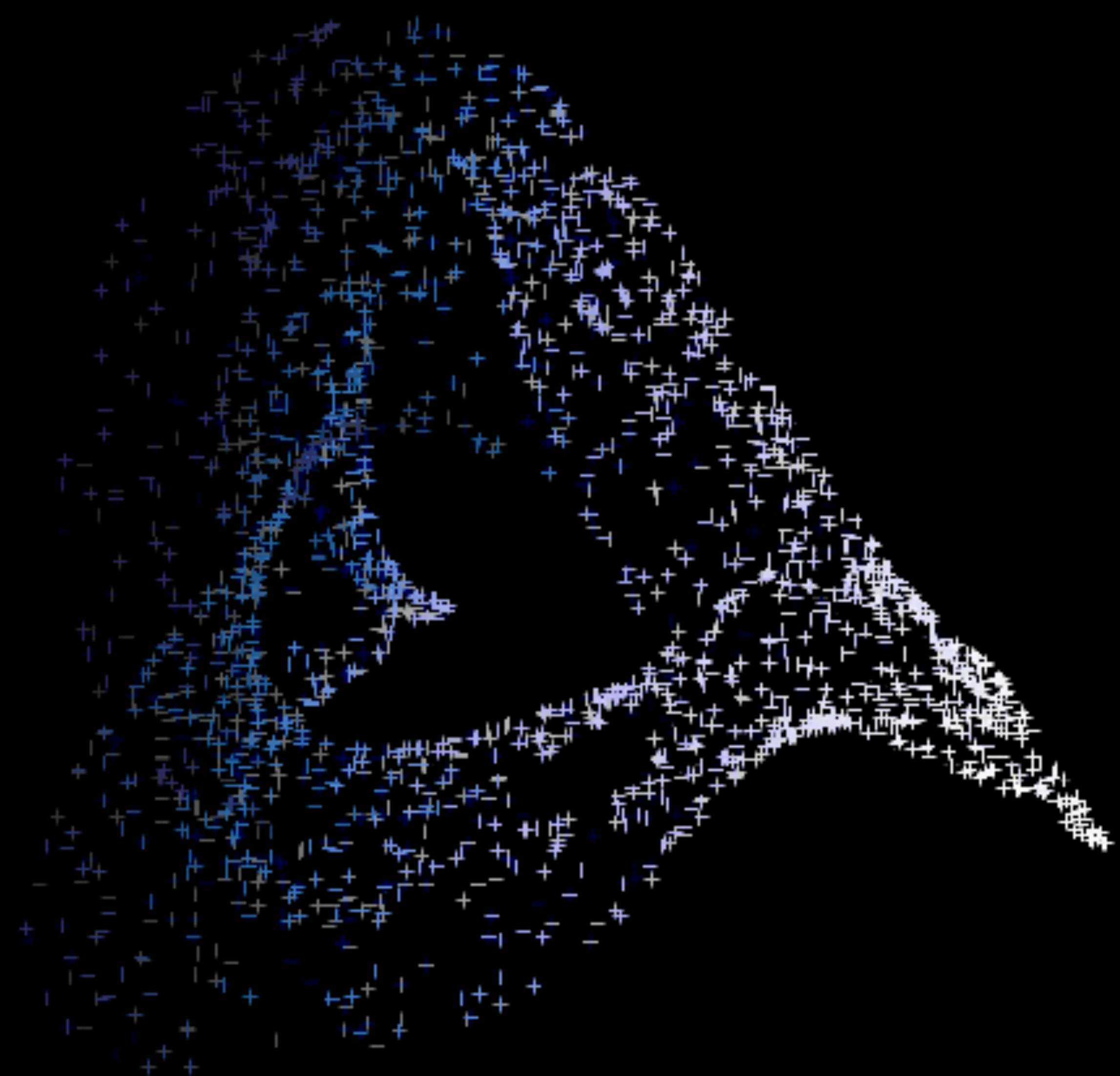
Distributional Control of Ensemble Systems

AFOSR D&C Program Review
Jr-Shin Li
Washington University in St. Louis

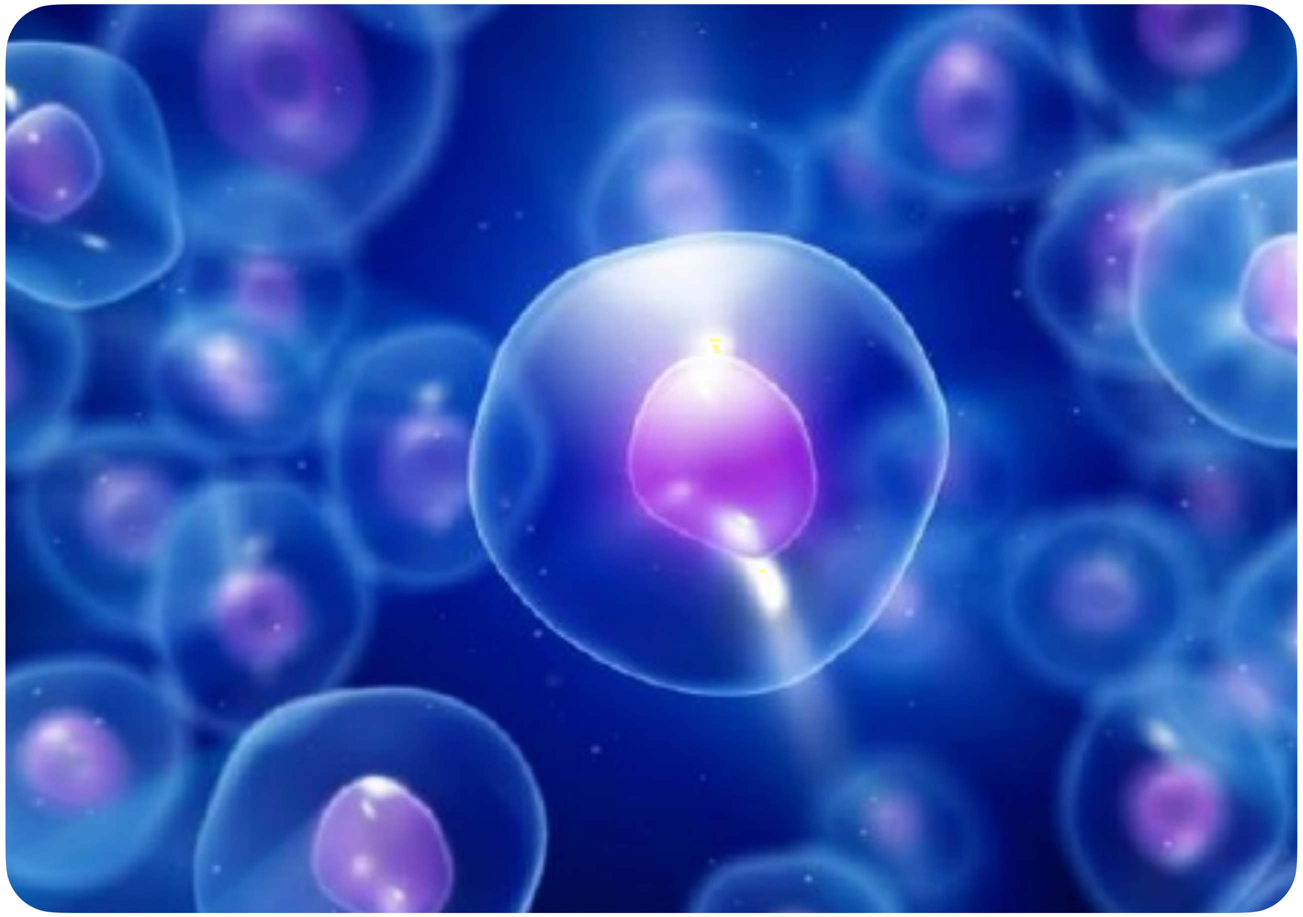
August 28, 2024

Ensemble Systems

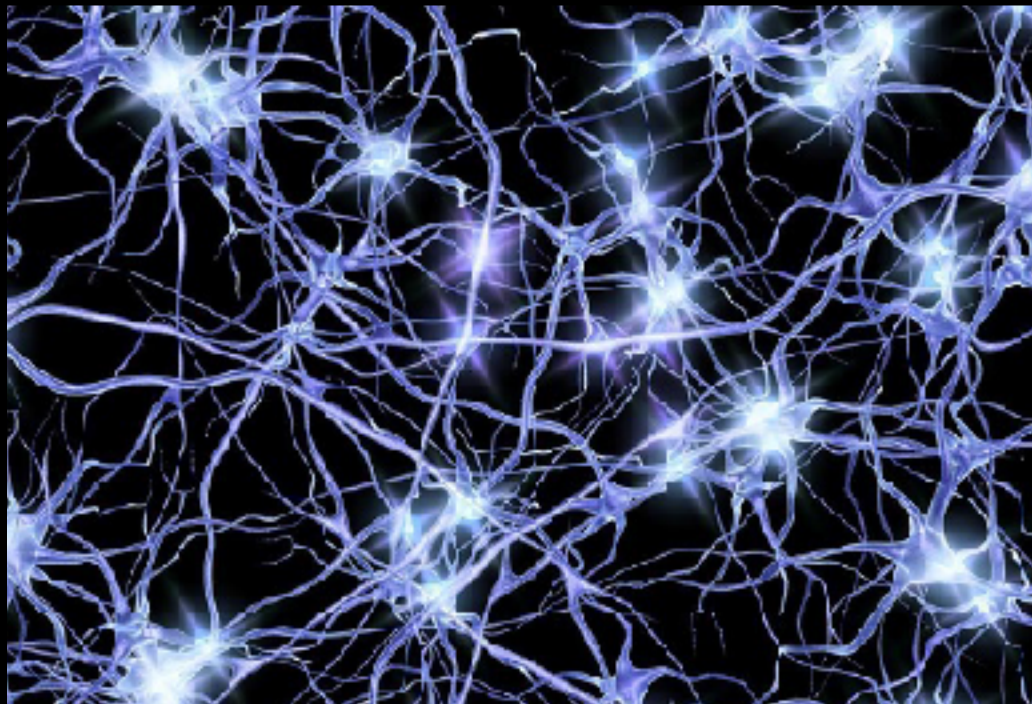


- Population systems consisting of a large number of structurally similar dynamic units
  - Finitely or infinitely many
  - Isolated or interconnected agents

