Machine Learning Methods for Neural Data Analysis

(Switching) Linear Dynamical Systems

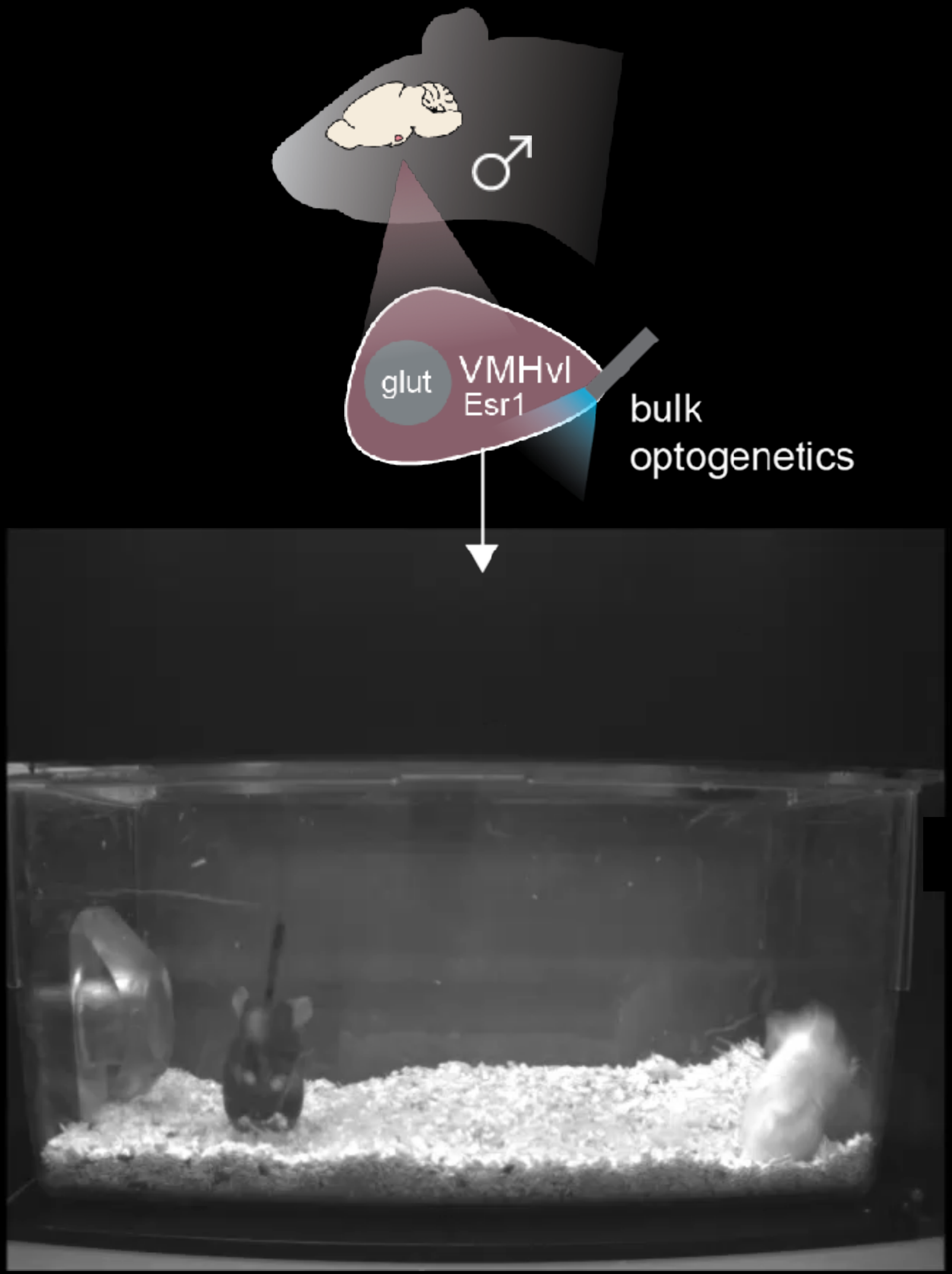
Scott Linderman

STATS 220/320 (NBIO220, CS339N).

Agenda

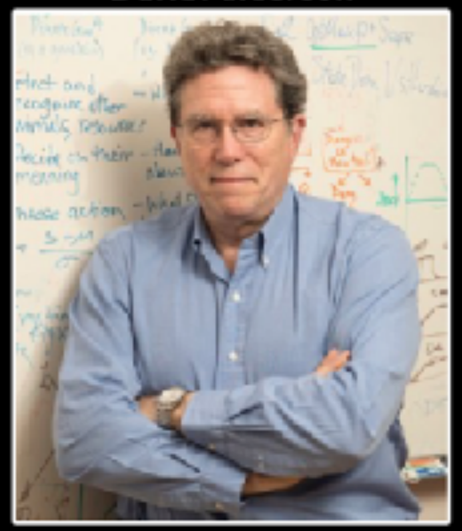
- Motivating example
- Linear dynamical systems (LDS)
- Switching linear dynamical systems (SLDS)
- Results of a recent scientific study

Optogenetic activation of neurons in the hypothalamus elicits attack behavior



Adi Nair

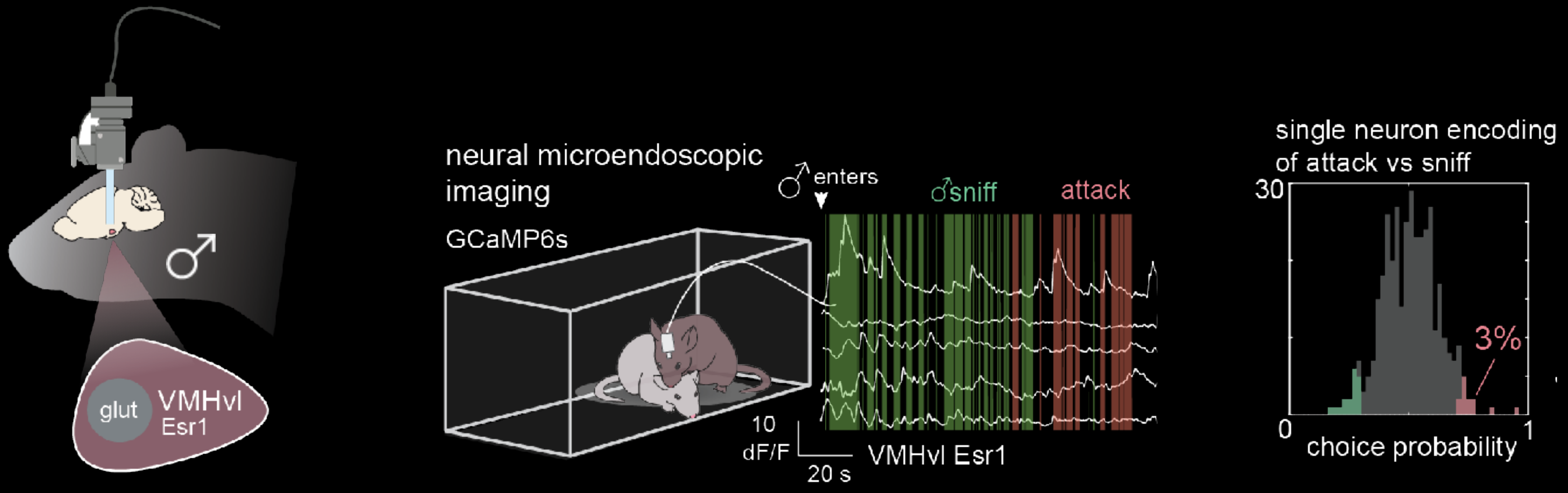
David Anderson



Ann Kennedy

Lee et al. (Nature, 2014)

Miniscope imaging in VMHvl during spontaneous aggression shows mixed selectivity



Most neurons in VMHvl are tuned to intruder sex and are active during both sniffing and attack.

Remedios, Kennedy et al. (Nature, 2019)
Karigo et al (Nature, 2021)

Miniscope imaging in VMHvl during spontaneous aggression shows mixed selectivity

Hypothesis

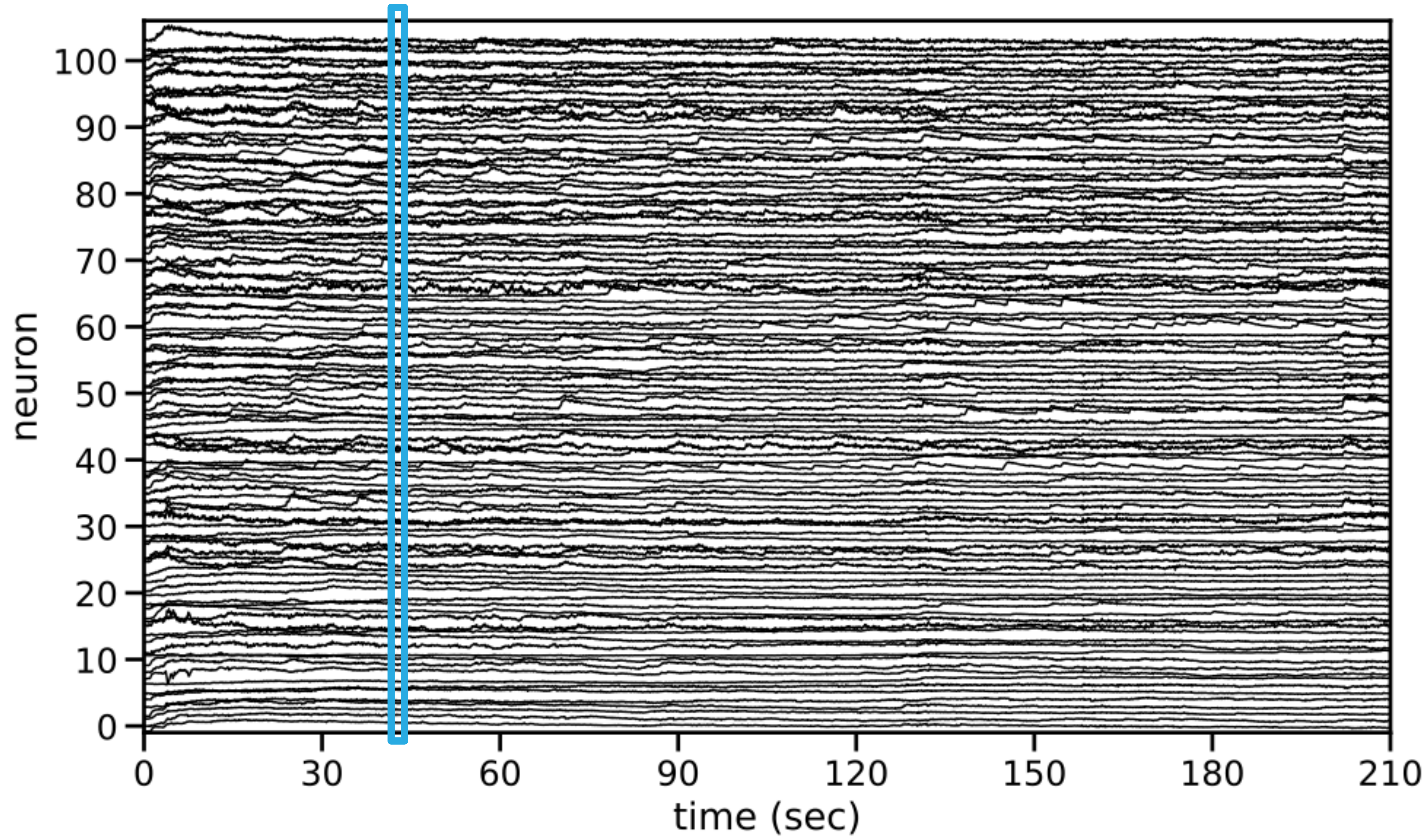
An **internal state of aggressiveness** is encoded in the **collective activity of neurons** in the VMHvl.

Most attack.

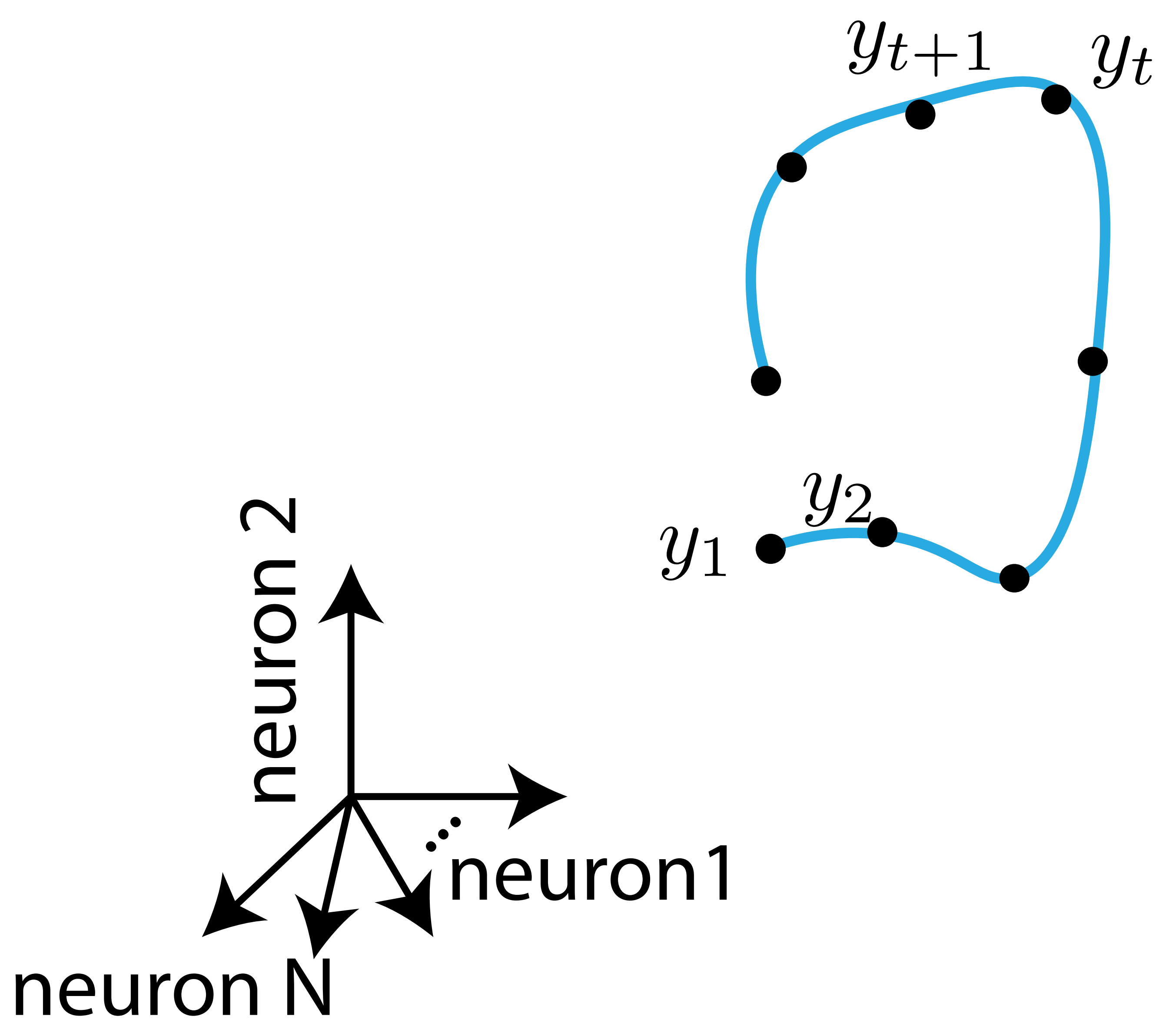
Remedios, Kennedy et al. (Nature, 2019)
Karigo et al (Nature, 2021)

Formalizing this hypothesis with a probabilistic model

$\boldsymbol{y}_t \in \mathbb{R}^N$: neural population activity at time t

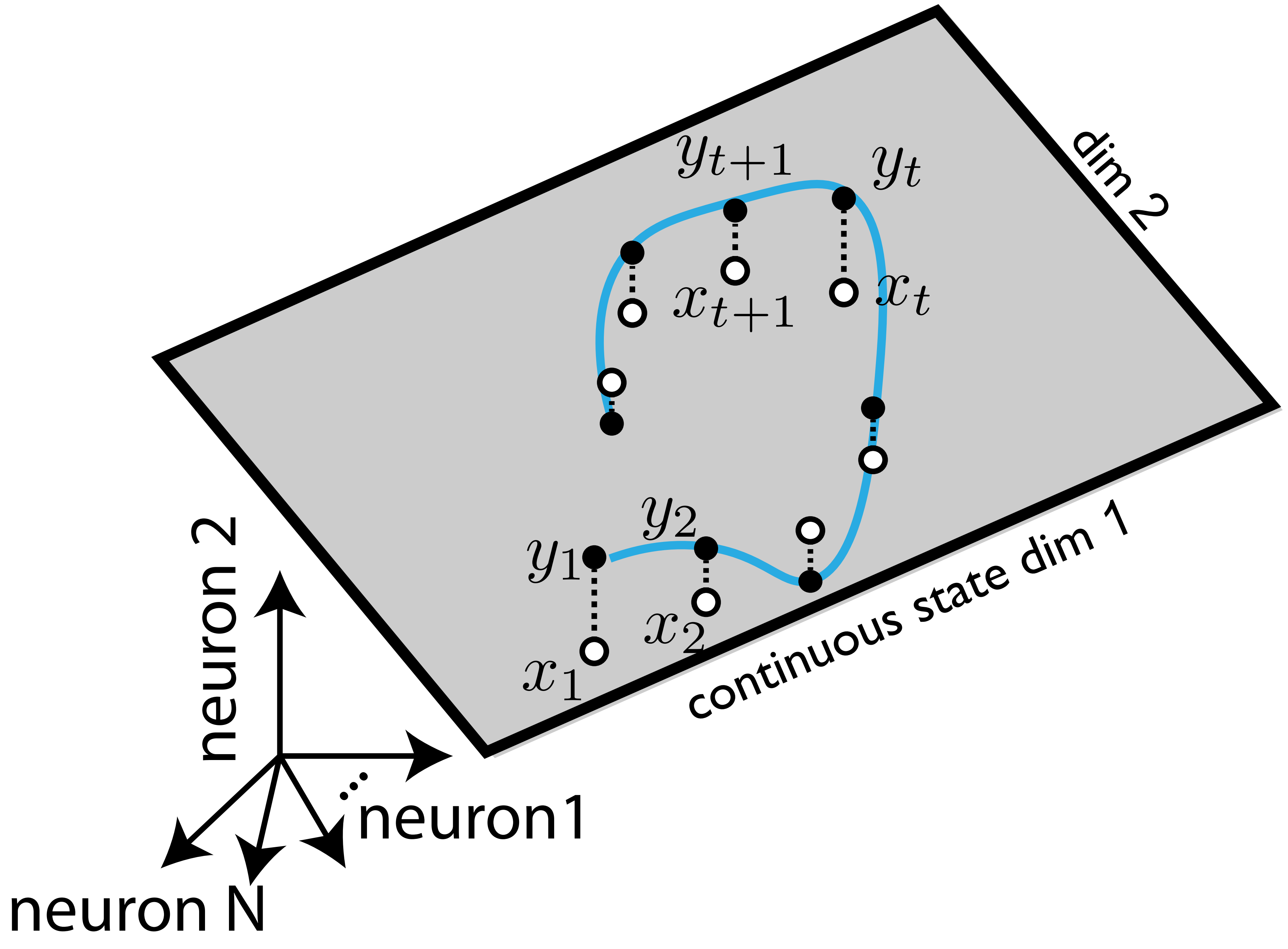


activity traces a trajectory through neural state space



Low-dimensional structure in neural data

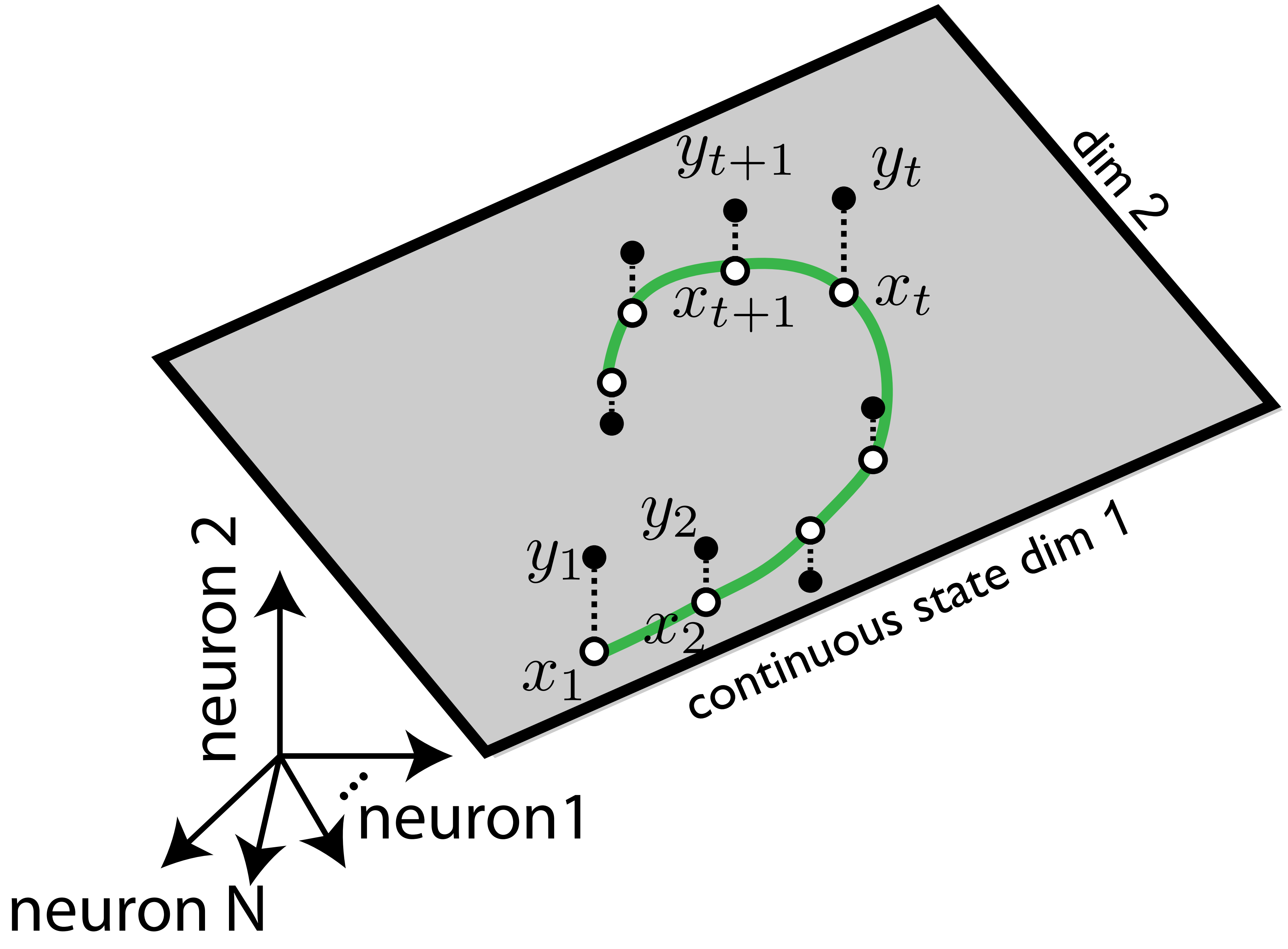
If collective activity encodes a low-dimensional state (e.g., "aggressiveness"), the data should lie near a low-dimensional manifold.



