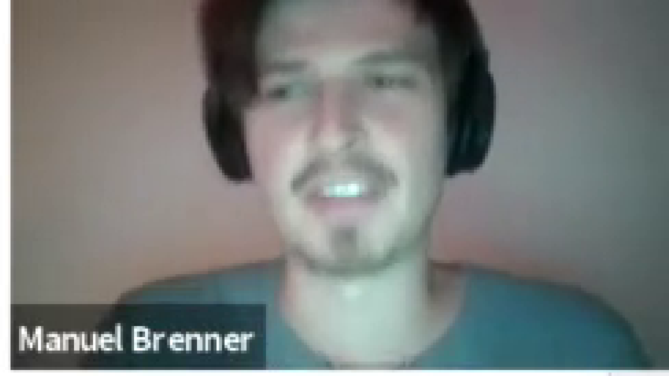
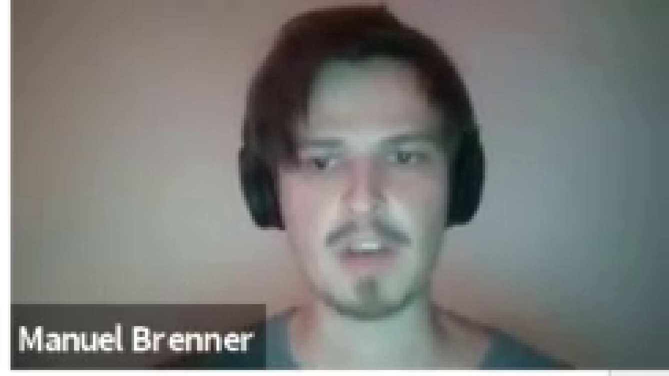


Fast and scalable learning of generative models for chaotic dynamical systems and neural data

Leonard Bereska, Po-Chen Kuo, Manuel Brenner, Daniel Durstewitz
Central Institute of Mental Health Mannheim, University of Heidelberg
28.10.2020, Neuromatch 3.0



Fast and scalable learning of generative models for chaotic dynamical systems and neural data

Leonard Bereska, Po-Chen Kuo, Manuel Brenner, Daniel Durstewitz
Central Institute of Mental Health Mannheim, University of Heidelberg

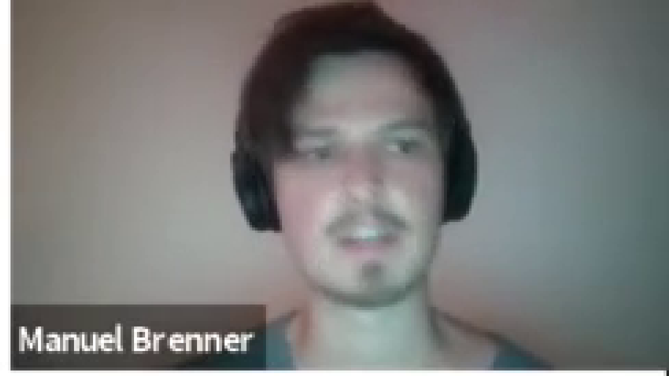
28.10.2020, Neuromatch 3.0

677,3 x 381,0 mm

Fast and scalable learning of generative models for chaotic dynamical systems and neural data

Leonard Bereska, Po-Chen Kuo, Manuel Brenner, Daniel Durstewitz
Central Institute of Mental Health Mannheim, University of Heidelberg