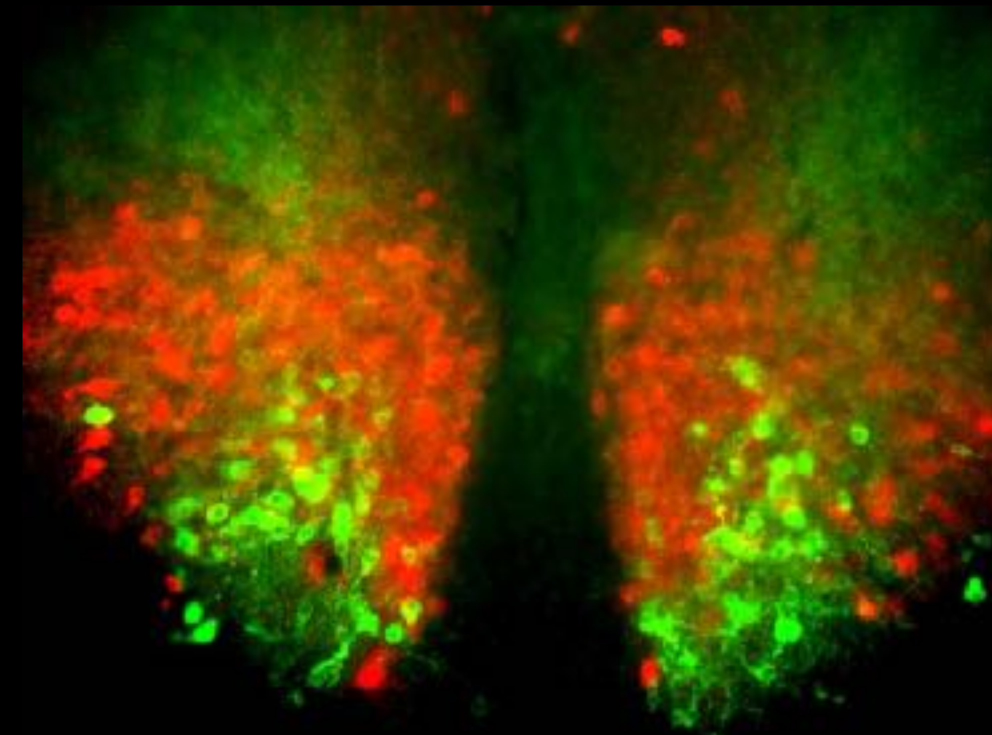
Large-Scale Dynamic Populations



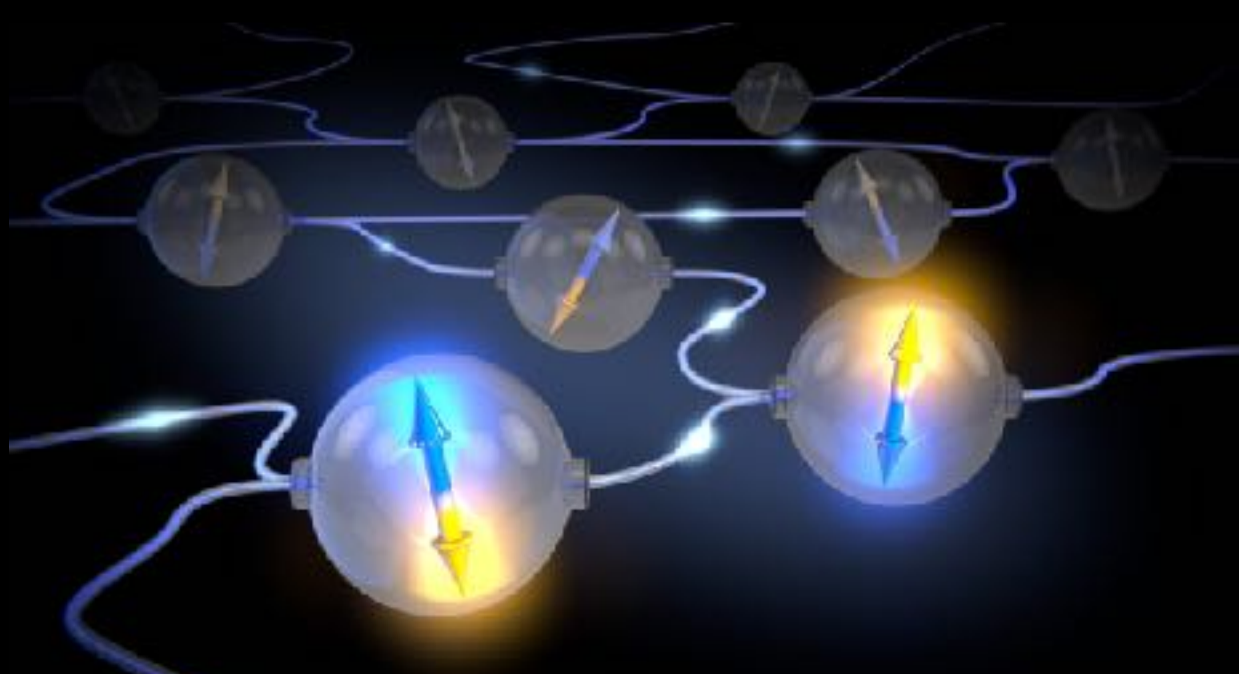
Large-Scale Dynamic Networks



Neuronal networks



Biological networks



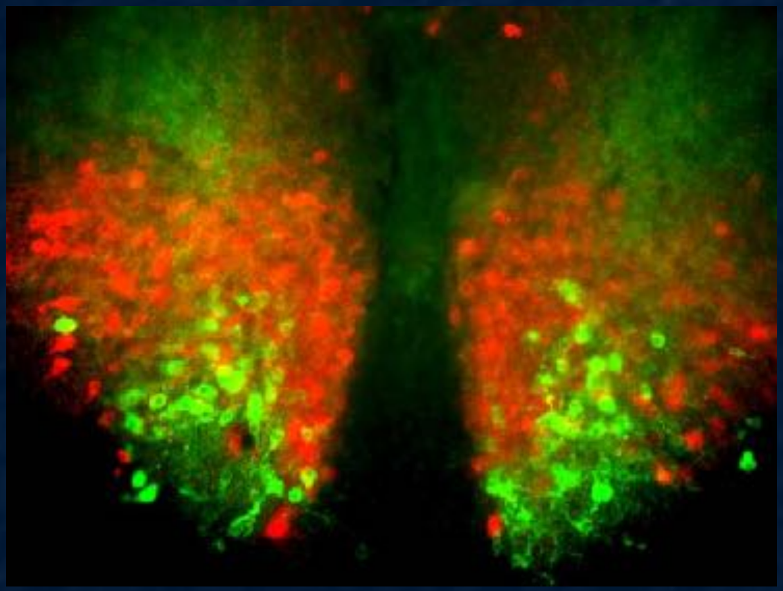
Quantum networks

(Source: QuTech)

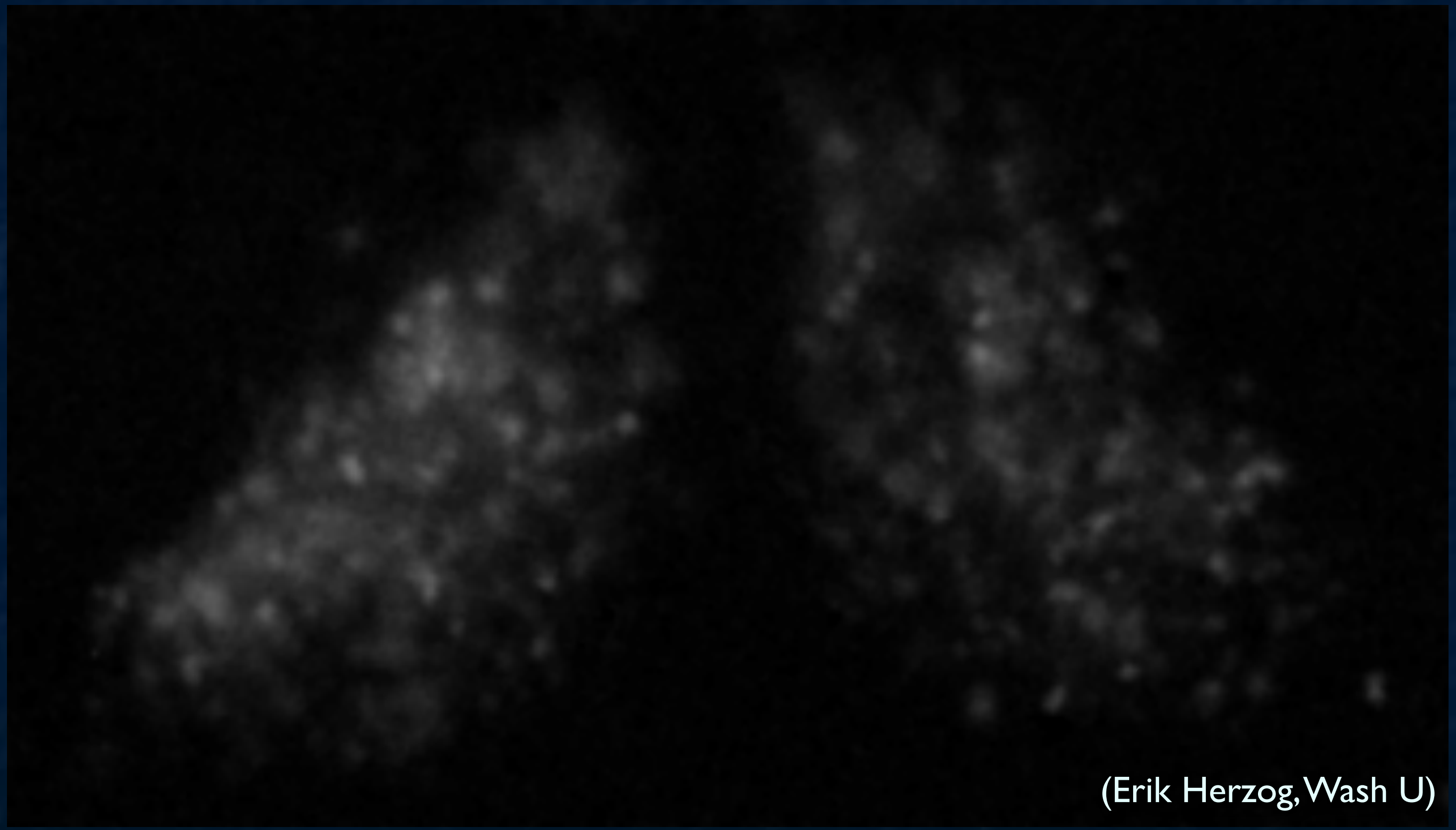
Social networks

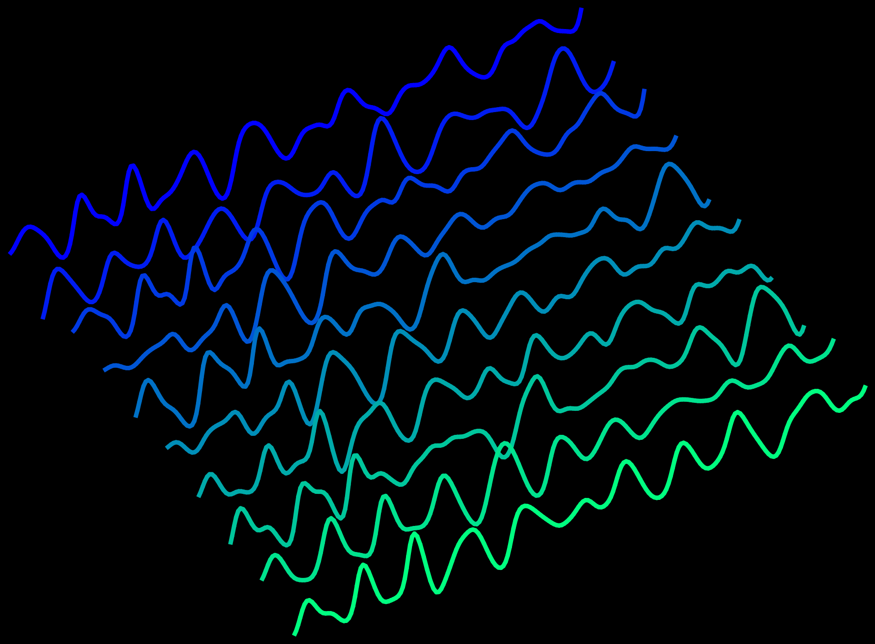
Population Data

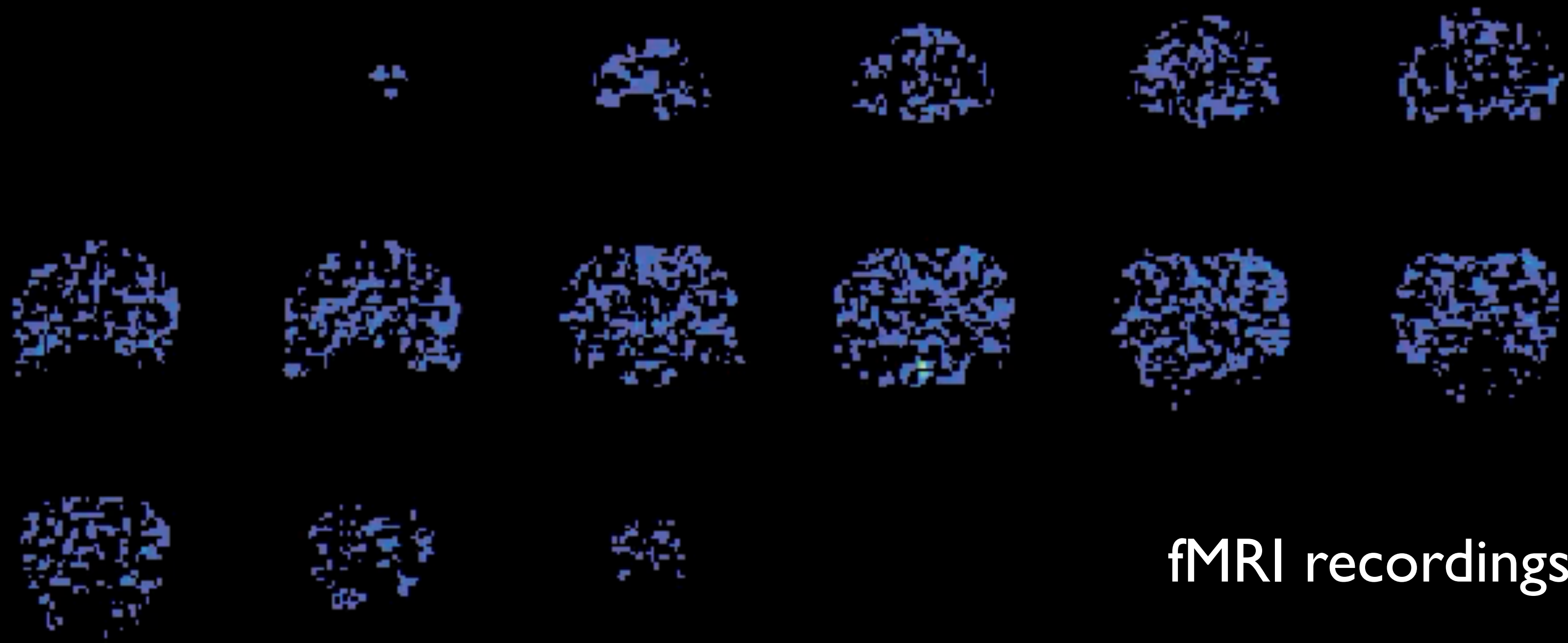
A clock in a dish: The SCN in vitro



Circadian clocks







fMRI recordings

Outline

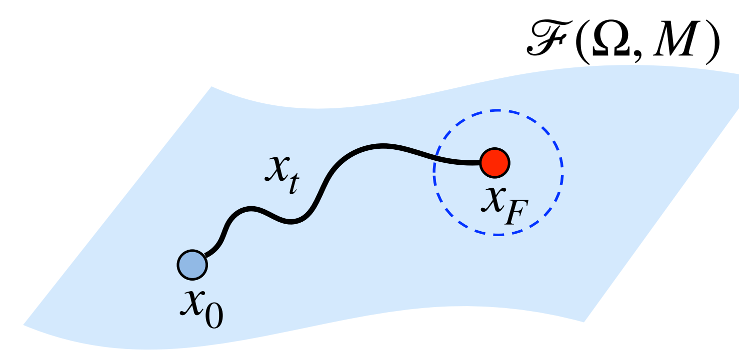
- Distributional control of ensemble systems
  - Distributional information inherent in aggregated measurements
  - Controlling output measures induced by aggregated measurements
- Dual representations generated by moment kernelization
- Link of ensemble control with optimal transport
- Distributional control on fibre space
- Ensemble control systems and representation learning