$\boldsymbol{x}_t \in \mathbb{R}^D$: *continuous latent state* (i.e., manifold coordinate)

Low-dimensional structure in neural data

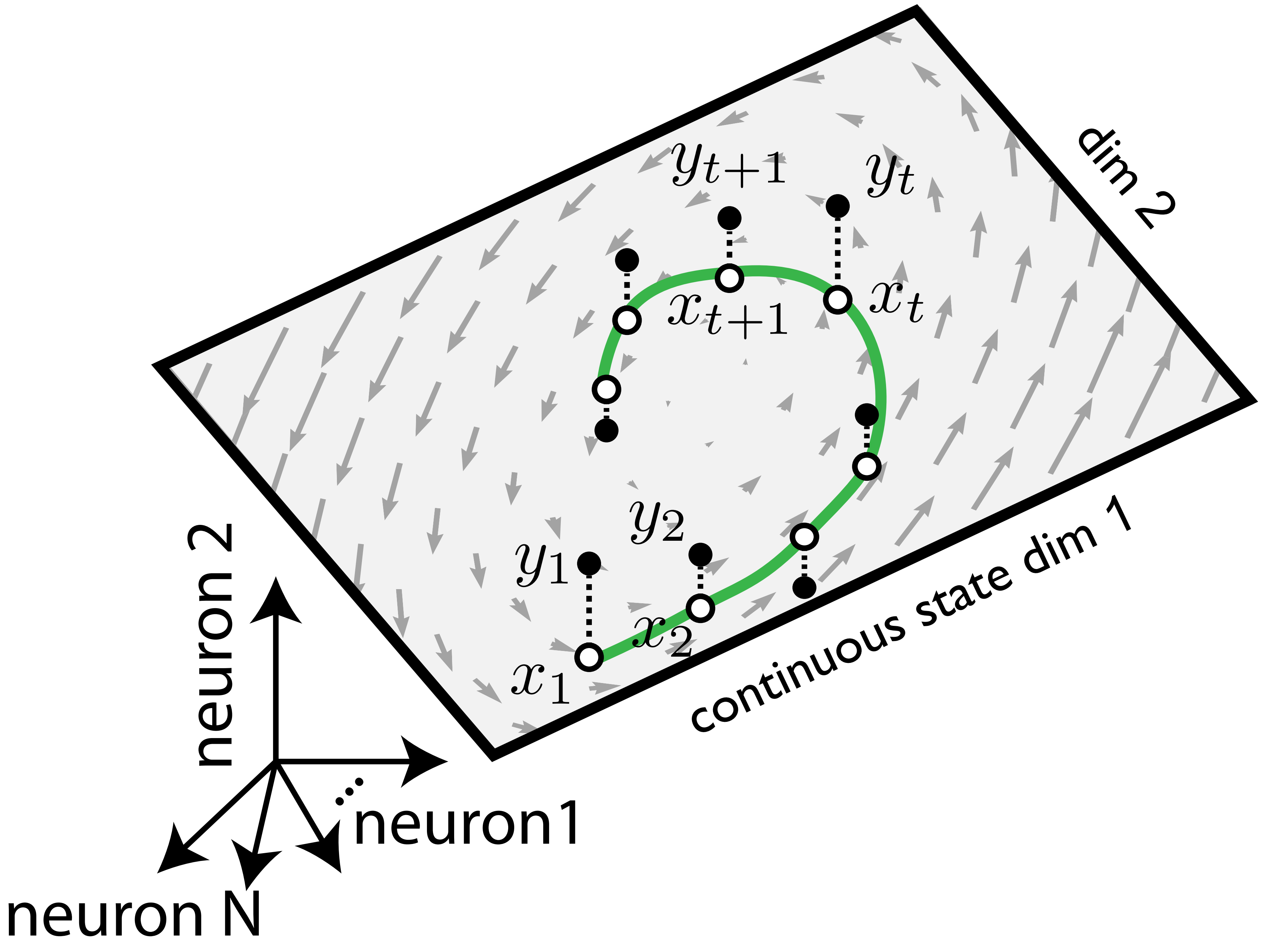
We think of neural activity as a noisy observation of a trajectory on the low-d manifold.

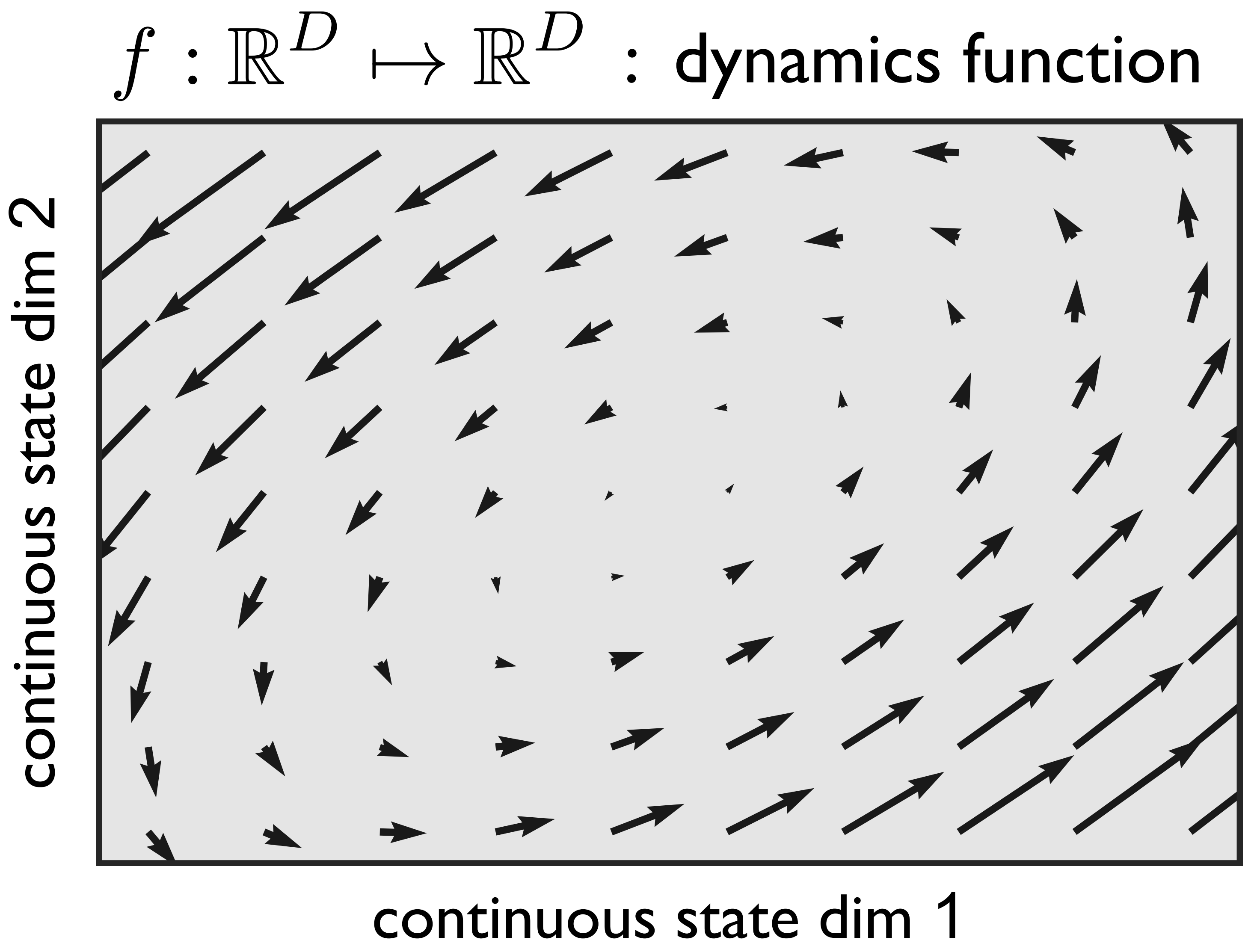


$\boldsymbol{x}_t \in \mathbb{R}^D$: *continuous latent state* (i.e., manifold coordinate)

Low-dimensional structure in neural data

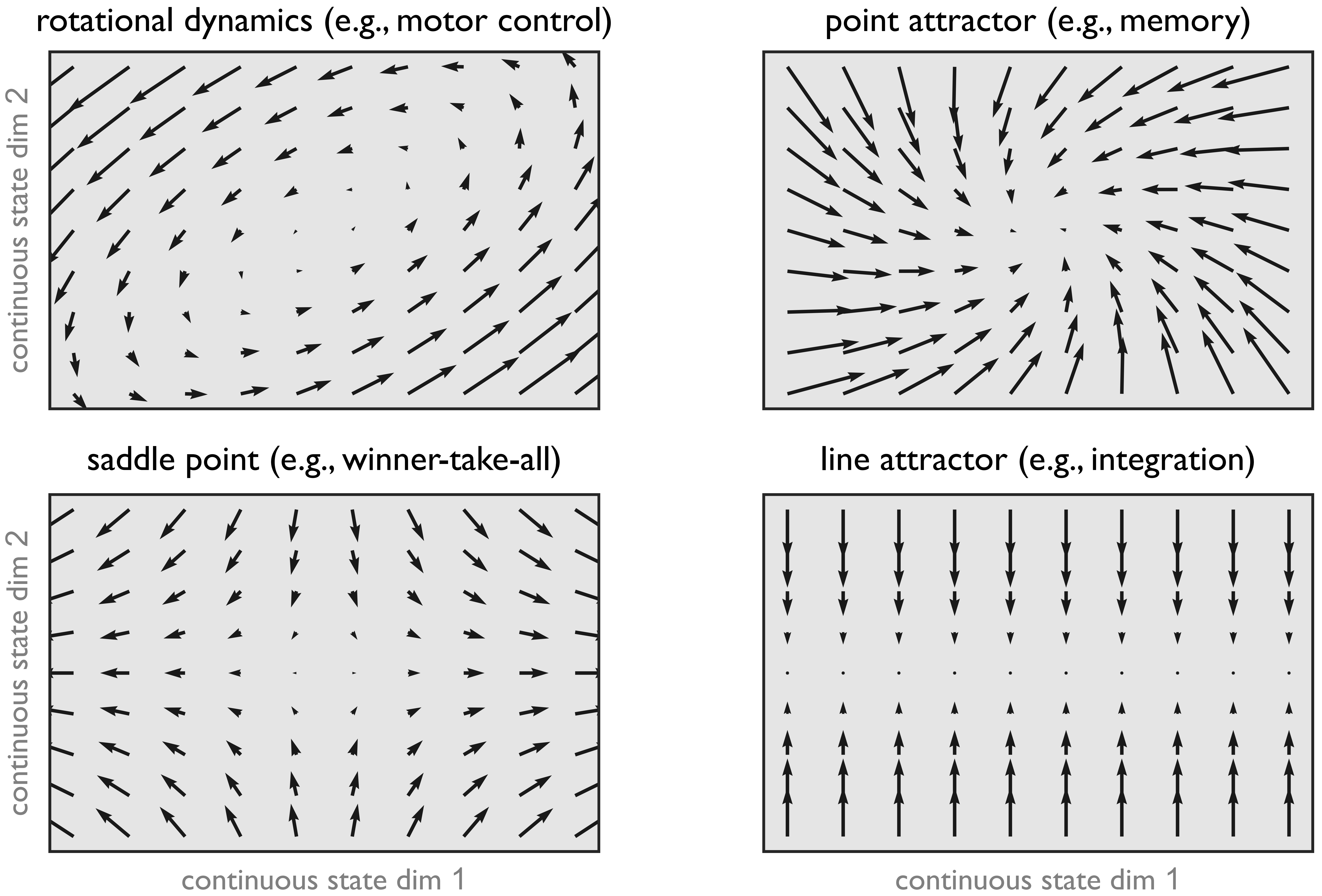
We want to learn the dynamics that govern how trajectories unfold.





Computation through neural dynamics

Dynamical motifs are hypothesized to underlie various forms of neural computation.

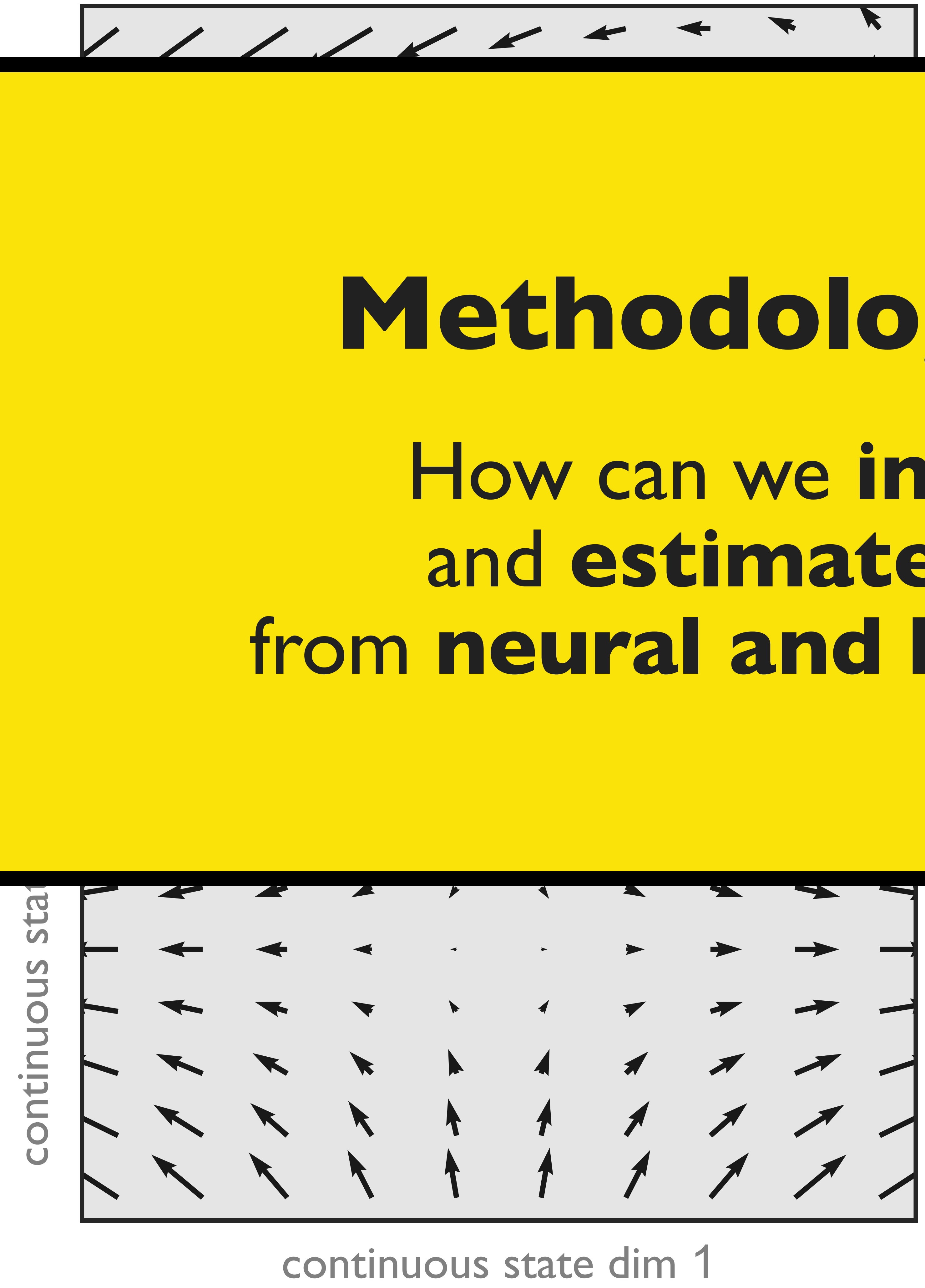


Adapted from Vyas et al. (2020)

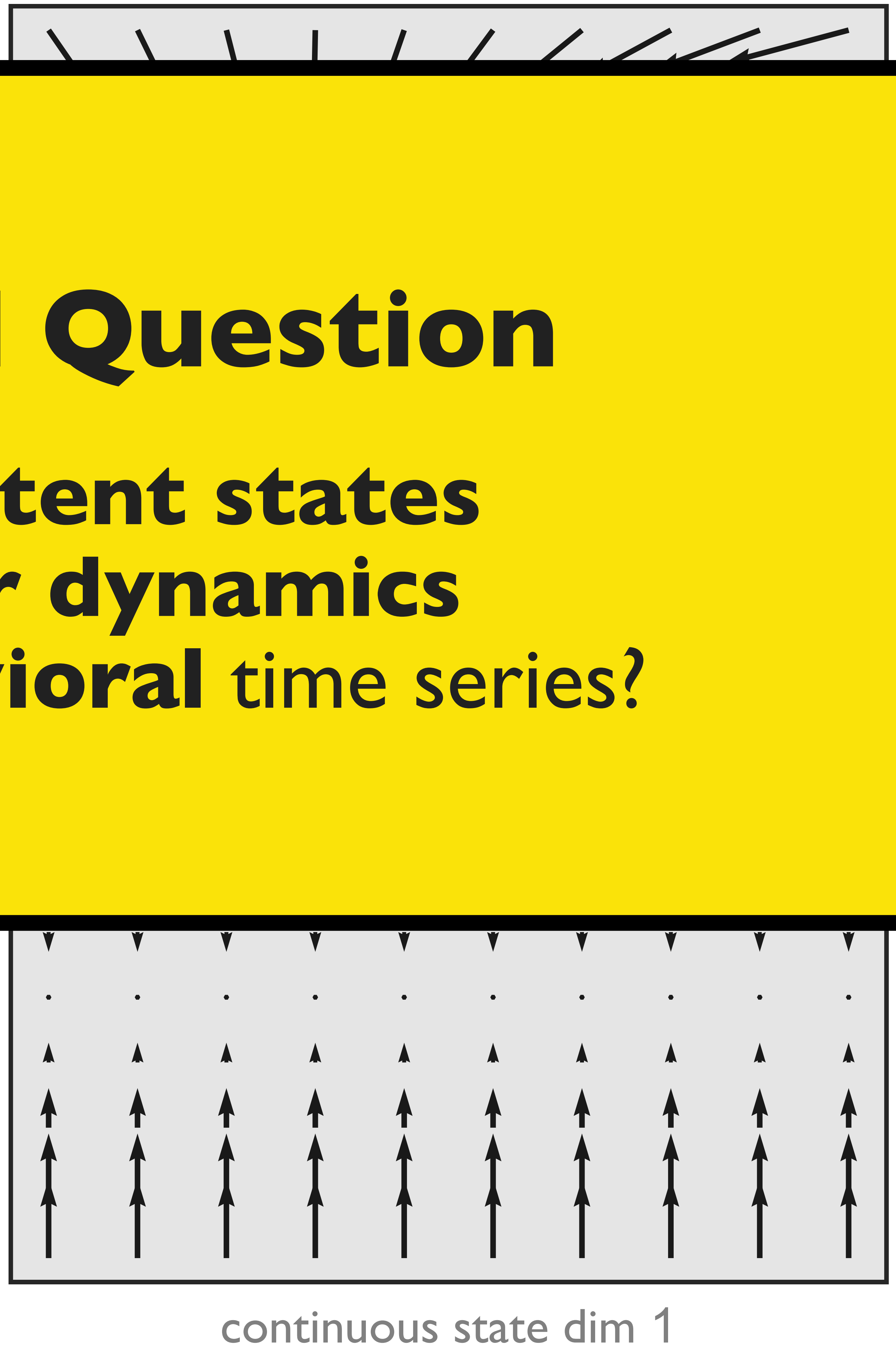
Computation through neural dynamics

Dynamical motifs are hypothesized to underlie various forms of neural computation.