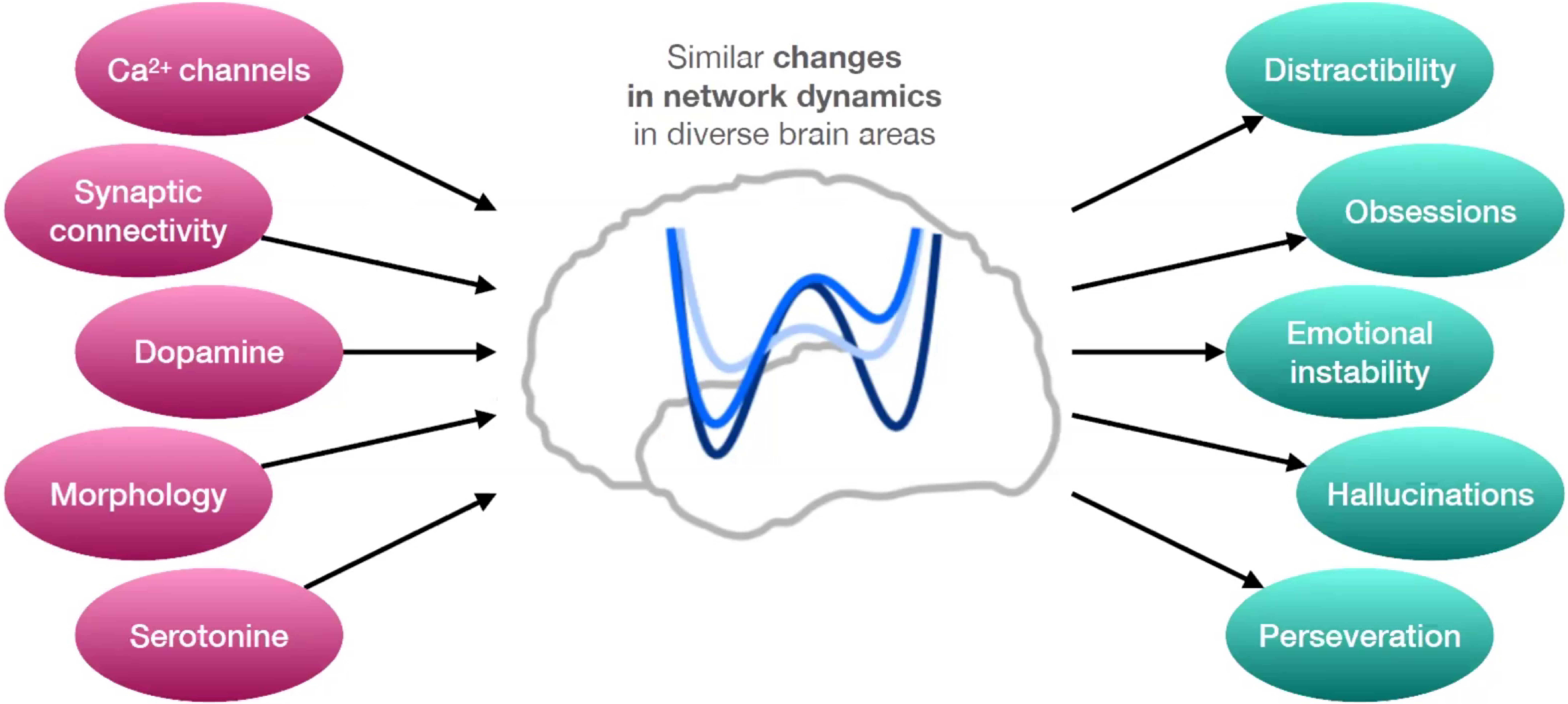
28.10.2020, Neuromatch 3.0



Motivation

Diverse biophysical and structural **causes**

Diverse **changes in** cognitive and emotional **experience**



Ca^{2+} channels

Synaptic connectivity

Dopamine

Morphology

Serotonine

Similar **changes in network dynamics** in diverse brain areas

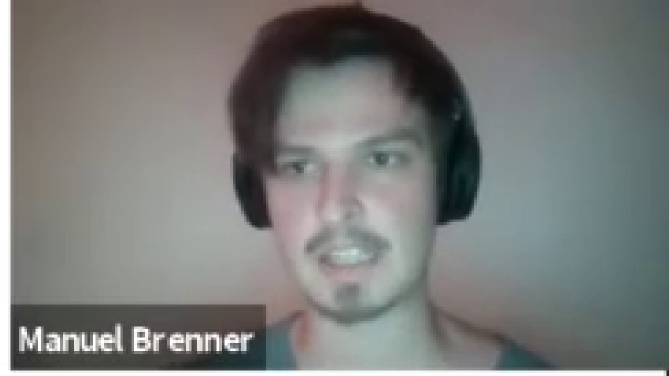
Distractibility

Obsessions

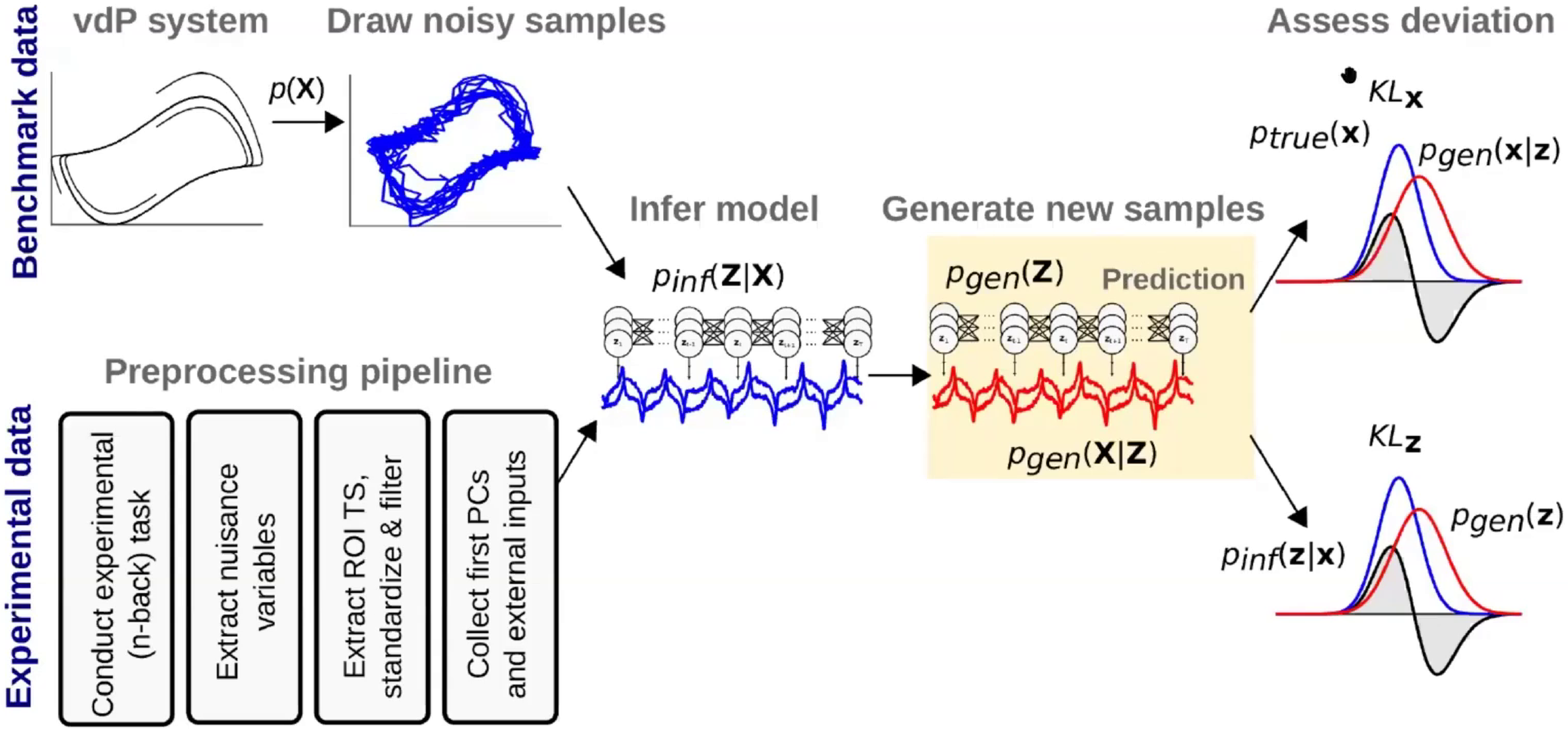
Emotional instability

Hallucinations

Perseveration



Inferring Generative Models from Data



Benchmark data

vdP system

Draw noisy samples

$p(\mathbf{X})$