Control of Ensemble Systems

Ensemble system

$$\frac{d}{dt}x(t,\beta) = F\big(t,\beta,x(t,\beta),u(t)\big), \qquad \beta \in \Omega \subset \mathbb{R}^d$$

$$x(t,\beta) \in M \subset \mathbb{R}^n, \quad u(t) \in \mathbb{R}^m, \qquad \Omega \text{ is compact}$$



$$x_t \doteq x(t,\cdot\,) \in \mathscr{F}(\Omega, M)$$

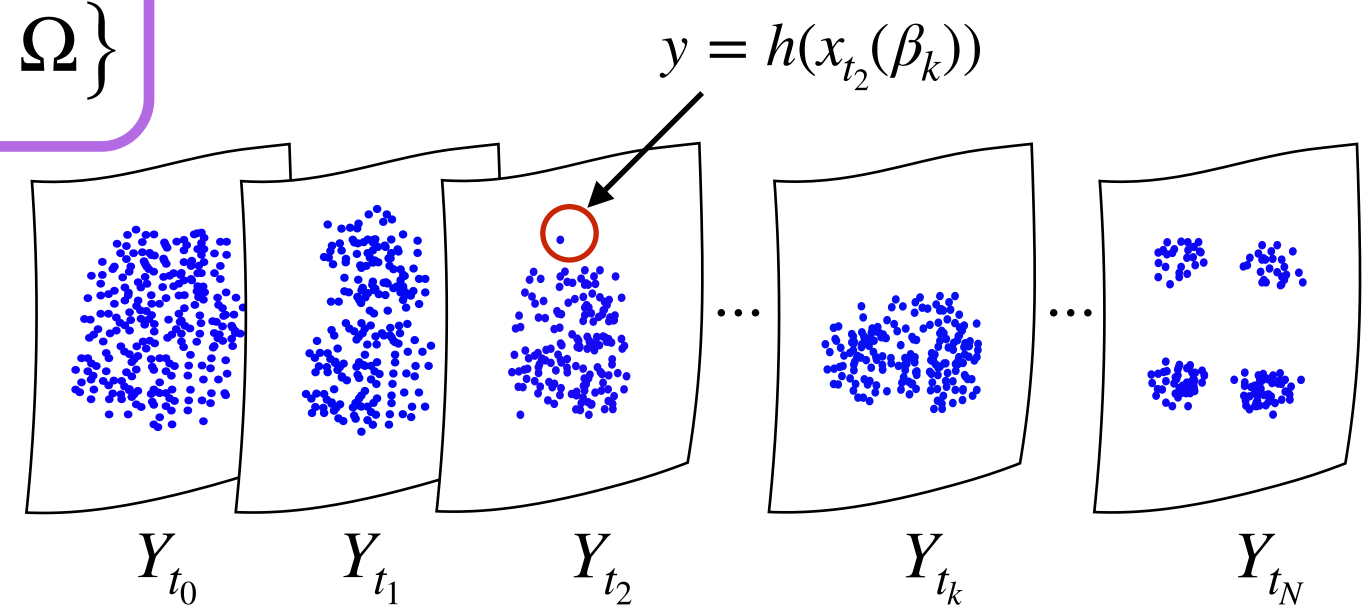
$$x_t : \Omega \to M$$

Aggregated measurements

$$\begin{aligned} Y_t &= y_t(\Omega) = h \circ x_t(\Omega) \\ &= \big\{ y \in \mathbb{R}^r : y = h(x(t,\beta)), \beta \in \Omega \big\} \end{aligned}$$

Set-valued measurements

$y_t \sim \mu_t$ Probability measure

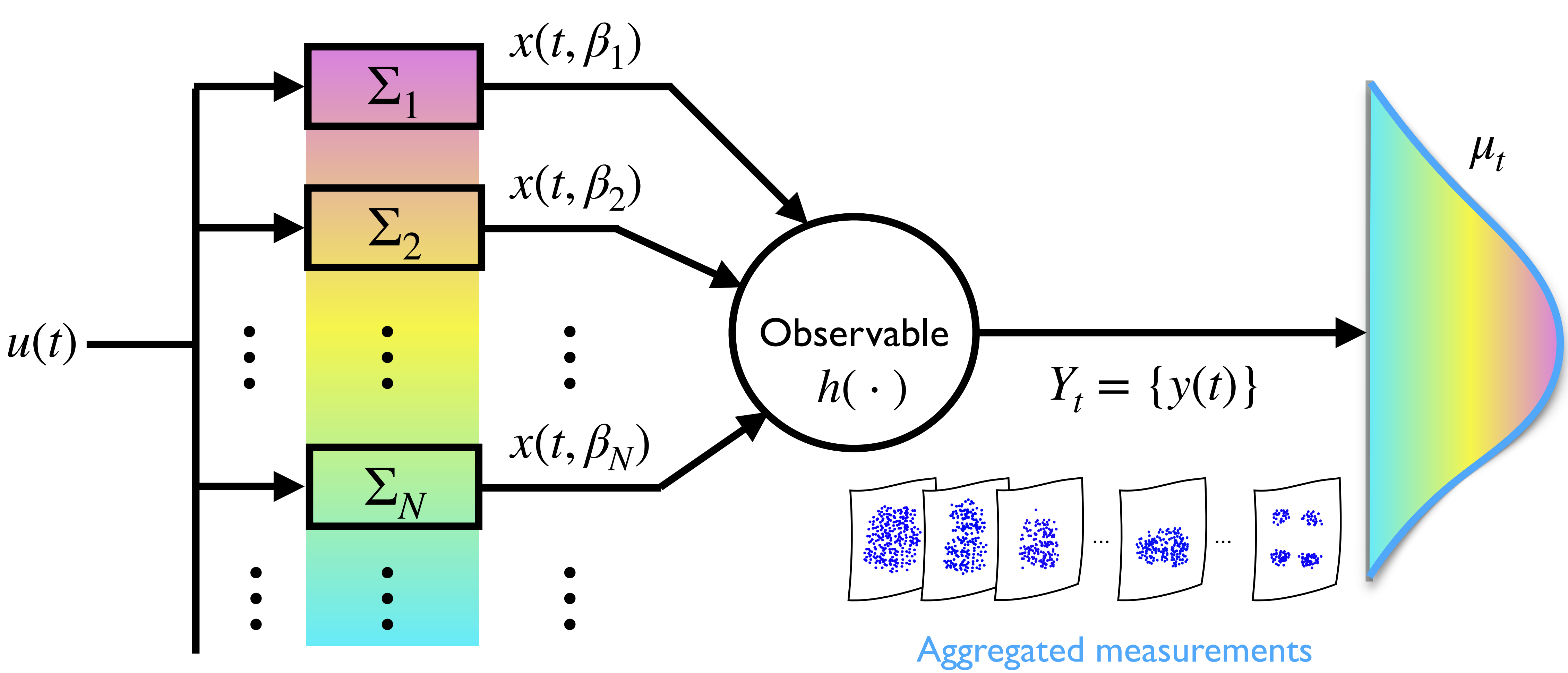


Control of Ensemble Distributions

- Shaping ensemble “distributions”

$$\frac{d}{dt}x(t,\beta) = F\big(t,\beta,x(t,\beta),u(t)\big), \quad \beta \in \Omega \subset \mathbb{R}^d$$

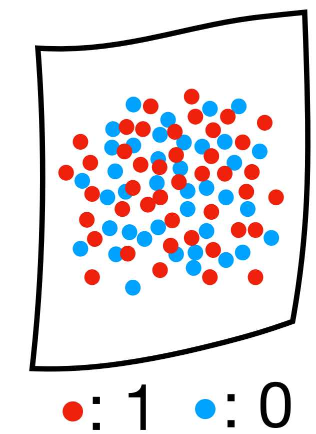
$$Y_t = y_t(\Omega) = h \circ x_t(\Omega)$$



Distributional Control of Ensembles

- A pedagogic example

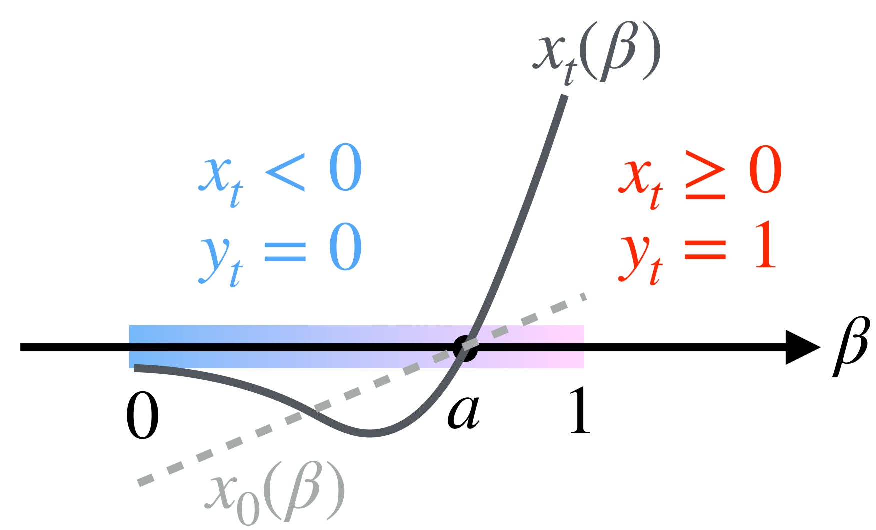
$$\begin{cases} \dfrac{d}{dt}x(t,\beta) = \beta x(t,\beta) + u(t), \quad \beta \in [0,1], \quad x(t,\beta) \in \mathbb{R} \\[2ex] y_t(\beta) = h(x_t(\beta)) = \begin{cases} 0, & \text{if } x_t(\beta) < 0 \\ 1, & \text{if } x_t(\beta) \geq 0 \end{cases} \end{cases} \quad \Rightarrow \quad Y_t = \{0,1\}$$



- $x_0(\beta) = \beta - a, \quad u(t) = 0$

$$x_t(\beta) = e^{t\beta}x_0(\beta) \begin{cases} < 0, & \beta \in [0,a) \\ \geq 0, & \beta \in [a,1] \end{cases}$$

$$\Rightarrow \quad y_t = \mathbf{1}_{[a,1]}$$



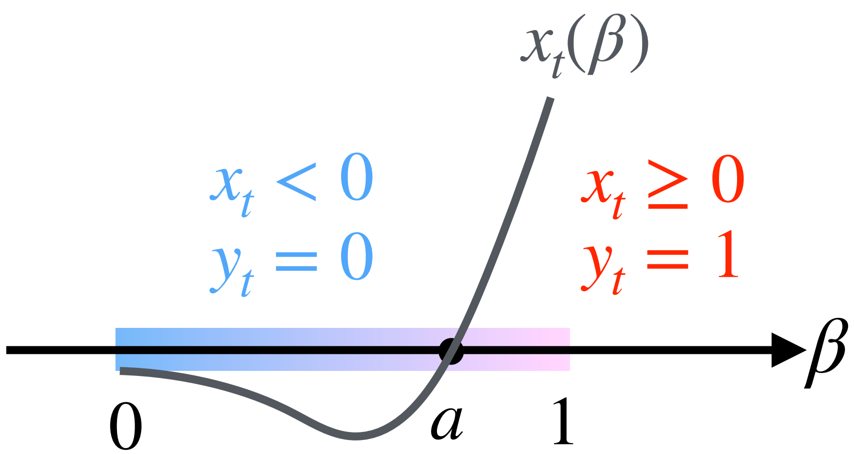
$$\mu_t(\{0\}) = a$$

$$\mu_t(\{1\}) = 1 - a$$

$$y_t \sim \mu_t : \text{Bernoulli}(a)$$

Distributional Control of Ensembles

- Mathematical formalism

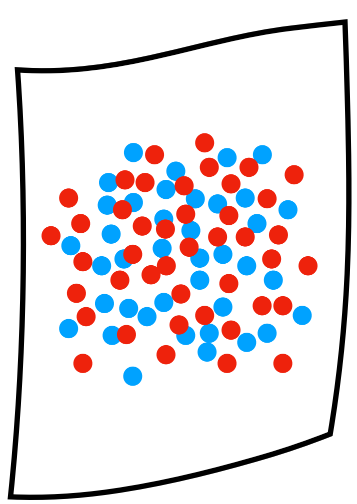


$Y_t = \{0,1\}, \quad y_t = \mathbf{1}_{[a,1]}$

$\Omega = [0,1], \quad \lambda$ Lebesgue measure

$$\mu_t(\{0\}) = \lambda([0,a)) = \lambda(y_t^{-1}(0)) = \big((y_t)_{\#}\lambda\big)(\{0\})$$