rotational dynamics (e.g., motor control)



point attractor (e.g., memory)

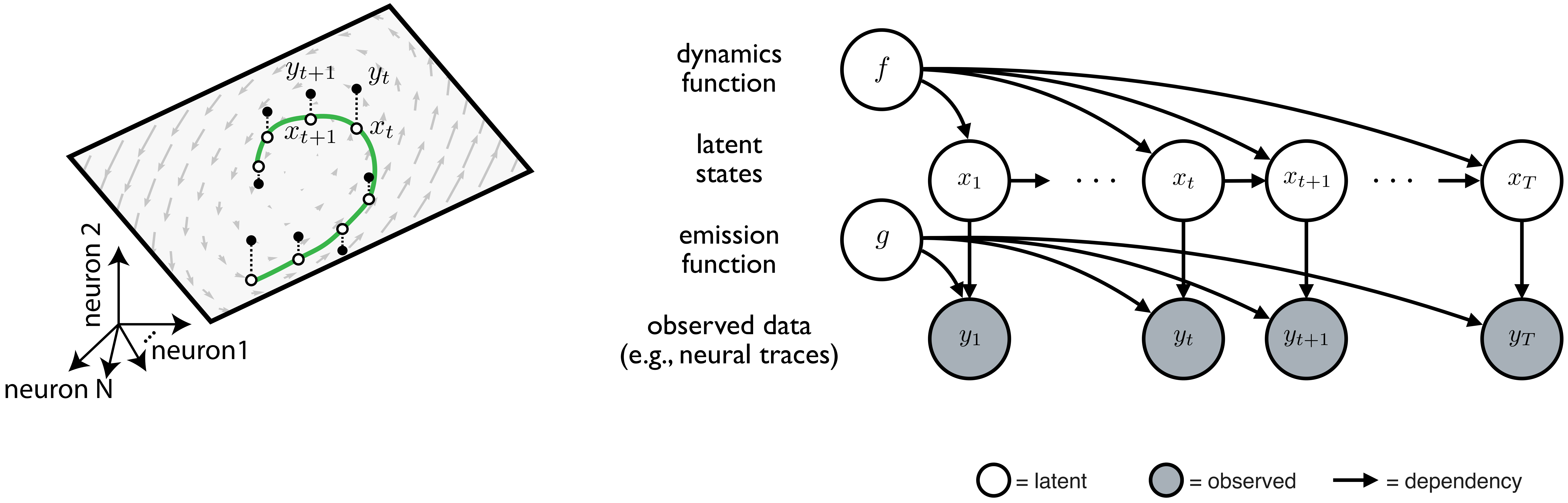


Methodological Question

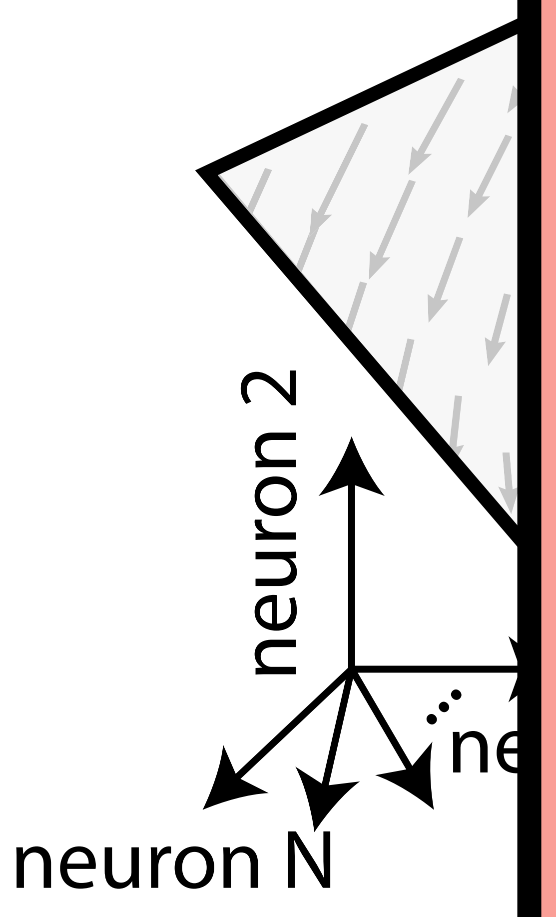
How can we **infer latent states**
and **estimate their dynamics**
from **neural and behavioral** time series?

Adapted from Vyas et al. (2020)

Probabilistic state space models



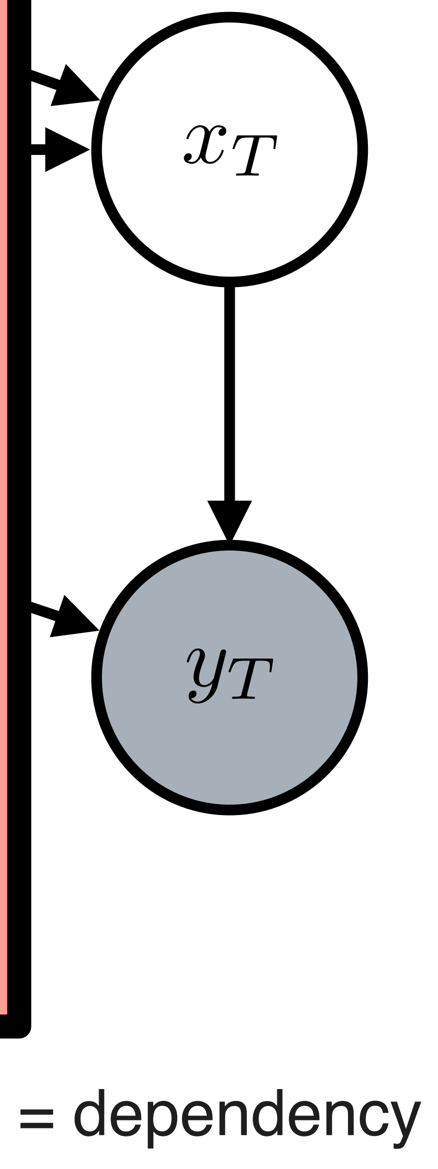
Probabilistic state space models



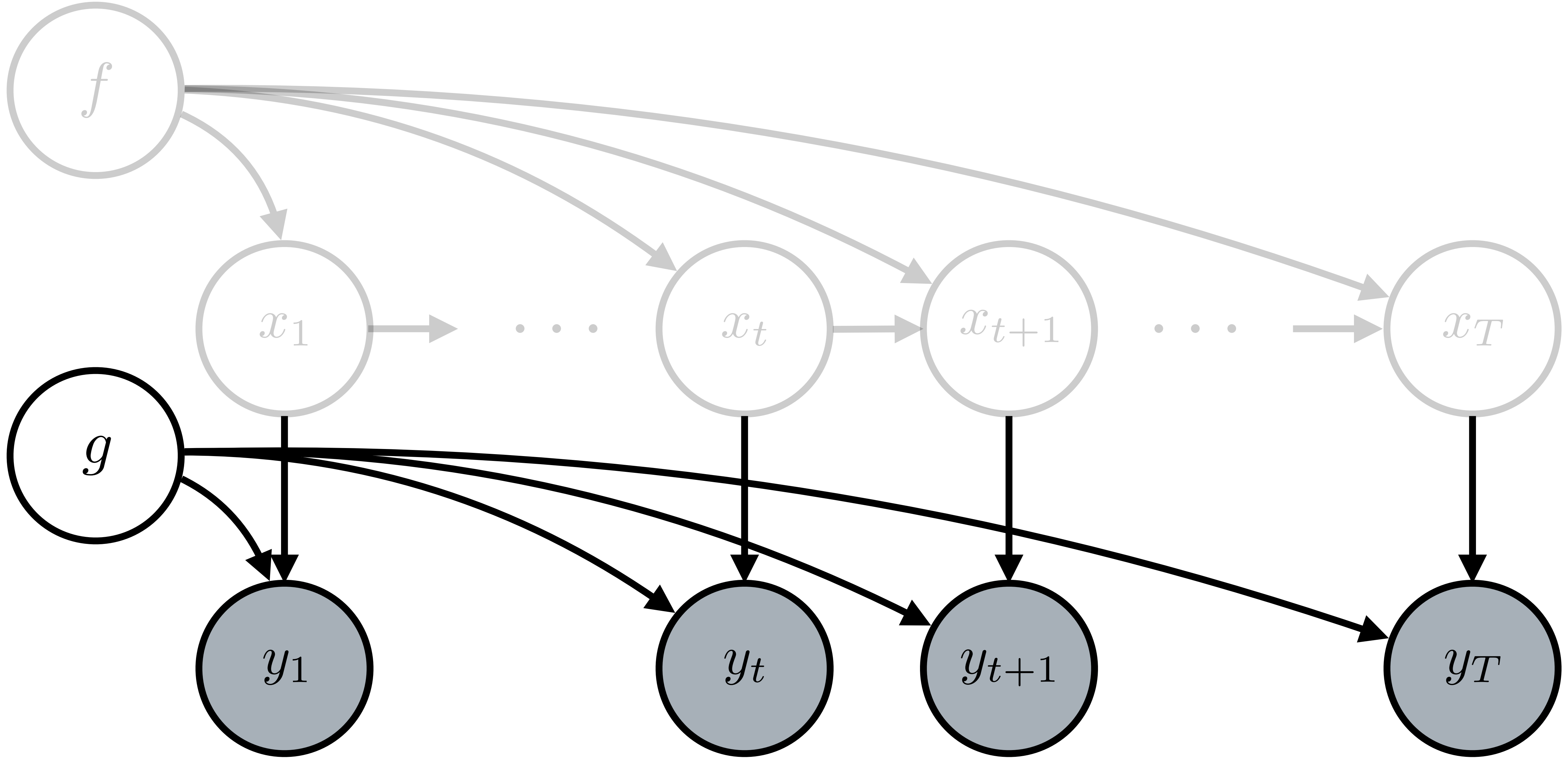
Assumptions

1. **Markovian dynamics**: next state is independent of previous states given the current state.

2. **Conditionally independent observations:** current observation is independent of others given the current state.



To start, assume a linear Gaussian observation model



For now, assume a **linear mapping** from latent states to observations.

$$g(x) = Cx + d$$

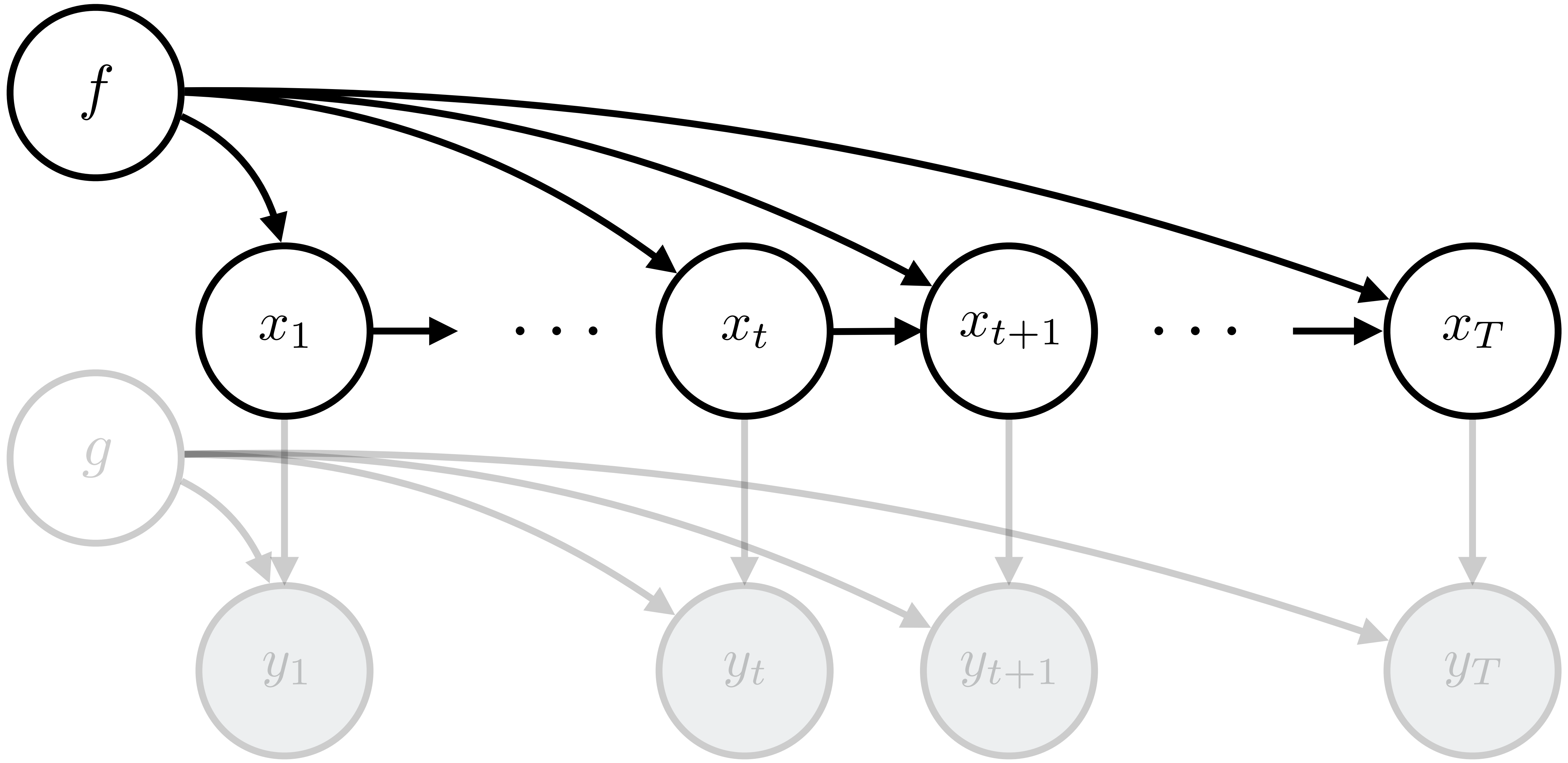
and a **Gaussian noise model**

$$y_t \mid x_t, g \sim \mathcal{N}(g(x_t), R)$$

parameterized by $\theta_{\text{obs}} = (C, d, R)$.

We can relax these assumptions later.

Desiderata for selecting a dynamics model

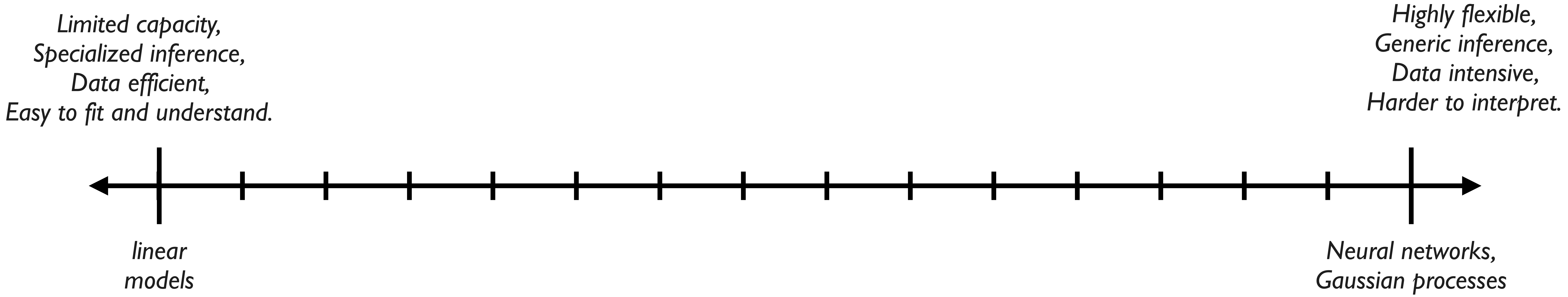


1. **Flexibility:** we need a rich enough family of models to capture a range of neural dynamics.

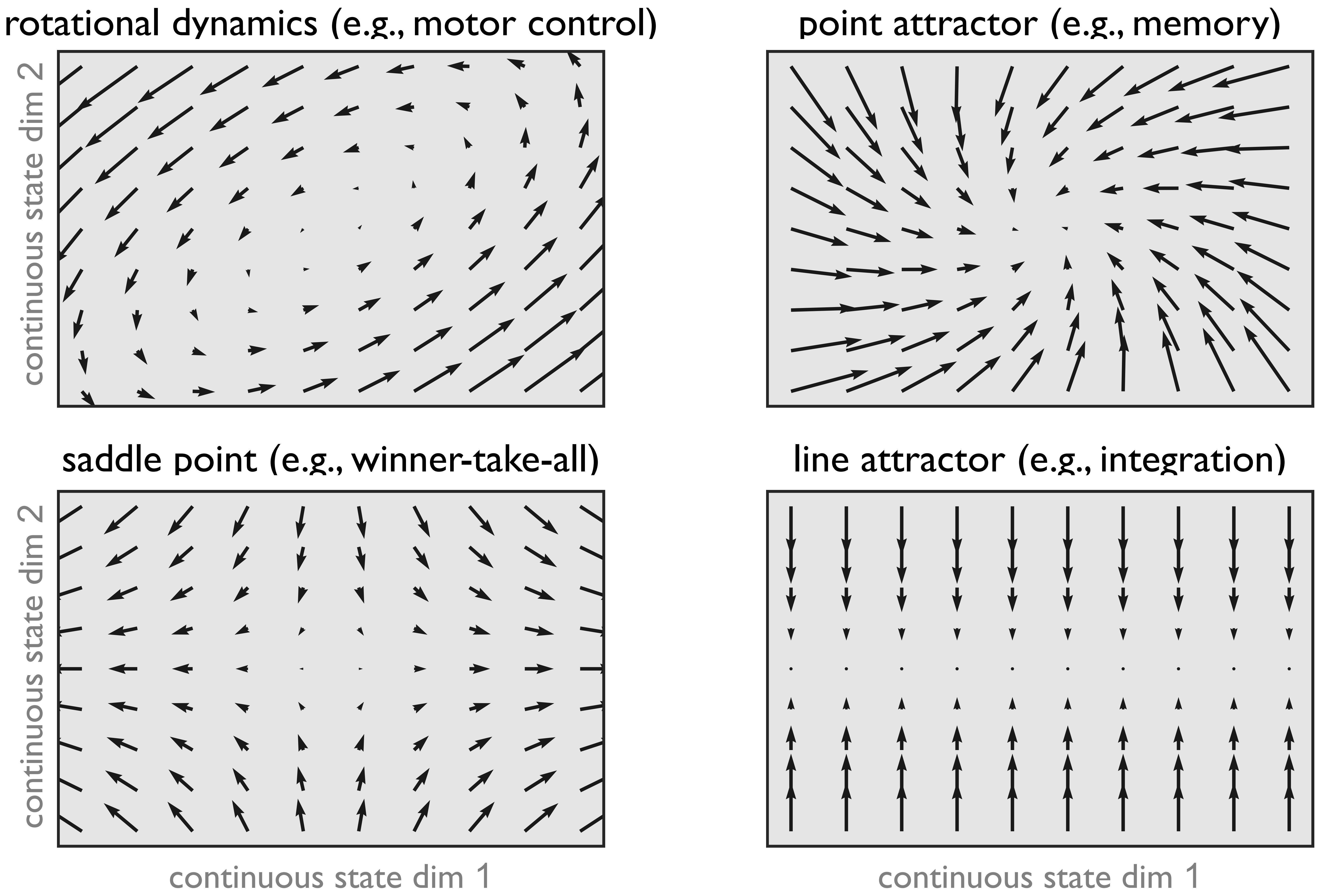
2. **Data efficiency:** we need to fit these models to a limited number of noisy recordings.