Infer model

$p_{inf}(\mathbf{Z}|\mathbf{X})$

Generate new samples

$p_{gen}(\mathbf{Z})$

Prediction

$p_{gen}(\mathbf{X}|\mathbf{Z})$

Assess deviation

$KL_{\mathbf{x}}$

$p_{true}(\mathbf{x})$

$p_{gen}(\mathbf{x}|\mathbf{z})$

$KL_{\mathbf{z}}$

$p_{inf}(\mathbf{z}|\mathbf{x})$

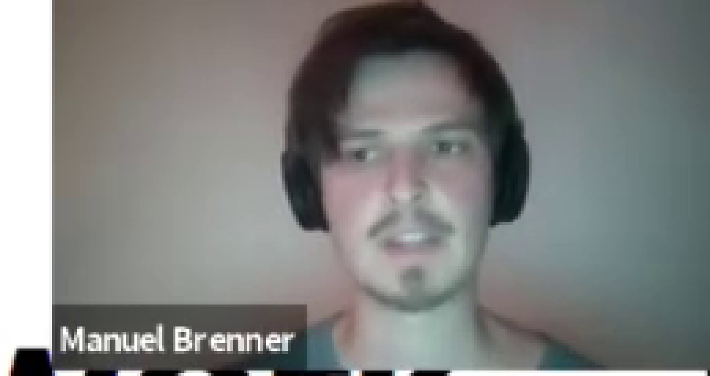
$p_{gen}(\mathbf{z})$

Experimental data

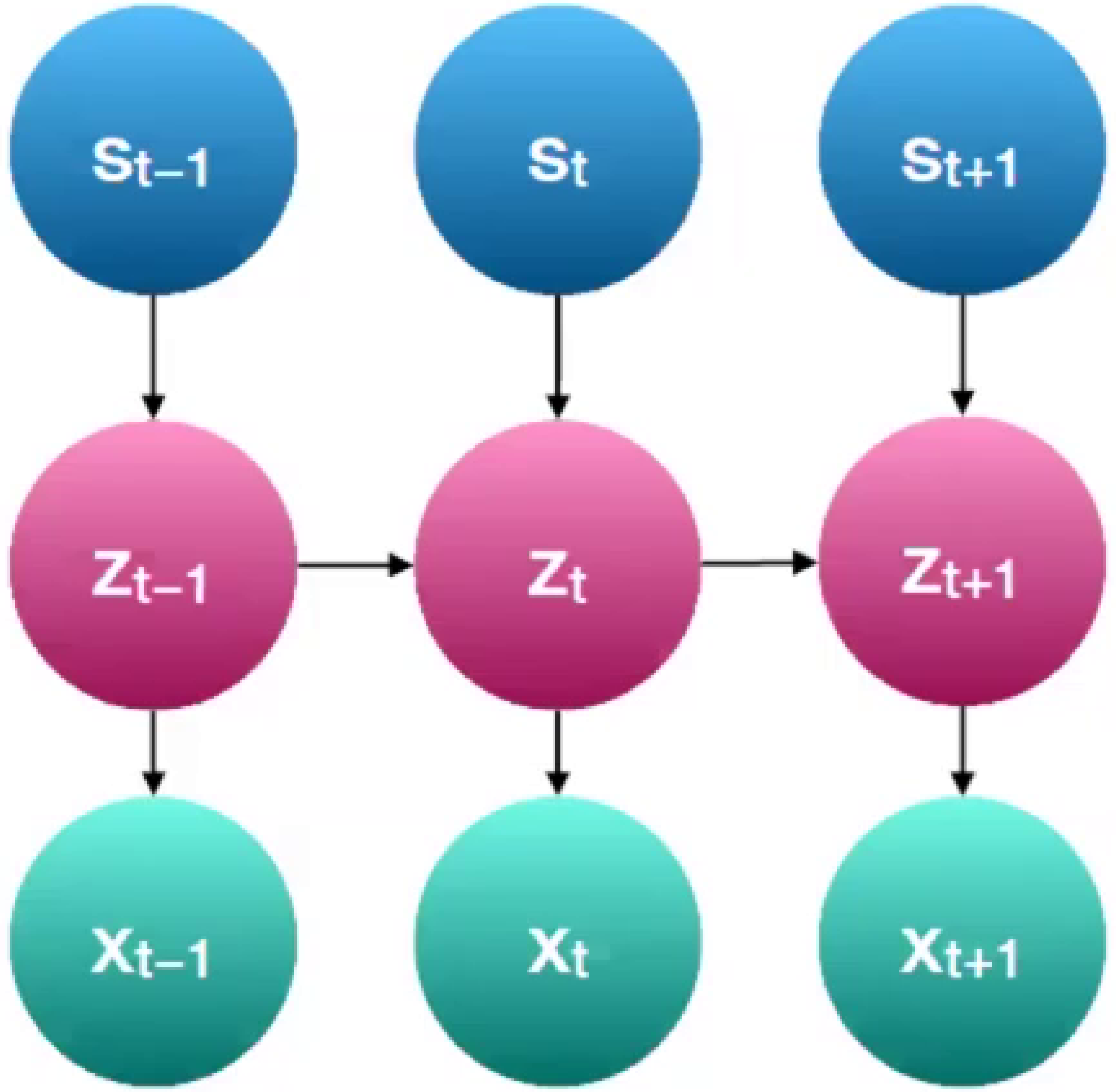
Preprocessing pipeline

- Conduct experimental (n-back) task
- Extract nuisance variables
- Extract ROI TS, standardize & filter
- Collect first PCs and external inputs