$$\mu_t(\{1\}) = \lambda([a,1]) = \lambda(y_t^{-1}(1)) = \big((y_t)_{\#}\lambda\big)(\{1\})$$



$$\boxed{\mu_t = (y_t)_{\#}\lambda}$$

$$d\lambda = \frac{\omega}{\mathrm{vol}(\Omega)}$$

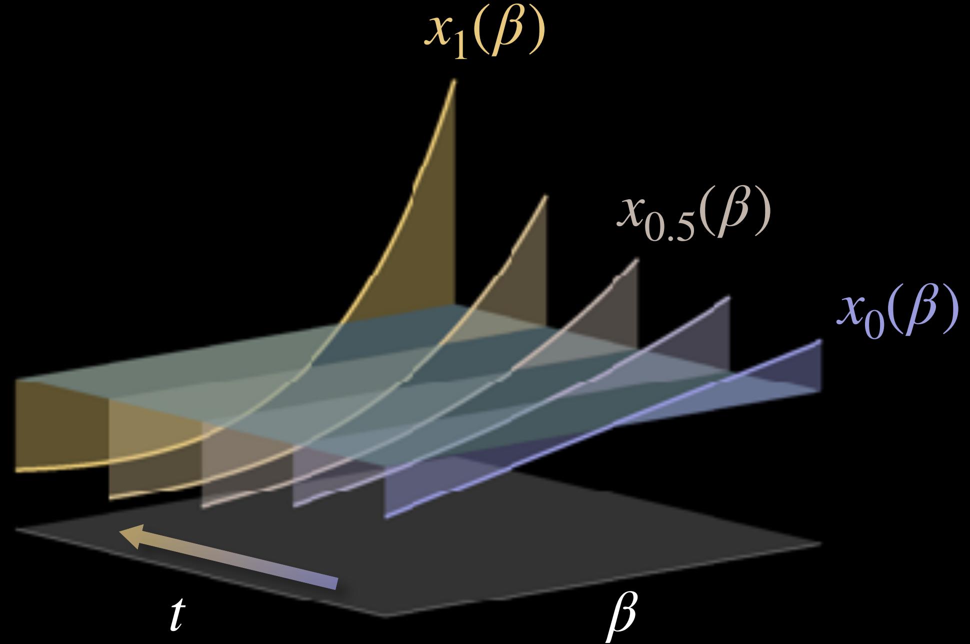
$$\begin{cases} \mu_t(B) = \lambda(y_t^{-1}(B)) \\ \displaystyle\int_{\mathbb{R}^r} f\,d\mu_t = \int_{\Omega} f \circ y_t\,d\lambda \end{cases}$$

Distributional control ⟷ Transport of probability measures

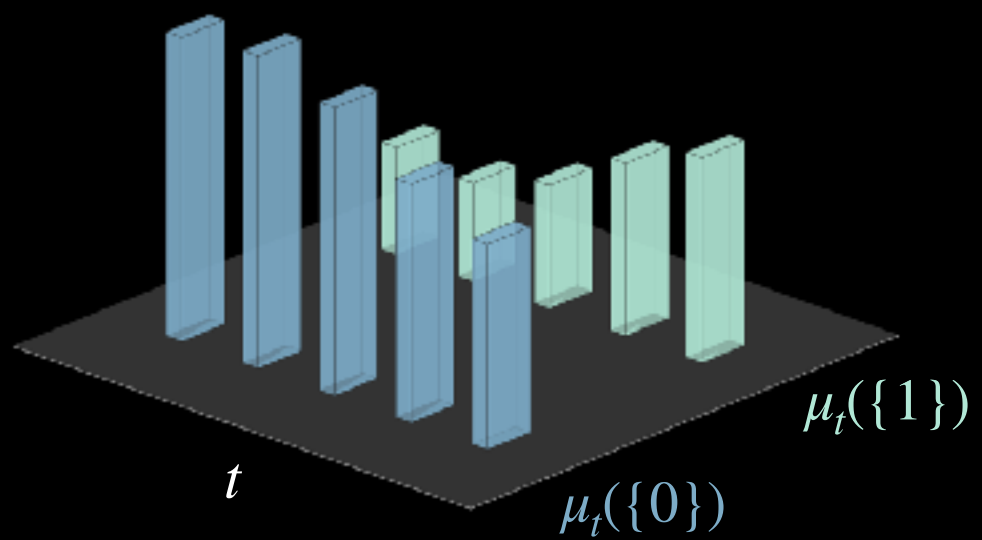
Distributional Control of Ensembles

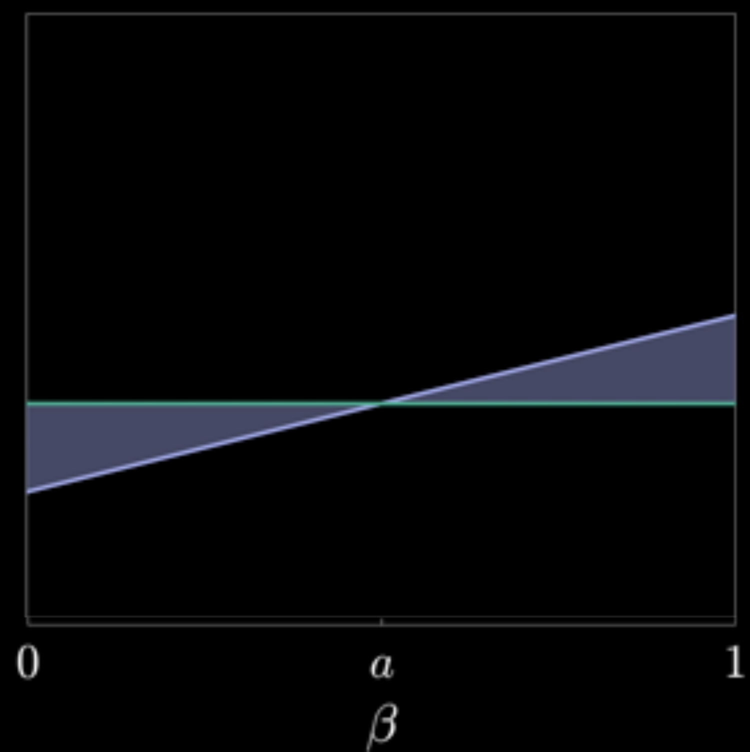
$$\frac{d}{dt}x_t(\beta) = \beta x_t(\beta) + u(t), \quad \beta \in [0,1]$$
$$y_t(\beta) = h(x(t,\beta))$$

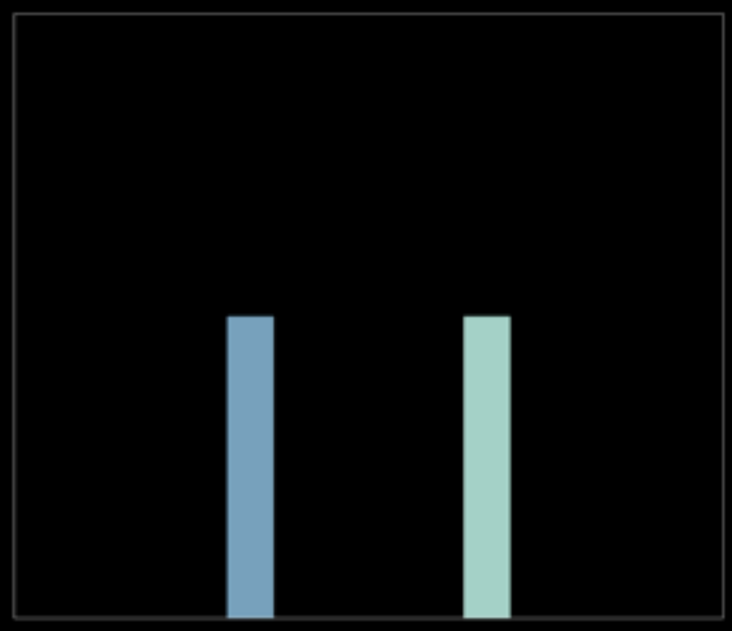
$$x_0(\beta) = \beta - 0.5$$
$$u(t) = -\sin(t)\cos(t)$$



