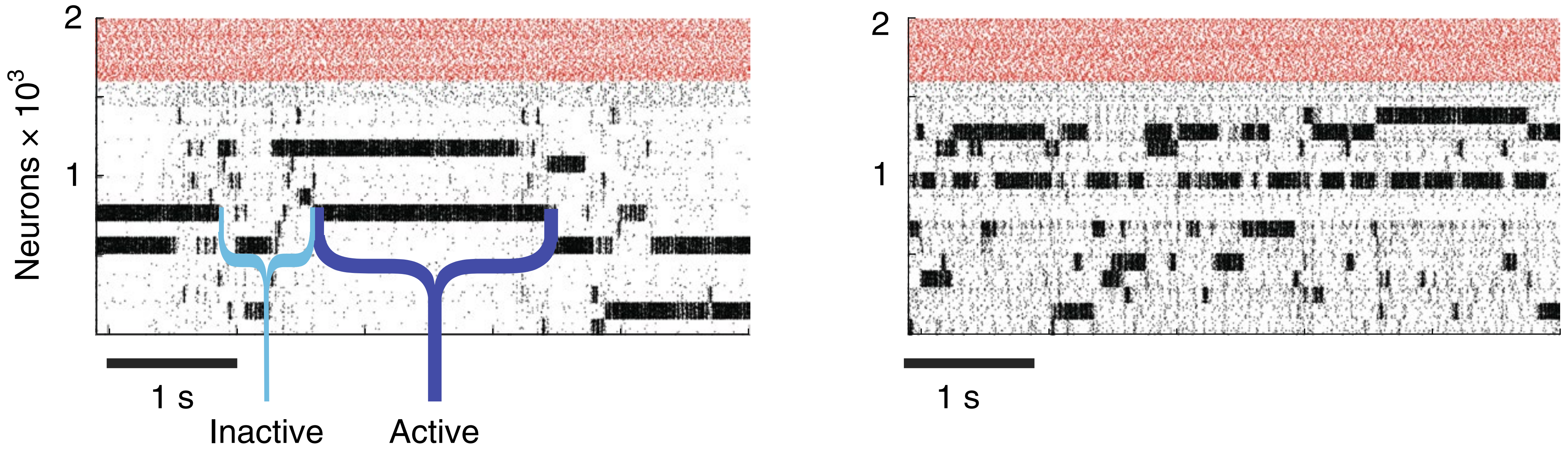



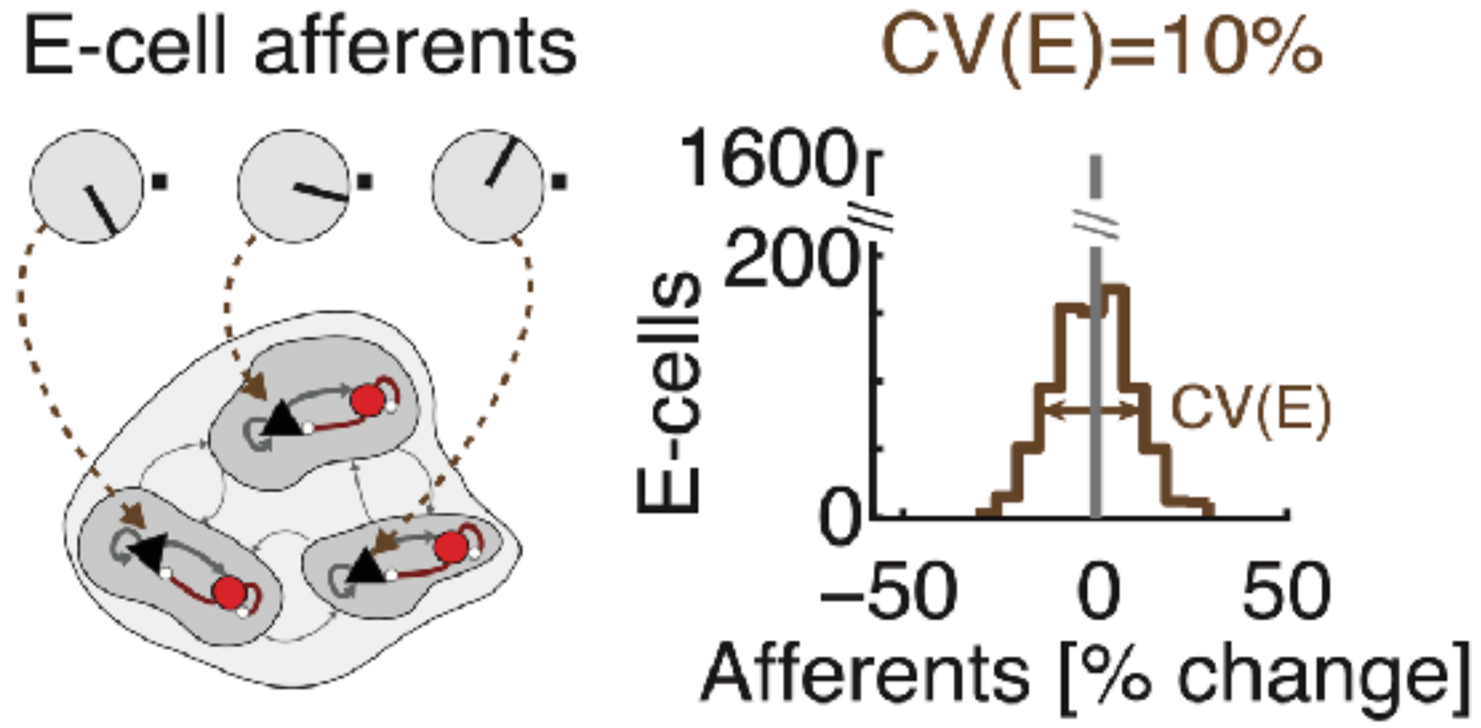

**Barrier height —> Transition time**

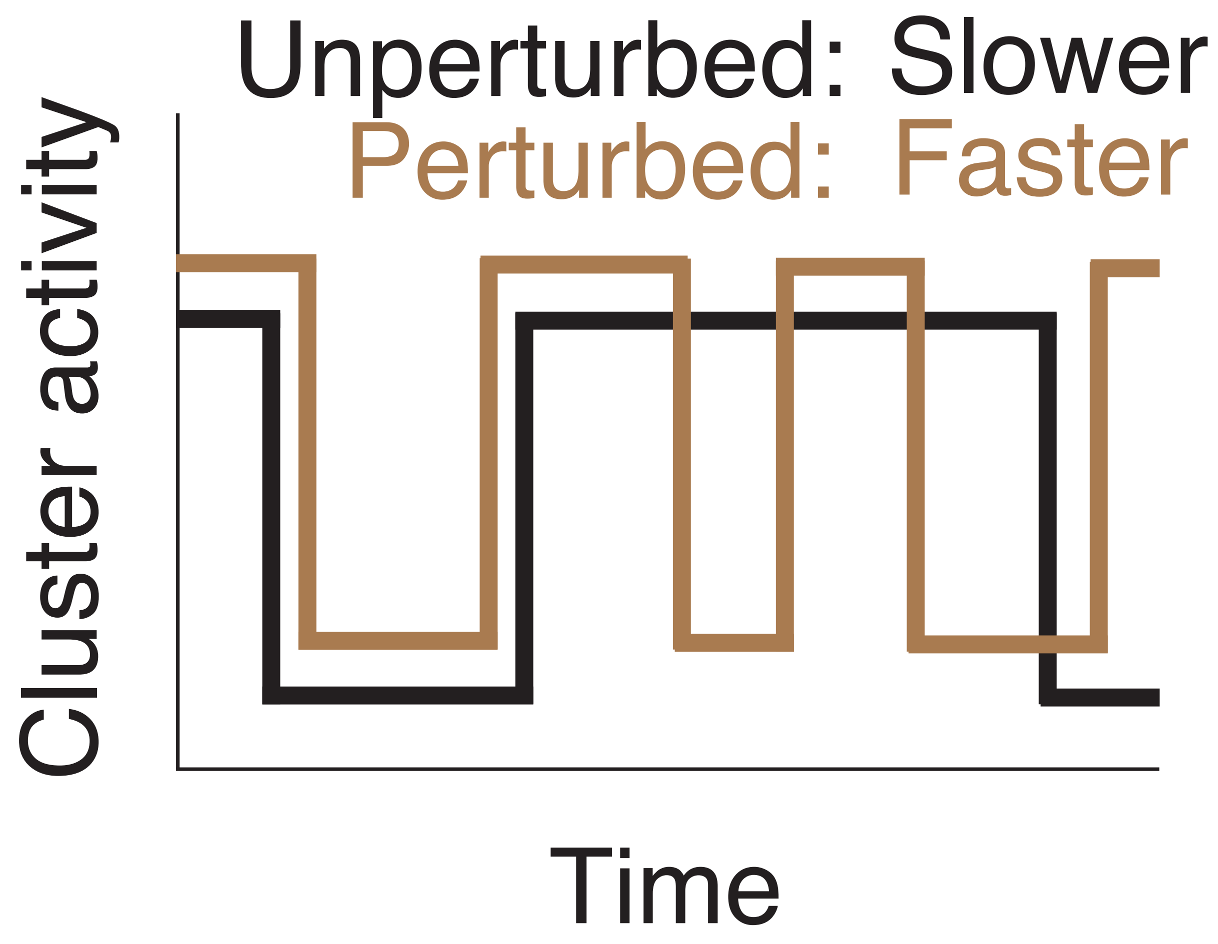
[LM et al, 2019]
[Wyrick & LM, 2021]

# Ongoing activity



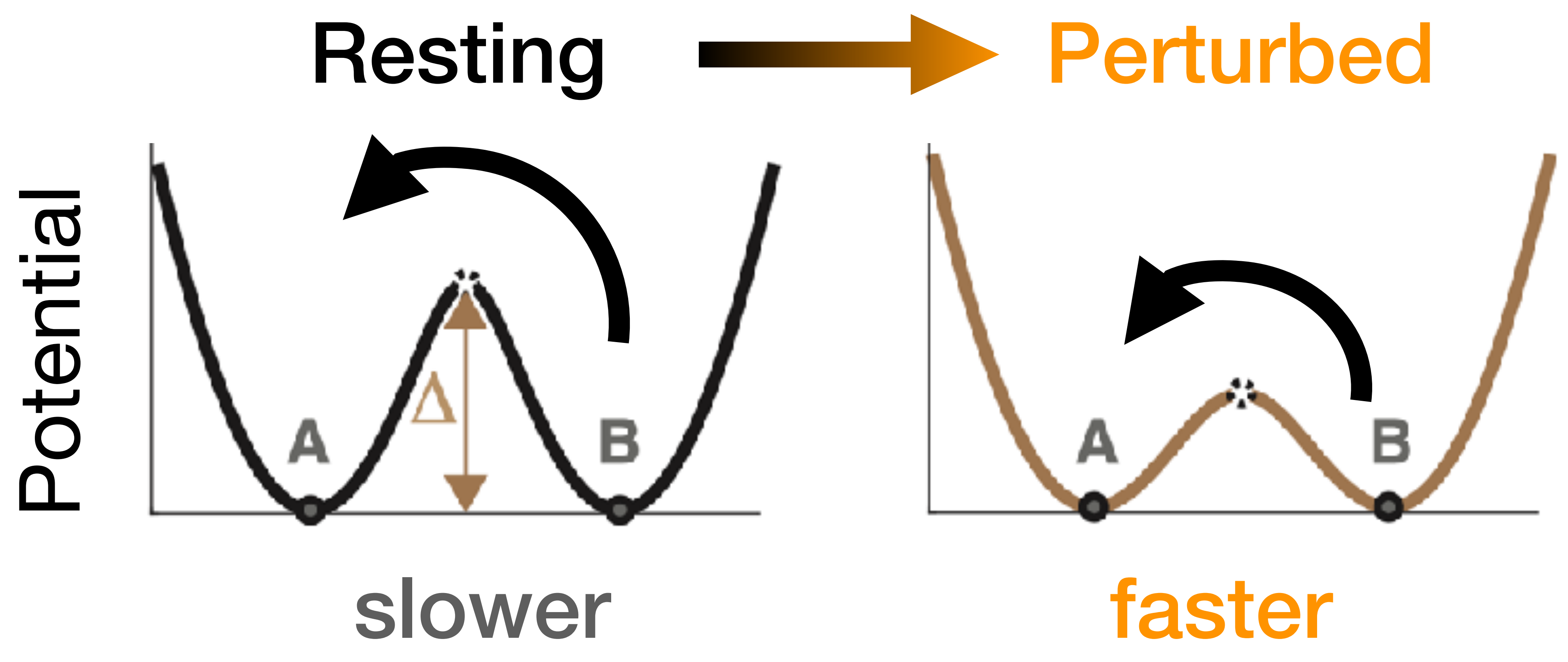



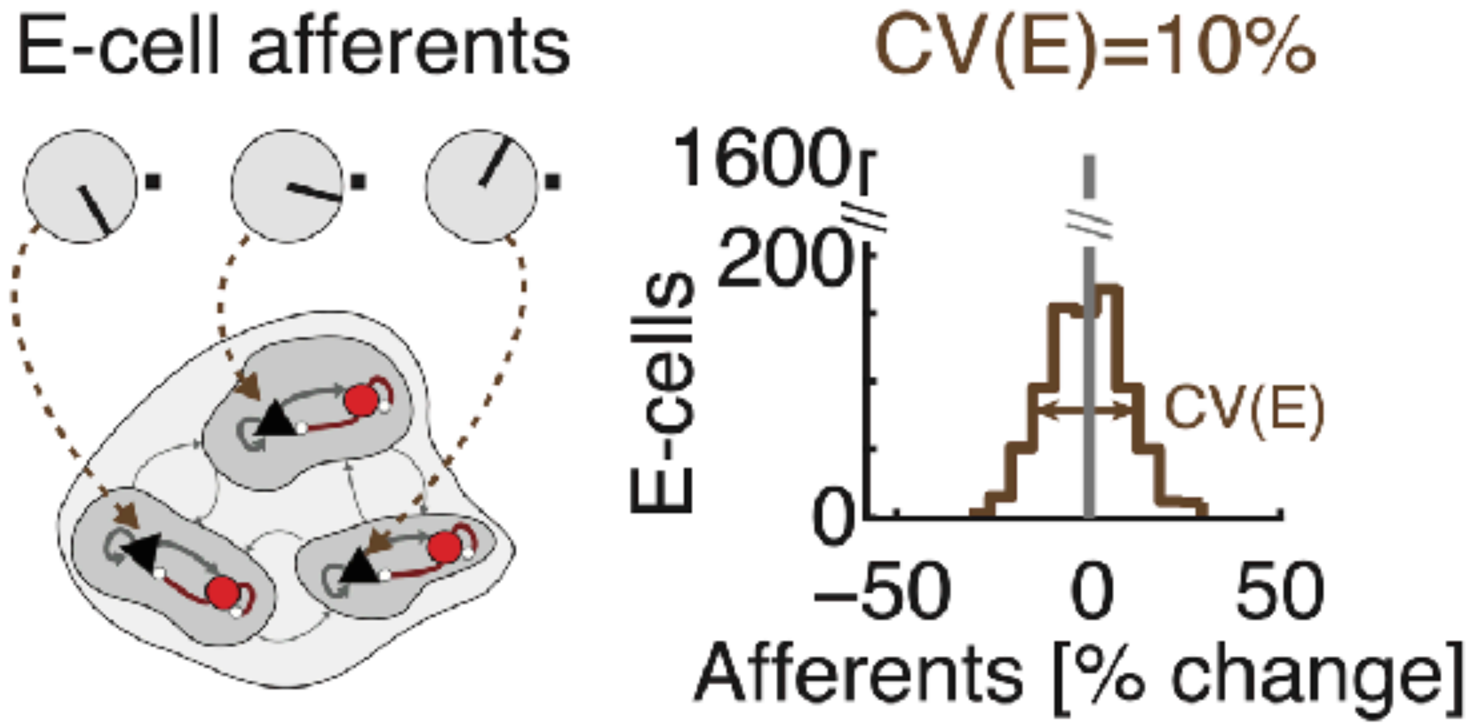



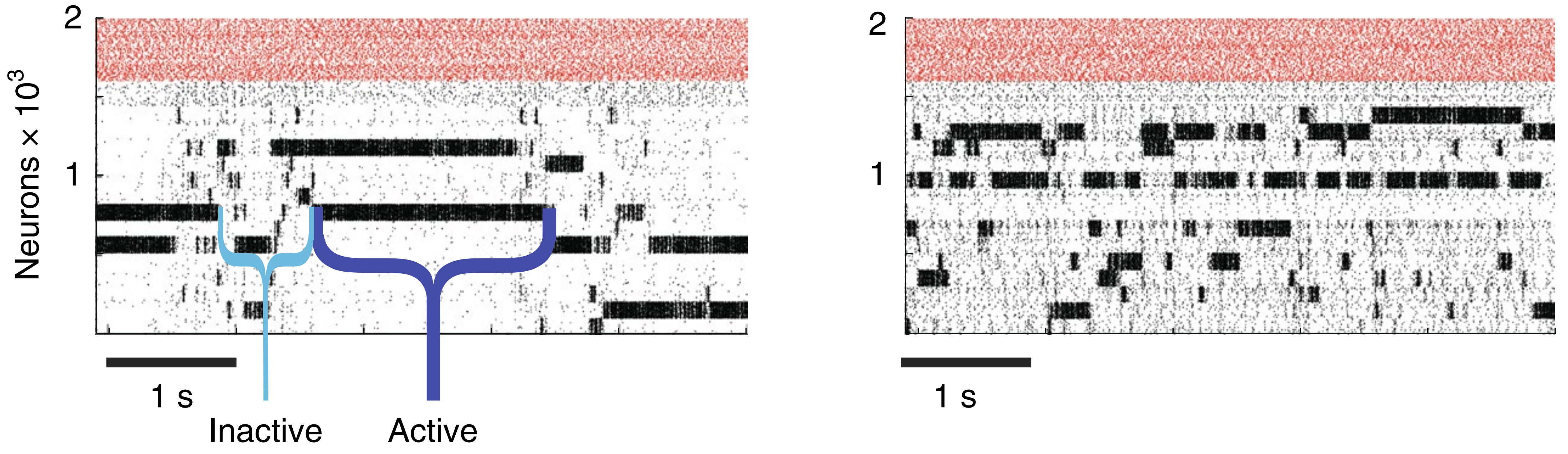

**Lower barrier —> Faster transition time**

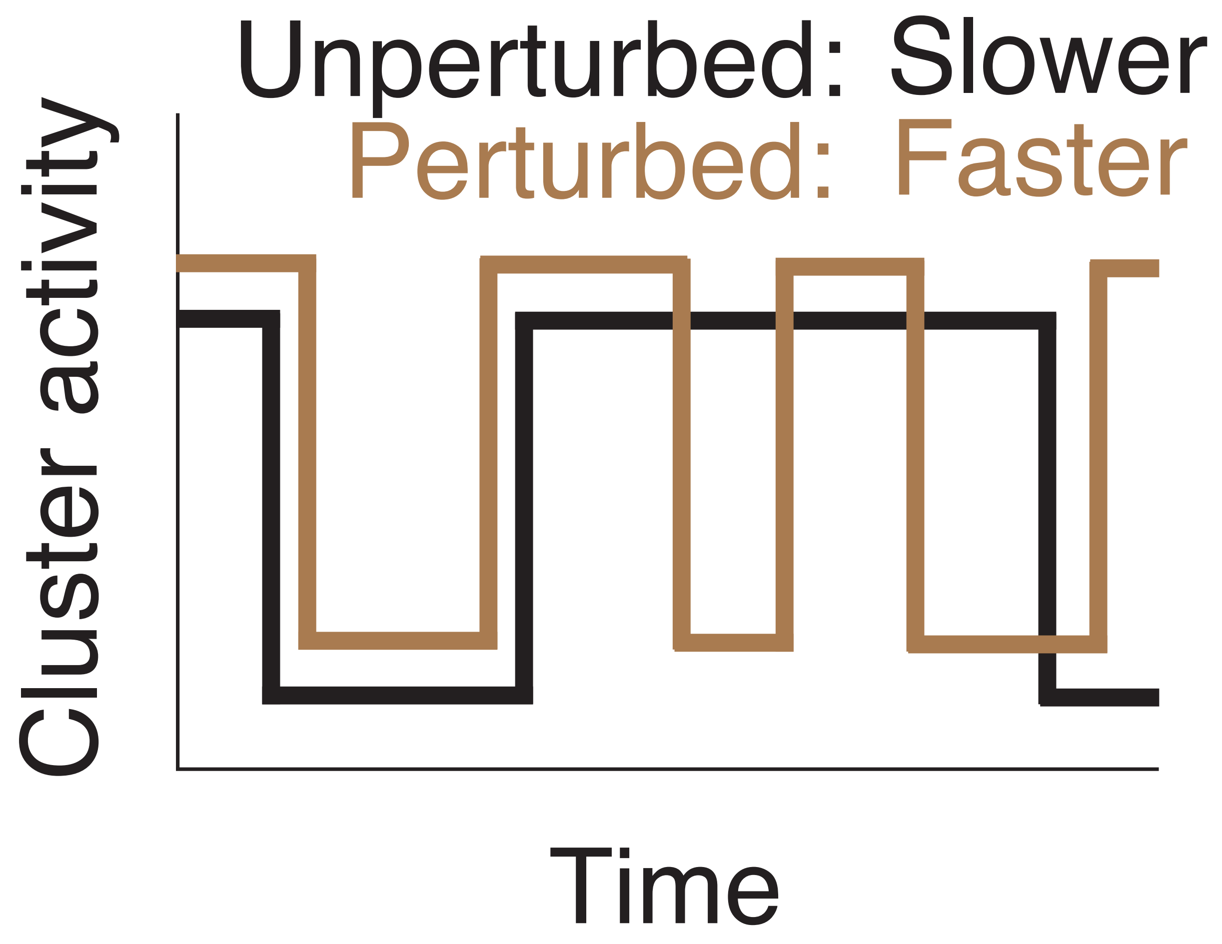
[LM et al, 2019]
[Wyrick & LM, 2021]

# Ongoing activity




**Challenge:**

- **Fit HMM to resting and perturbed conditions**
- **Compare state dwell times**

[LM et al, 2019]
[Wyrick & LM, 2021]

Metastable activity:

- is observed in many cortical areas and tasks
- is involved in sensory and cognitive processes
- can emerge spontaneously in clustered spiking networks

www.mazzulab.com

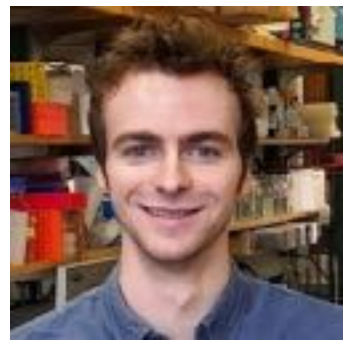

Stefano Recanatesi


Lia Papadopoulos


David Wyrick


Audra McNamee


Nicu Istrate


Peregrine Painter


Merav Stern
(Hebrew U.)