Figure from: Koppe, G., Toutounji, H., Kirsch, P., Lis, S., & Durstewitz, D. (2019). Identifying nonlinear dynamical systems via generative recurrent neural networks with applications to fMRI. PLoS Computational Biology, 15.



Piece-wise Linear Recurrent Neural Network



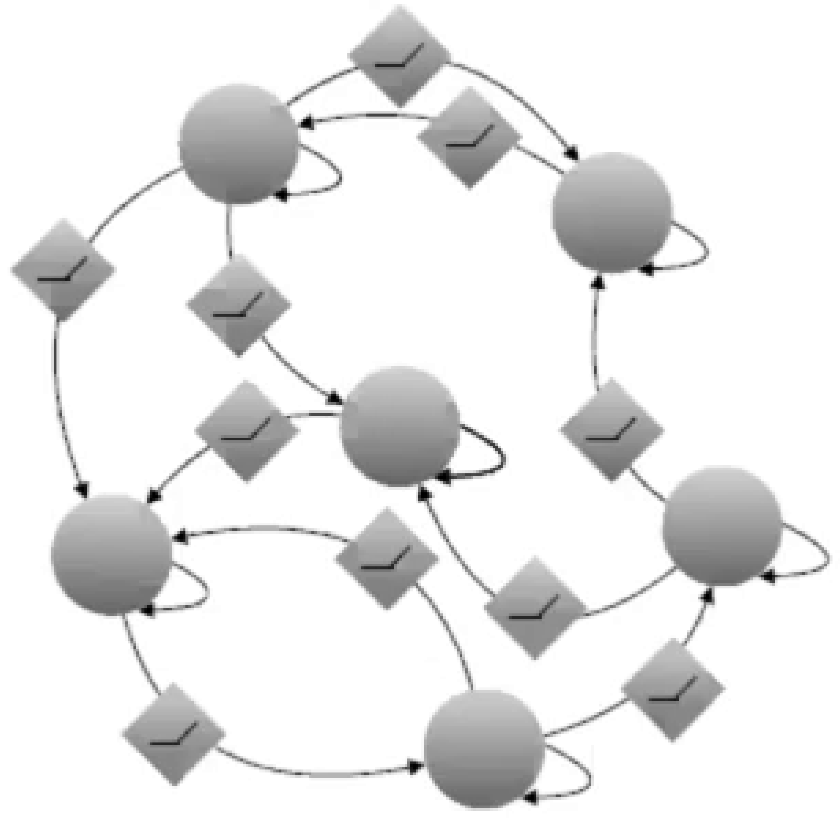
Latent Model $z_t \in \mathbb{R}^M$

$$z_t = \underset{\substack{\uparrow\\ \text{Diagonal}}}{\boldsymbol{A}} z_{t-1} + \underset{\substack{\uparrow\\ \text{Off-diagonal}}}{\boldsymbol{W}} \phi(z_{t-1}) + \boldsymbol{h}_0 + \boldsymbol{C} s_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma)$$