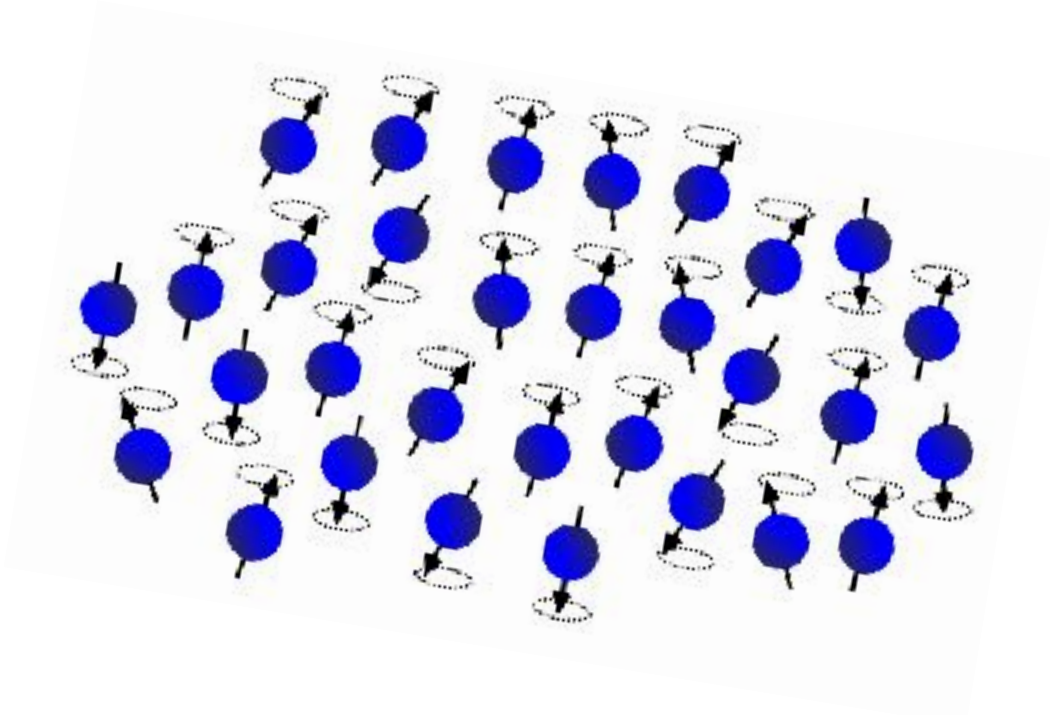
$(y_t)_{\#}\lambda$



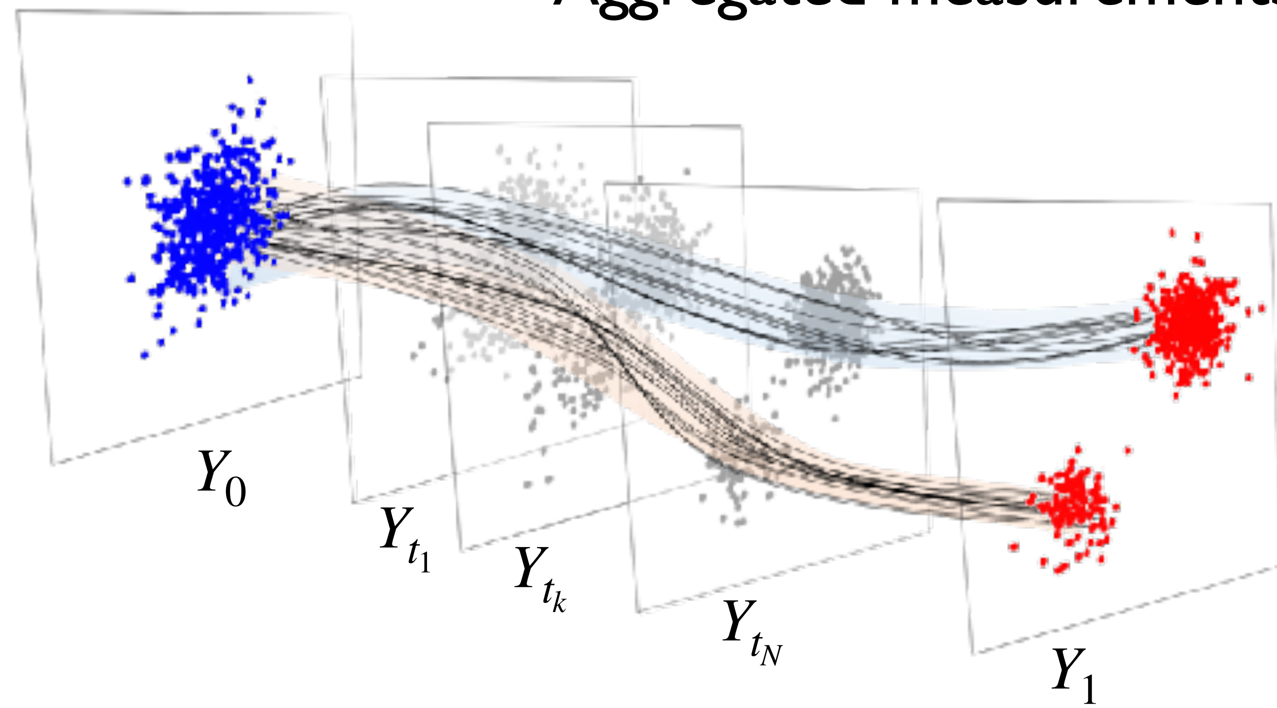




Distributional Control via Optimal Transport



Aggregated measurements



$$\frac{d}{dt}x(t,\beta) = F(t,\beta,x(t,\beta),u(t))$$

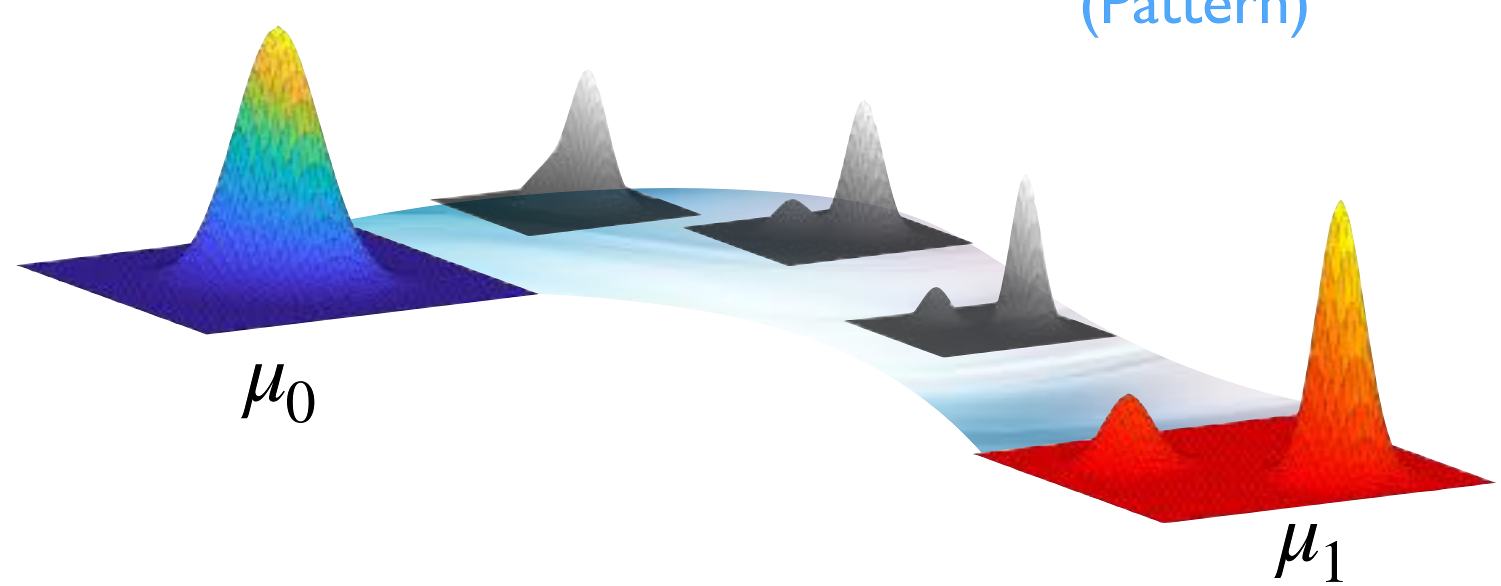
$$Y_t = y_t(\Omega) = h \circ x_t(\Omega)$$

μ_0 μ_1

$\mu_t = (y_t)_{\#}\lambda$

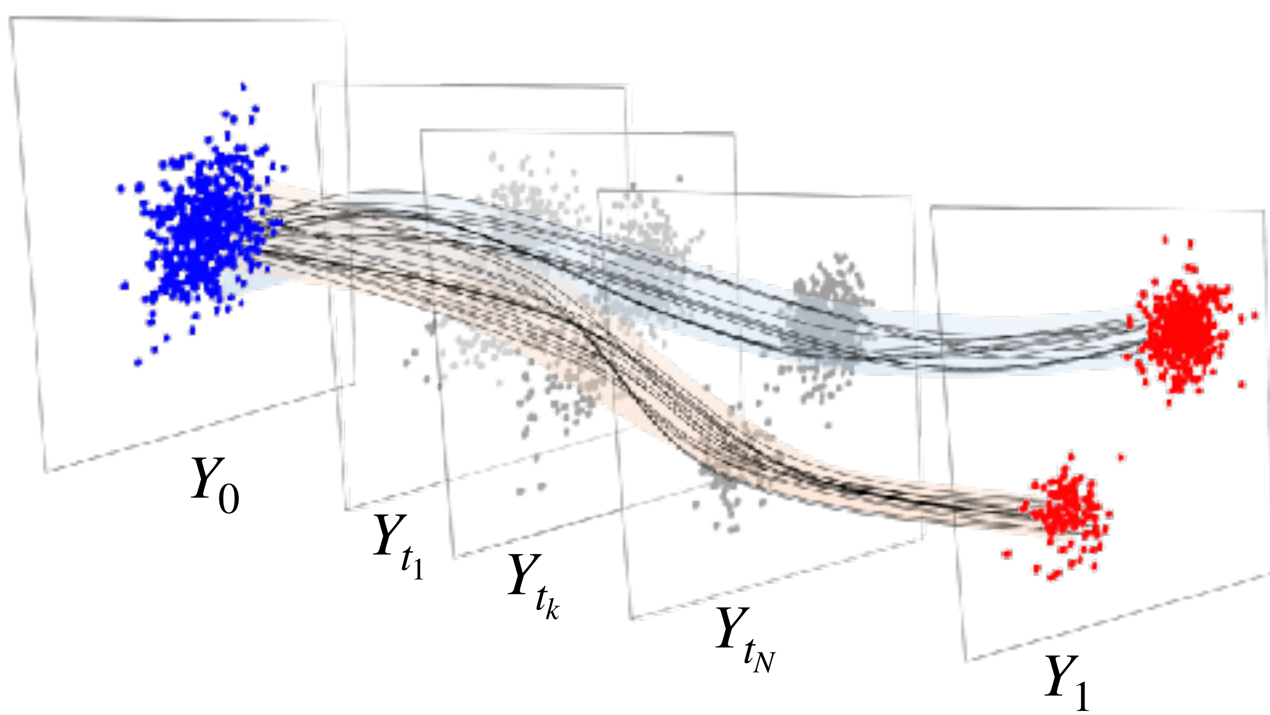
Output distribution
(Pattern)

Ensemble control ⟷ Optimal transport



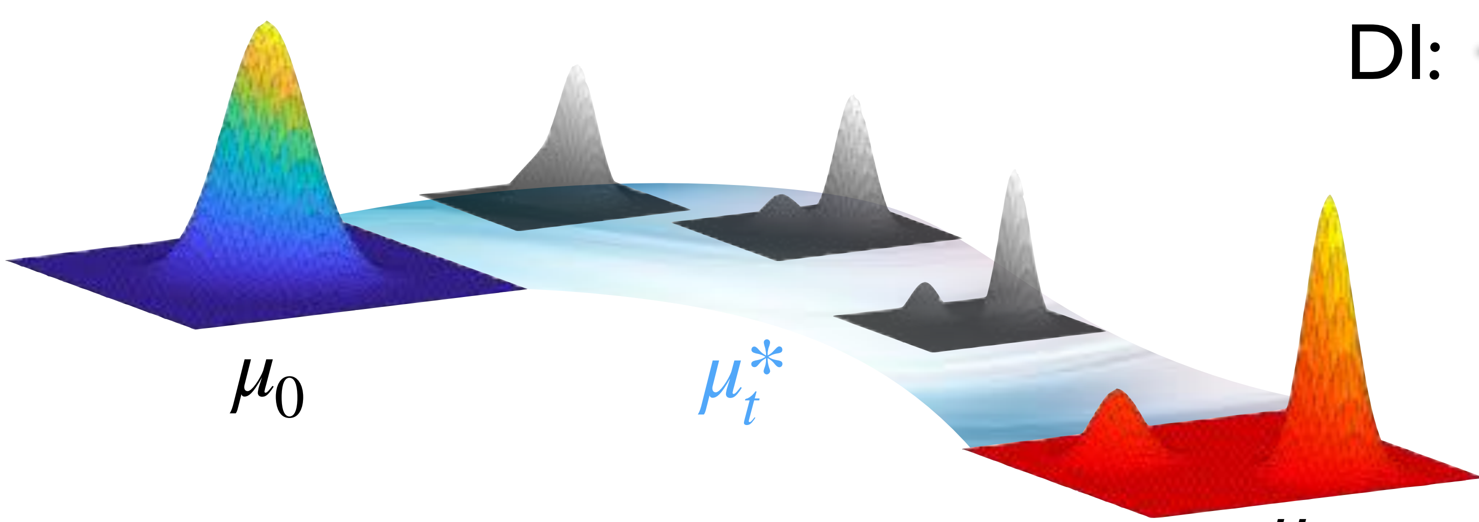
Time-Dependent Optimal Transport

- Displacement interpolation (DI)



$\mu_0 \longrightarrow \mu_1$

$\mu_t \approx \mu_t^*$



OT: $\begin{cases} \displaystyle\min_{\Phi:\mathbb{R}^r \to \mathbb{R}^r} \int_{\mathbb{R}^r} c(y, \Phi(y))d\mu_0(y) \\ \text{s.t.} \quad \mu_1 = \Phi_\# \mu_0 \end{cases}$

Interpolating (μ_0, μ_1)

DI: $\begin{cases} \displaystyle\min_{\{\Phi_t:\mathbb{R}^r \to \mathbb{R}^r\}_{0 \le t \le 1}} \int_{\mathbb{R}^r} C(\Phi_t(y))d\mu_0(y) \\ \text{s.t.} \quad \Phi_0 = I, \quad (\Phi_1)_\# \mu_0 = \mu_1 \end{cases}$

$$\mu_t^* = \big((1-t)I + t\Phi\big)_\# \mu_0$$

Optimal Transport for Ensemble Control

- Distributional control as a tracking problem

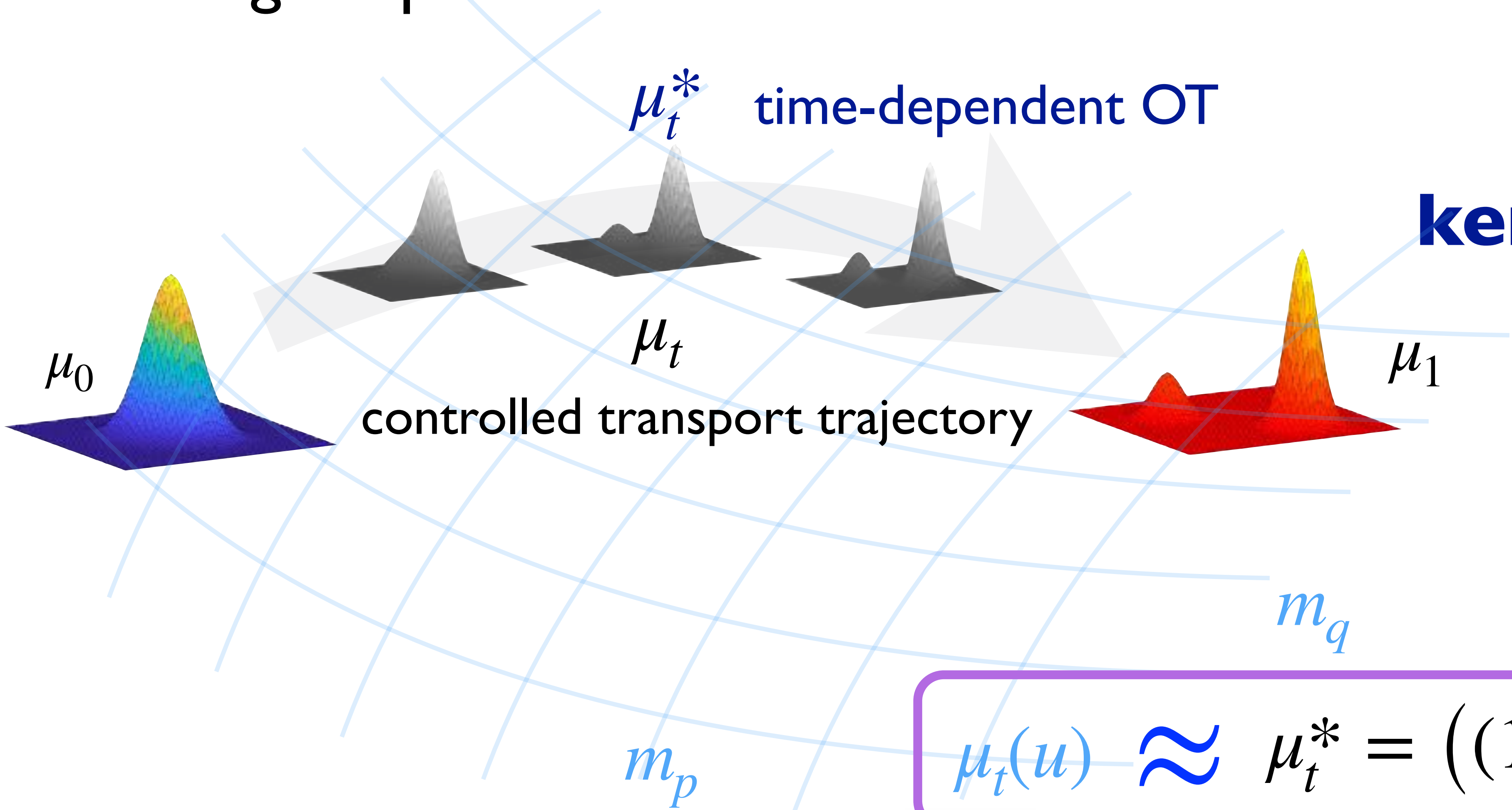
$$\frac{d}{dt}x(t,\beta) = F(t,\beta,x(t,\beta),u(t))$$