Alireza Tavanfar
(Champalimaud)


# Funding

NIH National Institute on Deafness and Other Communication Disorders

BRAIN INITIATIVE

www.mazzulab.com

Stefano Recanatesi

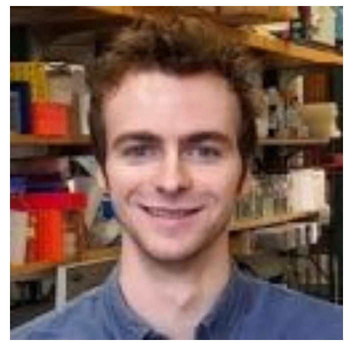
David Wyrick

Nicu Istrate

Merav Stern
(Hebrew U.)

Lia Papadopoulos

Audra McNamee



Peregrine Painter

Alireza Tavanfar
(Champalimaud)

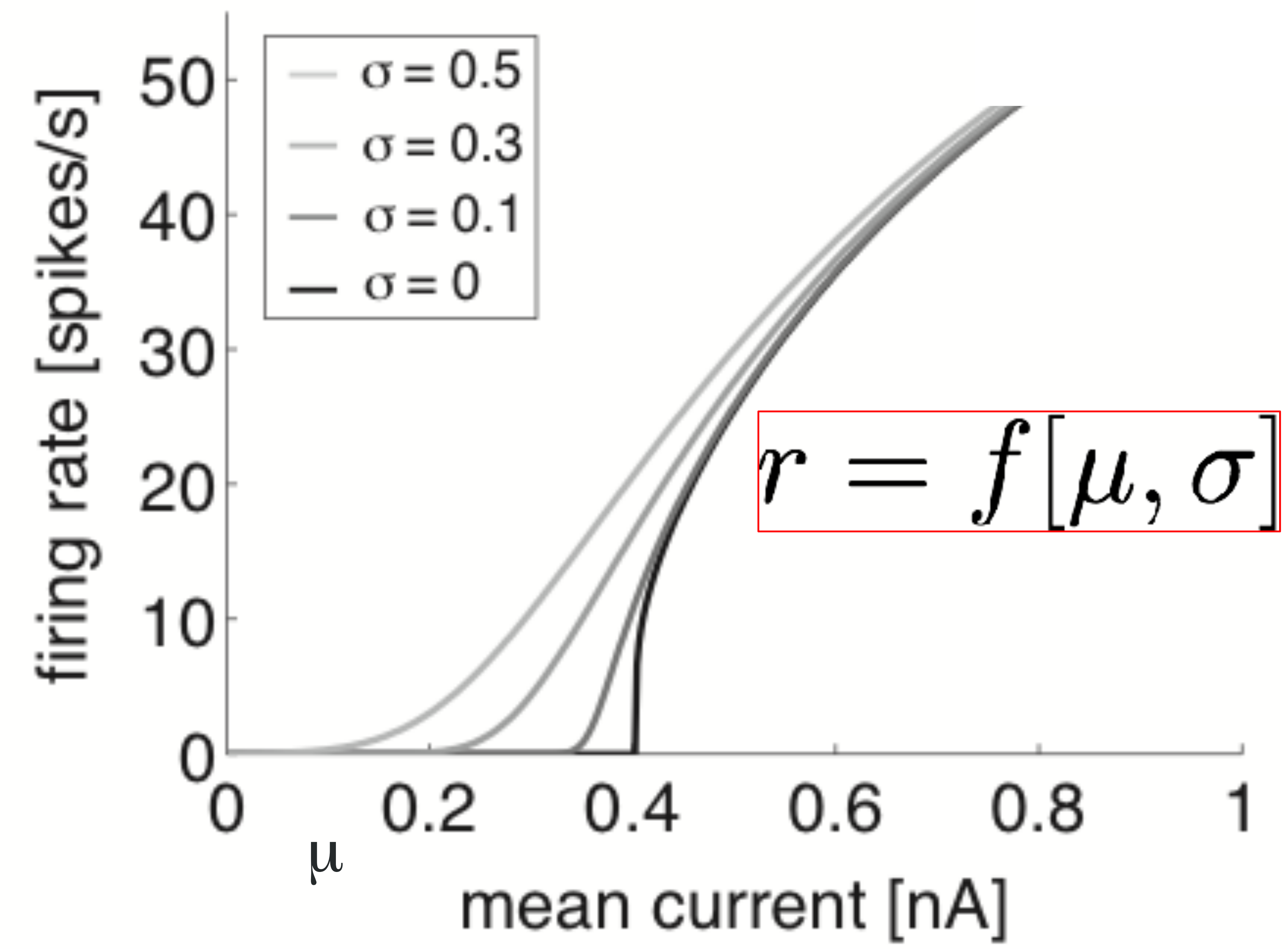

# Funding

NIH National Institute on Deafness and Other Communication Disorders

BRAIN INITIATIVE


# Supplementary details

# Transfer function




$$r = f[\mu, \sigma]$$

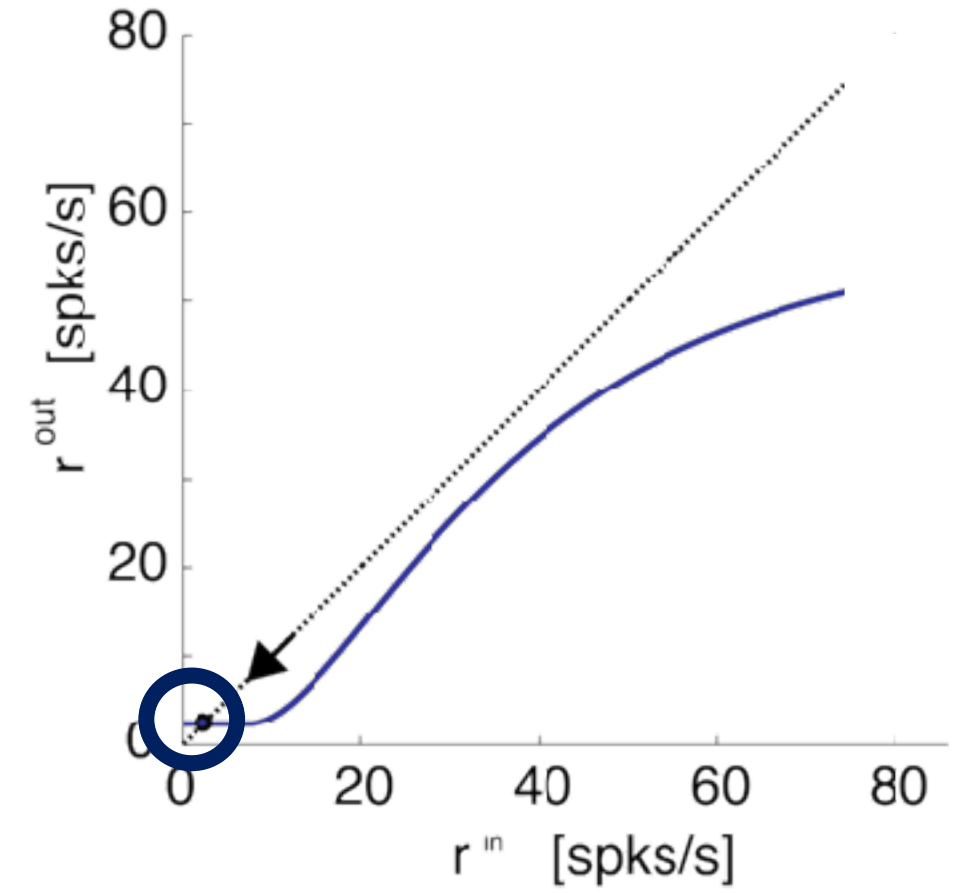
$$\mu = \mu(r)$$
$$\sigma = \sigma(r)$$

[Amit & Brunel, 1997]

# Model – Attractors

$$r^{out} = f\left[\mu(r^{in}), \sigma(r^{in})\right]$$

$$r^{out} = r^{in}$$ fixed point

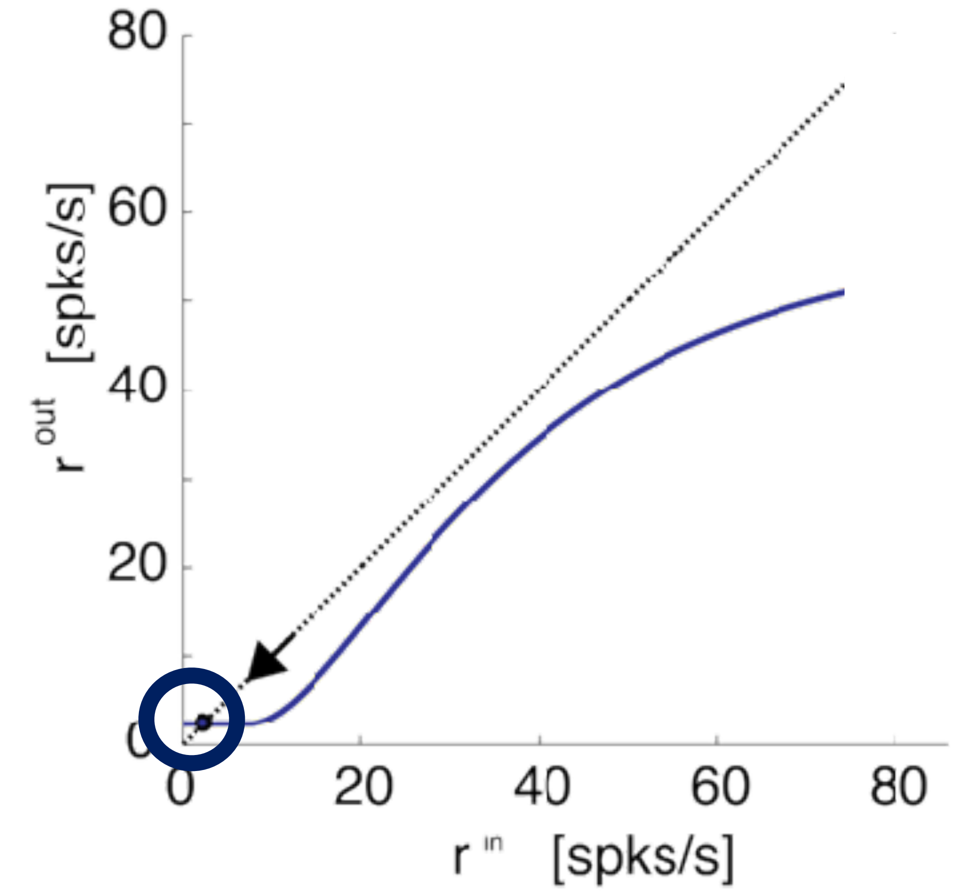



[Amit & Brunel, 1997]

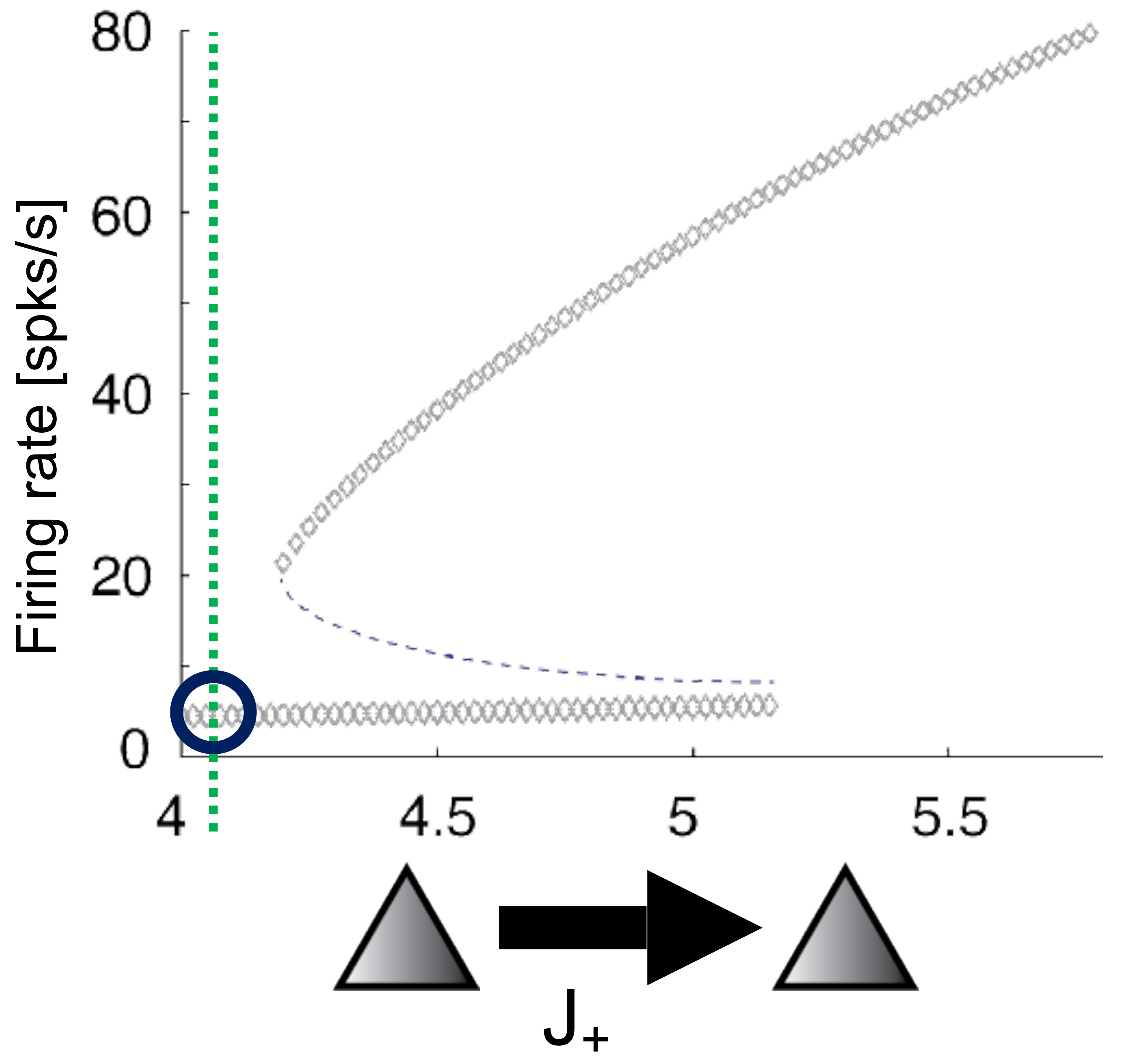
# Model – Attractors

$$r^{out} = f\left[\mu(r^{in}), \sigma(r^{in})\right]$$

$$r^{out} = r^{in}$$ fixed point




Network attractors (stable states)

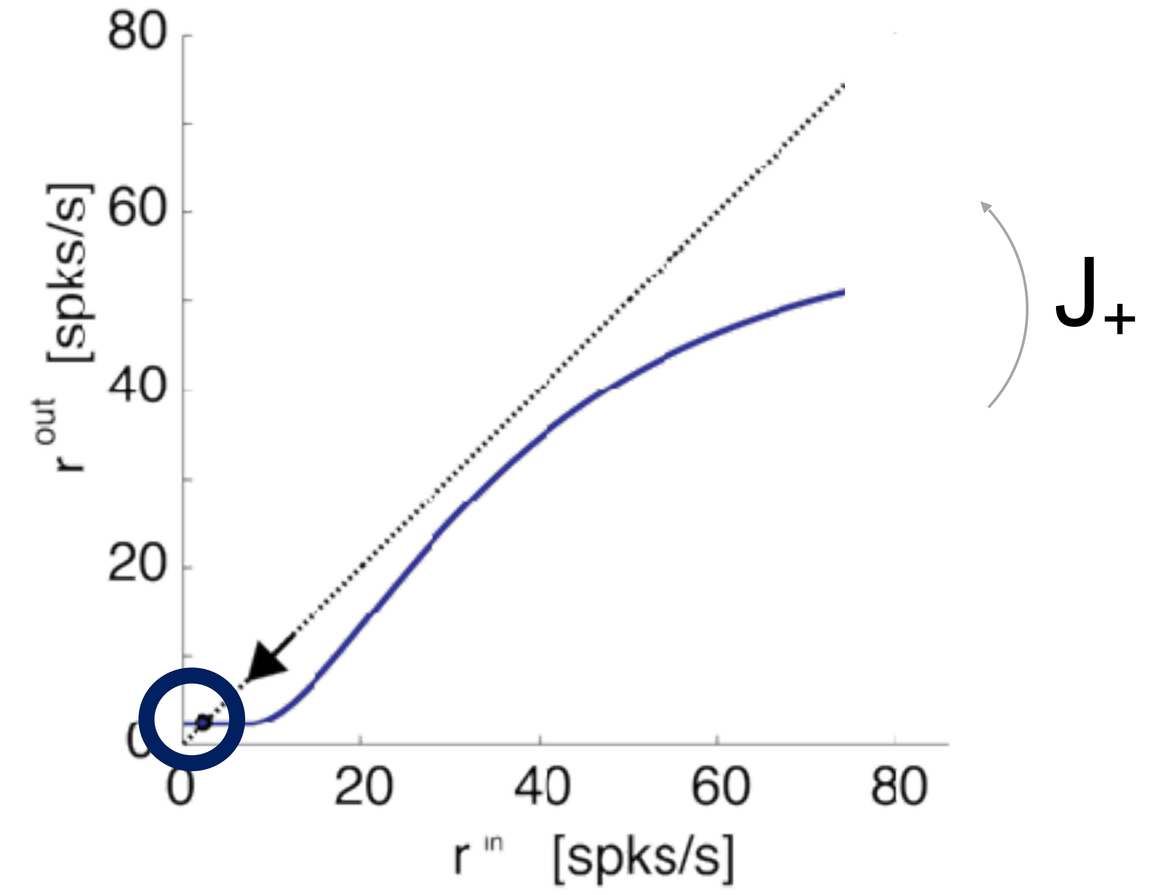



[Amit & Brunel, 1997]