3. **Interpretability:** we want to be able to explain how these dynamics support neural computation.

A spectrum of dynamics models



What can linear models do?



A lot! E.g., the motifs from before were all linear models, $f(x) = Ax + b$.

Moreover, linear systems are interpretable.

We can find analytical solutions for:

- fixed points and stability
- dynamics along eigenmodes
- posterior distribution over latent states (with the Kalman filter/smoother)
- optimal control (with dynamic programming)

Fixed points of a (continuous-time) linear dynamical system

Consider a **continuous-time** linear dynamical system

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax + b$$

We can obtain a **discrete-time** LDS with a first-order Euler approximation,

$$\begin{aligned} x_{t+\delta} &= x_t + \delta(Ax_t + b) \\ &= (I + \delta A)x_t + \delta b \end{aligned}$$

The **fixed points** of the continuous-time system are where the time-derivative is zero,

$$\{x : Ax = -b\}$$

If A is **non-singular,** then there is a **unique fixed point** $x^\star = A^{-1}b$.

Dynamics along the eigenmodes

We can understand the system by studying its dynamics along each **eigenmode**.

Take the eigendecomposition of A,

$$A = V \Lambda V^{-1}$$

where the columns of V are the eigenvectors and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_D)$ are the eigenvalues.

For now, assume $b = 0$ and let $z = V^{-1}x$. In terms of z, the dynamics are,

$$\begin{aligned} \frac{\mathrm{d}z}{\mathrm{d}t} &= V^{-1} \frac{\mathrm{d}x}{\mathrm{d}t} \\ &= V^{-1} V \Lambda V^{-1} x \\ &= \Lambda z. \end{aligned}$$

Eigenmodes with real eigenvalues produce exponential growth or decay

Since Λ is diagonal, this is a collection of **separable, scalar linear dynamical systems**,

$$\frac{\mathrm{d}z_d}{\mathrm{d}t} = \lambda_d z_d$$

The solution of these systems is

$$z_d(t) = z_d(0)e^{\lambda_d t}$$

Since A is real-valued, its eigenvalues are either real-valued or they come in complex conjugate pairs.

First, suppose $\lambda_d \in \mathbb{R}$. Then there are two cases to consider:

1. $\lambda_d > 0 \Rightarrow z_d(t)$ **grows exponentially**, and this mode is **unstable.**
2. $\lambda_d < 0 \Rightarrow z_d(t)$ **decays exponentially**, and this mode is **stable.**

Complex eigenmodes produce oscillations

Now, consider a complex eigenvalue, $\lambda_d = \mathrm{Re}[\lambda_d] + j\mathrm{Im}[\lambda_d]$ where j is the imaginary unit.

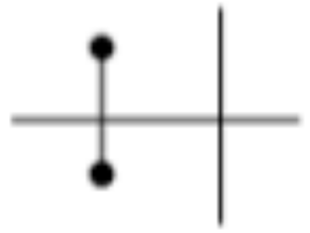
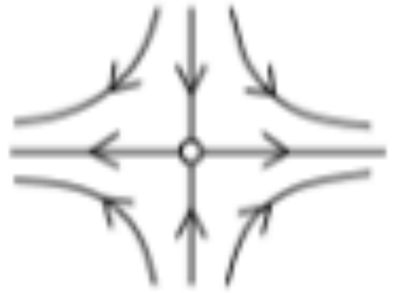
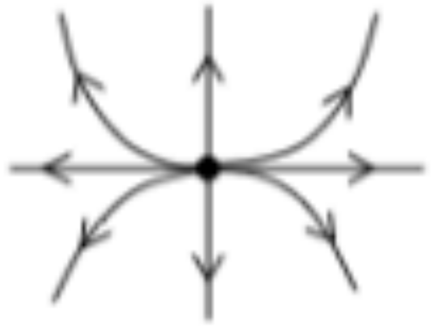
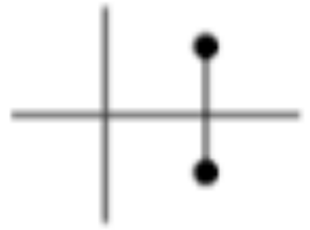
We can write the solution using **Euler's formula**,

$$
\begin{aligned}
z_d(t) &= z_d(0)e^{\lambda_d t} \\
&= z_d(0)e^{\mathrm{Re}[\lambda_d]t + j\mathrm{Im}[\lambda_d]t} \\
&= z_d(0)e^{\mathrm{Re}[\lambda_d]t}\left[\cos(\mathrm{Im}[\lambda_d]t) + j\sin(\mathrm{Im}[\lambda_d]t)\right]
\end{aligned}
$$

The **real part** of the eigenvalue determines the exponential growth or decay, and the **imaginary part** produces an **oscillation**.

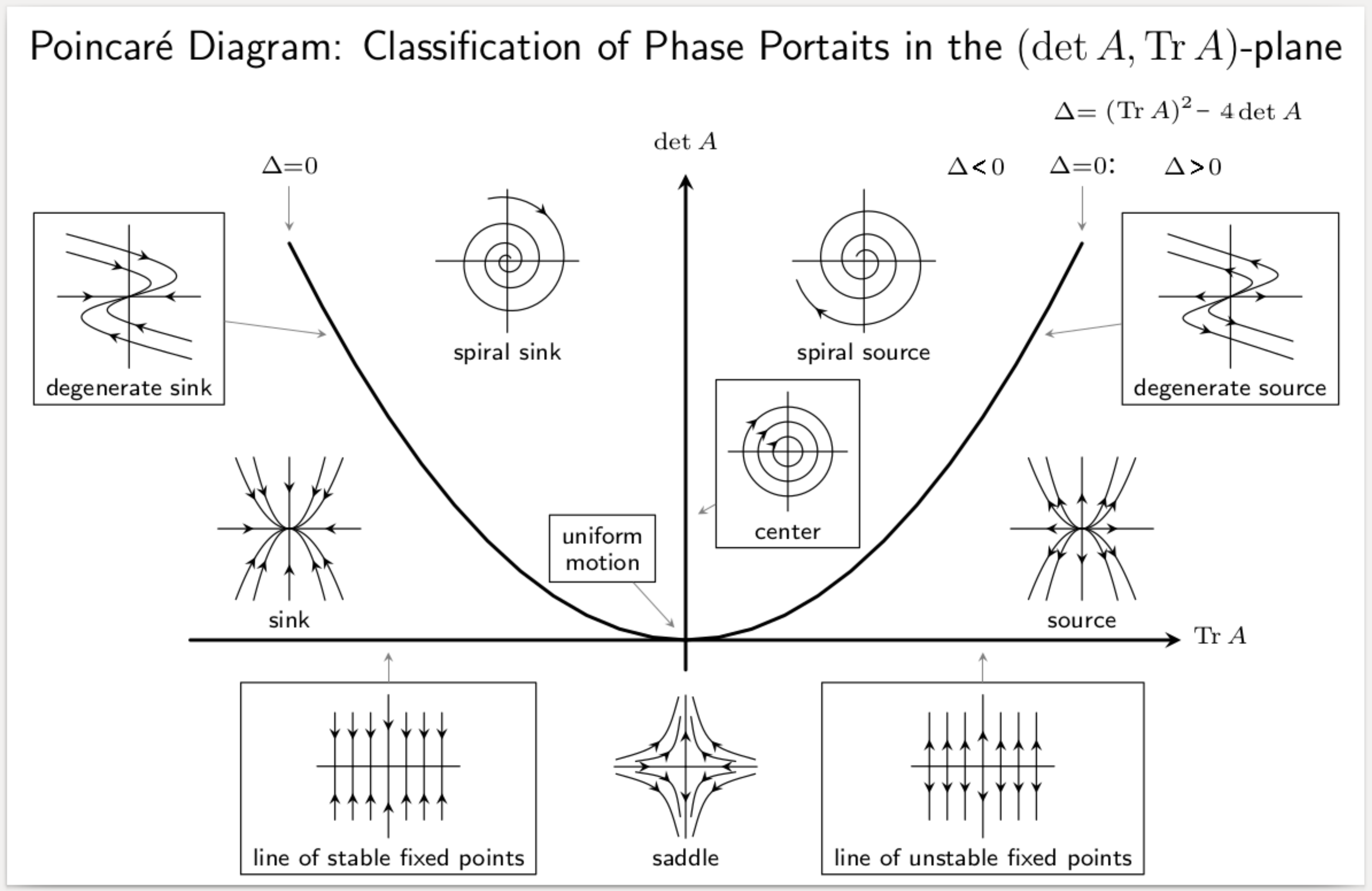
What happens to the complex part of the state?! Remember that the eigenvalues come in **complex conjugate pairs**, and so do the corresponding eigenvectors. The complex parts cancel out when we map $z(t)$ back to $x(t)$.

Linear dynamical system phase portraits as a function of the eigenvalues

Eigenvalues	Phase portrait		Stability
		node	stable
		focus	stable
		saddle	unstable
		node	unstable
		focus	unstable

http://wpage.unina.it/p.maffettone/Lezione01PLM.pdf

Linear dynamical system phase portraits as a function of the trace and determinant of A

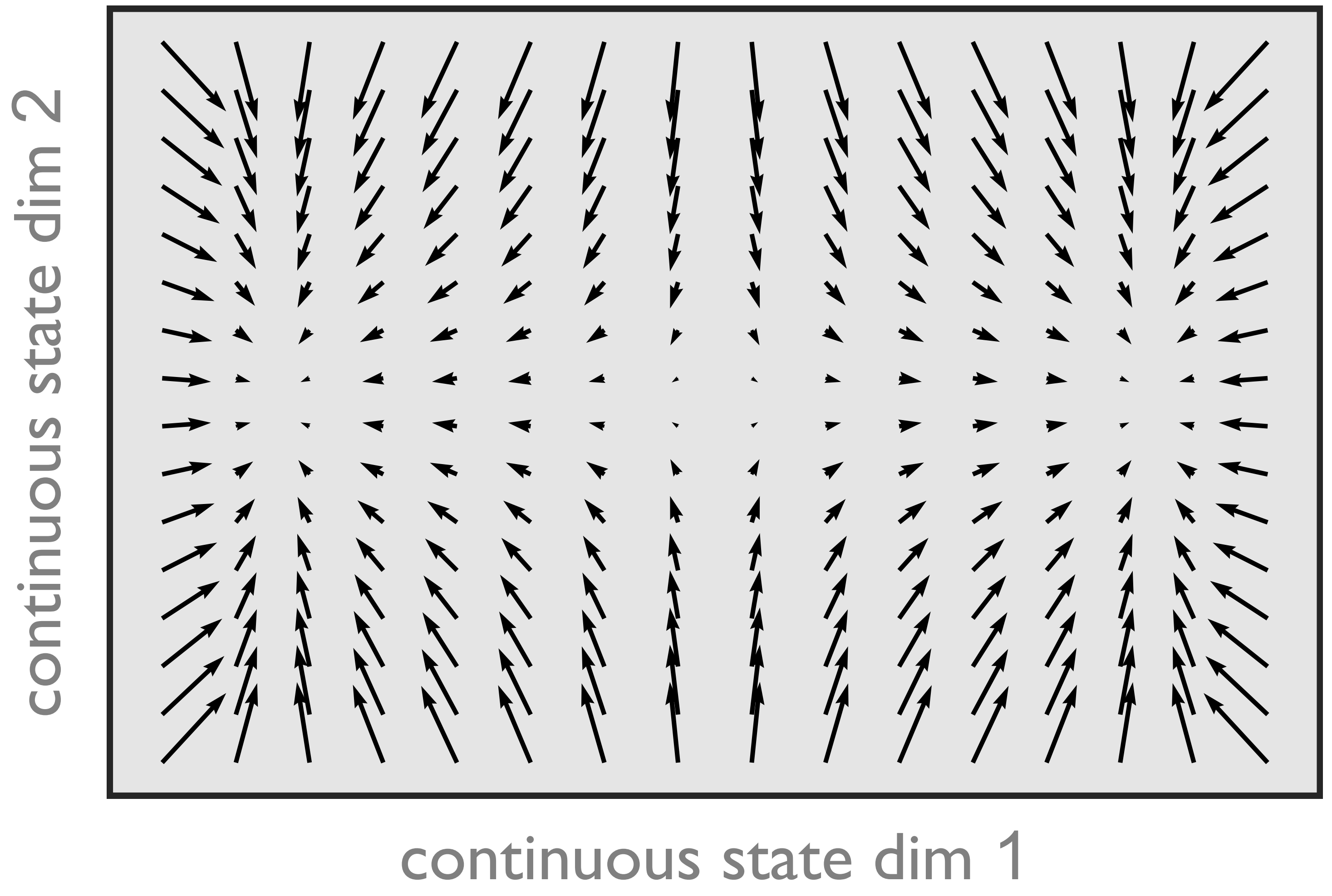


https://en.wikipedia.org/wiki/Stability_theory

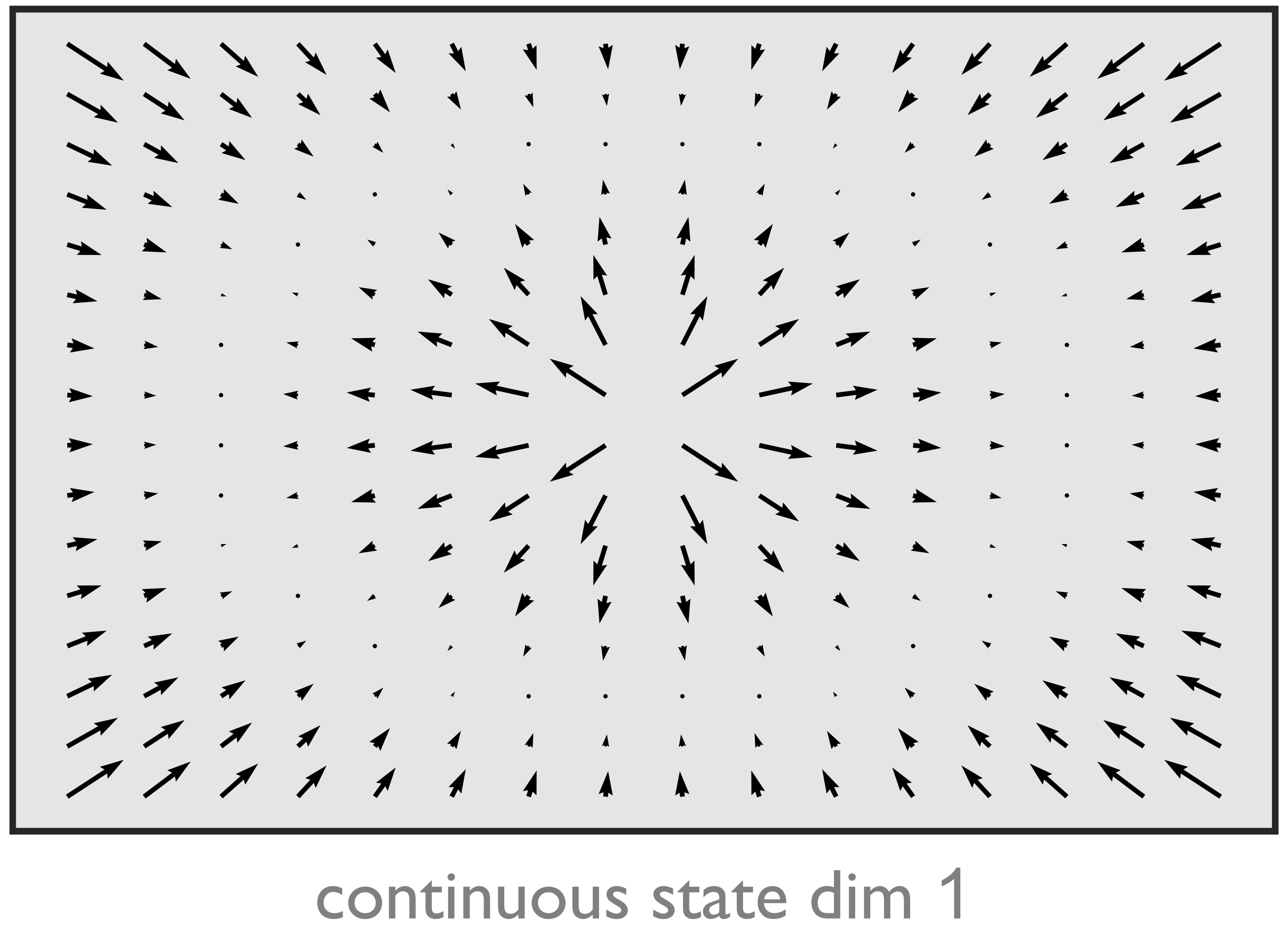
What **can't** linear models do?