Observation Model

$$x_t = \boldsymbol{B} z_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \Gamma)$$

Activation Function



$$\phi(z_{t-1}) = \max(0, z_{t-1})$$

Complete Likelihood

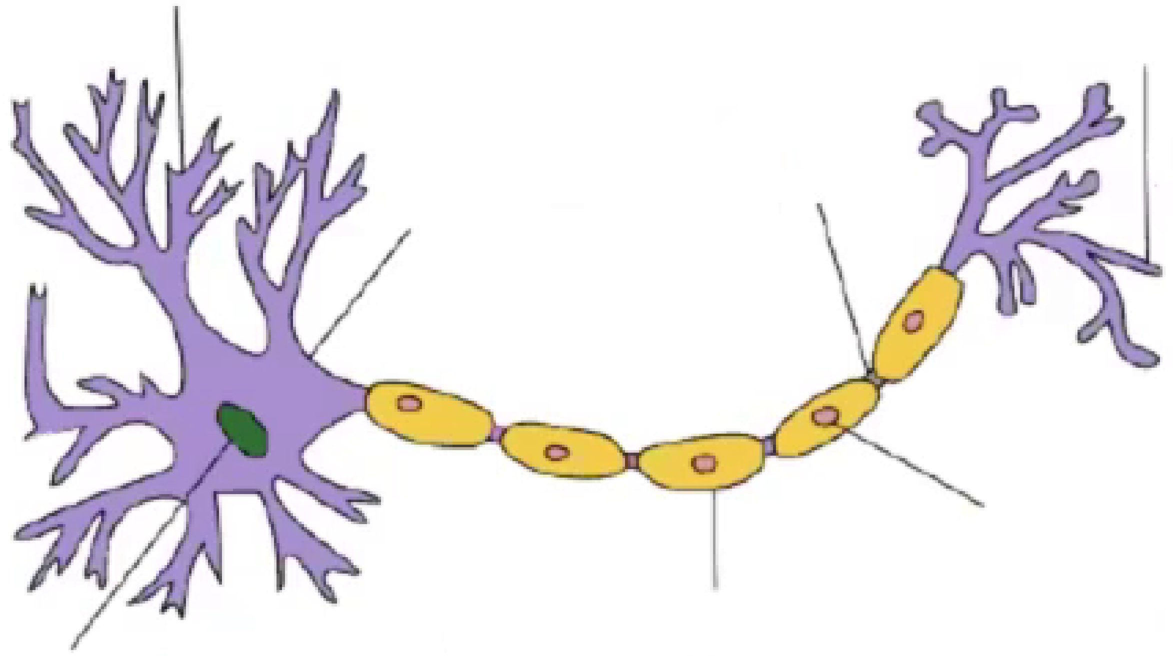
$$\begin{aligned}\log p_\theta(\boldsymbol{x}, \boldsymbol{z}) = & -\frac{1}{2}(z_1 - \boldsymbol{\mu}_0 - \boldsymbol{s}_1)^T \boldsymbol{\Sigma}^{-1}(z_1 - \boldsymbol{\mu}_0 - \boldsymbol{s}_1) \\ & -\frac{1}{2}\sum_{t=2}^{T}(z_t - \boldsymbol{A}z_{t-1} - \boldsymbol{W}\phi(z_{t-1}) - \boldsymbol{s}_t)^T \boldsymbol{\Sigma}^{-1}(z_t - \boldsymbol{A}z_{t-1} - \boldsymbol{W}\phi(z_{t-1}) - \boldsymbol{s}_t) \\ & -\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{x}_t - \boldsymbol{B}z_t)^T \boldsymbol{\Gamma}^{-1}(\boldsymbol{x}_t - \boldsymbol{B}z_t)\end{aligned}$$

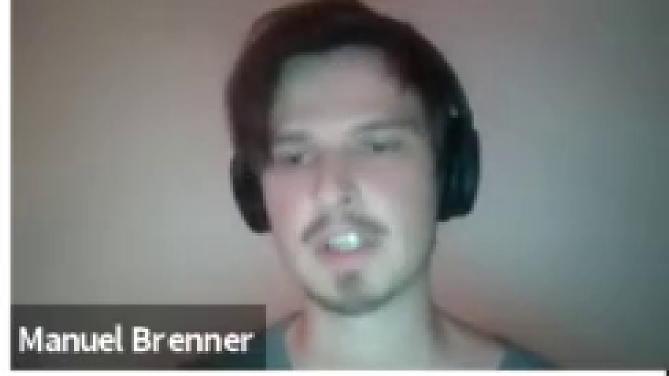
For **fMRI observation model** refer to: Koppe, G., Toutounji, H., Kirsch, P., Lis, S., & Durstewitz, D. (2019). Identifying nonlinear dynamical systems via generative recurrent neural networks with applications to fMRI. PLoS Computational Biology, 15.