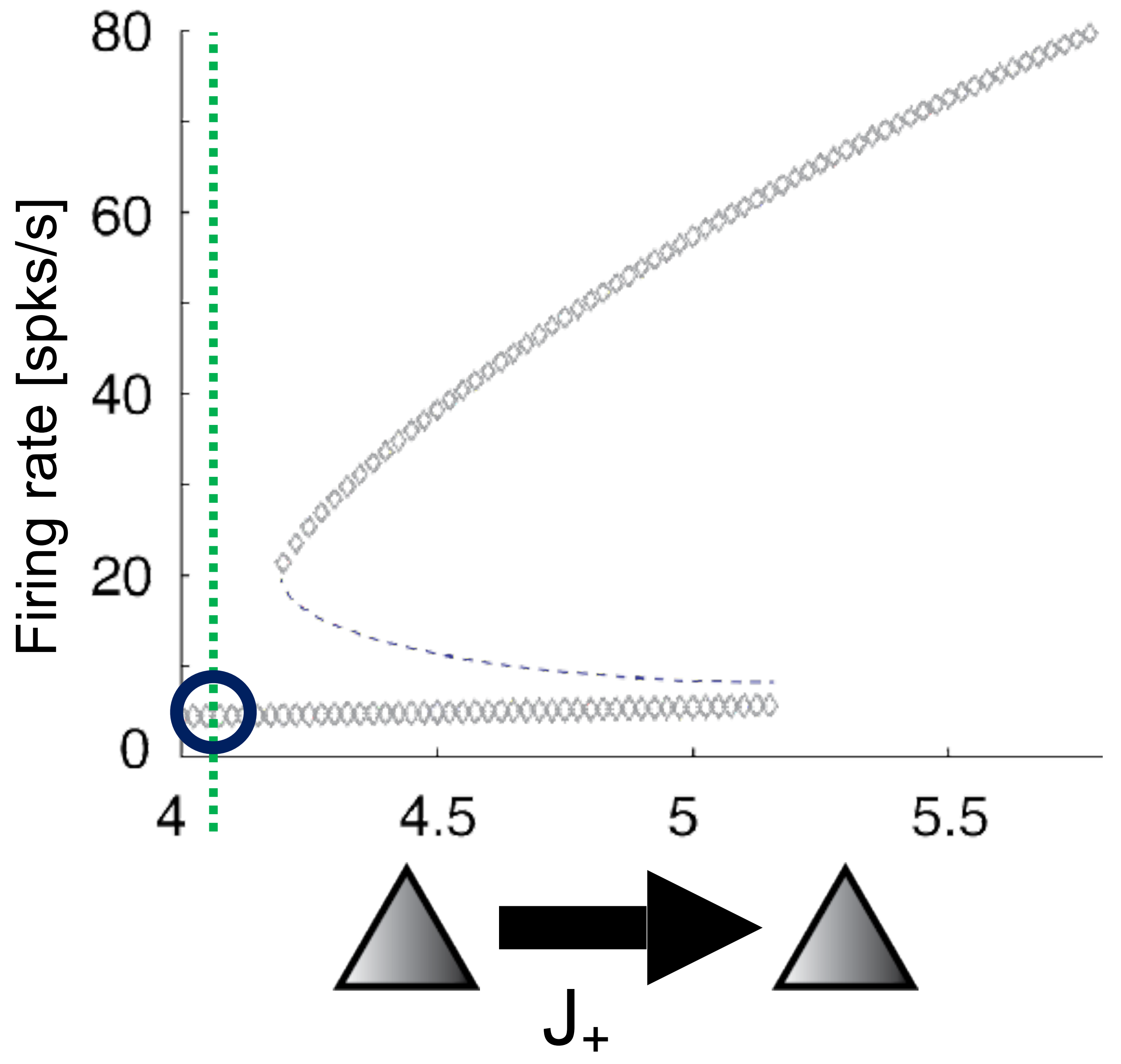
# Model – Attractors

$$r^{out} = f\left[\mu(r^{in}), \sigma(r^{in})\right]$$

$r^{out} = r^{in}$ fixed point




Network attractors (stable states)

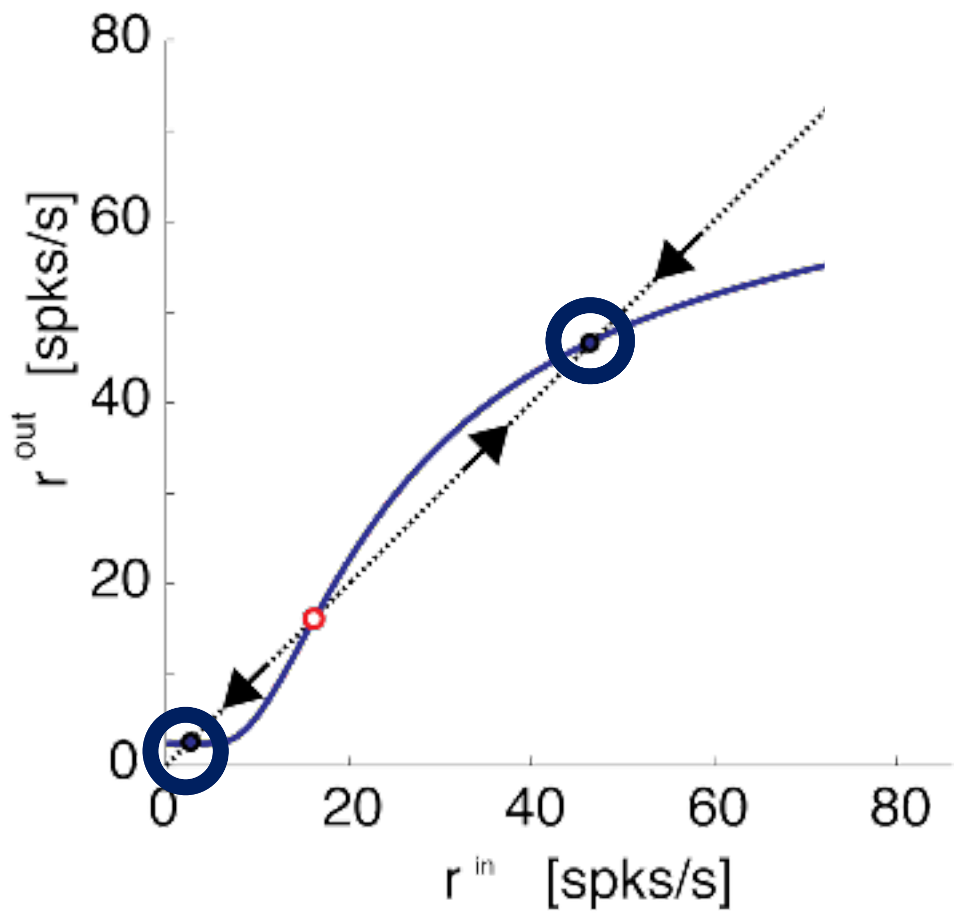


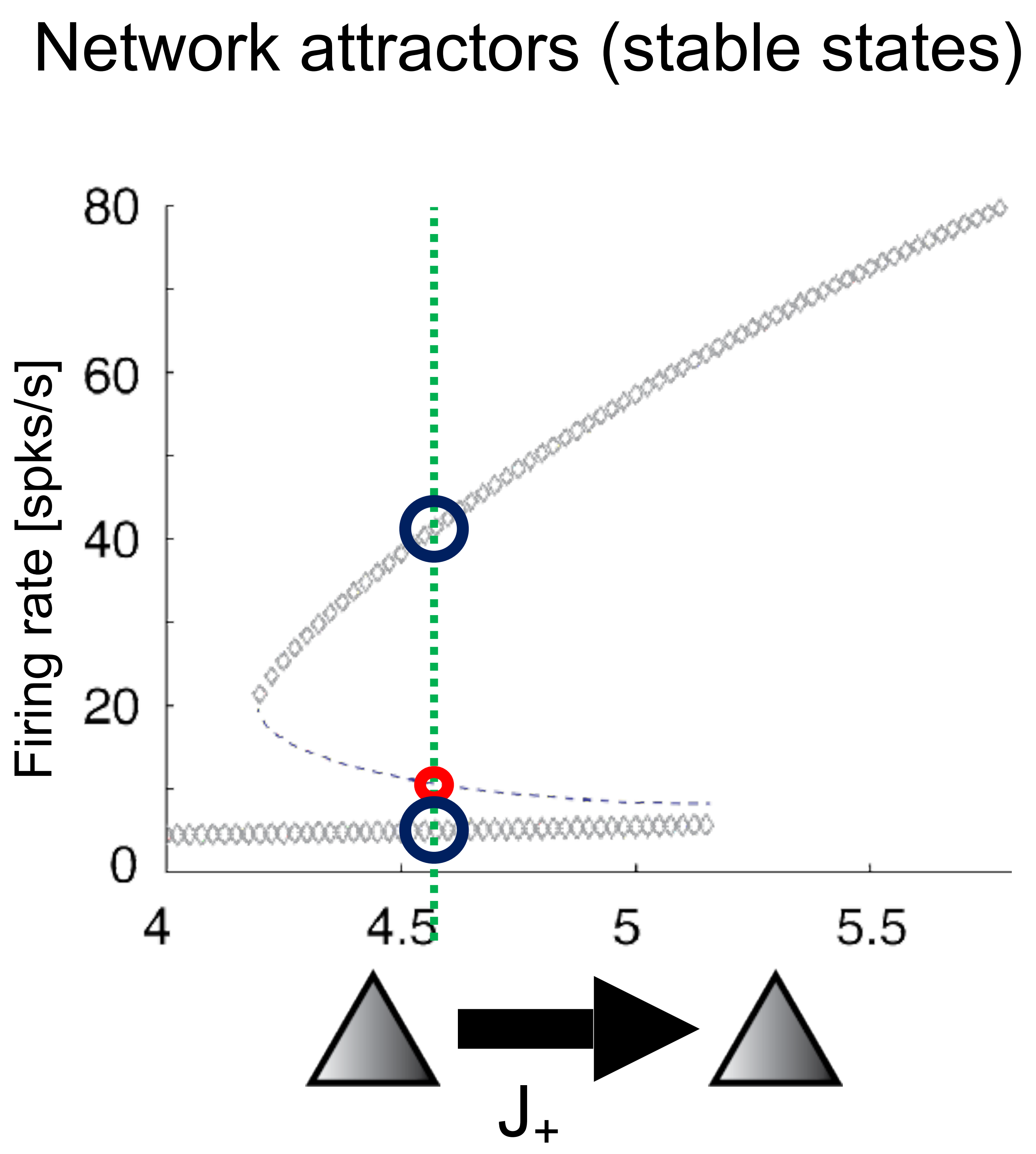

[Amit & Brunel, 1997]

# Model – Attractors

$$r^{out} = f\left[\mu(r^{in}), \sigma(r^{in})\right]$$

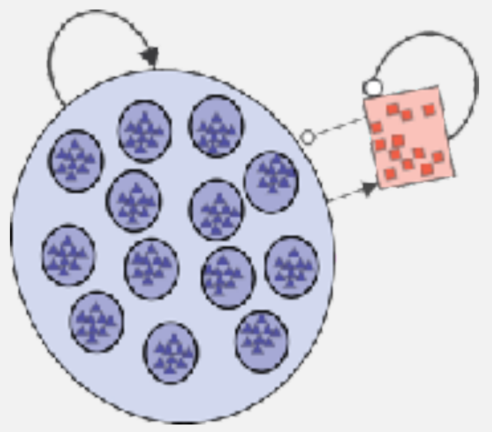
$$r^{out} = r^{in}$$ fixed point

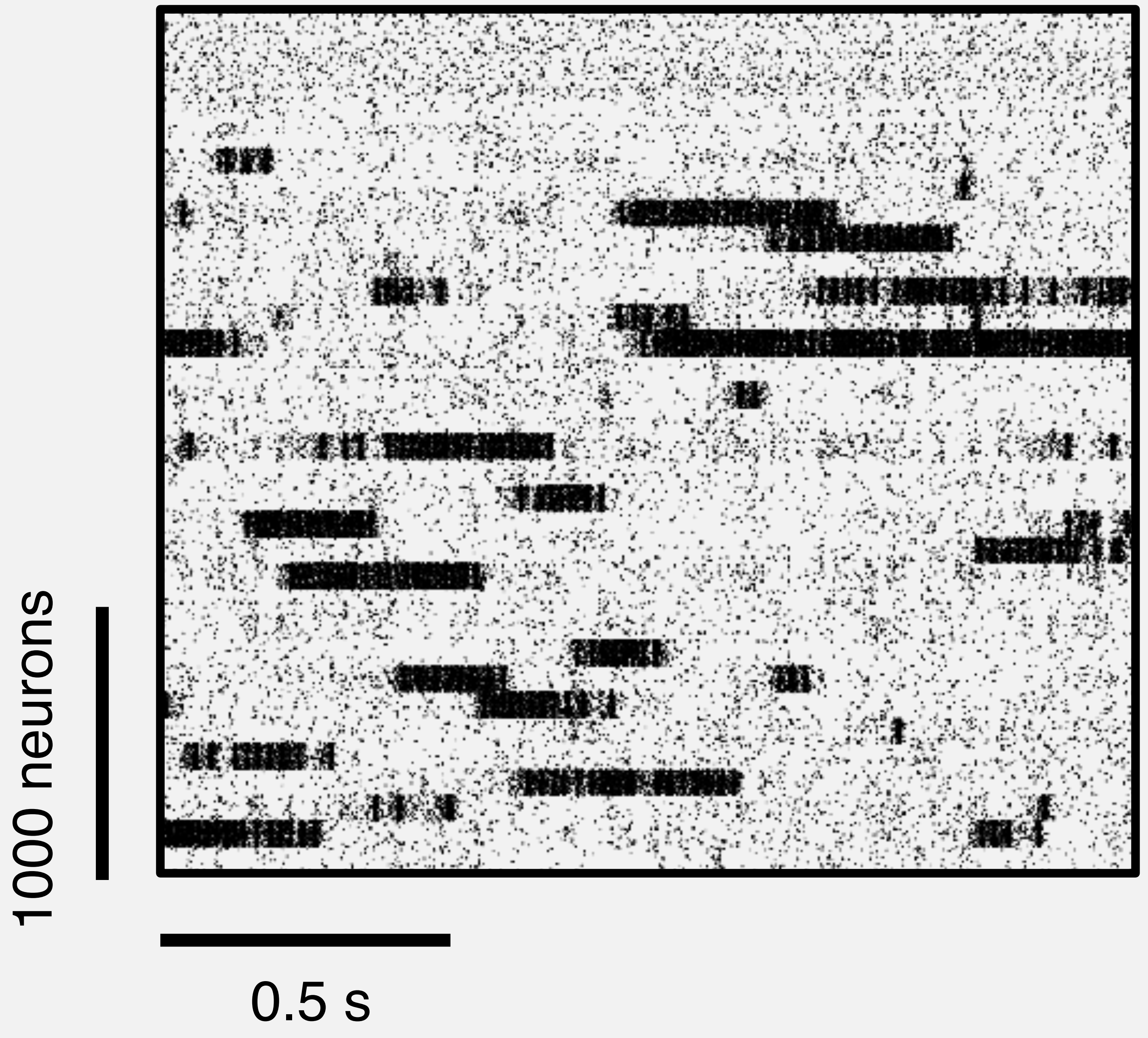



[Amit & Brunel, 1997]