Still, most computations require nonlinear dynamics.

bistability (e.g., decision making)



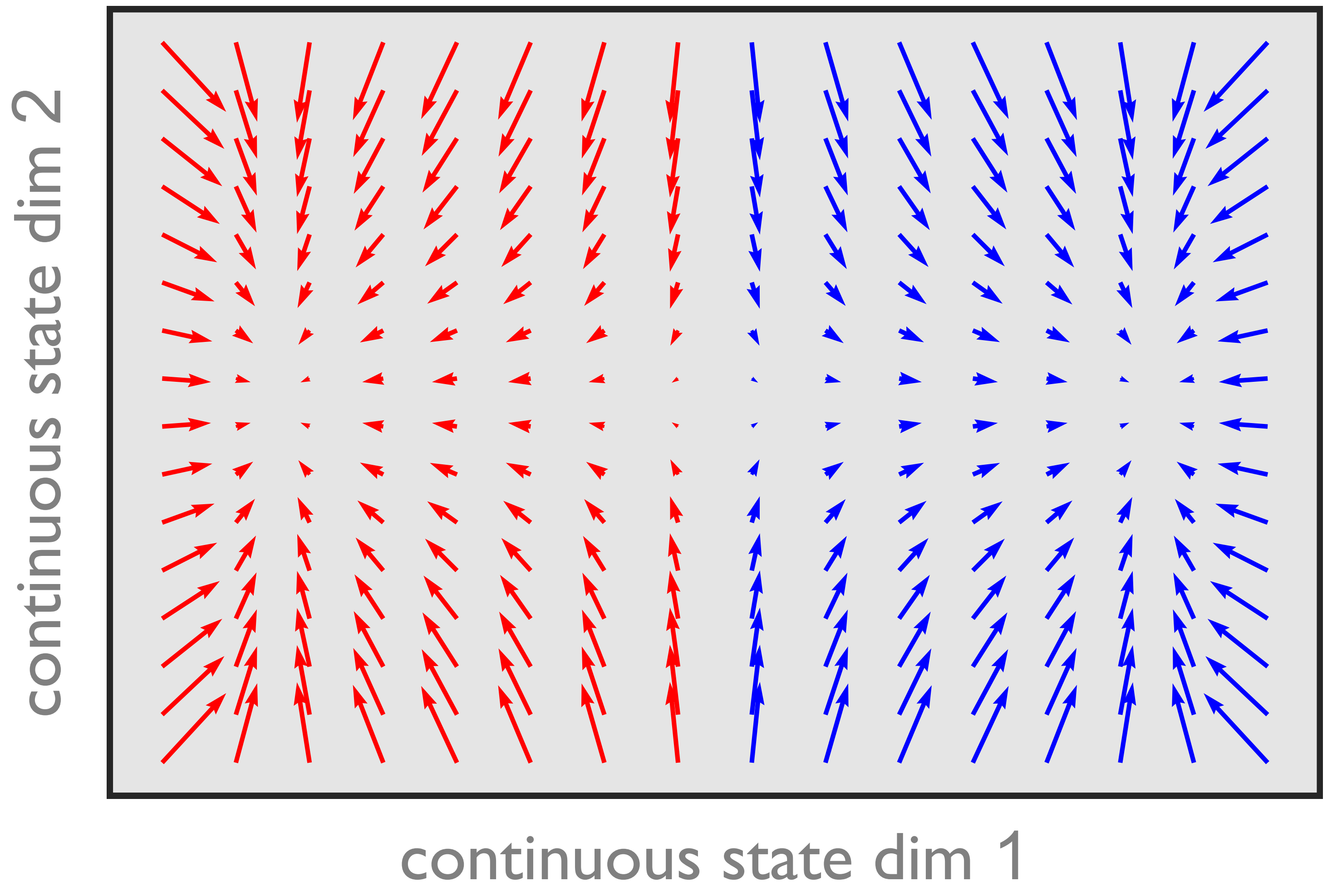
ring attractor (e.g., head direction)



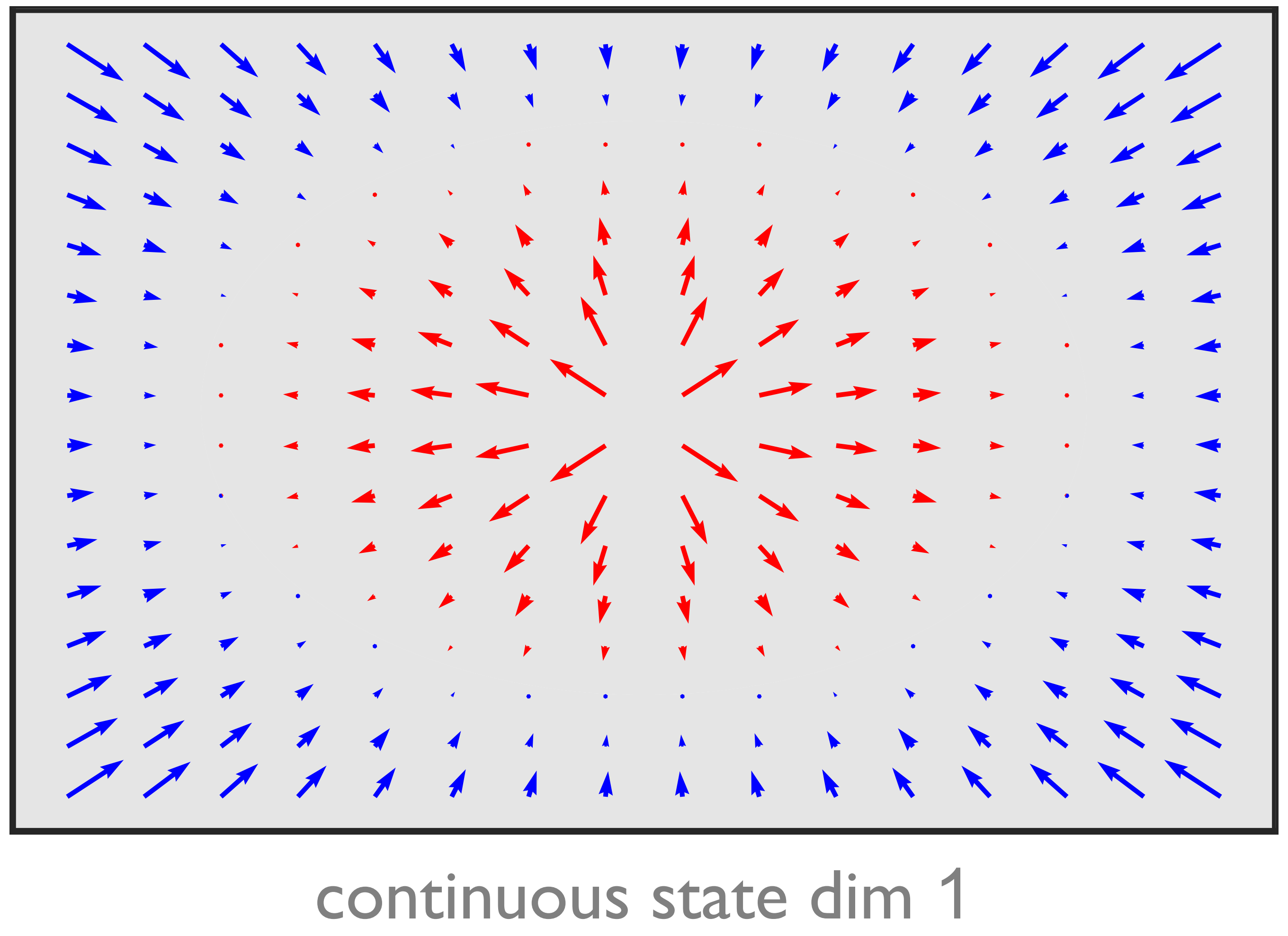
Key idea: nonlinear dynamics can often be approximated as piecewise-linear

Indeed, that's often how we analyze nonlinear dynamical systems!

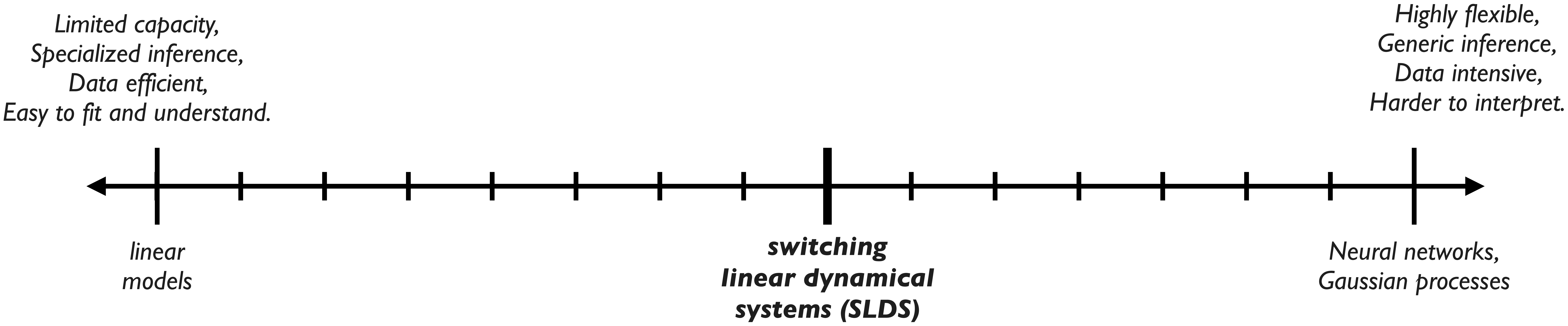
bistability (e.g., decision making)



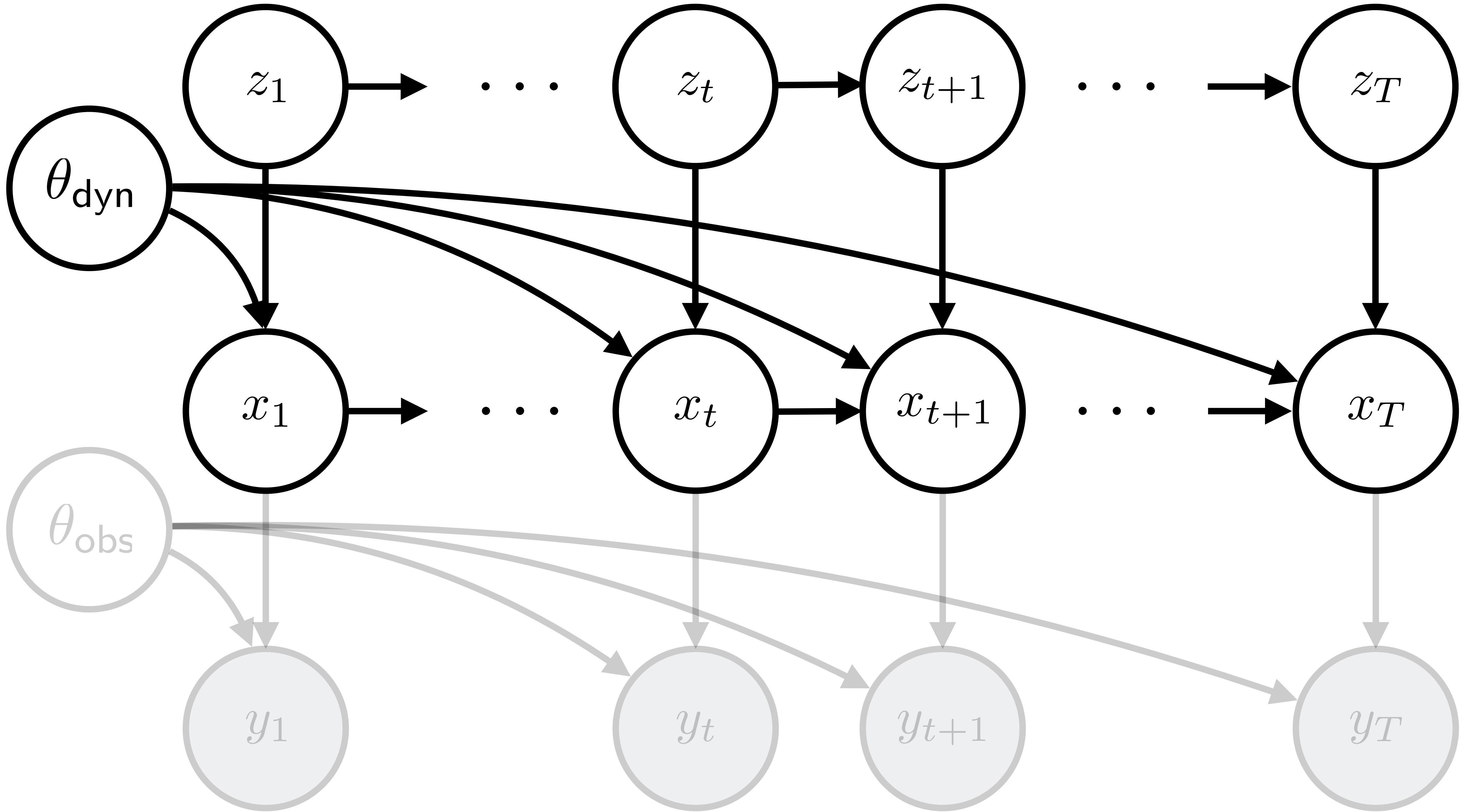
ring attractor (e.g., head direction)



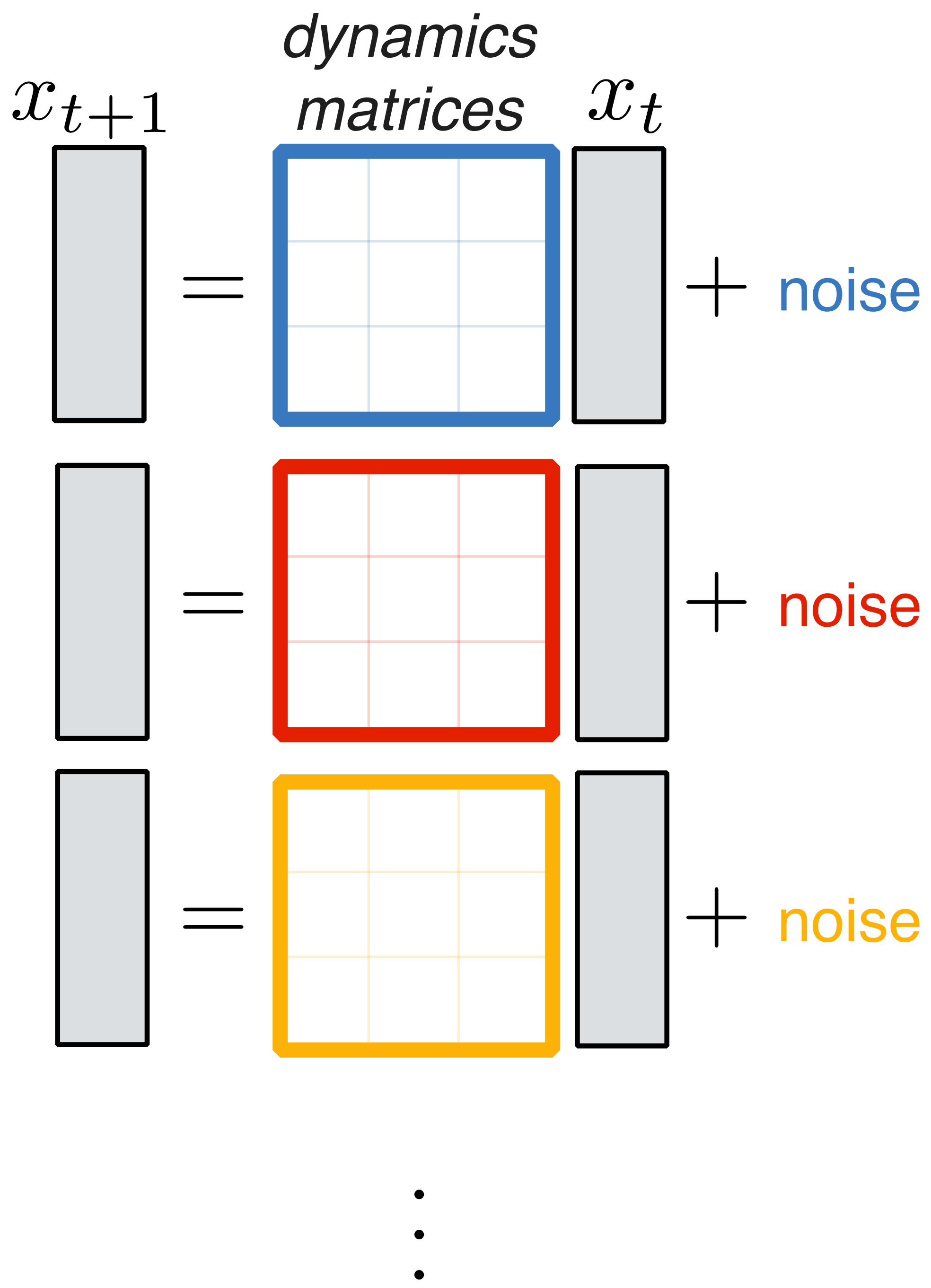
A spectrum of dynamics models



Switching linear dynamical systems (SLDS)



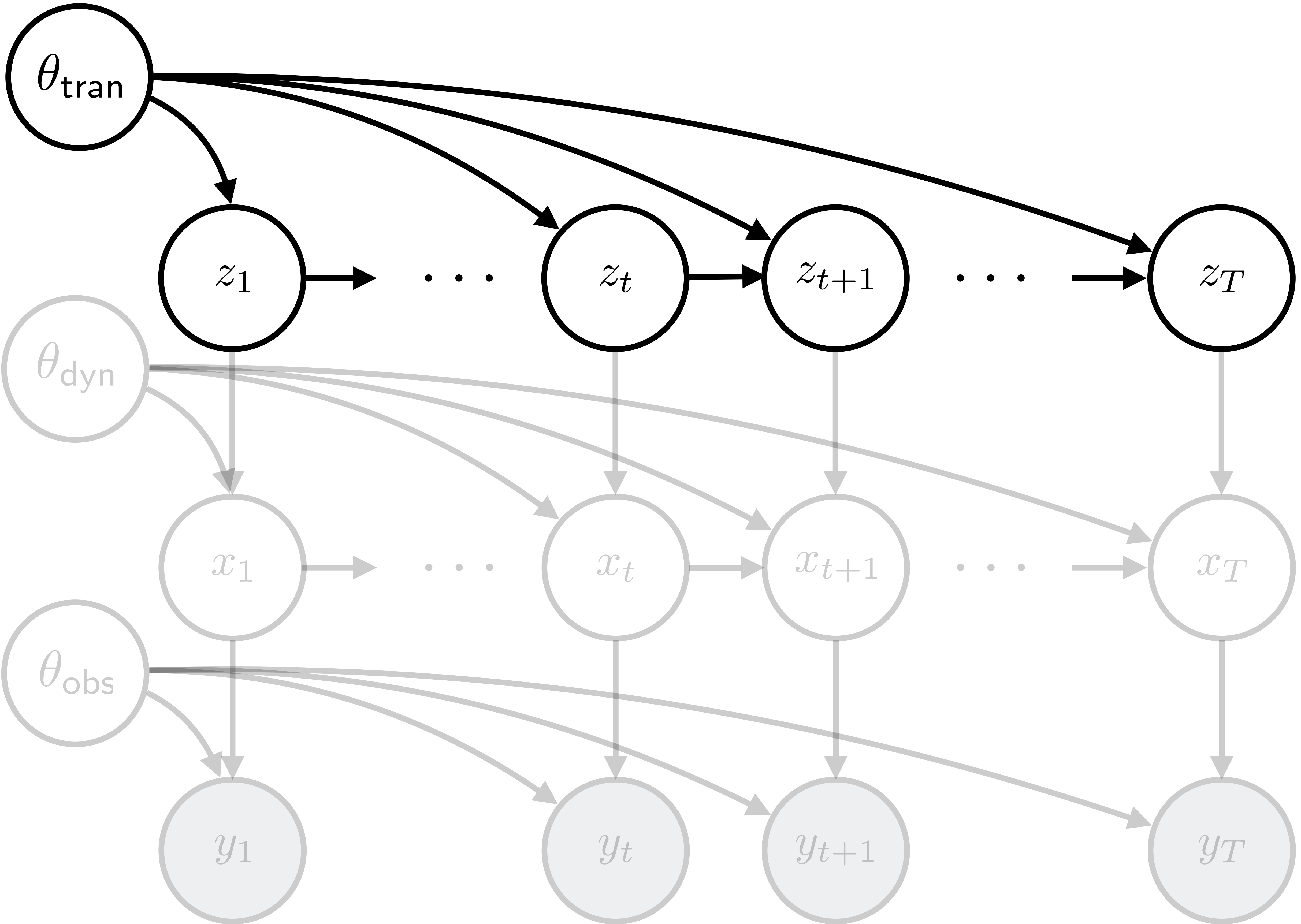
Different **linear dynamics**
in each discrete state



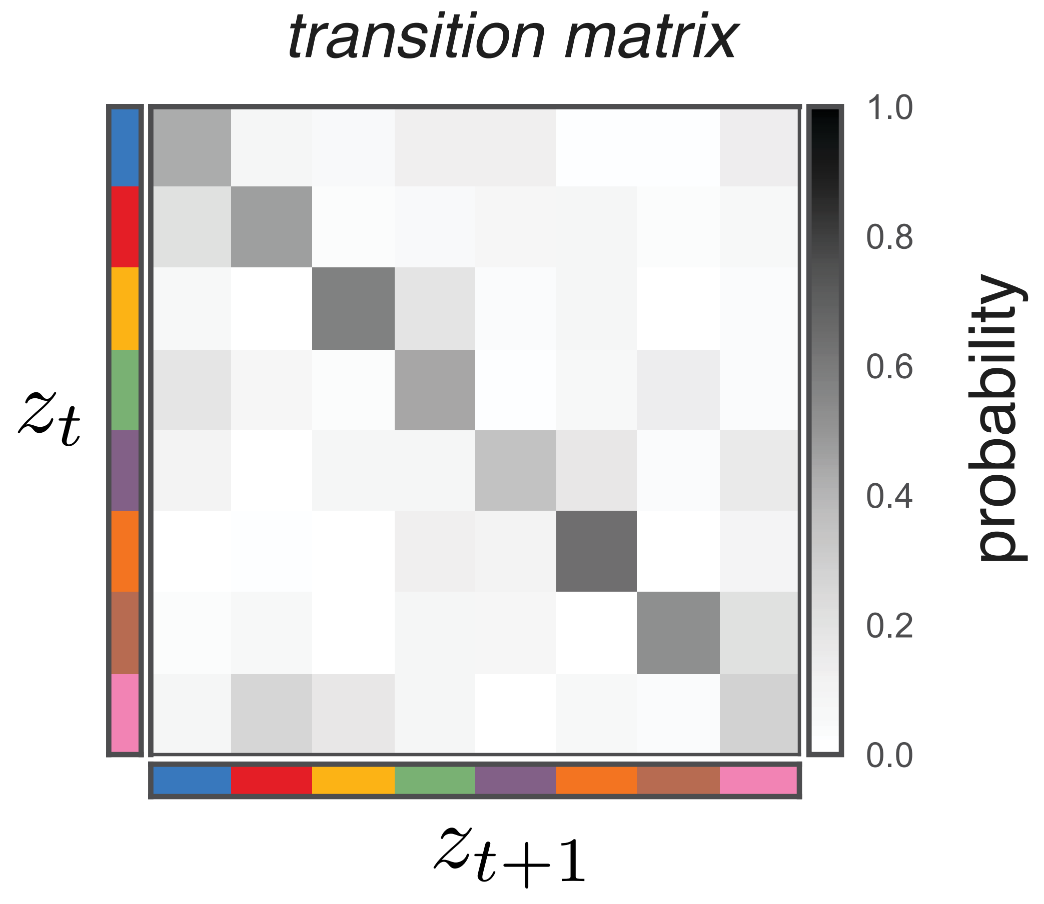
Ackerson and Fu (1970)
Chang and Athans (1978)
Hamilton (1990)
Ghahramani and Hinton (1996)
Murphy (1998)
Fox et al (2009)