Increasing Computational Capacity

- $\boldsymbol{z_t = A z_{t-1} + W \phi(z_{t-1}) + h_0 + C s_t + \epsilon_t}, \;\; \boldsymbol{\epsilon_t \sim \mathcal{N}(0, \Sigma)}$

- Can we reduce the dimensionality of latent space?
- Retain piece-wise linear form
- Neurophysiological analogy: dendritic computation or neuronal diversity

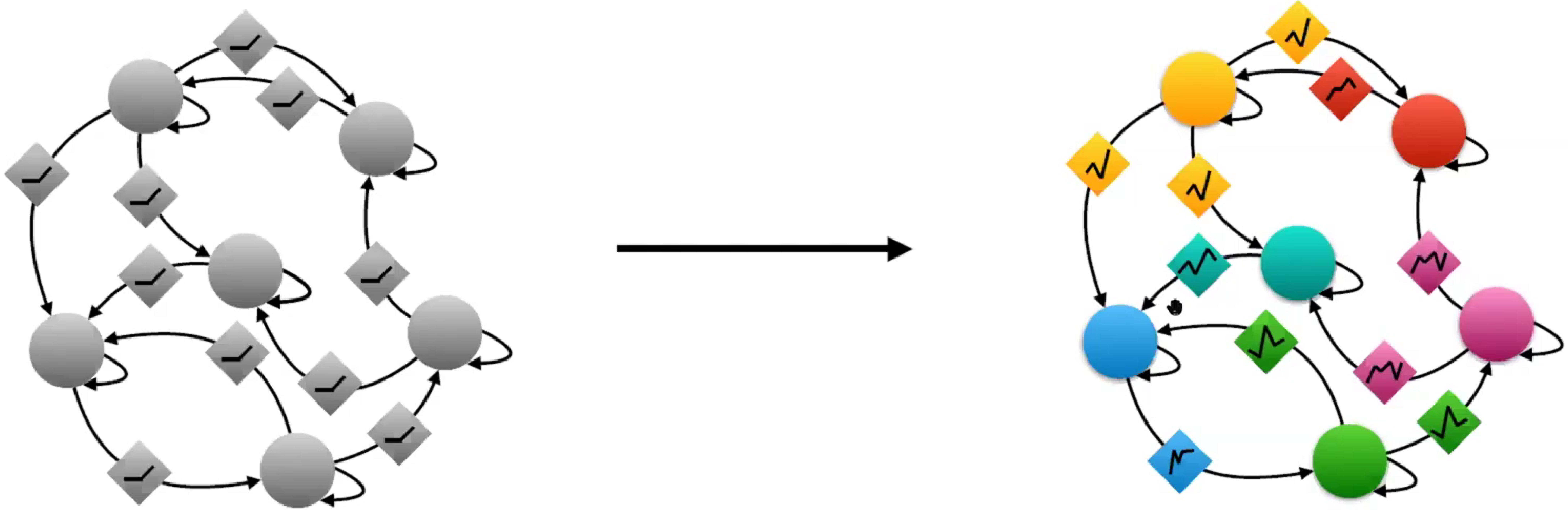




Basis Expansion

New activation function in latent model

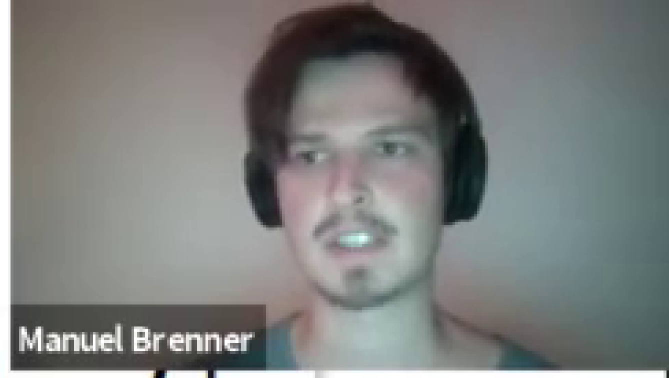
$$z_t = \boldsymbol{A} z_{t-1} + \boldsymbol{W}\phi(z_{t-1}) + \boldsymbol{h}_0 + \boldsymbol{C}\boldsymbol{s}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$



$$\phi(z_{t-1}) = \max(0, z_{t-1})$$

$$\phi(z_{t-1}) = \sum_{b=1}^{B} \alpha_b \max(0, z_{t-1} - \boldsymbol{h}_b)$$

Variational Inference



$$\boldsymbol{x} = \{\boldsymbol{x}_t \,|\, t = 1 \ldots T\}$$
$$z = \{z_t \,|\, t = 1 \ldots T\}$$