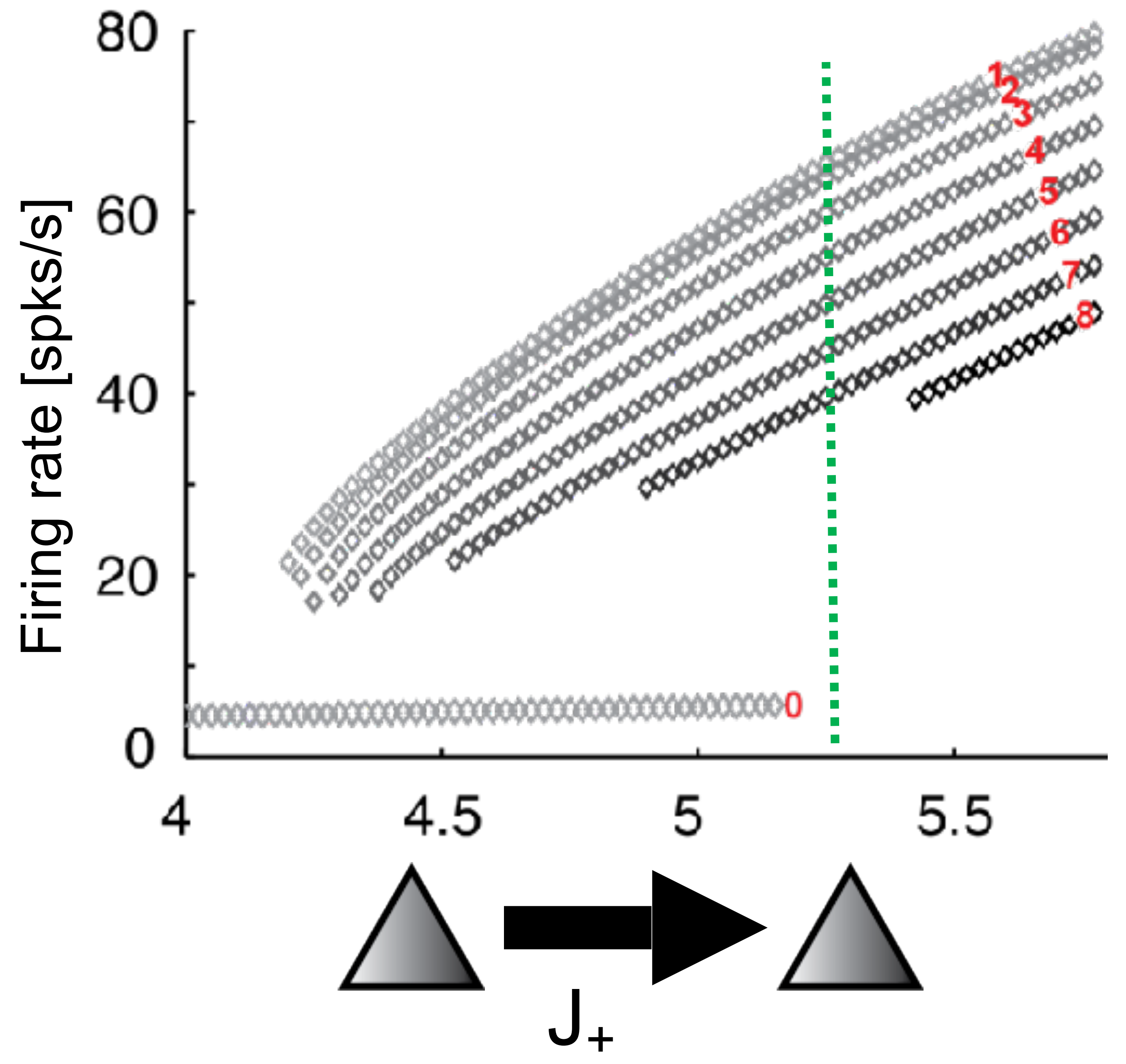
# Model – Spontaneous activity






$$r_i = f\left[\mu_i(\overrightarrow{r}), \sigma_i(\overrightarrow{r})\right]$$

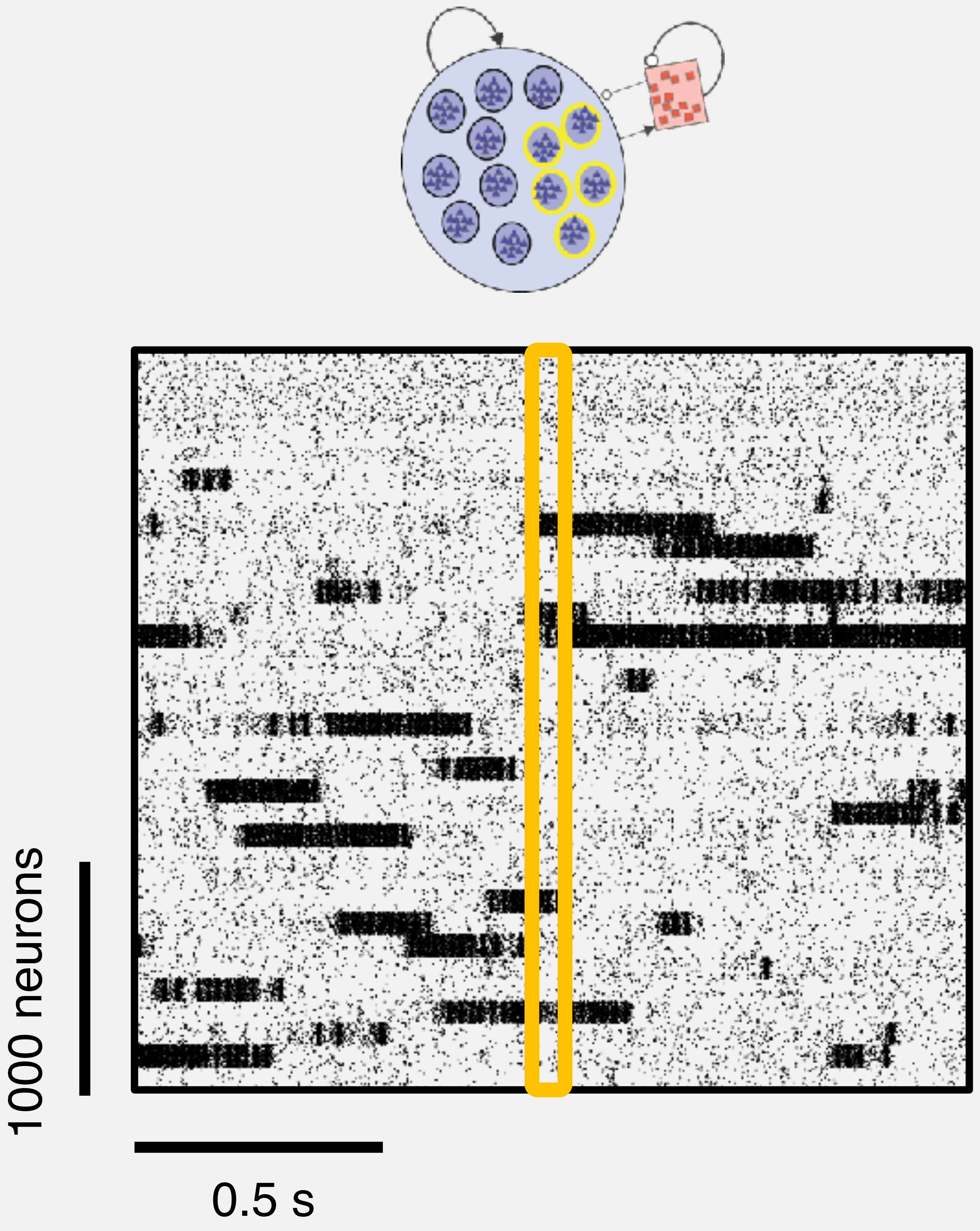
Attractor landscape



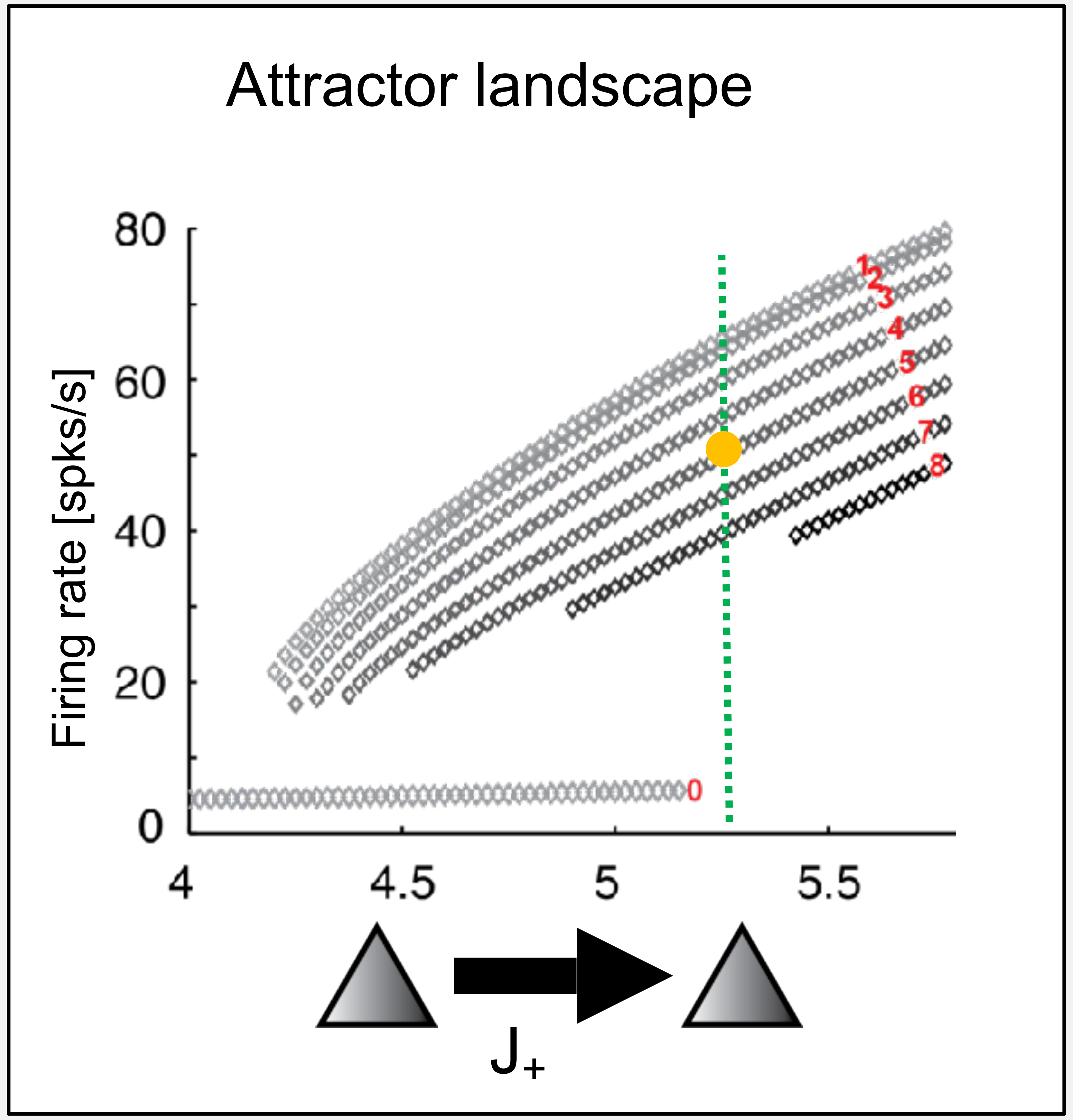

MFT details ▶

[Mazzucato et al., 2015, 2016]

# Model – Spontaneous activity



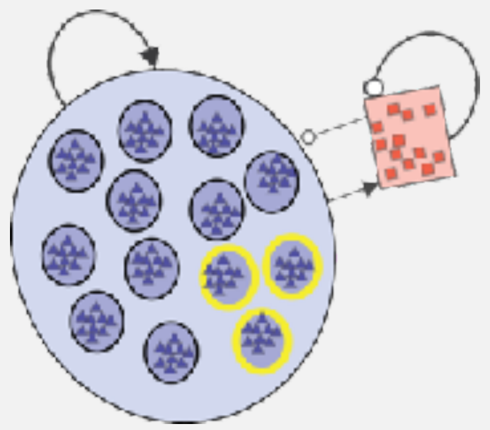

$$r_i = f\left[\mu_i(\overrightarrow{r}), \sigma_i(\overrightarrow{r})\right]$$

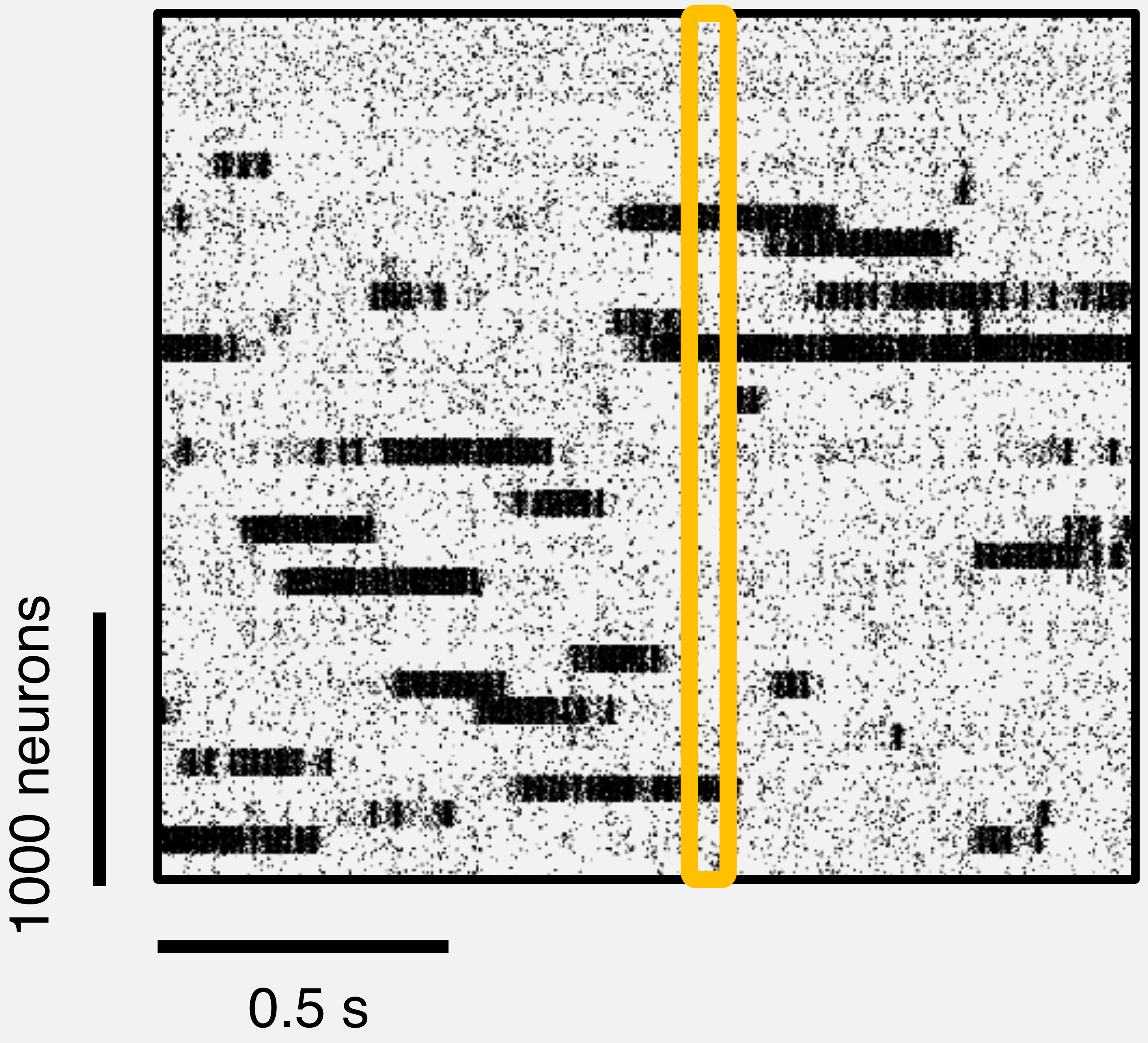



MFT details

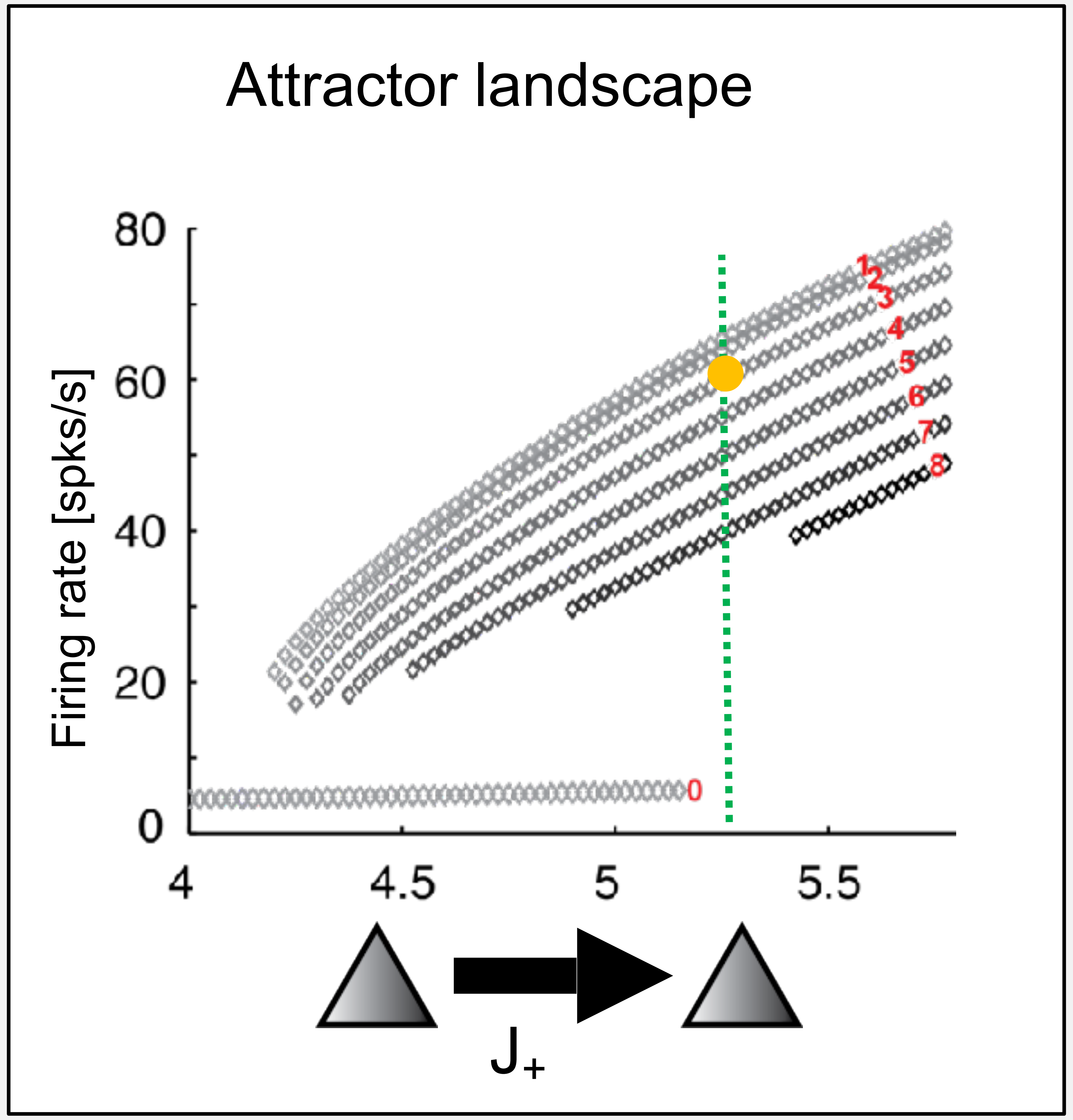
[Mazzucato et al., 2015, 2016]

# Model – Spontaneous activity






$$r_i = f\left[\mu_i(\overrightarrow{r}), \sigma_i(\overrightarrow{r})\right]$$

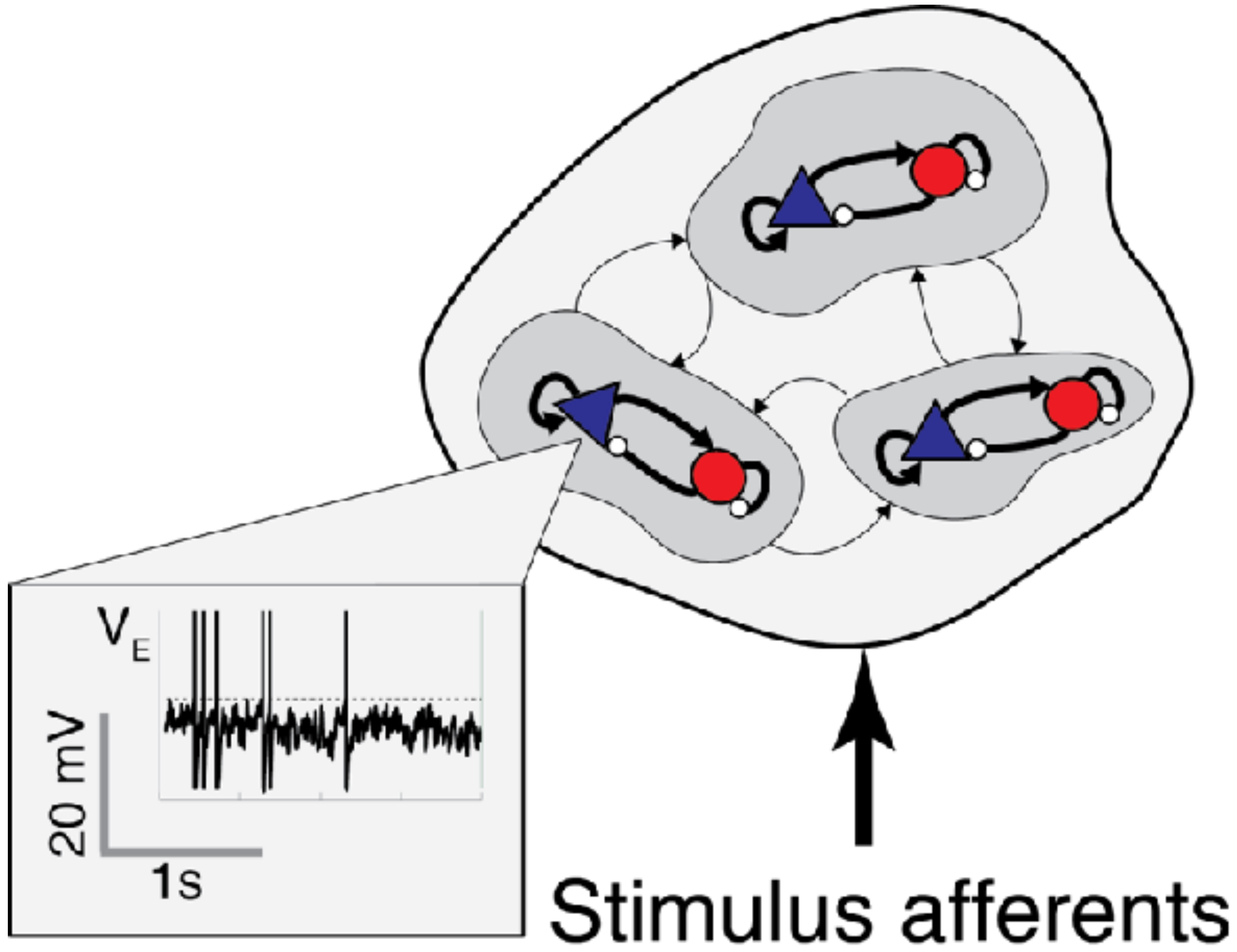



MFT details ▶

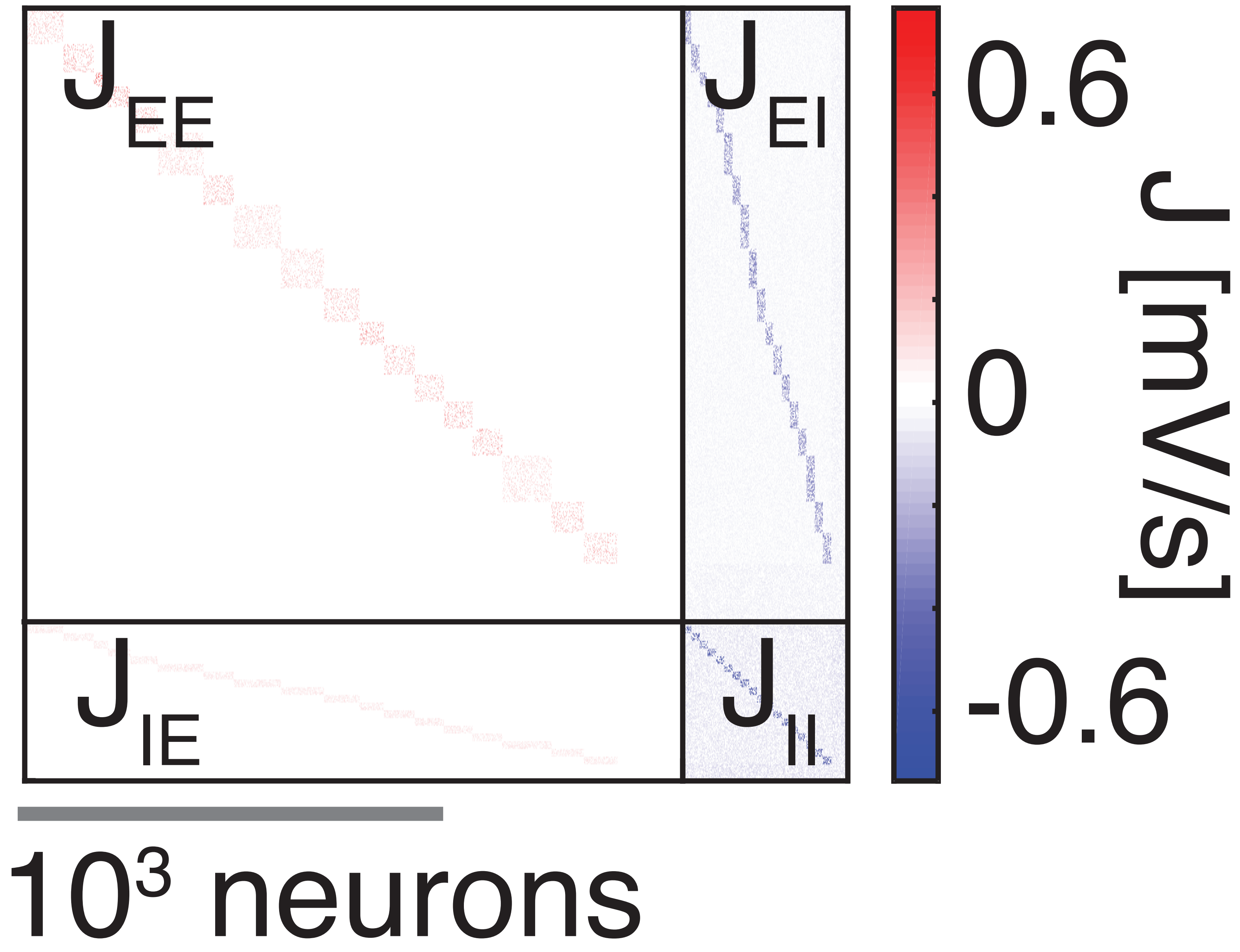
[Mazzucato et al., 2015, 2016]