$$Y_t = y_t(\Omega) = h \circ x_t(\Omega)$$

Ensemble distribution

$$\mu_t = (y_t)_{\#}\lambda$$

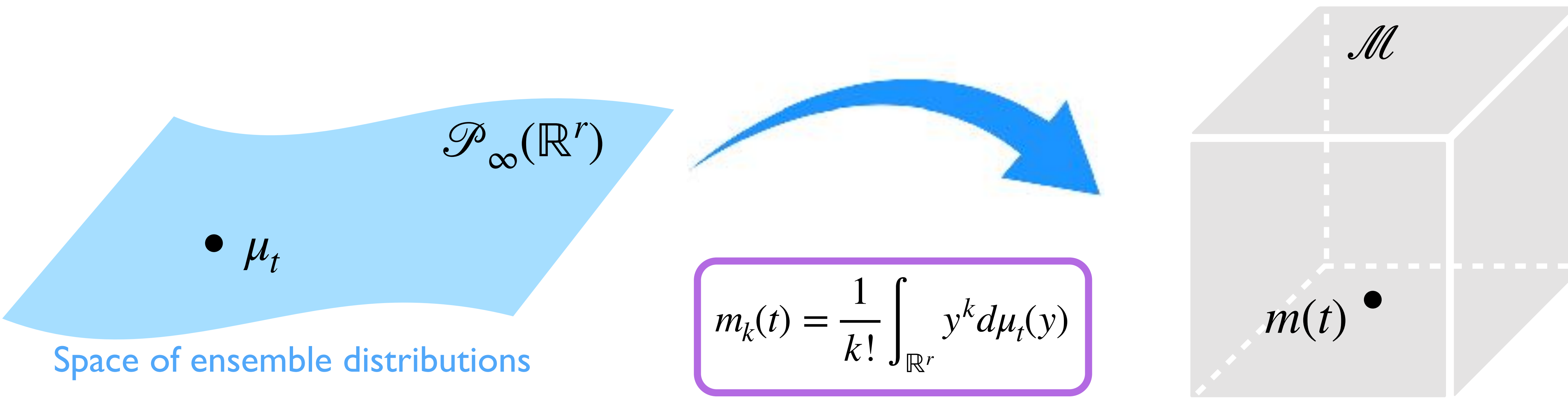
- Steering output distributions



Moment kernelization

$$\mu_t(u) \approx \mu_t^* = \big((1-t)I + t\Phi\big)_{\#}\mu_0$$

Moment Kernel



Monomial moments

Ensemble distributions

$$\mu_t = (y_t)_{\#}\lambda, \quad y_t = h \circ x_t$$

Moment space

$$m(t) : \mathbb{N}^r \to \mathbb{R}$$

- k-th ensemble output moment

$$m_k(t) = \langle \varphi_k, \mu_t \rangle, \quad \{\varphi_k\}_{k \in \mathbb{N}^r} \text{ dual set of a basis of } \mathscr{P}_\infty(\mathbb{R}^r)$$

- Reproducing kernel induced by the moment transform

$$\mathscr{K} : \mathbb{N}^r \times \mathbb{N}^r \to \mathbb{R}, \ (i,j) \mapsto k_j(i) = \langle k_i, k_j \rangle$$

Moment Kernelization

- Moment transform

$$d\mu(y) = f(y)dy$$

$$f(y) = \frac{1}{\sqrt{2\sigma^2}} e^{-\frac{y^2}{2\sigma^2}}$$

Probability measure

$$m_k = \int_{\mathbb{R}} y^k d\mu(y)$$

$$m = \begin{bmatrix} 0 \\ \sigma^2 \\ 0 \\ 3\sigma^4 \\ \vdots \end{bmatrix}$$

Moment sequence

$$\frac{d}{dt}x(t,\beta) = F(t,\beta,x(t,\beta),u(t))$$

$$Y_t = y_t(\Omega), \quad \mu_t = (y_t)_{\#}\lambda$$

$$\boxed{m_k(t) = \int_{\mathbb{R}^r} y^k d\mu_t(y)}$$

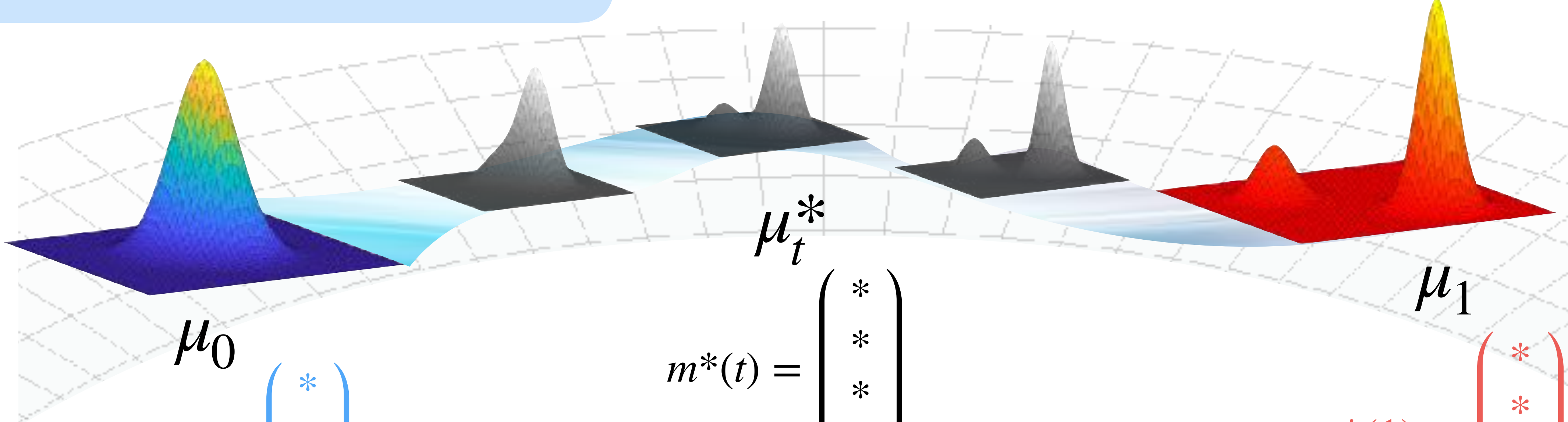
$$m(t) = \big(m_k(t)\big)_{k\in\mathbb{N}^r}$$

Moment Kernelized Ensemble control

- Ensemble control as time-dependent optimal transport

$$\frac{d}{dt}x(t,\beta) = F(t,\beta,x(t,\beta),u(t))$$
$$Y_t = y_t(\Omega) = h \circ x_t(\Omega)$$

$$\frac{d}{dt}m(t) = \bar{F}(t,m(t),u(t))$$



$$m^*(0) = \begin{pmatrix} * \\ * \\ * \\ \vdots \end{pmatrix} \qquad m^*(t) = \begin{pmatrix} * \\ * \\ * \\ \vdots \end{pmatrix} \qquad m^*(1) = \begin{pmatrix} * \\ * \\ * \\ \vdots \end{pmatrix}$$

$$m_k^*(t) = \int_\Omega y^k d\mu_t^*(y)$$

$$\frac{d}{dt}m^*(t) = \bar{F}^*(t,m^*(t))$$

OT dynamics in moment coordinates

Tracking

$$m(t) \approx m^*(t)$$

$$\frac{d}{dt}m(t) = \bar{F}(t,m(t),u^*(t))$$

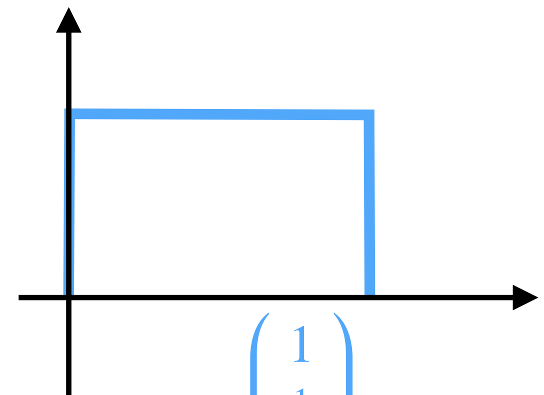
Controlled moment system

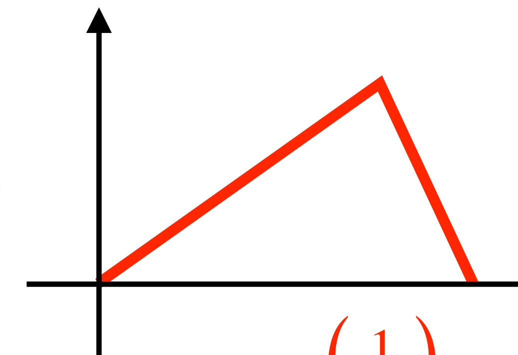
Moment Kernelized Ensemble control

- Transport from a square to a triangle wave

$$\frac{d}{dt}x(t,\beta) = \beta x(t,\beta) + \sum_{i=1}^{p} \beta^{i-1} u_i(t)$$

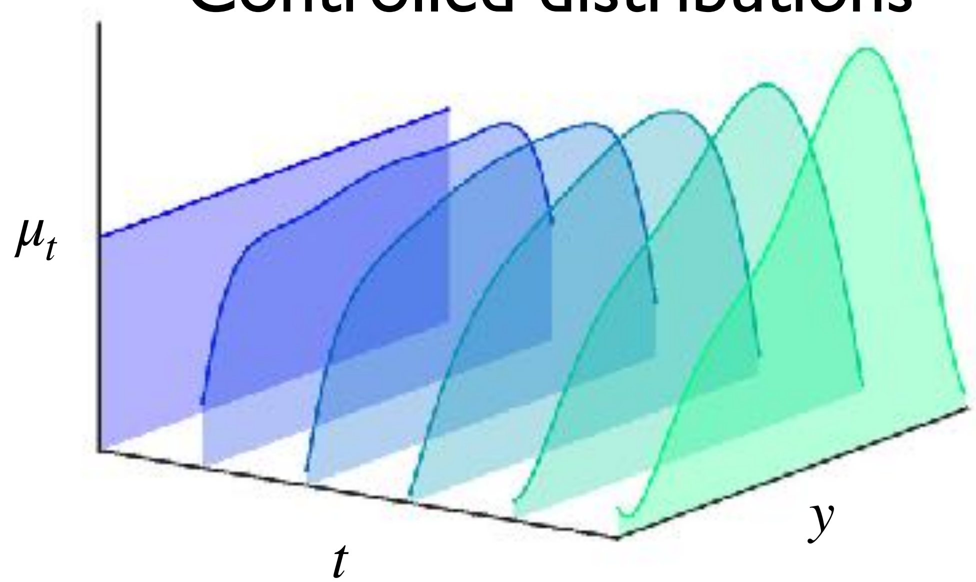
$$y_t(\beta) = h(x_t(x)) = x_t(\beta) \quad \beta \in [0,1]$$





$$m(0) = \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{3} \\ \vdots \end{pmatrix} \qquad m_k(t) = \int_0^1 y^k d\mu_t(y) \qquad m(T) = \begin{pmatrix} 1 \\ \frac{7}{12} \\ \frac{37}{96} \\ \vdots \end{pmatrix}$$

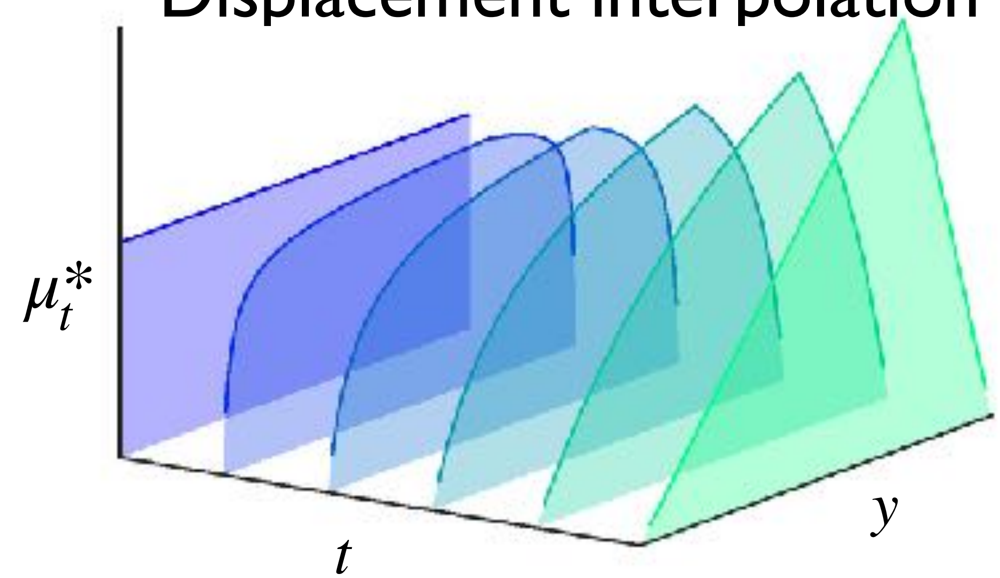
Controlled distributions



$$\frac{d}{dt}m_k(t) = m_{k+1}(t) + \sum_{i=1}^{p} \frac{u_i(t)}{k+i}$$

Output moment dynamics

Displacement interpolation



$$\frac{d}{dt}m_k^*(t) = \int_0^1 k\big((1-t)y + t\Phi_1(y)\big)^{k-1}(\Phi_1(y) - y)dy$$