Note: here z is a **discrete latent variable!*

Switching linear dynamical systems (SLDS)

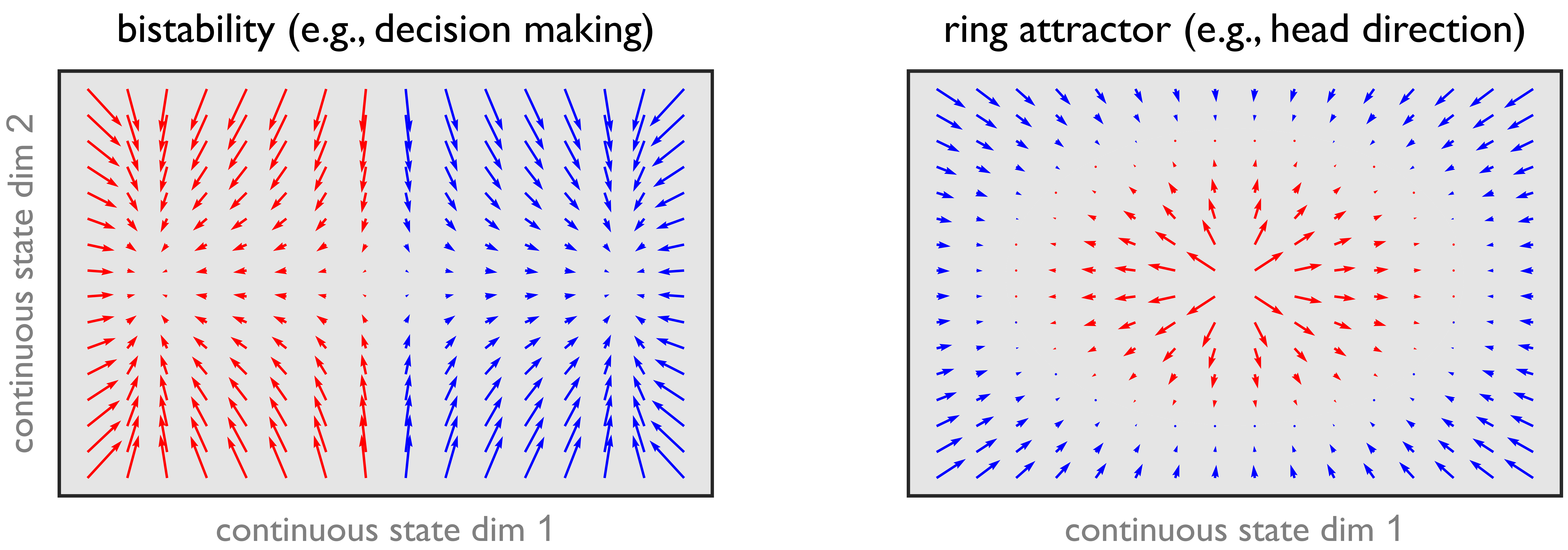


State-dependent switching probabilities

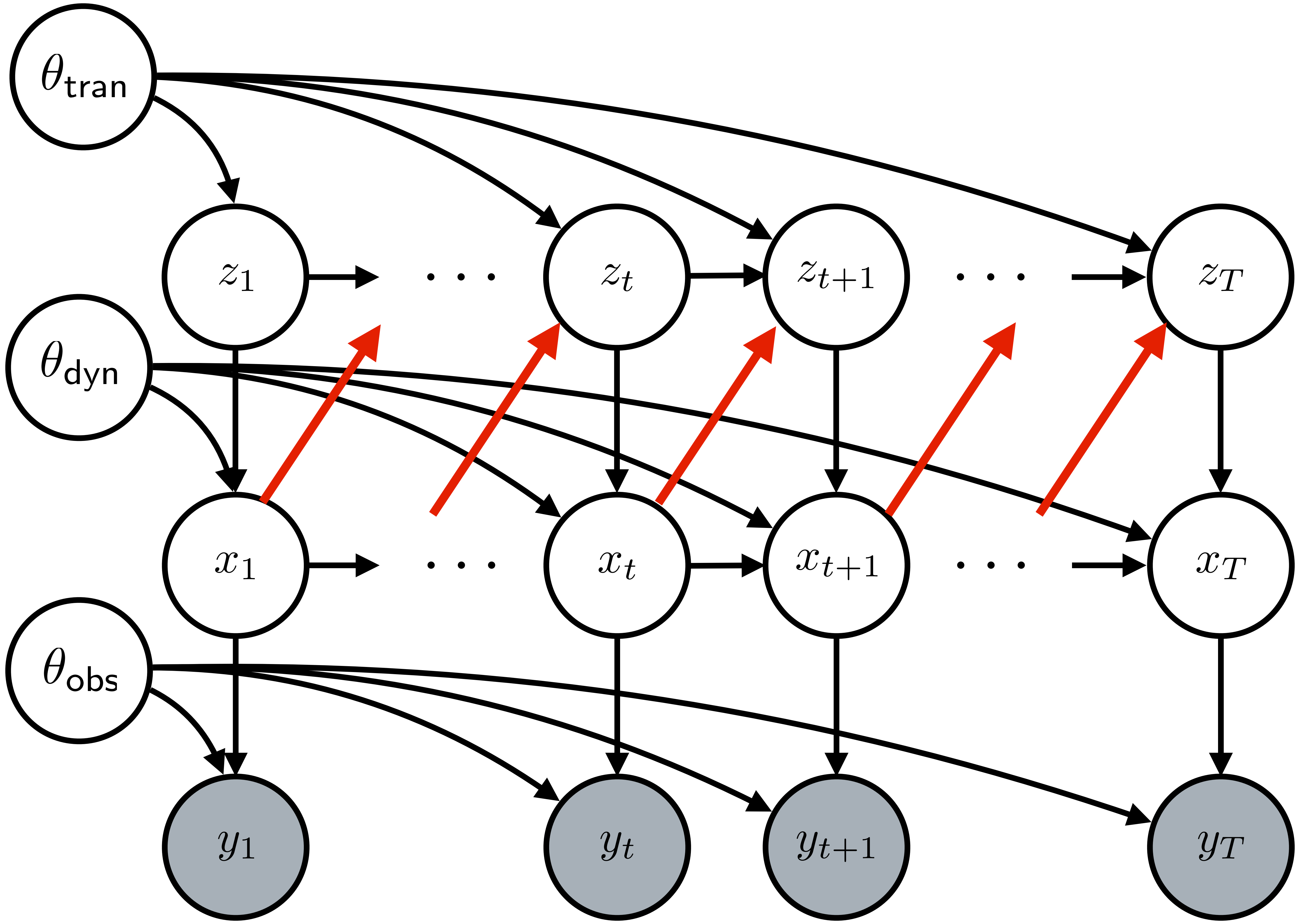


Ackerson and Fu (1970)
Chang and Athans (1978)
Hamilton (1990)
Ghahramani and Hinton (1996)
Murphy (1998)
Fox et al (2009)

Problem: in an SLDS, discrete state transitions are independent of location!

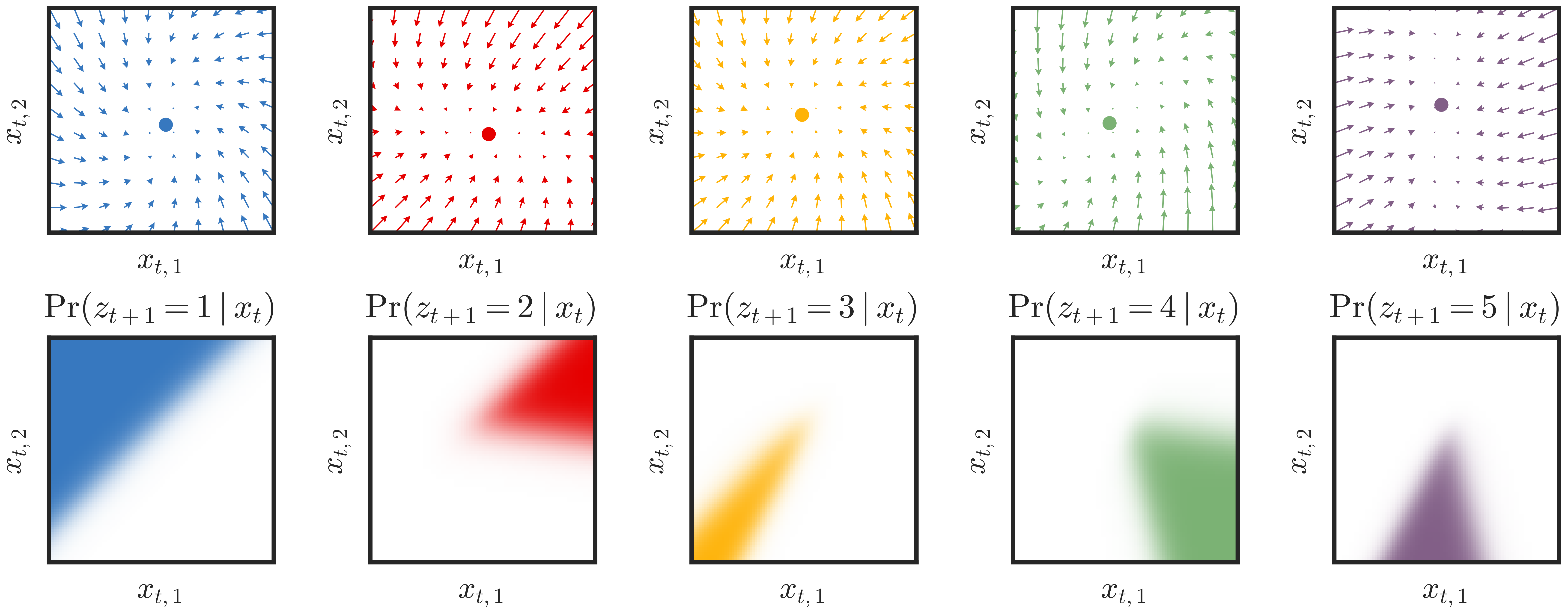


Recurrent switching linear dynamical systems (**r**SLDS)



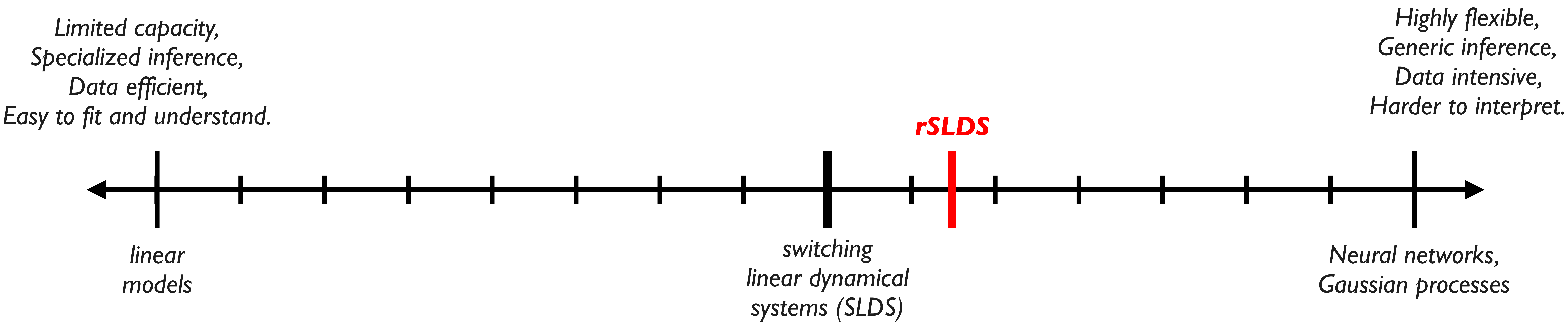
Linderman et al. (AISTATS, 2017)
Zoltowski, Pillow, & Linderman (2020)
…and now many more

Recurrent SLDS partition continuous state space into regions with linear dynamics

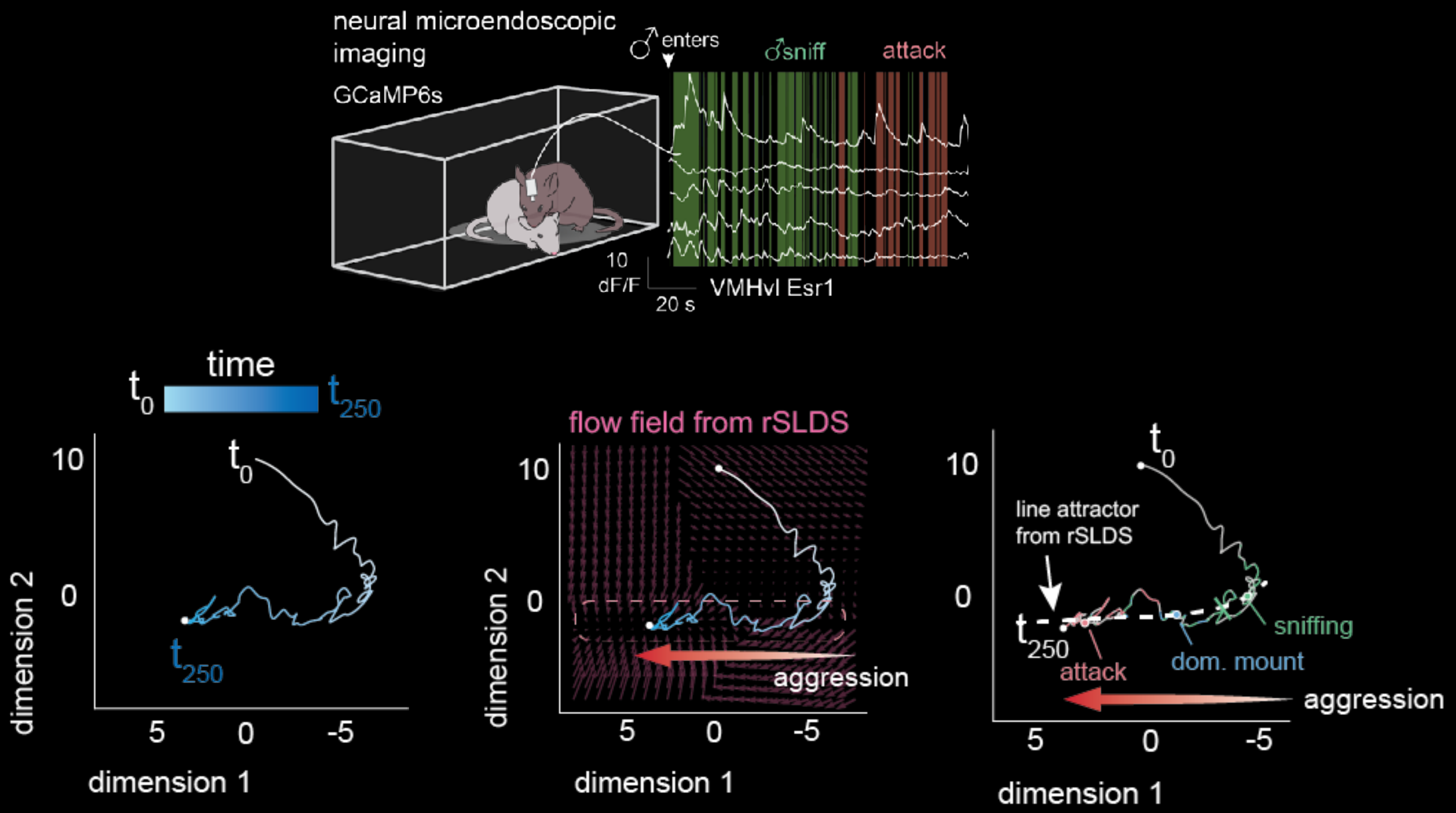


Linderman et al. (AISTATS, 2017)
Zoltowski, Pillow, & Linderman (2020)
…and now many more

A spectrum of dynamics models

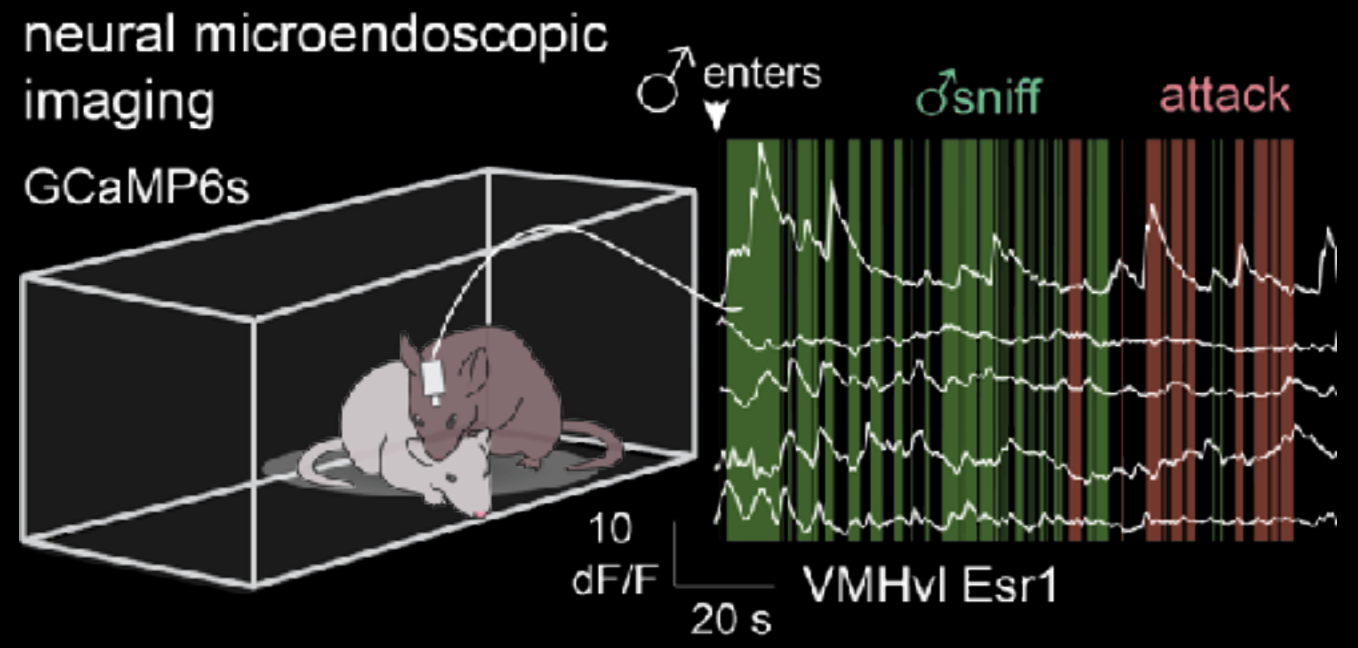


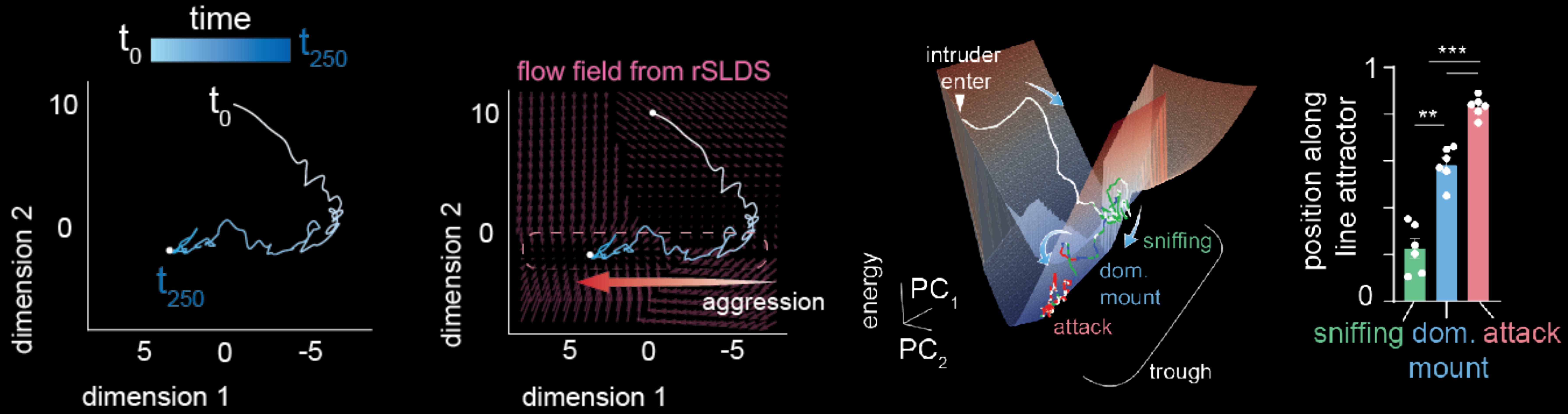
rSLDS analysis reveals line attractor-like dynamics in VMHvl