# Attractor networks

E & I clusters

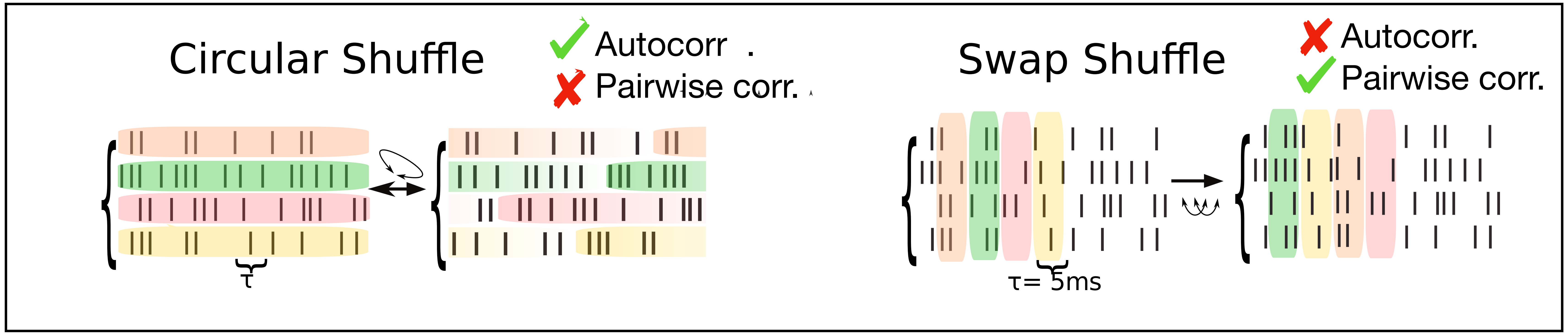



Synaptic weights J




[Deco & Hugues, 2012]
[Litwin-Kumar & Doiron, 2012]
[Mazzucato et al, 2019]
[Wyrick et al, 2021]

## Pattern sequences capture dynamics beyond auto- and pairwise




K-fold cross-validation: train HMM on K-1 folds

test on hold-out fold

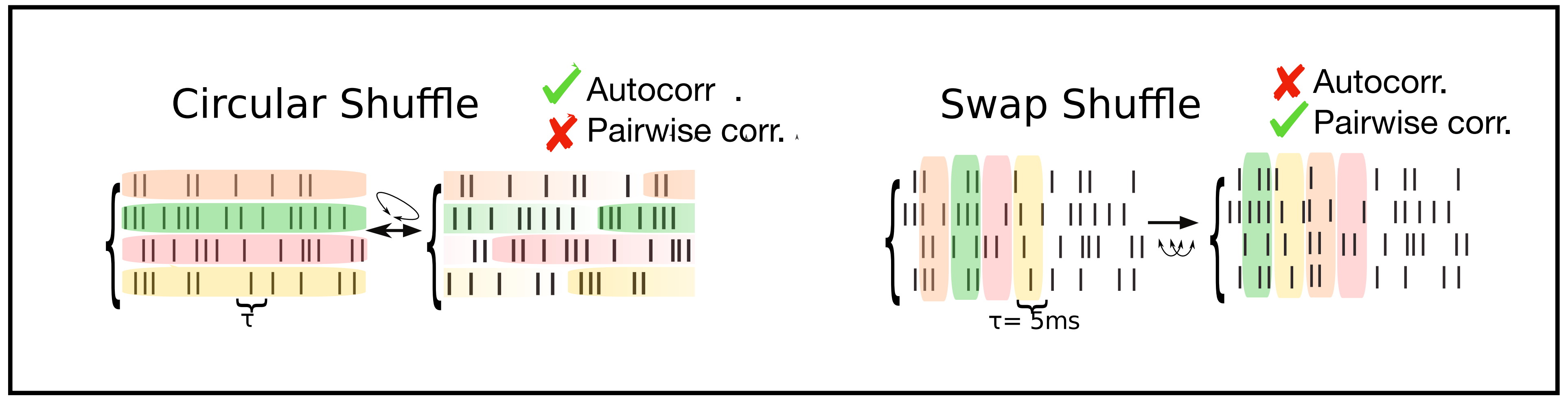
Data

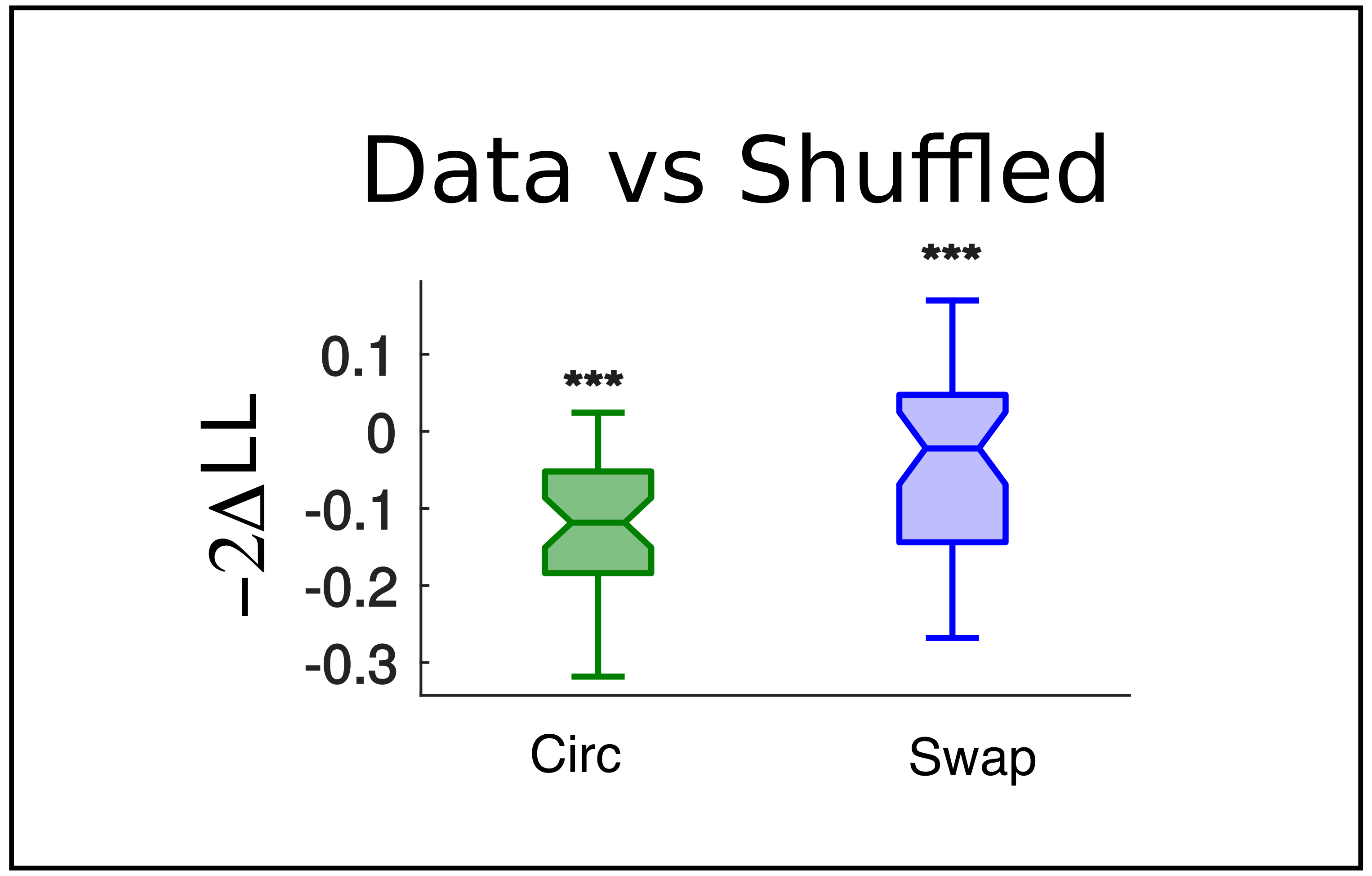
‘Circular’ shuffle

‘Swap’ shuffle

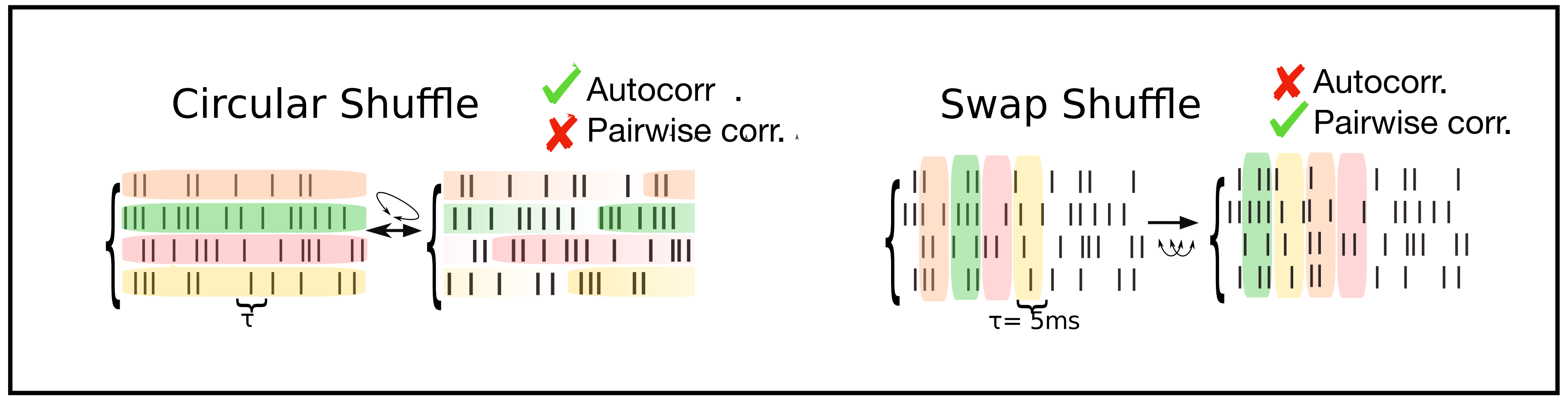
[Maboudi et al., 2018]

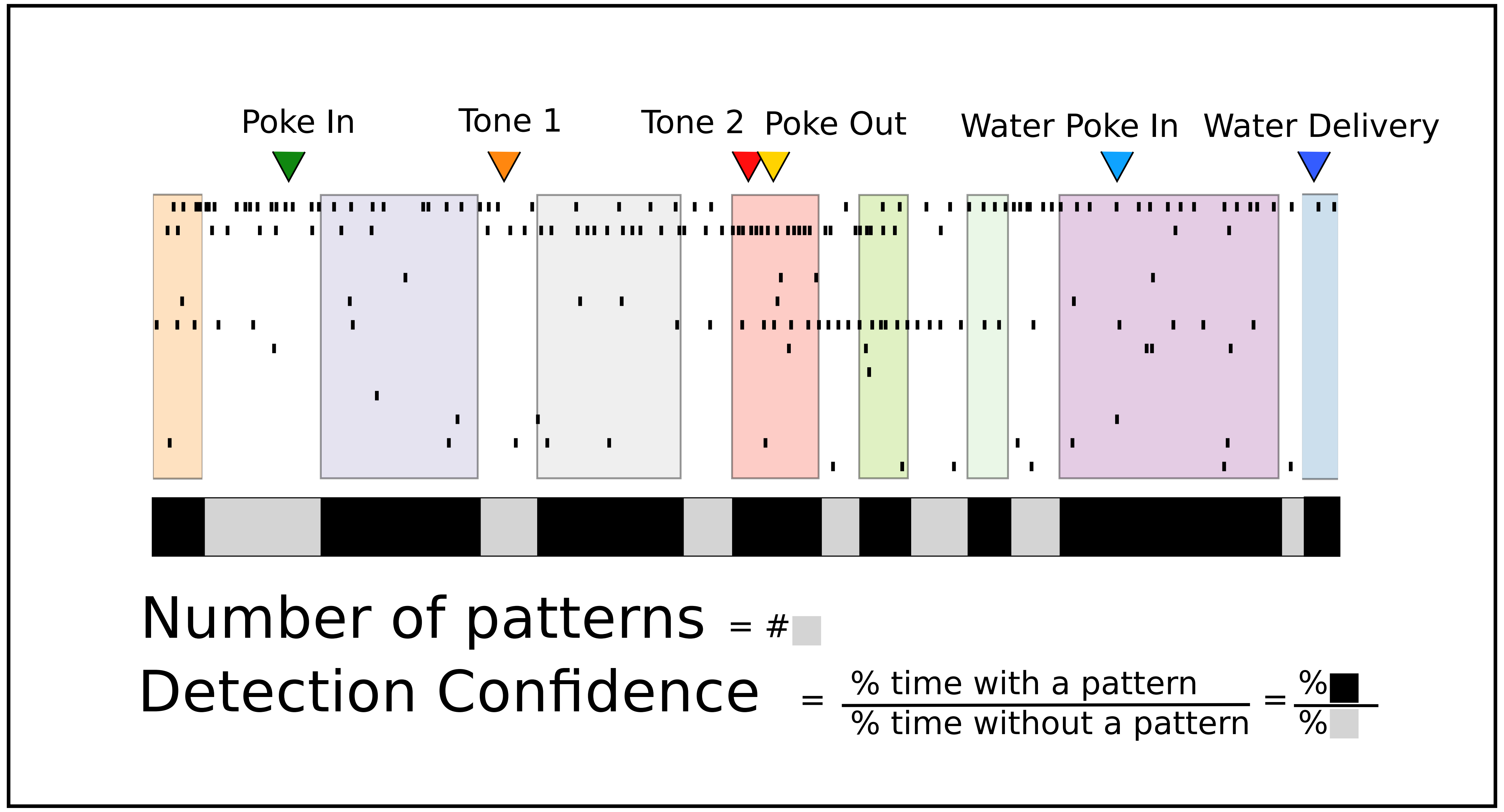
# Pattern sequences capture dynamics beyond auto- and pairwise



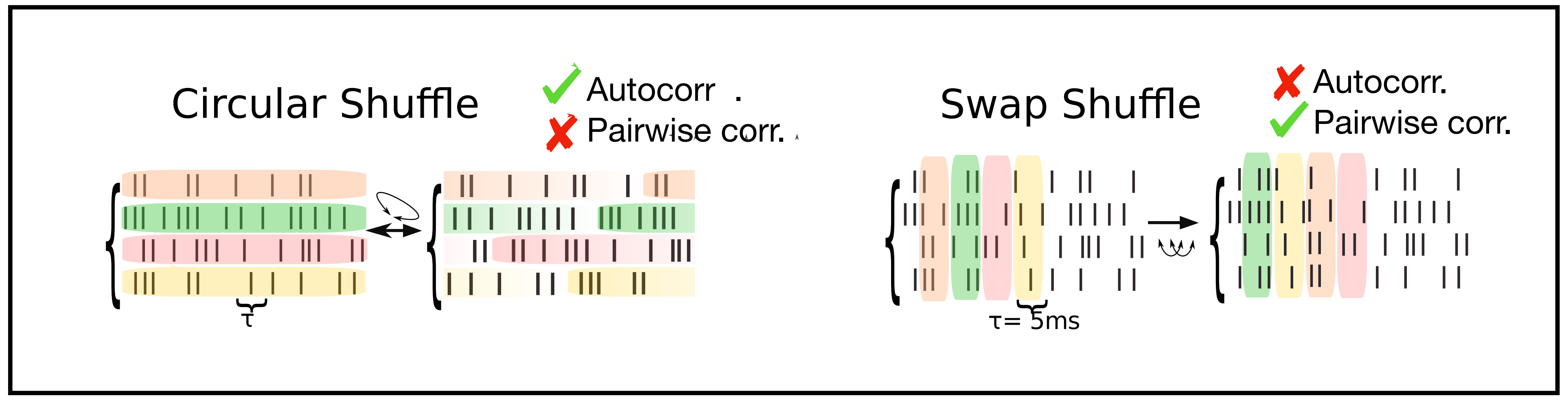



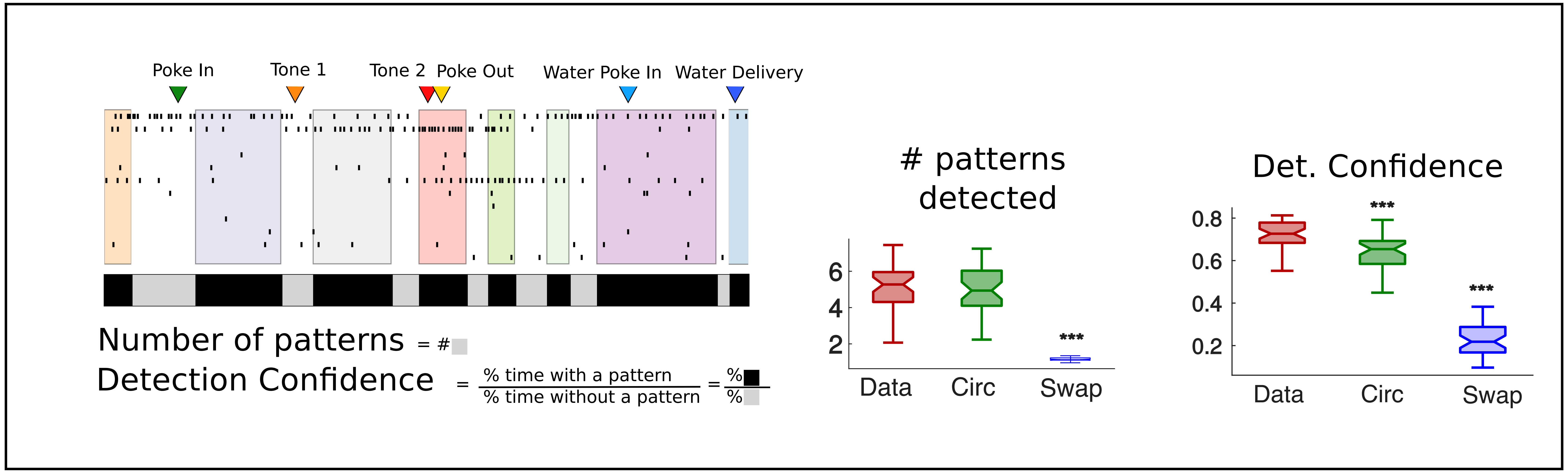
# Pattern sequences capture dynamics beyond auto- and pairwise