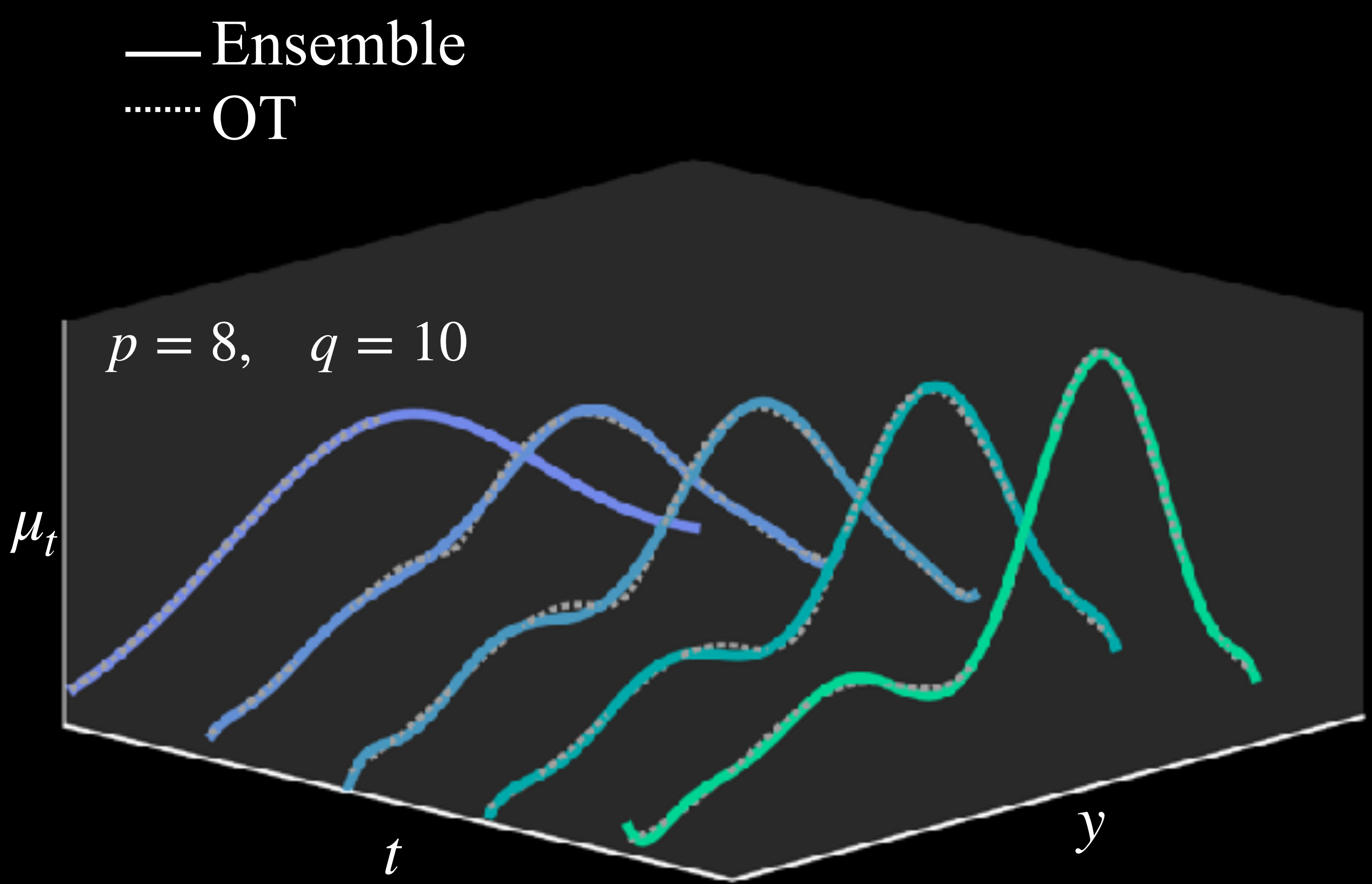
OT moment dynamics

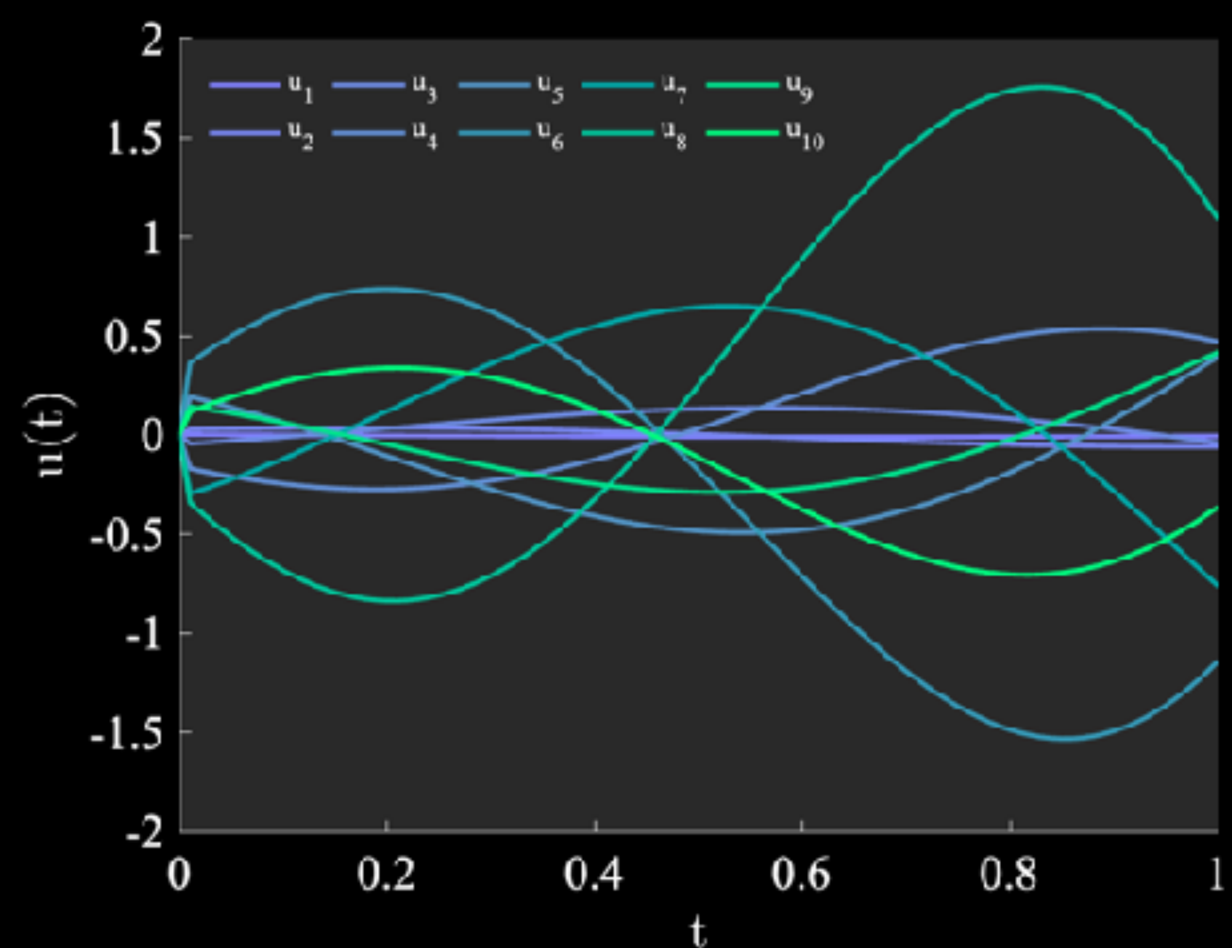
Moment Kernelized Ensemble control

$$\frac{d}{dt}x(t,\beta) = \beta x(t,\beta) + \sum_{i=1}^{p} \beta^{i-1} u_i(t)$$

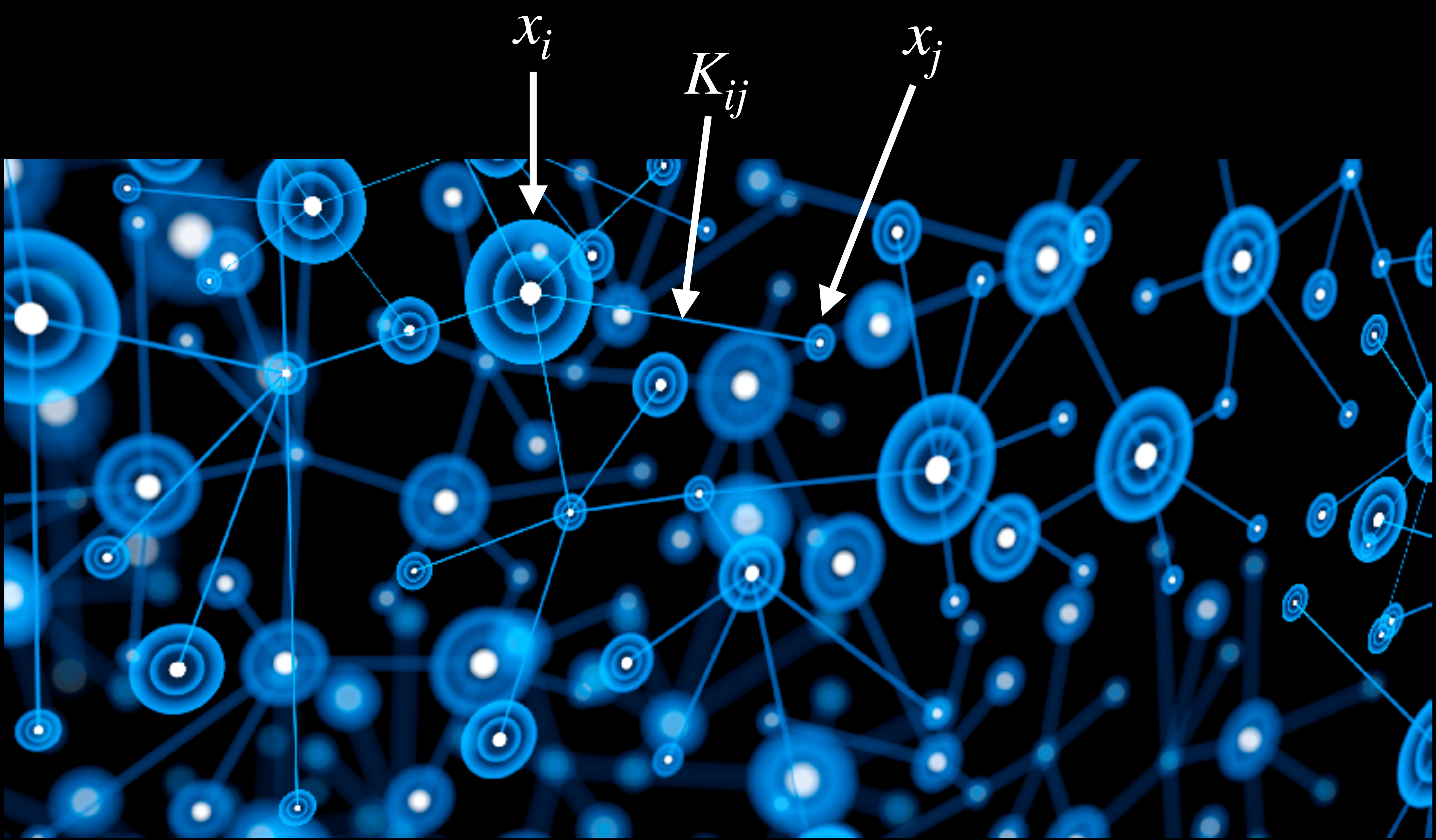
$$y_t(\beta) = h(x_t(x)) = x_t(\beta)$$

p: number of inputs
q: order of moments





Infinite Networks



$$\dot{x}_i(t) = f(x_i(t)) + \sum_{i,j=1}^{N} K_{ij}(x_i, x_j)$$

arbitrarily large

$N \to \infty$

(Infinite network)

Infinite Dynamic Networks

- Integral-differential system representation

$$\frac{d}{dt}x_i(t) = f_i(x_i(t)) + \boxed{\frac{1}{N}\sum_{j=1}^{N}\left(x_j(t) - x_i(t)\right)}$$