Variational Lower Bound

$$p(\boldsymbol{x}) \geq \mathscr{L}(\theta, \phi, \boldsymbol{x}) = \mathbb{E}_{q_\phi(z|\boldsymbol{x})}[\log p_\theta(\boldsymbol{x}, z)] + \mathbb{E}_{q_\phi(z|\boldsymbol{x})}[\log q_\phi(z \,|\, \boldsymbol{x})]$$

Complete Likelihood

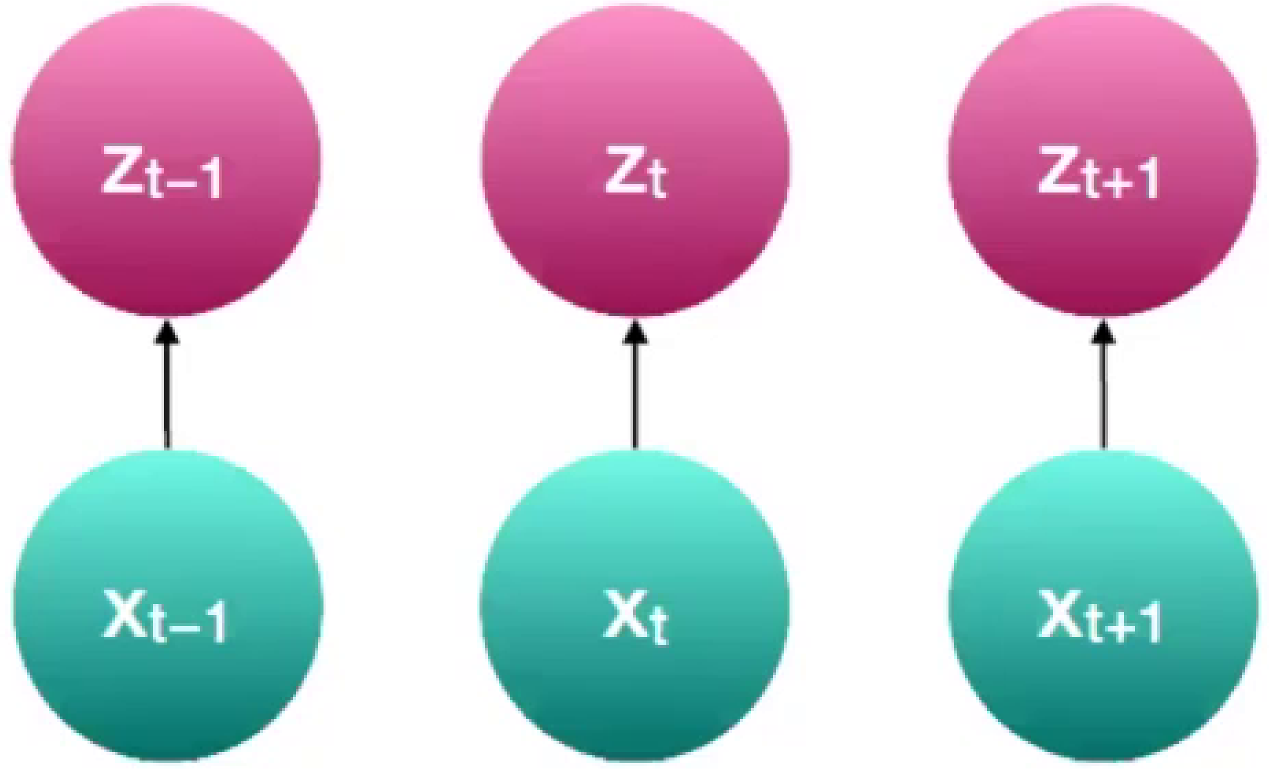
$$\begin{aligned}\log p_\theta(\boldsymbol{x}, z) = & -\frac{1}{2}(z_1 - \boldsymbol{\mu}_0 - s_1)^T \boldsymbol{\Sigma}^{-1}(z_1 - \boldsymbol{\mu}_0 - s_1) \\ & -\frac{1}{2}\sum_{t=2}^{T}(z_t - \boldsymbol{A}z_{t-1} - \boldsymbol{W}\phi(z_{t-1}) - s_t)^T \boldsymbol{\Sigma}^{-1}(z_t - \boldsymbol{A}z_{t-1} - \boldsymbol{W}\phi(z_{t-1}) - s_t) \\ & -\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{x}_t - \boldsymbol{B}z_t)^T \boldsymbol{\Gamma}^{-1}(\boldsymbol{x}_t - \boldsymbol{B}z_t)\end{aligned}$$