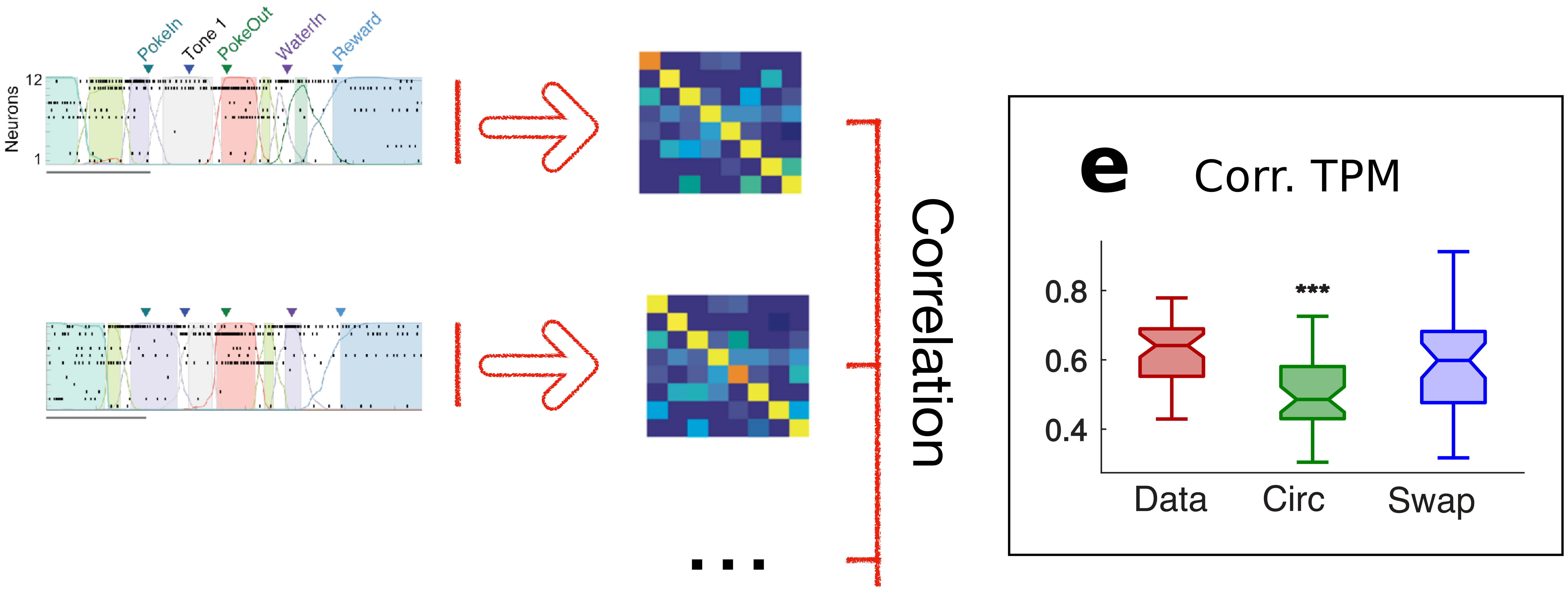
# Pattern sequences capture dynamics beyond auto- and pairwise



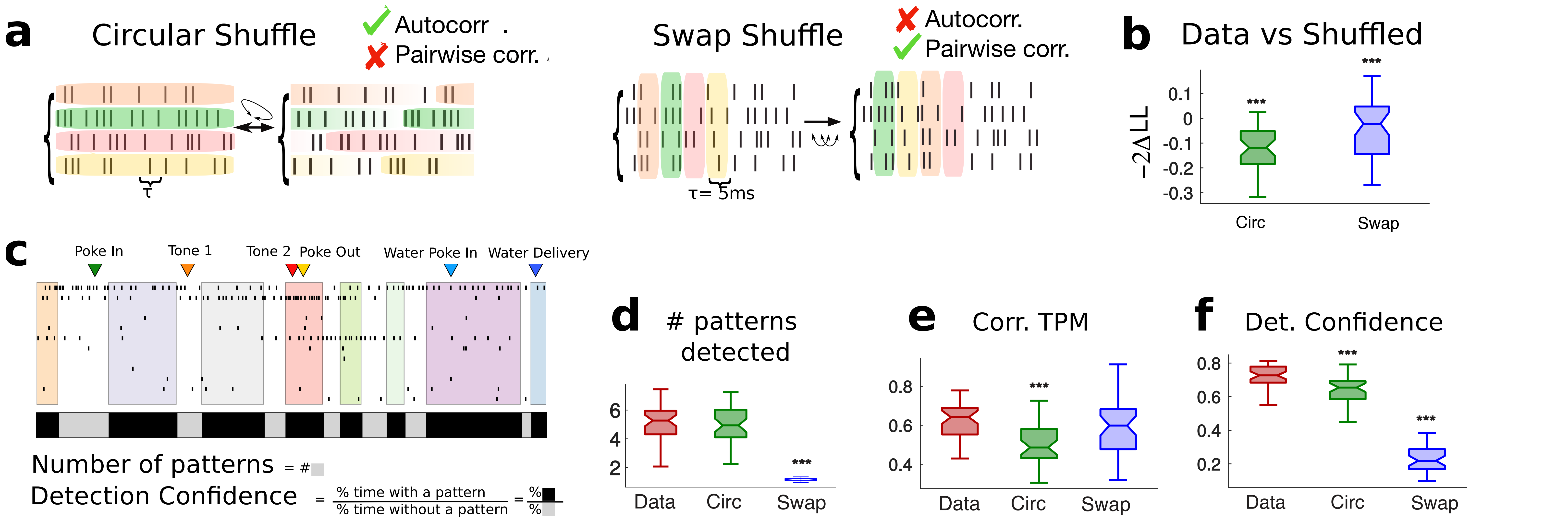
# Pattern sequences capture dynamics beyond auto- and pairwise

## Highly reliable sequences



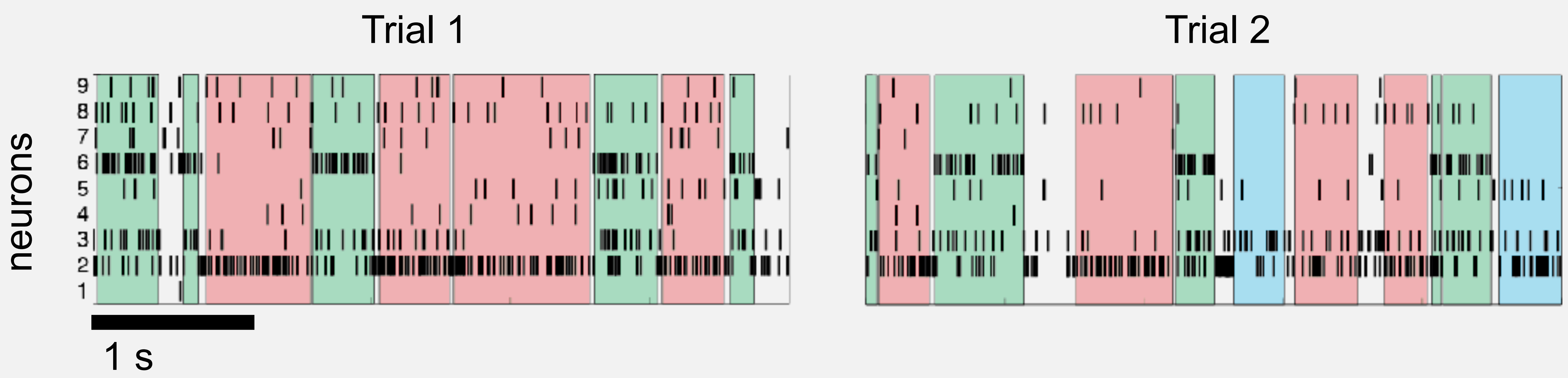
# Pattern sequences capture dynamics beyond auto- and pairwise

41 sessions - 7 animals

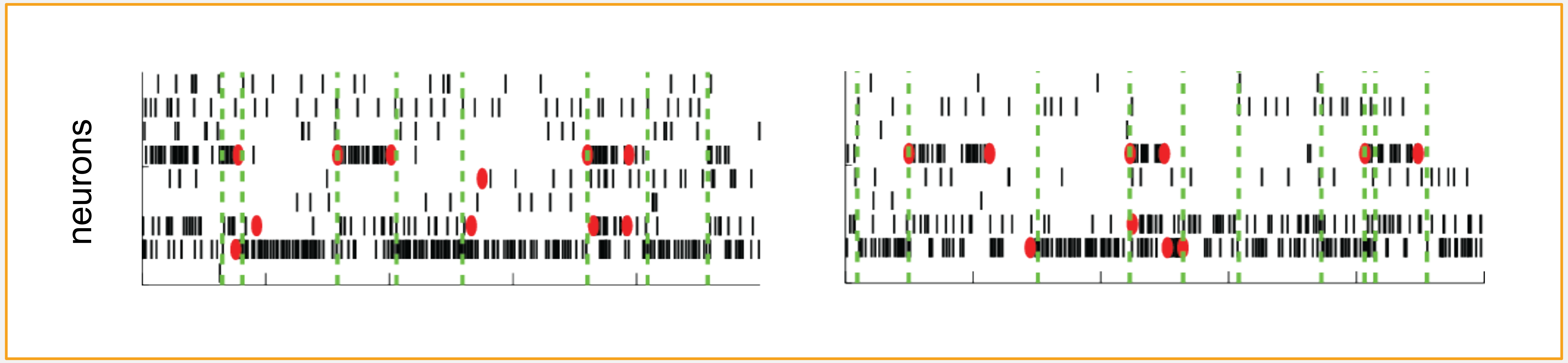


# Data – Spontaneous activity

- Post hoc sanity check tests



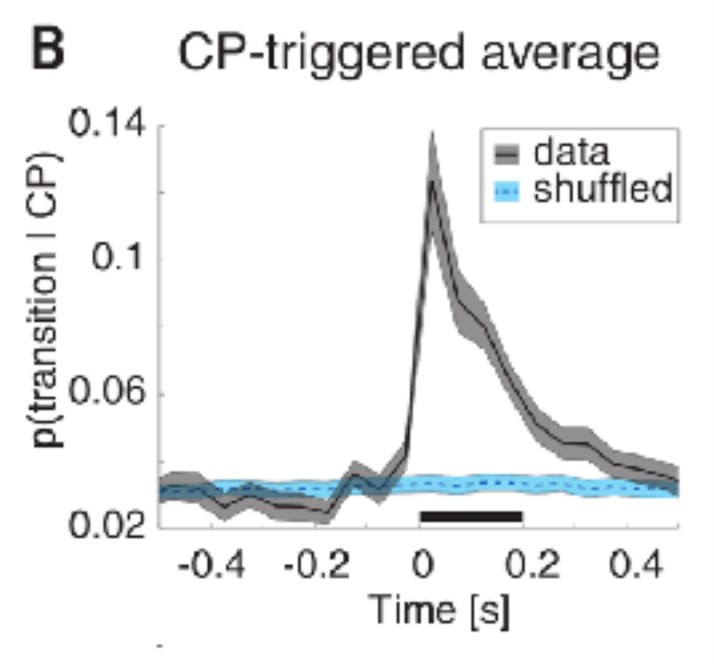

Compare single units Change-Points with HMM state transitions

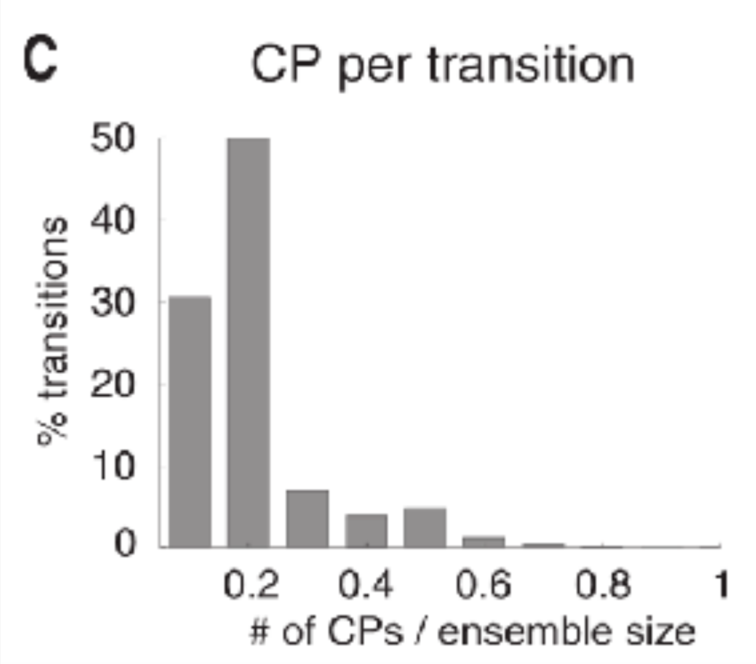



[Mazzucato et al., 2015]

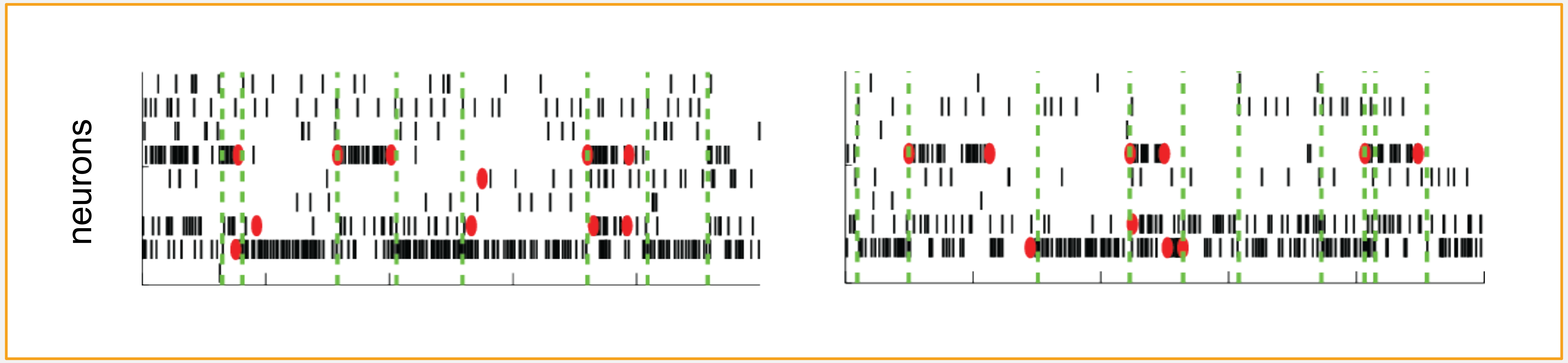
# Data – Spontaneous activity

- Post hoc sanity check tests




Compare single units Change-Points with HMM state transitions




[Mazzucato et al., 2015]

# Data – Spontaneous activity

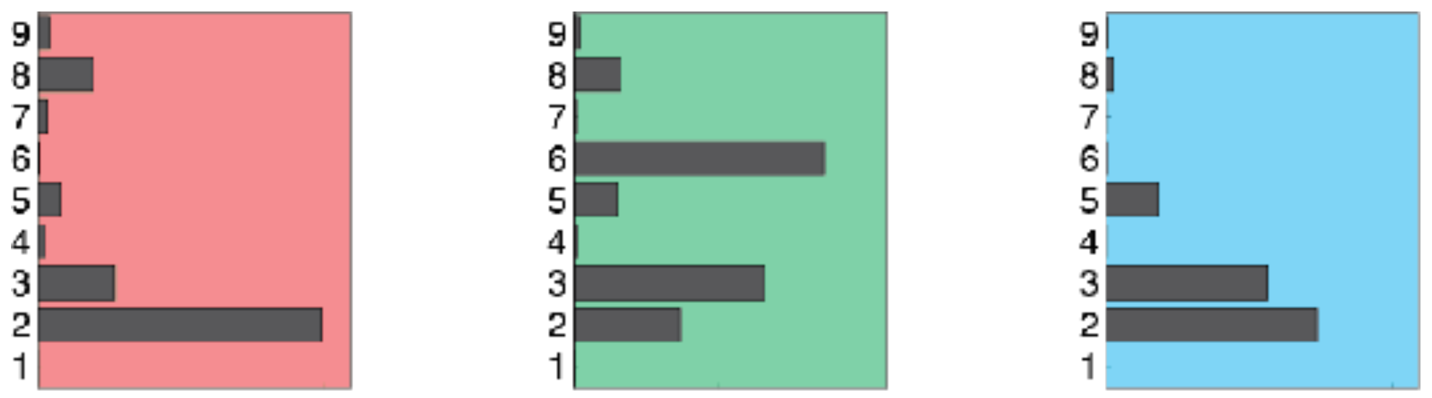
- Firing rate across states


1. Firing rate estimate *r* for neuron *i* in state *q* in single trial: *M-step of EM fit*

observations

$$r_i(q) = -\frac{1}{\Delta t} \ln \left( 1 - \frac{\sum_{t=1}^{T} p(S_t = q) y_i(t)}{\sum_{t=1}^{T} p(S_t = q)} \right)$$

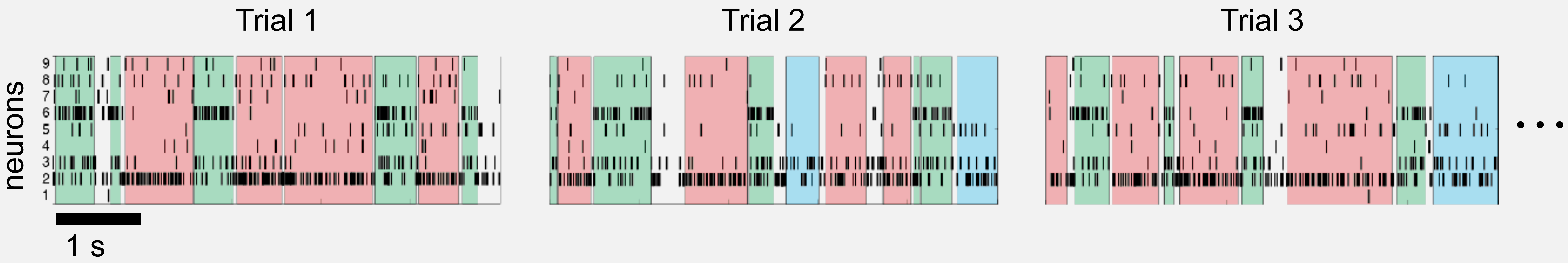
posterior prob. of state q

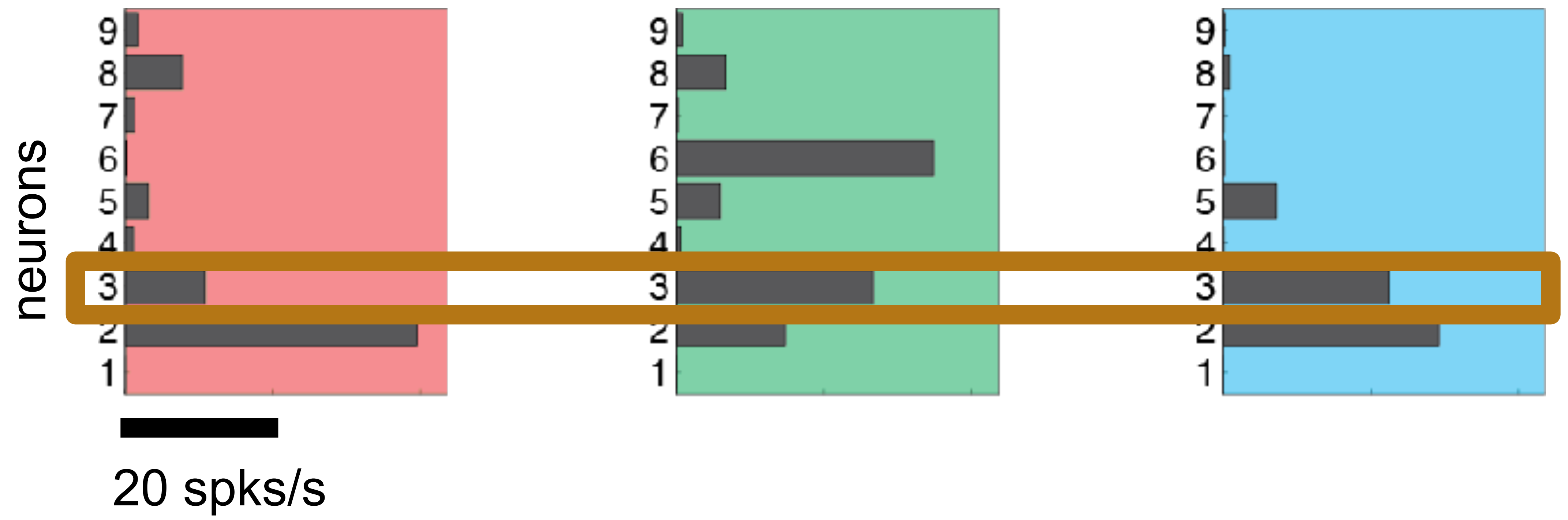


[Mazzucato et al., 2015]