Node dynamics

Diffusive coupling

$N \to \infty$

$$\frac{d}{dt}x(t,\beta) = f(\beta, x(t,\beta)) + \boxed{\int_{\Omega}\left(x(t,\beta') - x(t,\beta)\right)d\beta'}$$

$$\beta \in \Omega \subset \mathbb{R}, \qquad |\Omega| = 1$$

Infinite Dynamic Networks

- Moment kernelized infinite network

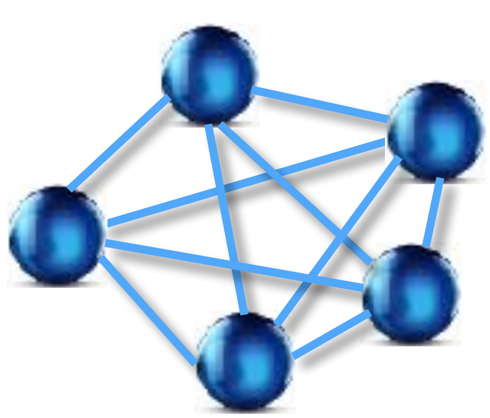
$$\frac{d}{dt}x(t,\beta) = f(\beta, x(t,\beta)) + \int_{\Omega} \big(x(t,\beta') - x(t,\beta)\big)\,d\beta'$$

Primal network

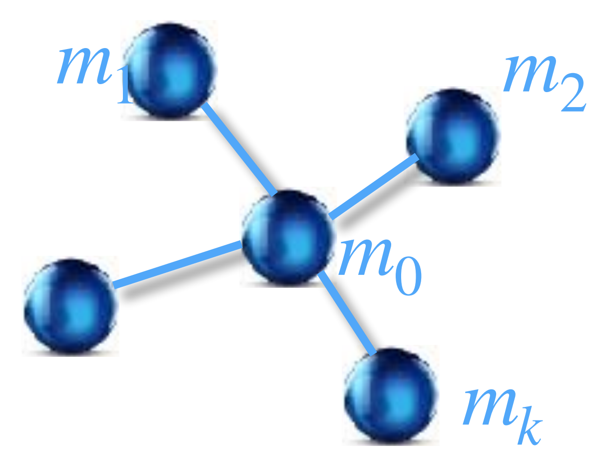
$$m(t) = (\mathscr{L}x)(t)$$

$$\frac{d}{dt}m_k(t) = \bar{f}_k(m(t)) + \big[c_k m_0(t) - m_k(t)\big]$$

Dual network



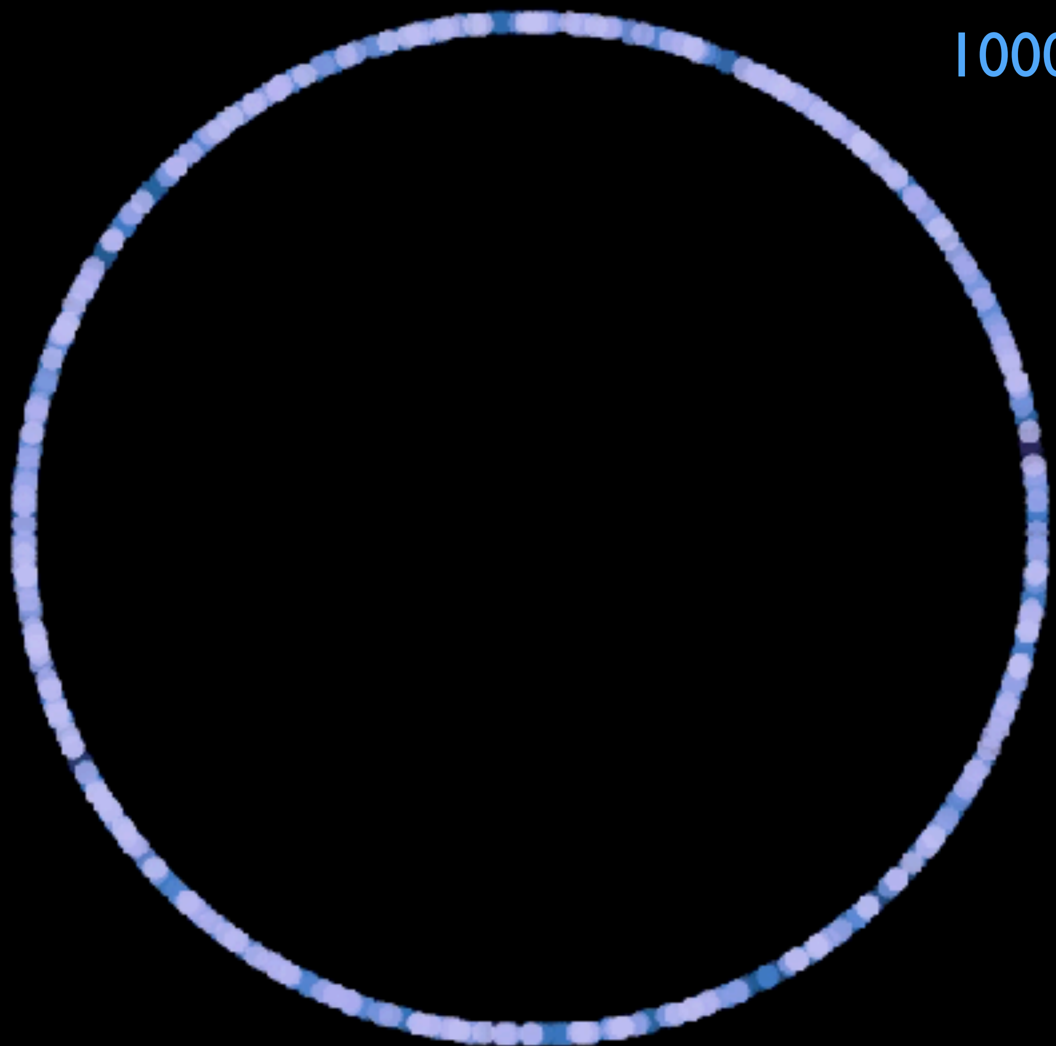
Infinite network
(all-to-all coupled)

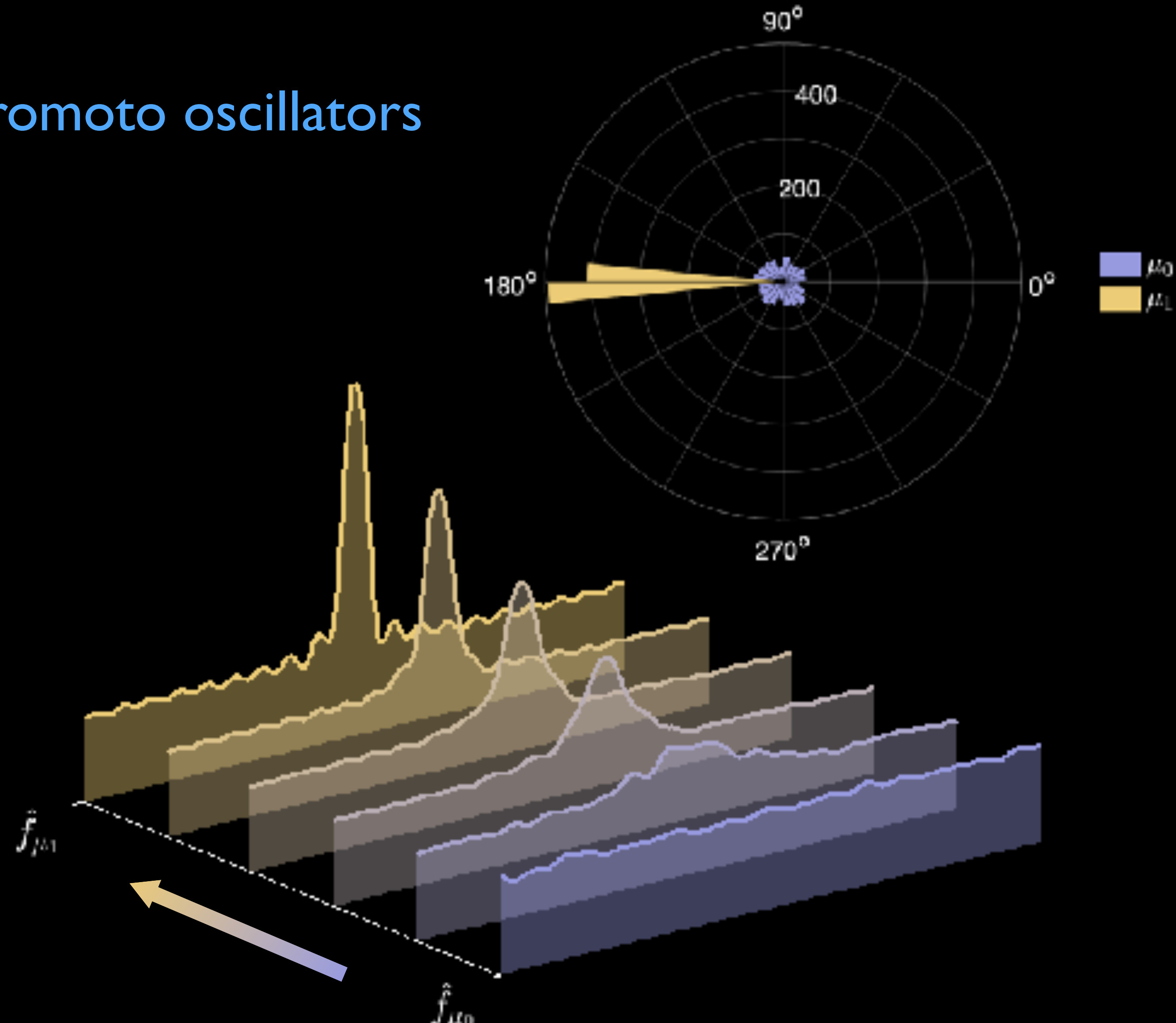


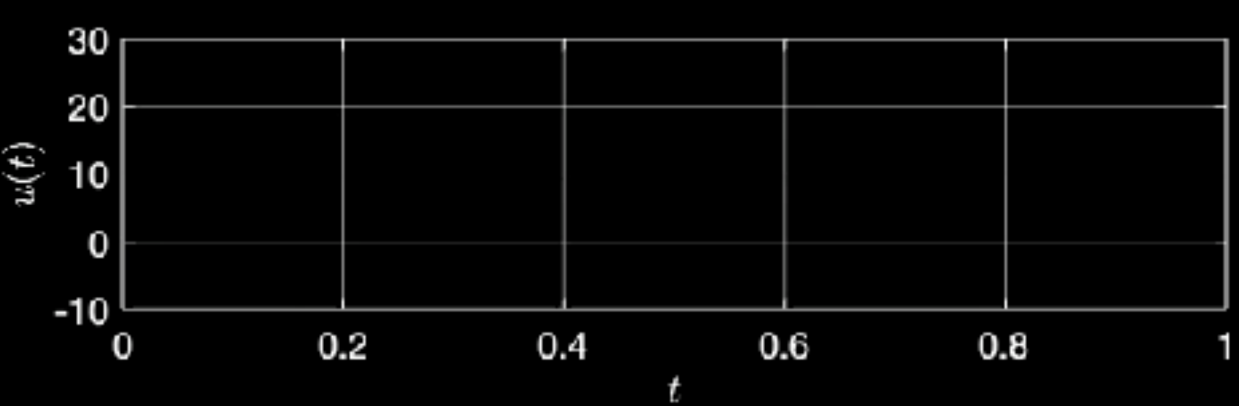
Star network
(truncatable)

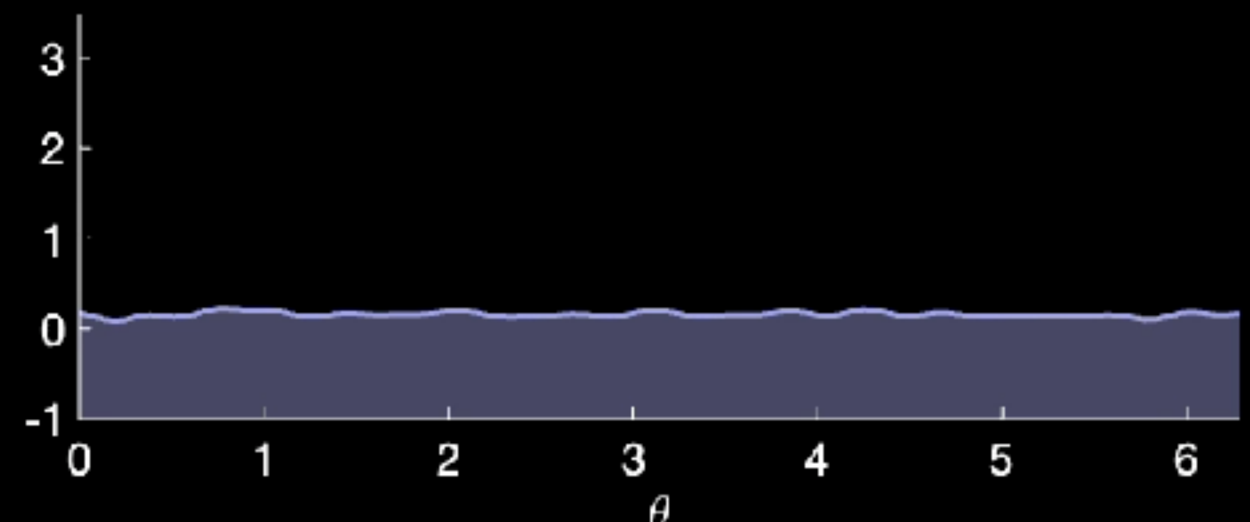
Synchronization Pattern Formation

1000 Kuromoto oscillators









Geometry of Distributional Control

- Fibration over distribution space