Nair et al. (Cell, 2023)

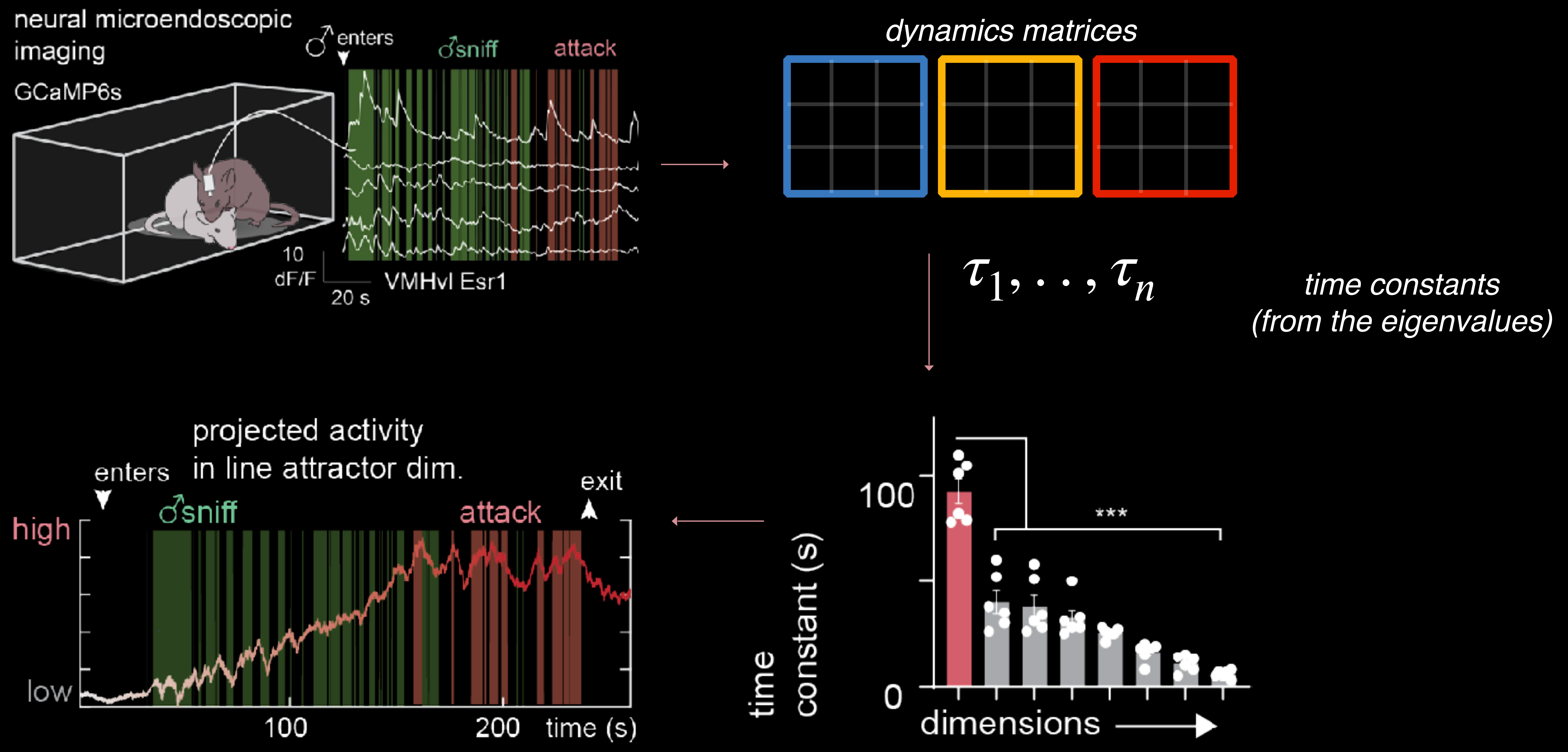
rSLDS analysis reveals line attractor-like dynamics in VMHvl





Nair et al. (Cell, 2023)

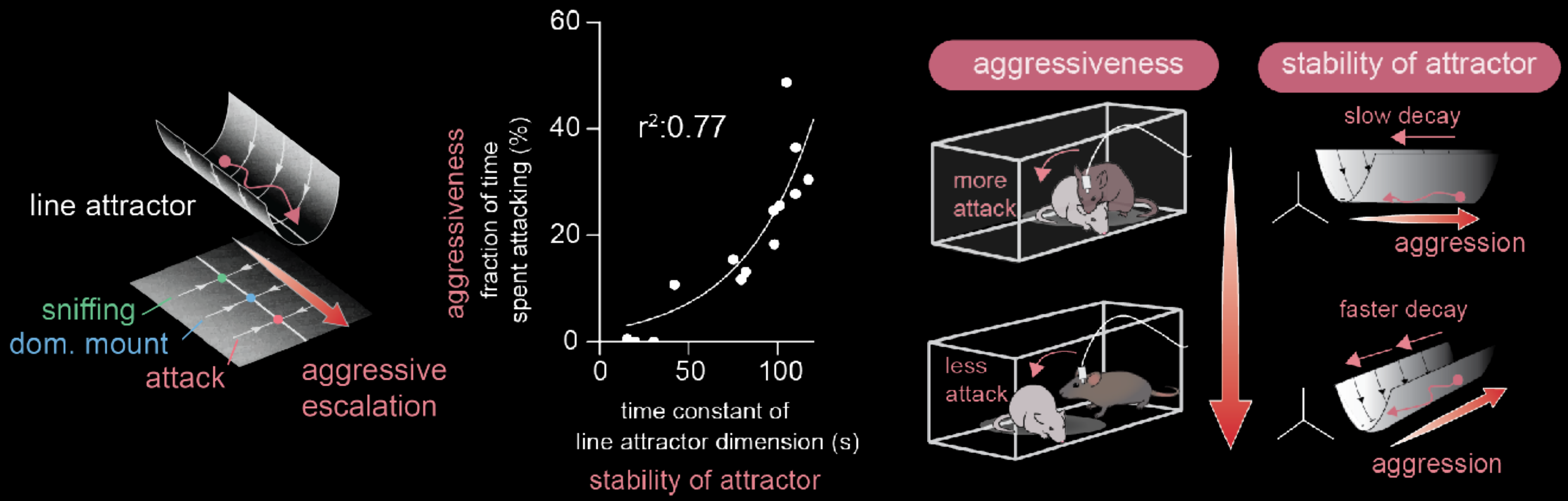
rSLDS analysis reveals line attractor-like dynamics in VMHvl



Importantly, this is not true of all hypothalamic nuclei, e.g., MPOA.

Nair et al. (Cell, 2023)

Dynamical systems explain individual differences in aggressiveness

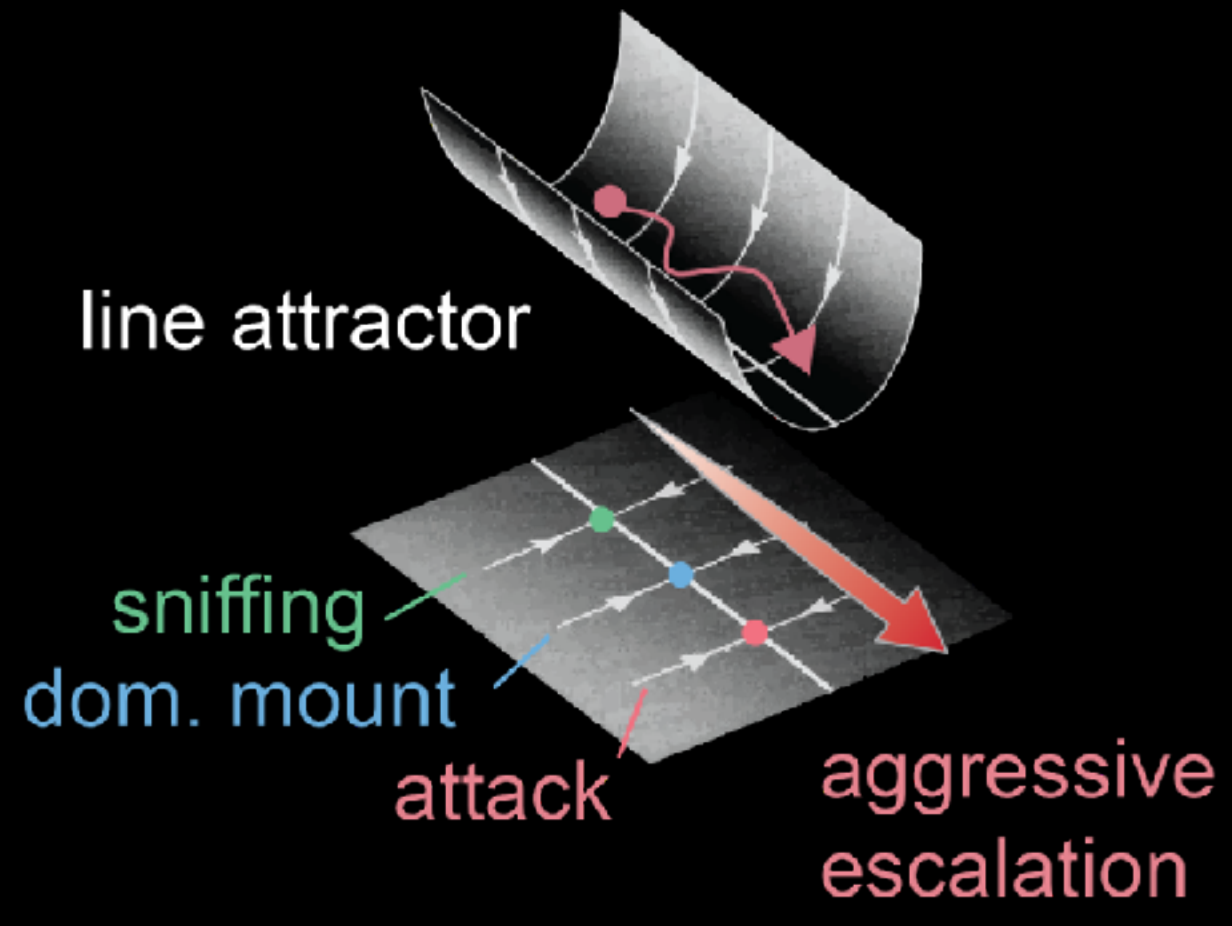


the stability of the attractor is enhanced in mice that are more aggressive

Nair et al. (Cell, 2023)

Are these dynamics intrinsic to VMHvl or a read-out of an upstream region?

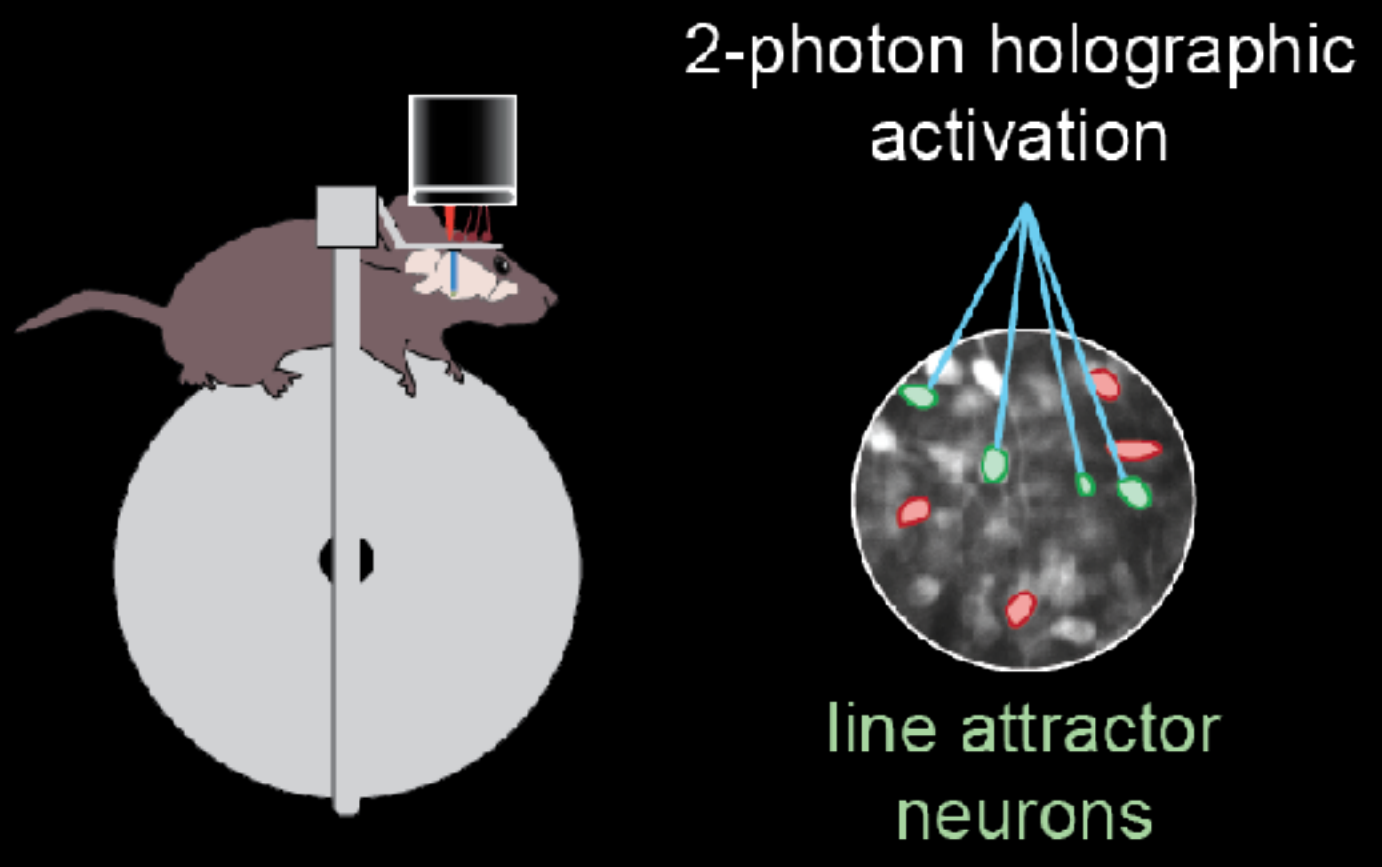
No study has causally demonstrated the existence of intrinsic line attractor dynamics in mammals.



Amit Vinograd

Vinograd, Nair et al. (Nature, 2024)

How can we gain access to the line attractor for perturbation?

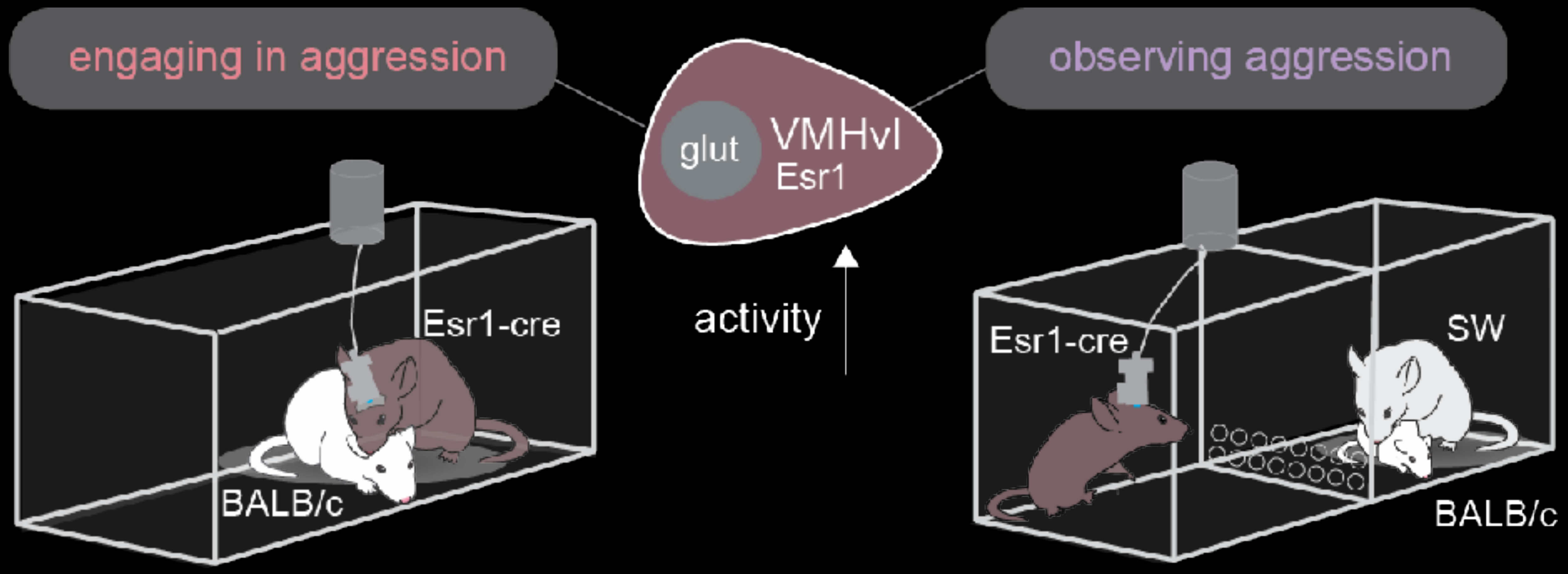


"dream experiment"

Unfortunately, head-fixation results in loss of attack behavior.

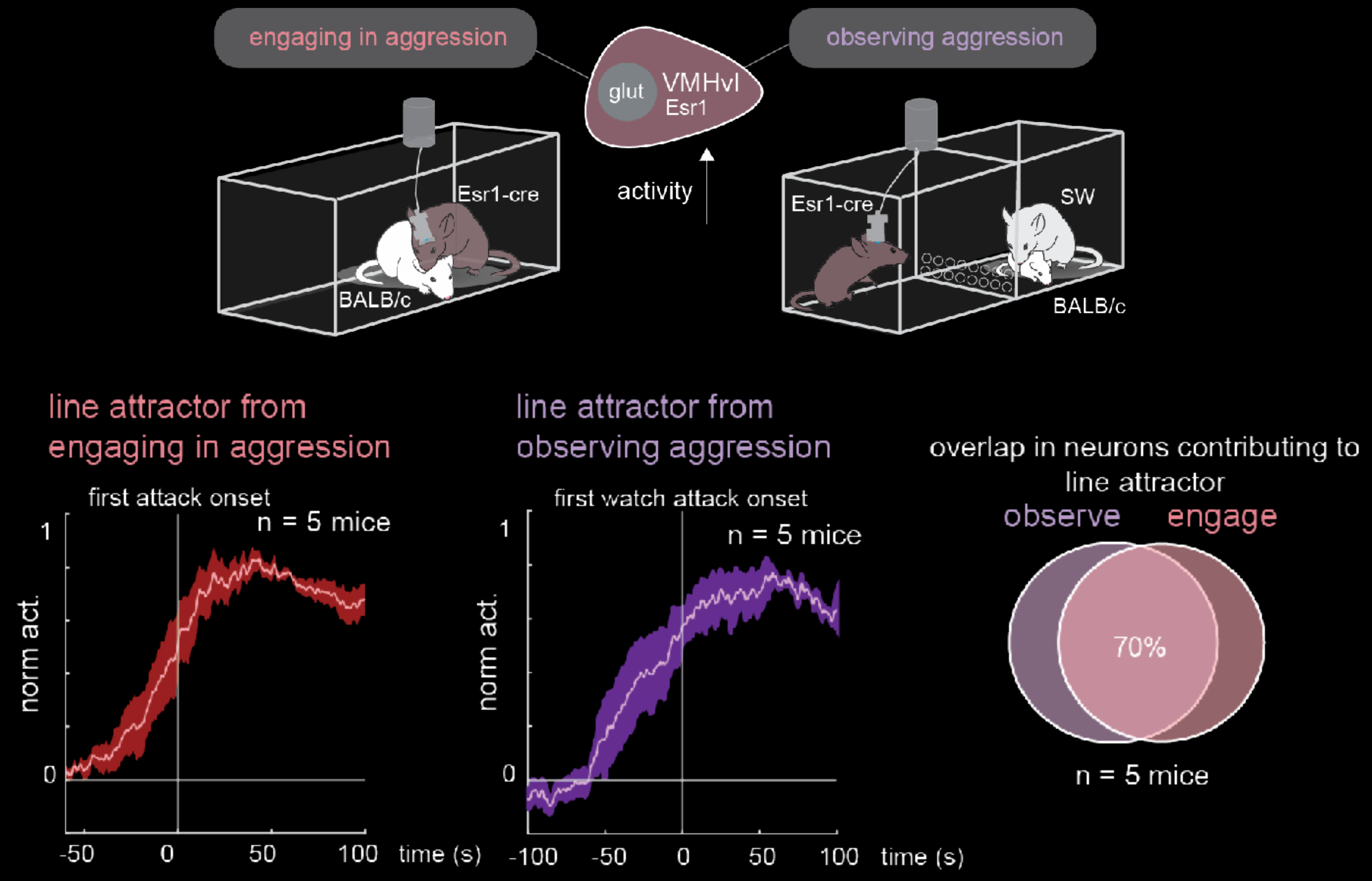
Vinograd, Nair et al. (Nature, 2024)

How can we gain access to the line attractor for perturbation?