# Data – Spontaneous activity

- Firing rate across states




Neuron #3: two rates across states→ **“bistable”**

[Mazzucato et al., 2015]

# Data – Spontaneous activity

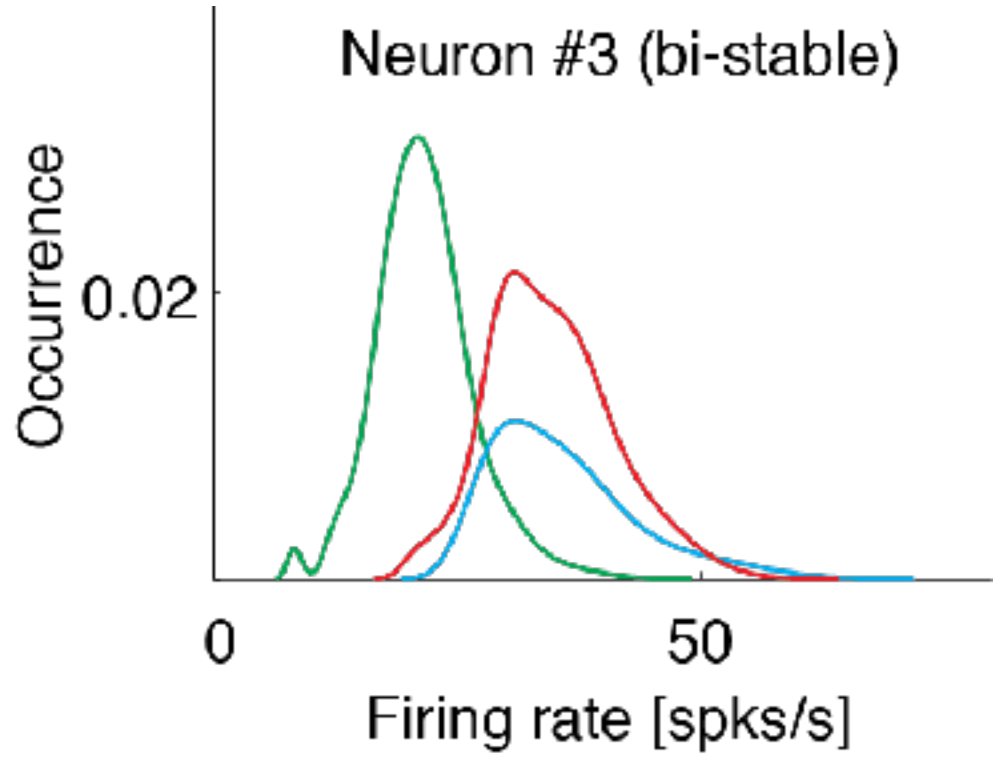
- Firing rate across states


1. Firing rate estimate *r* for neuron *i* in state *q* in single trial: *M-step of EM fit*

observations

$$r_i(q) = -\frac{1}{\Delta t} \ln \left( 1 - \frac{\sum_{t=1}^{T} p(S_t = q) y_i(t)}{\sum_{t=1}^{T} p(S_t = q)} \right)$$

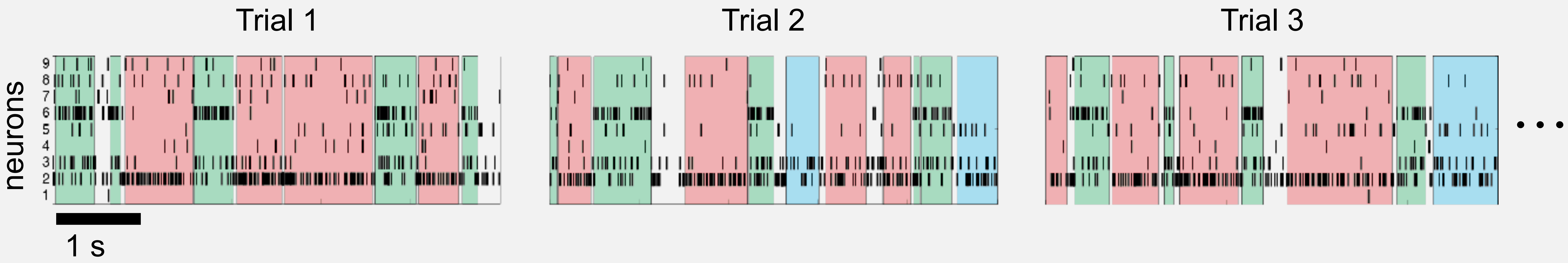
posterior prob. of state q

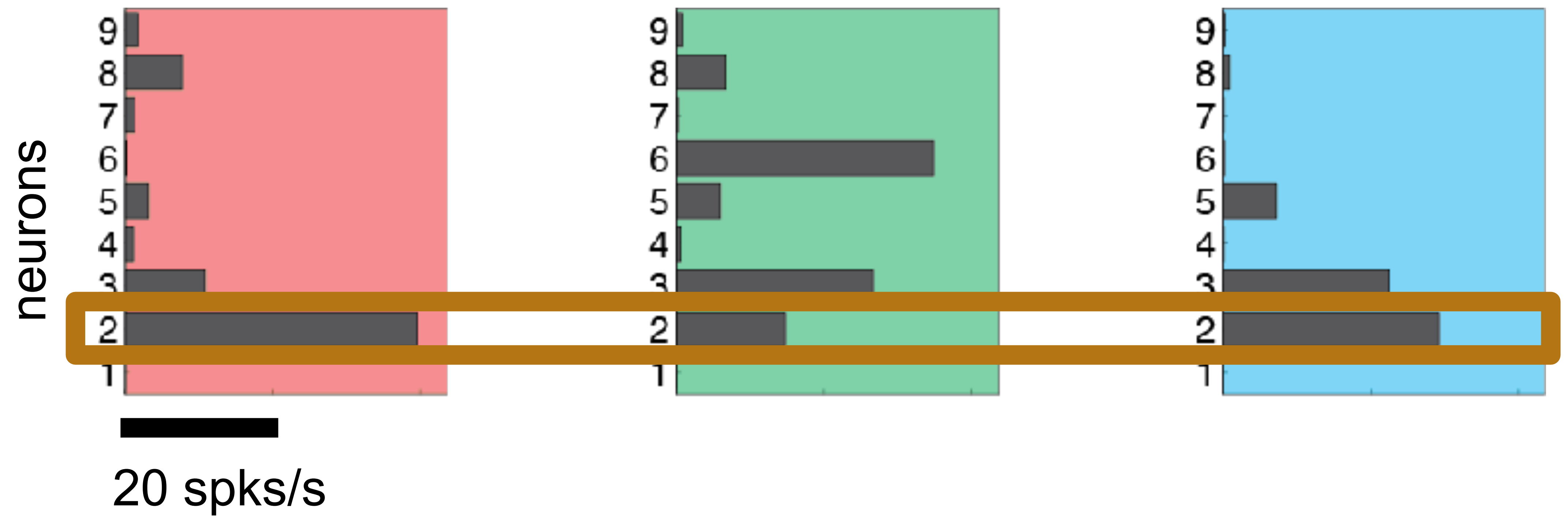



[Mazzucato et al., 2015]

# Data – Spontaneous activity

- Firing rate across states



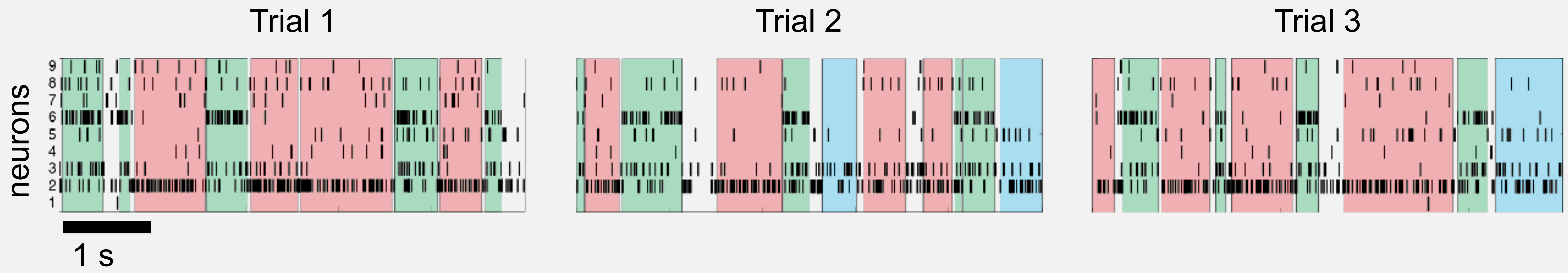




Neuron #2: three rates across states→ **“multistable”**

[Mazzucato et al., 2015]

# Data – Spontaneous activity

- Firing rate across states

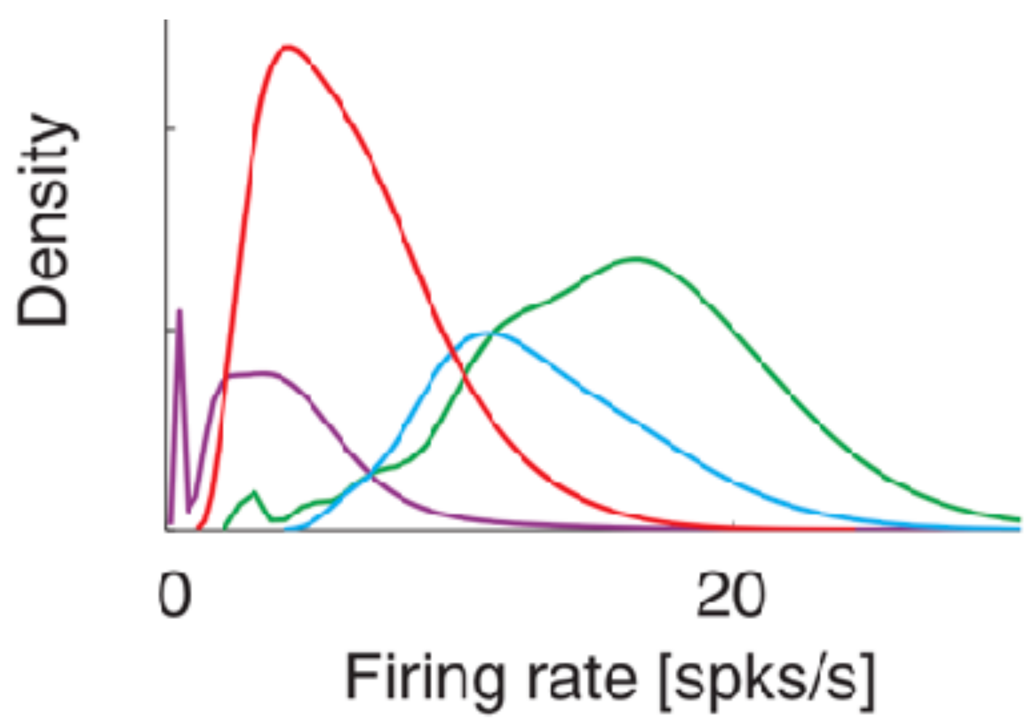



1. Firing rate estimate *r* for neuron *i* in state *q* in single trial: *M-step of EM fit*

observations

$$r_i(q) = -\frac{1}{\Delta t}\ln\left(1 - \frac{\sum_{t=1}^{T} p(S_t = q)\, y_i(t)}{\sum_{t=1}^{T} p(S_t = q)}\right)$$

posterior prob.
of state q

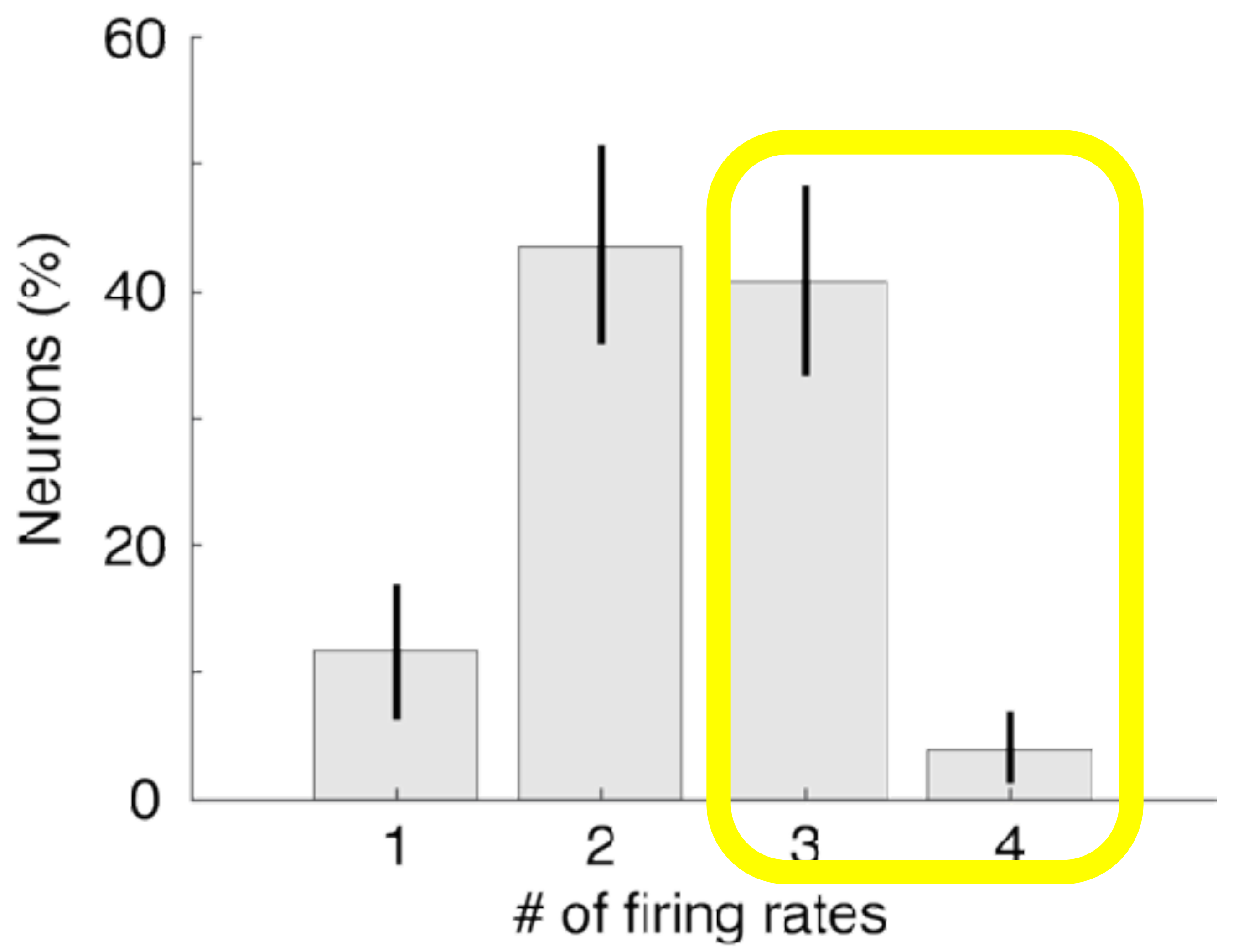



[Mazzucato et al., 2015]

# Data – Spontaneous activity

## Multistability

42% of neurons have 3 or more firing rates across states:

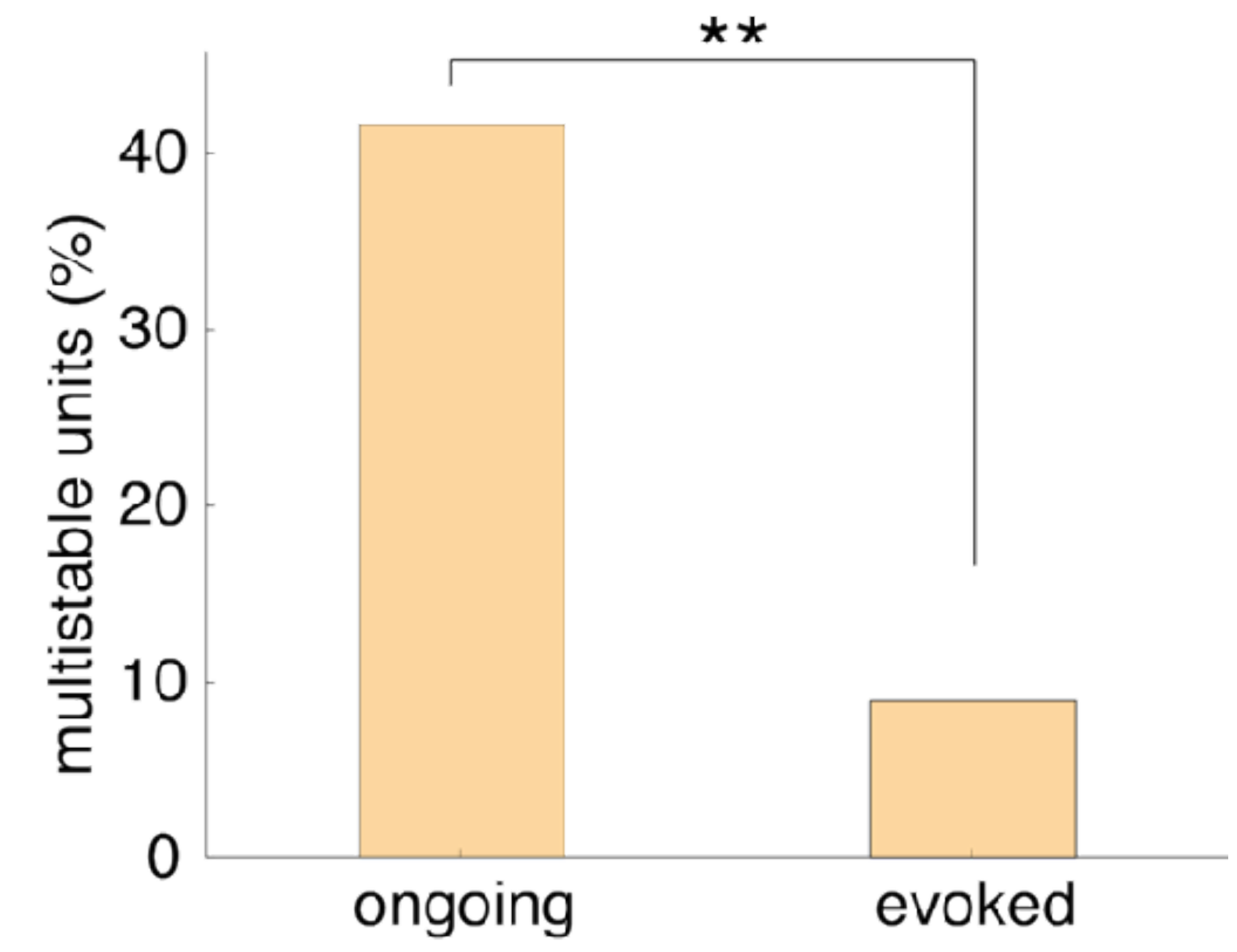



[Mazzucato et al., 2015]

# Stimuli reduce multistability

- Only **bistable** neurons are observed in response to stimulation

DATA




[Mazzucato et al., 2015]

# Circuit models of attractor dynamics

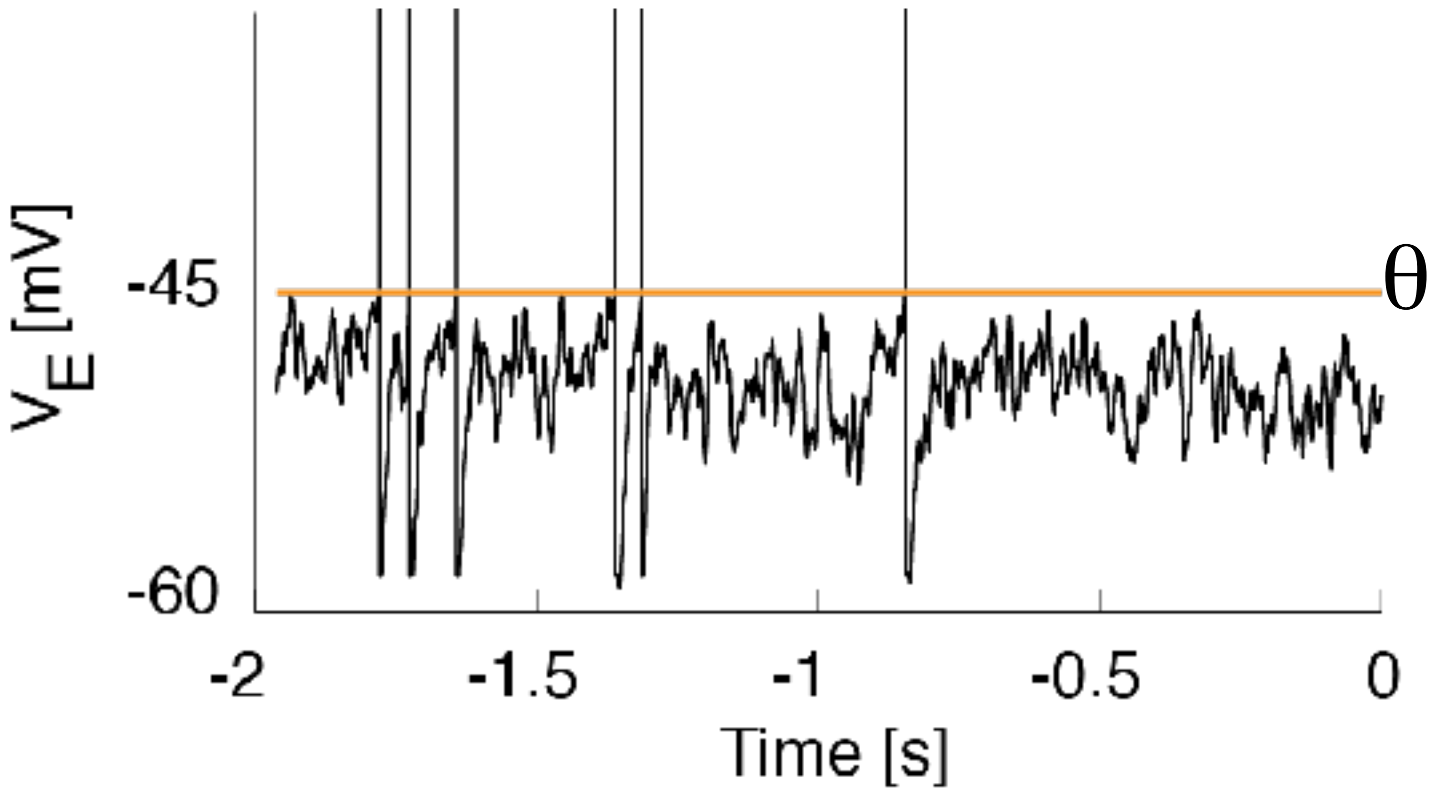
# Model – Neuron

PSC: $$\tau_s \frac{dI_i}{dt} = -I_i + I_{ext}(t) + \sum_j^{pre} J_{ij} \sum_n^{spikes} \delta(t - t_j^{(n)})$$

Potential: $$\tau_m \frac{dV_i}{dt} = V_L - V_i + I_i(t)$$

b.c.: $V(t^*) = \theta$

→ spike




[van Vreeswijk & Sompolinsky, 1996]
[Amit & Brunel, 1997]
[Haider et al., 2006]
[Renart et al., 2007]

# Model – Neuron

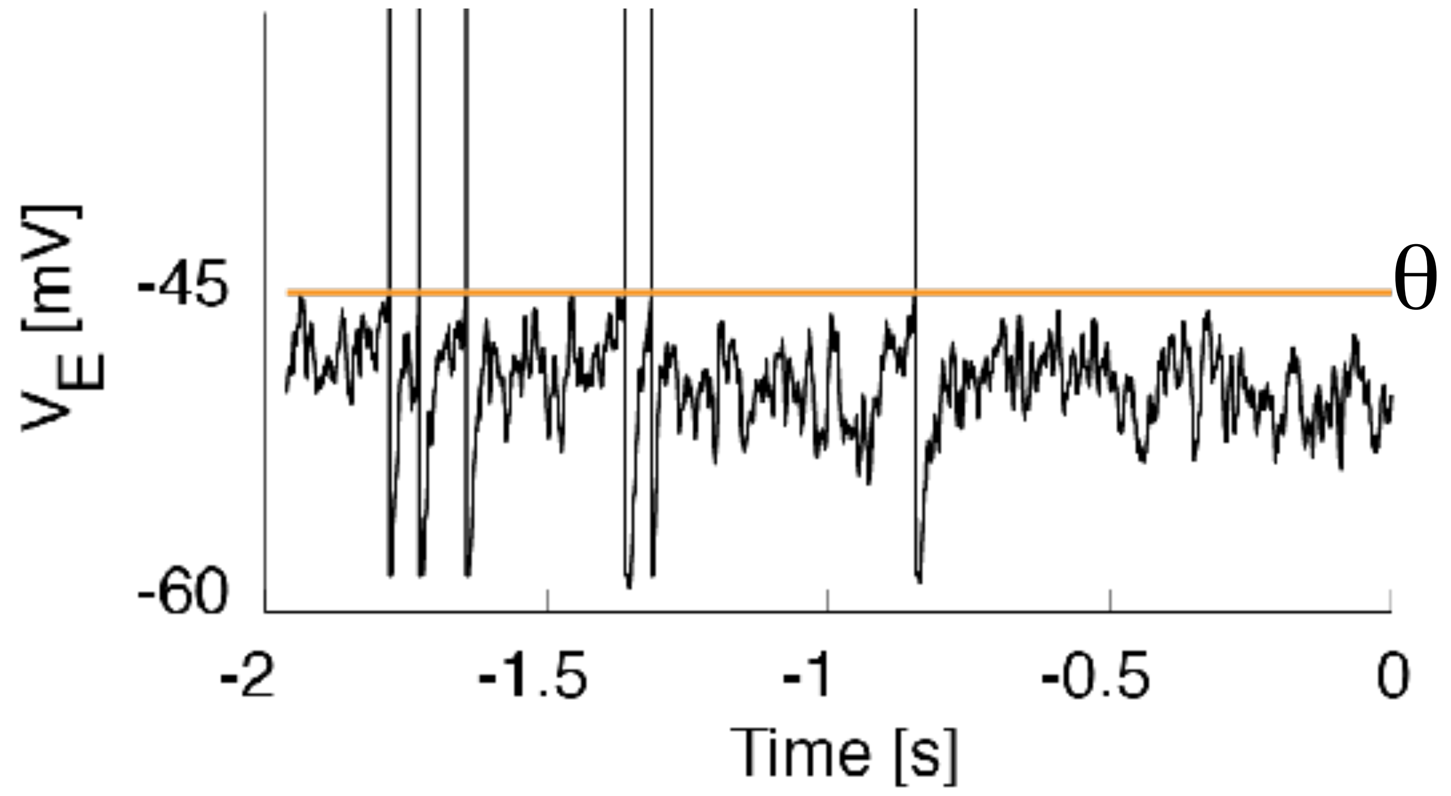
Ornstein–Uhlenbeck process

PSC: $$\tau_s \frac{dI_i}{dt} = -I_i + I_{ext}(t) + \boxed{\sum_j^{pre} J_{ij} \sum_n^{spikes} \delta(t - t_j^{(n)})}$$

Potential: $$\tau_m \frac{dV_i}{dt} = V_L - V_i + I_i(t)$$

b.c.: $V(t^*) = \theta$

→ spike




[Amit & Brunel, 1997]

# Model – Neuron

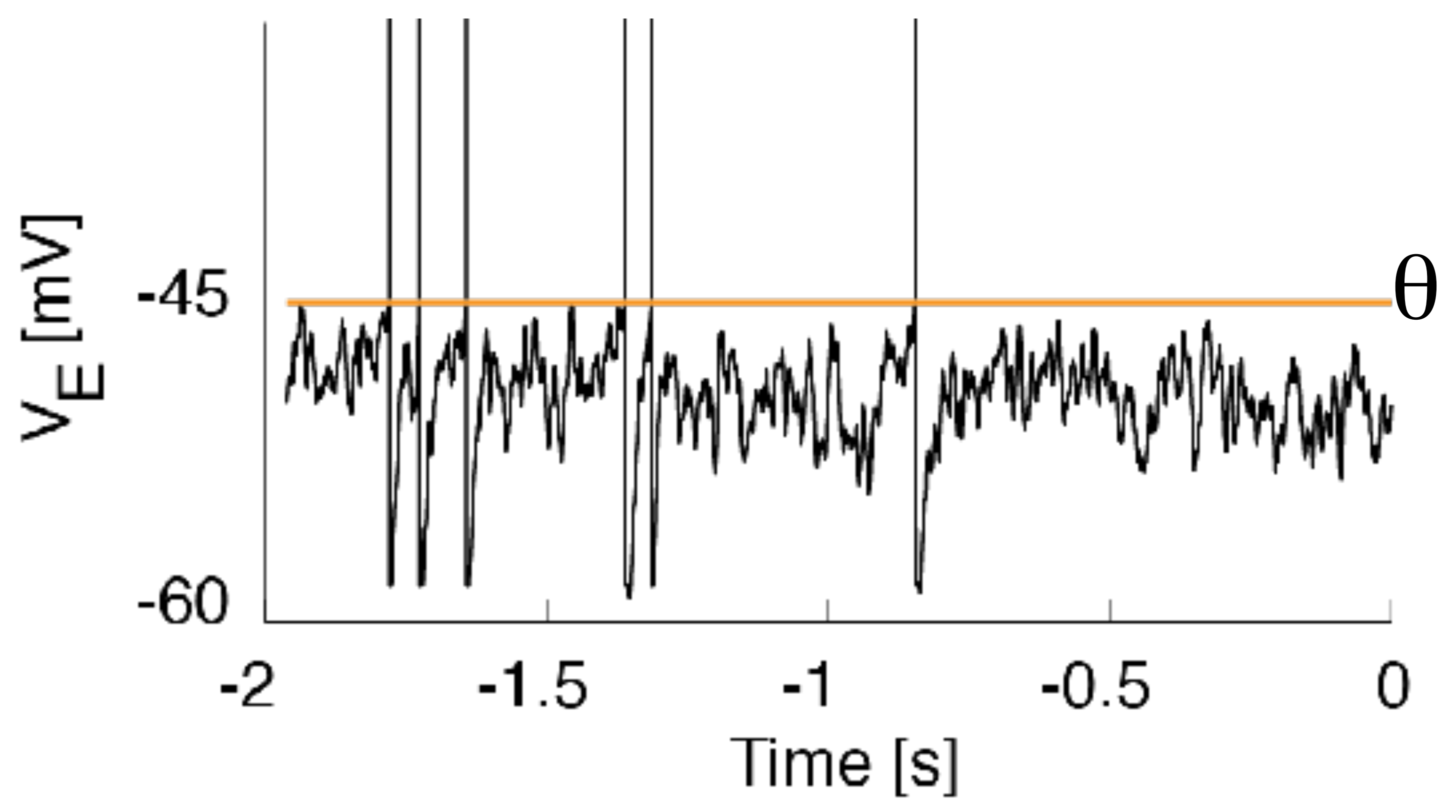
Ornstein–Uhlenbeck process

PSC: $$\tau_s \frac{dI_i}{dt} = -I_i + I_{ext}(t) + \boxed{\mu + \sigma\eta(t)}$$

Potential: $$\tau_m \frac{dV_i}{dt} = V_L - V_i + I_i(t)$$

b.c.: $V(t^*) = \theta$

→ spike



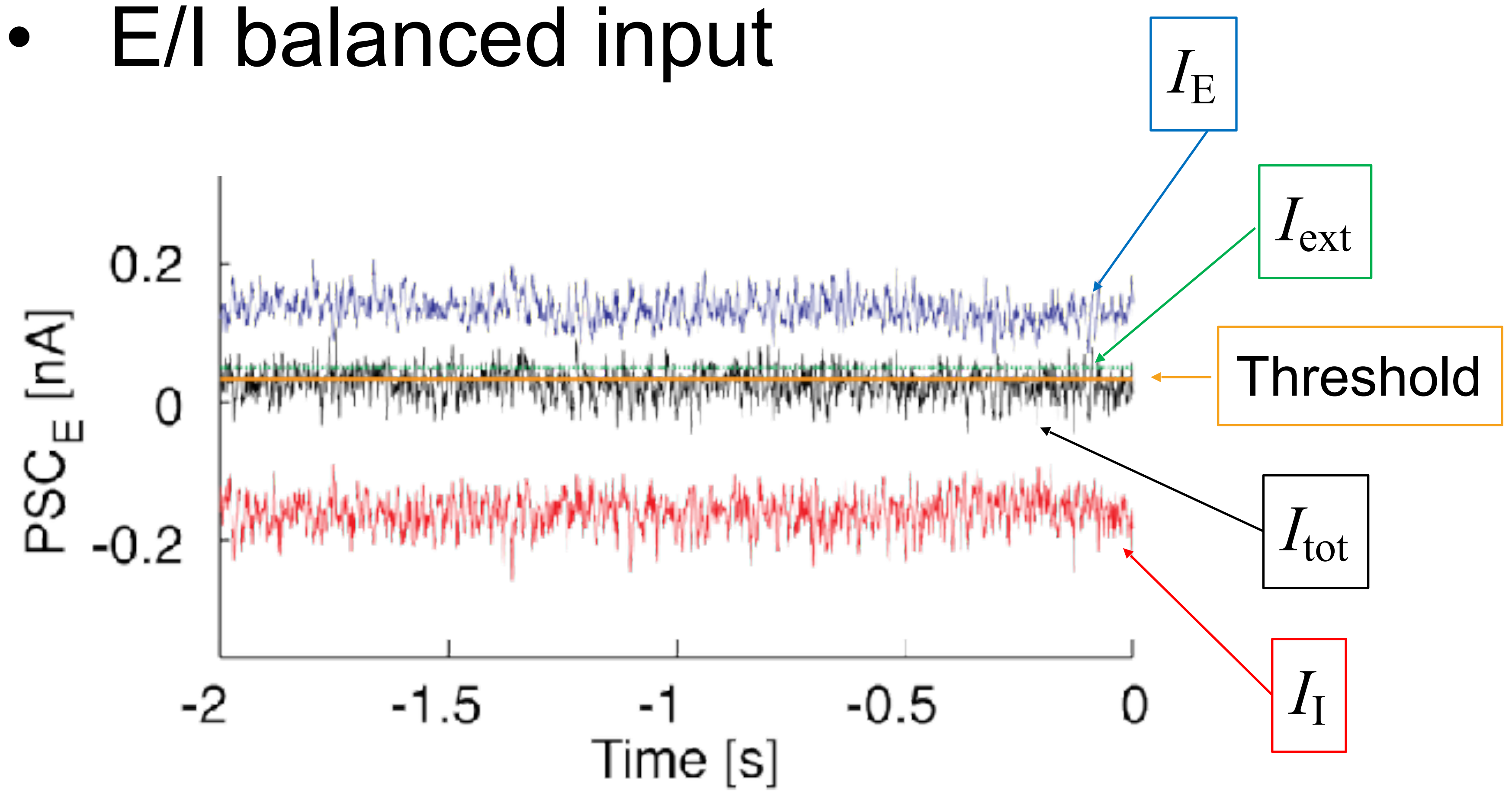

[Amit & Brunel, 1997]

# Model – Spontaneous activity

## Leaky Integrate–and–Fire neuron

$$\tau_m \frac{dV}{dt} = V_L - V + (I_E + I_I + I_{ext})$$

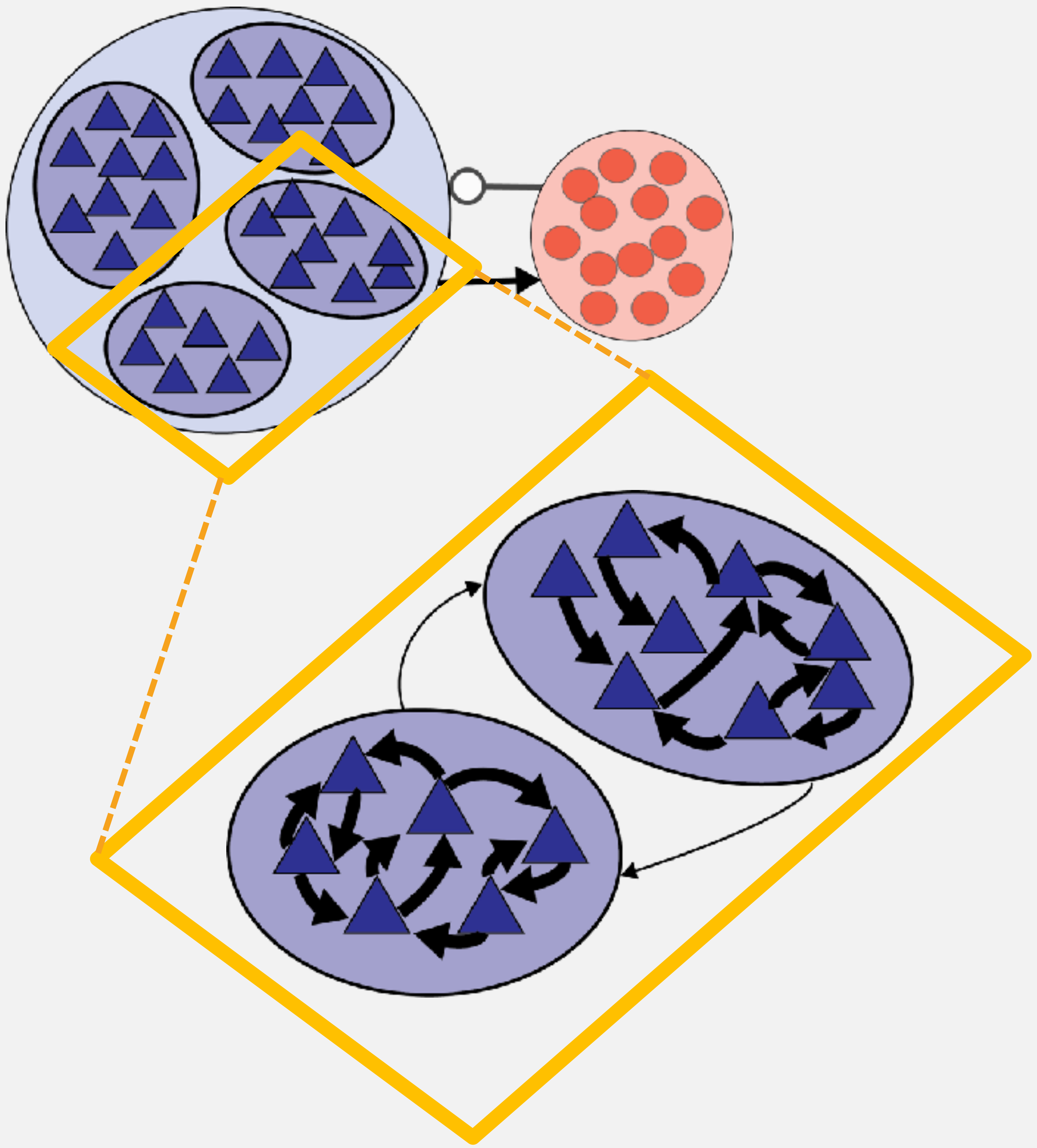
- E/I balanced input





[Haider et al., 2006]
[van Vreeswijk & Sompolinsky, 1996]
[Renart et al., 2007]

# Model – Network architecture



Recurrent E → E synaptic weights:

same cluster

$J_+ > 1$

different clusters

$J_- < 1$

[Amit & Brunel, 1997]
[Deco & Hugues, 2012]
[Litwin–Kumar & Doiron, 2012]
[Mazzucato et al., 2015]