VMHvl-Esr1 neurons are also active during observation of aggression
(Yang et al., Cell 2023)

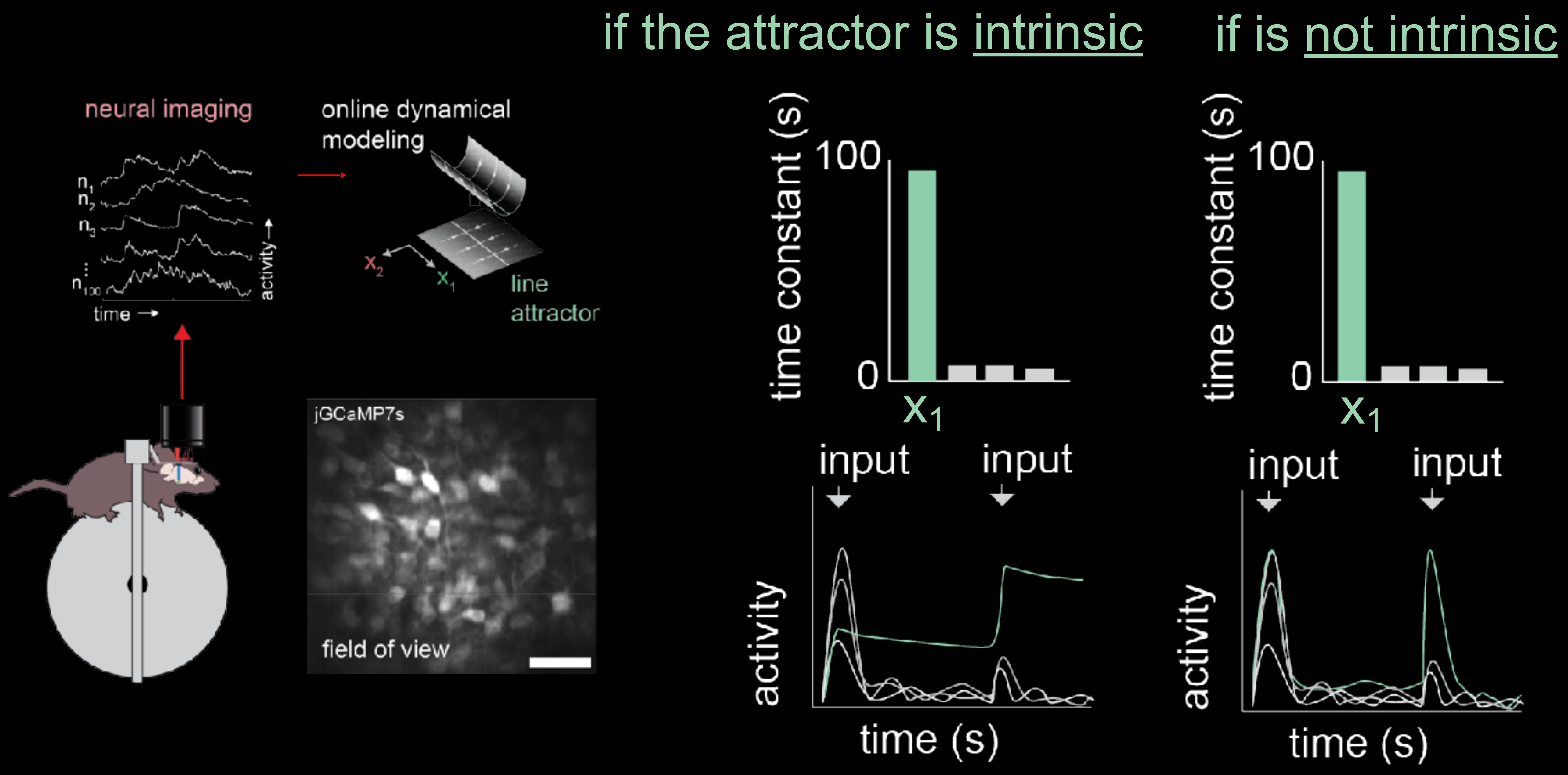
How can we gain access to the line attractor for perturbation?



VMHvl-Esr1 neurons are show line attractor dynamics during observation of aggression

Vinograd, Nair et al. (Nature, 2024)

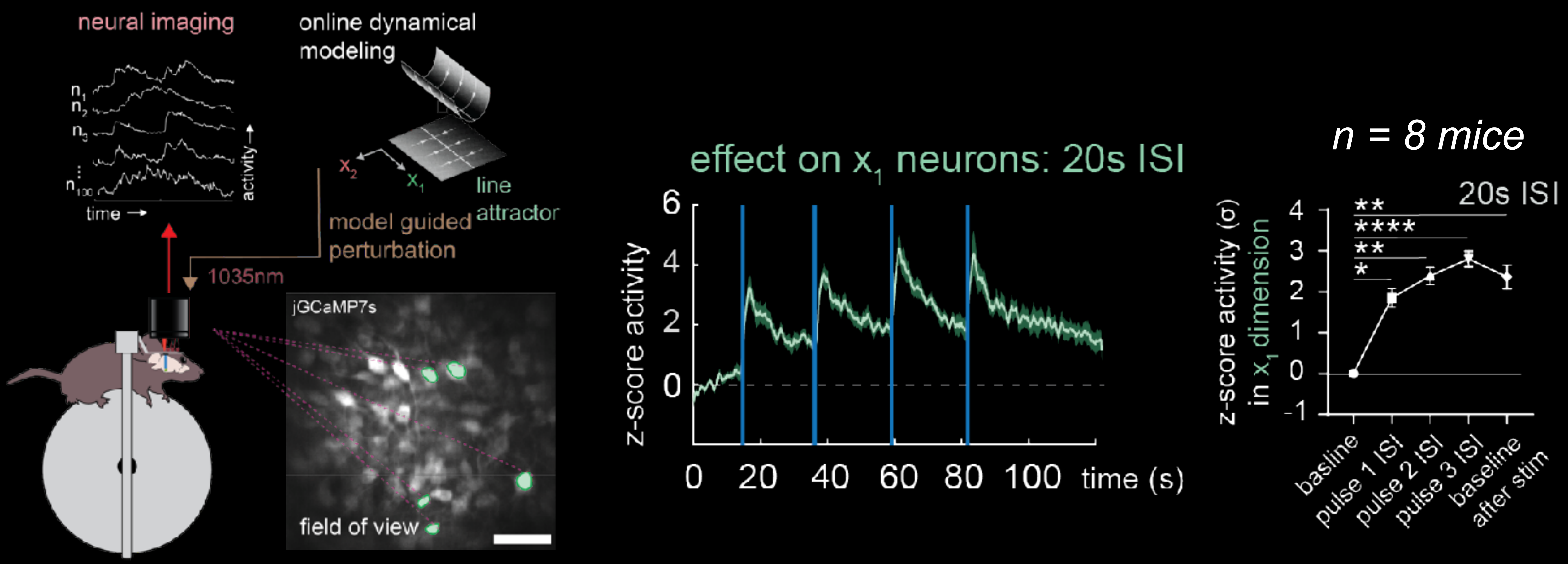
Closed-loop perturbation of dynamics in VMHvl



Activation of x_1 neurons should lead to integration if the line attractor is intrinsic.

Vinograd, Nair et al. (Nature, 2024)

Closed-loop perturbation of dynamics in VMHvl

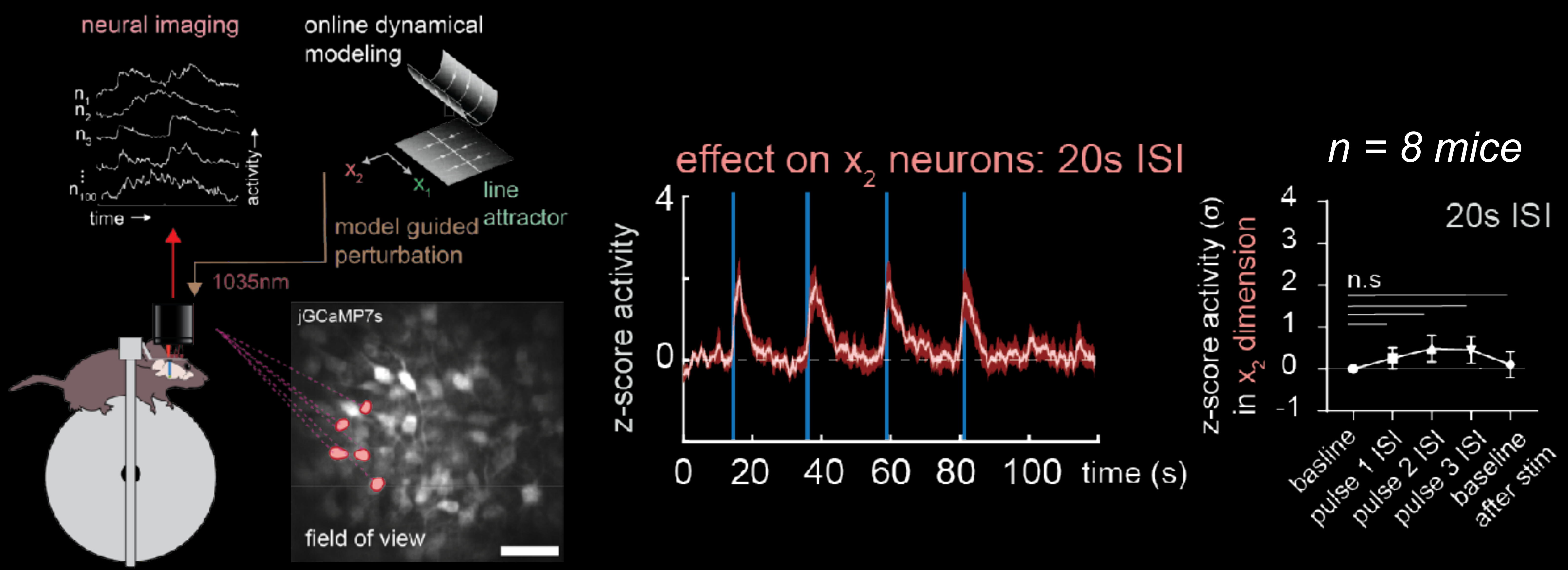


Holographic on-manifold activation* of line-attractor aligned x_1 neurons leads to integration.

**Note: this requires fitting an rSLDS online, during the session, to design the perturbation.*

Vinograd, Nair et al. (Nature, 2024)

Closed-loop perturbation of dynamics in VMHvl

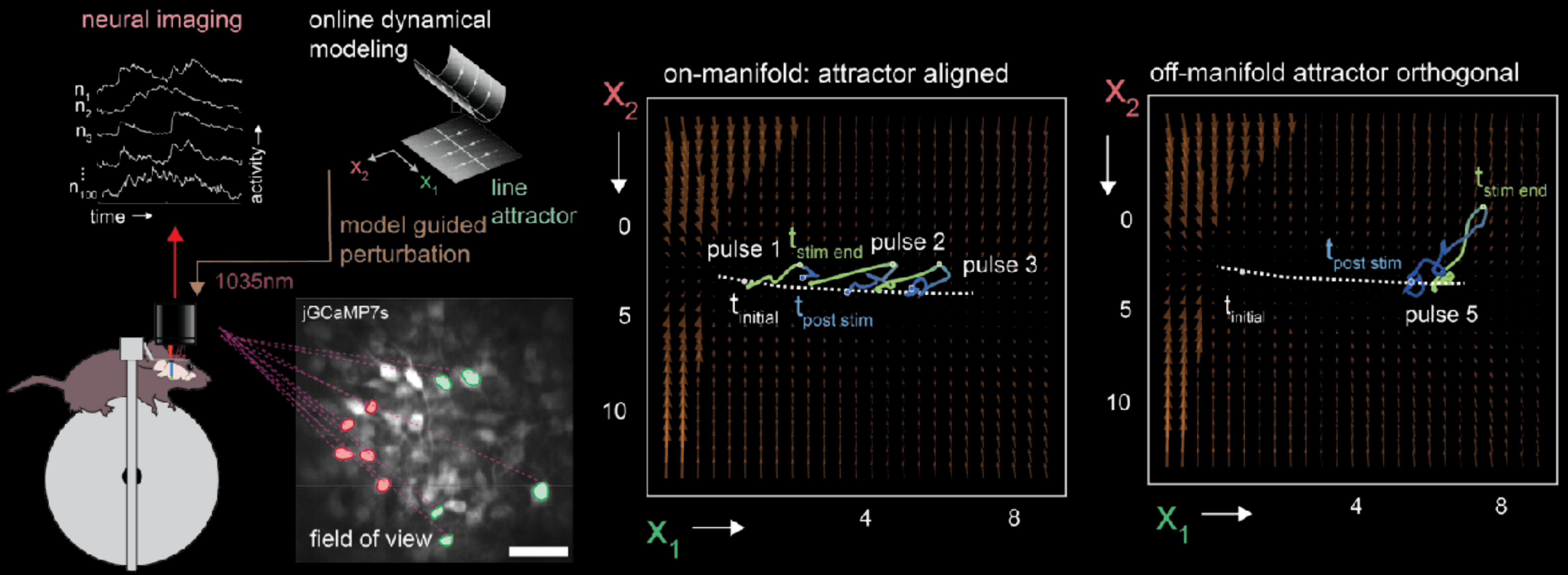


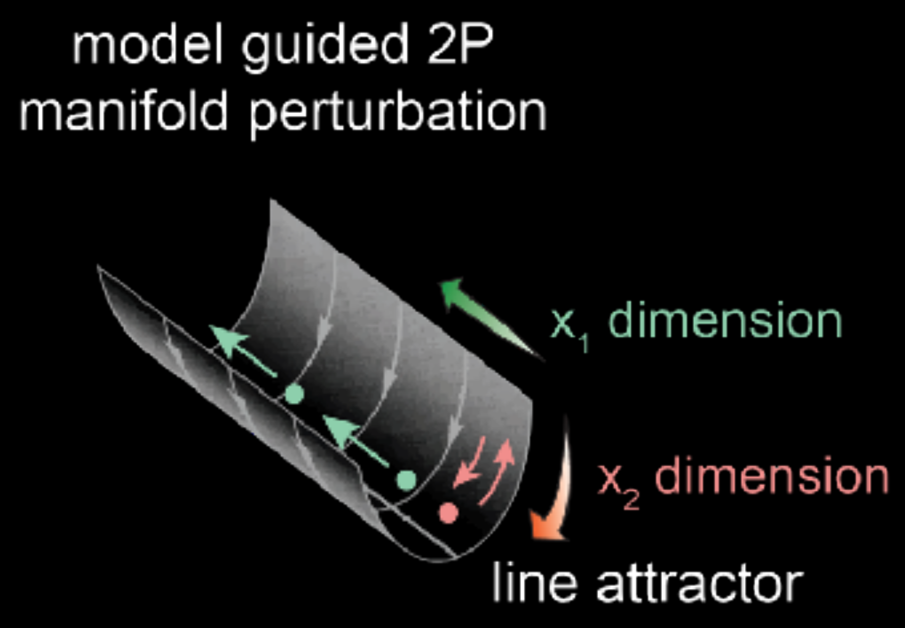
Holographic off-manifold activation of line-orthogonal x_2 neurons does not lead to integration.

Note: this requires fitting an rSLDS online, during the session, to design the perturbation.

Vinograd, Nair et al. (Nature, 2024)

Closed-loop perturbation of dynamics in VMHvl

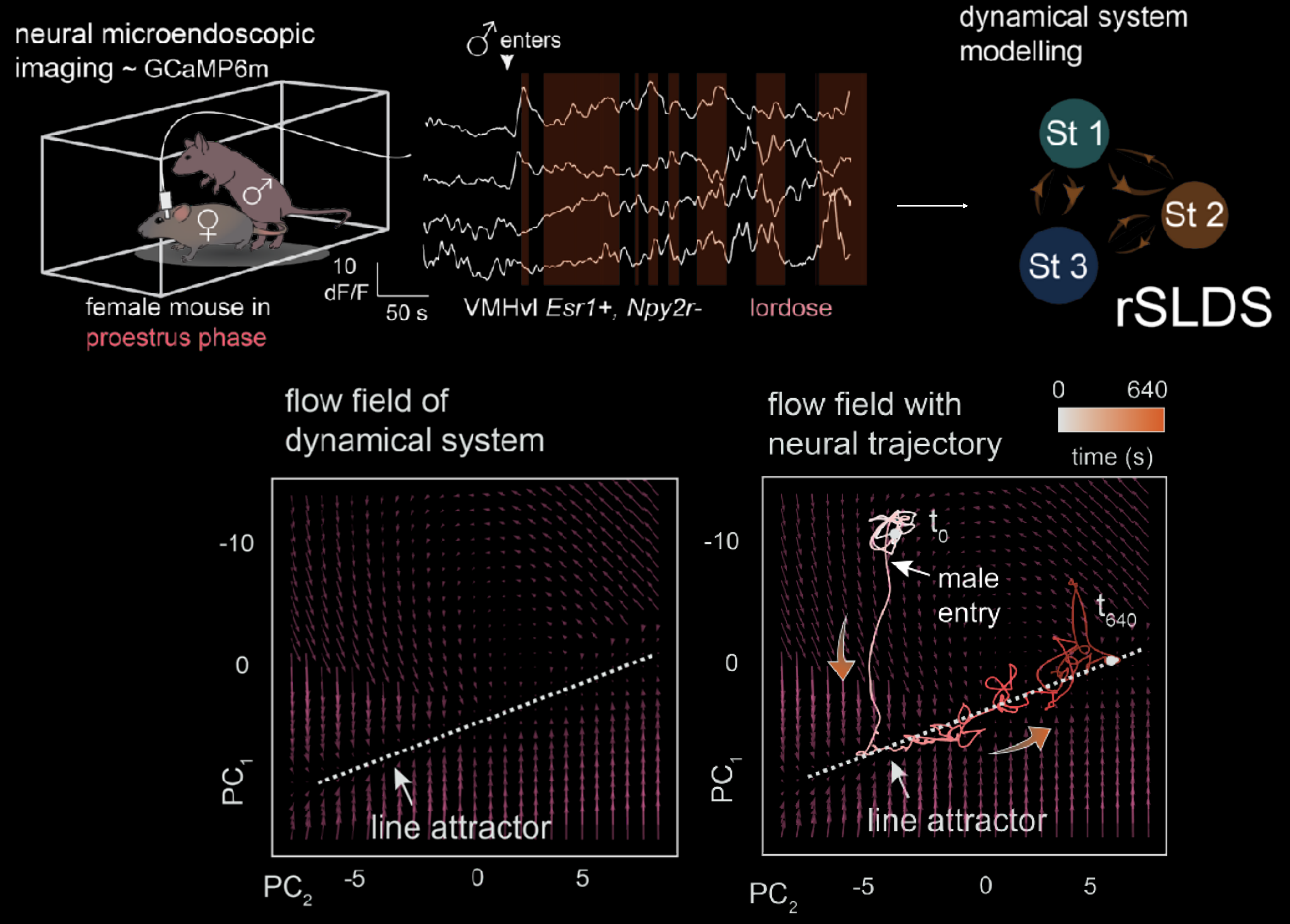




on- and off-manifold perturbations provide first evidence of an intrinsic mammalian line attractor

Vinograd, Nair et al. (Nature, 2024)

Do these attractor dynamics generalize to other internal state computations?



VMHvl shows attractor dynamics in female mice during mating, but only in proestrus.

Liu, Nair et al. (Nature, 2024)