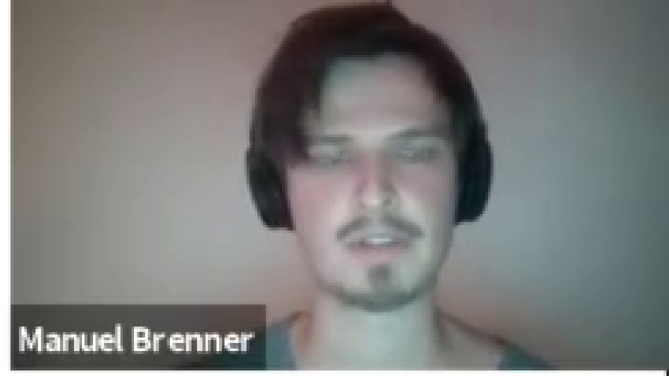
Mean-Field Approximation

$$q_\phi(z \,|\, \boldsymbol{x}) = \prod_{t=1}^{T} q_\phi(z_t \,|\, \boldsymbol{x}_t) = \prod_{t=1}^{T} \mathcal{N}(\mu_\phi(\boldsymbol{x}_t), \sigma_\phi(\boldsymbol{x}_t)^2)$$

$$\mu_\phi(x_t) = \mathrm{NN}_\mu(x_t)$$

$$\sigma_\phi(x_t) = \mathrm{NN}_\sigma(x_t)$$





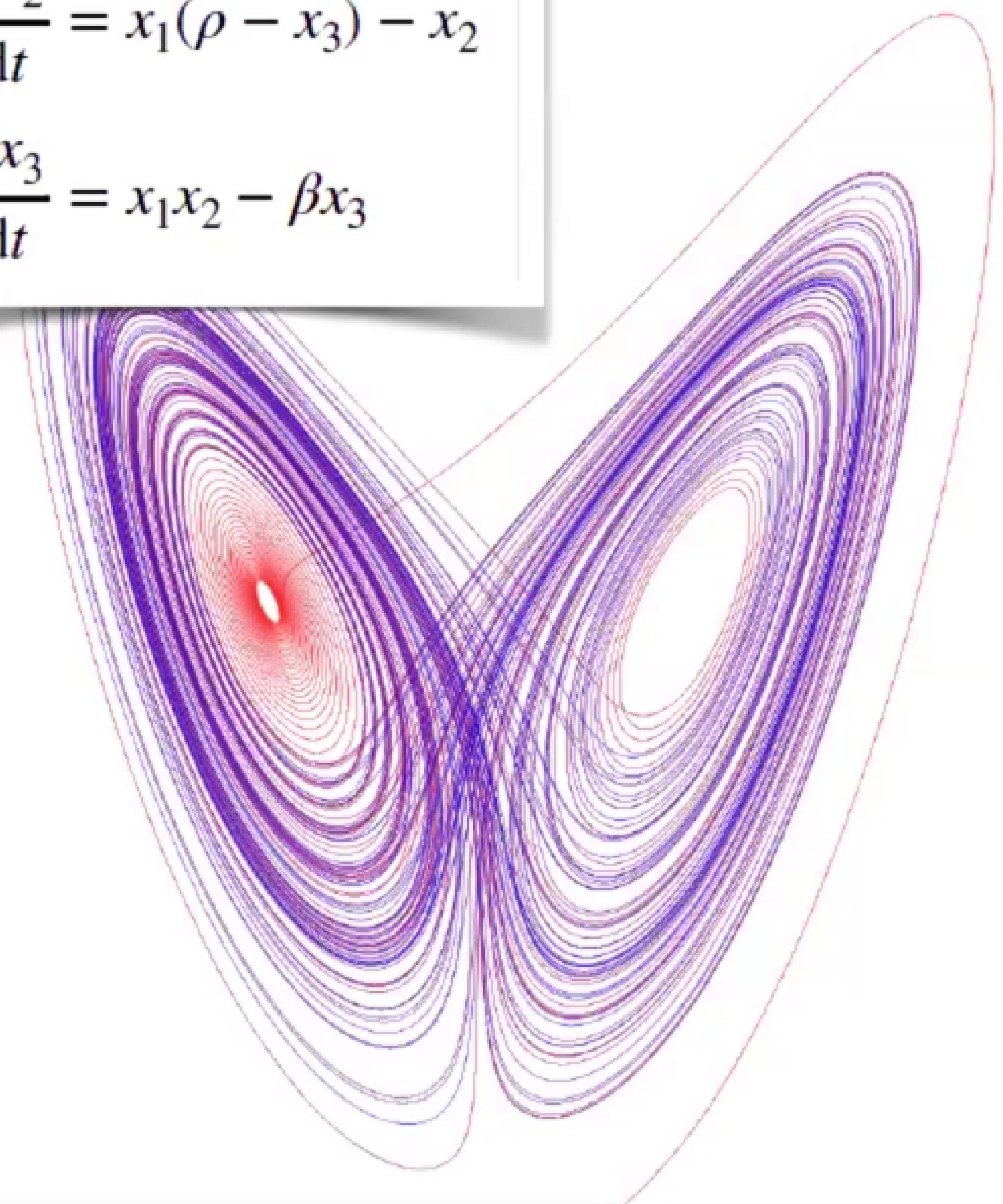
Benchmark Dynamical Systems

Lorenz Attractor

$$\frac{dx_1}{dt} = \sigma(x_2 - x_1)$$

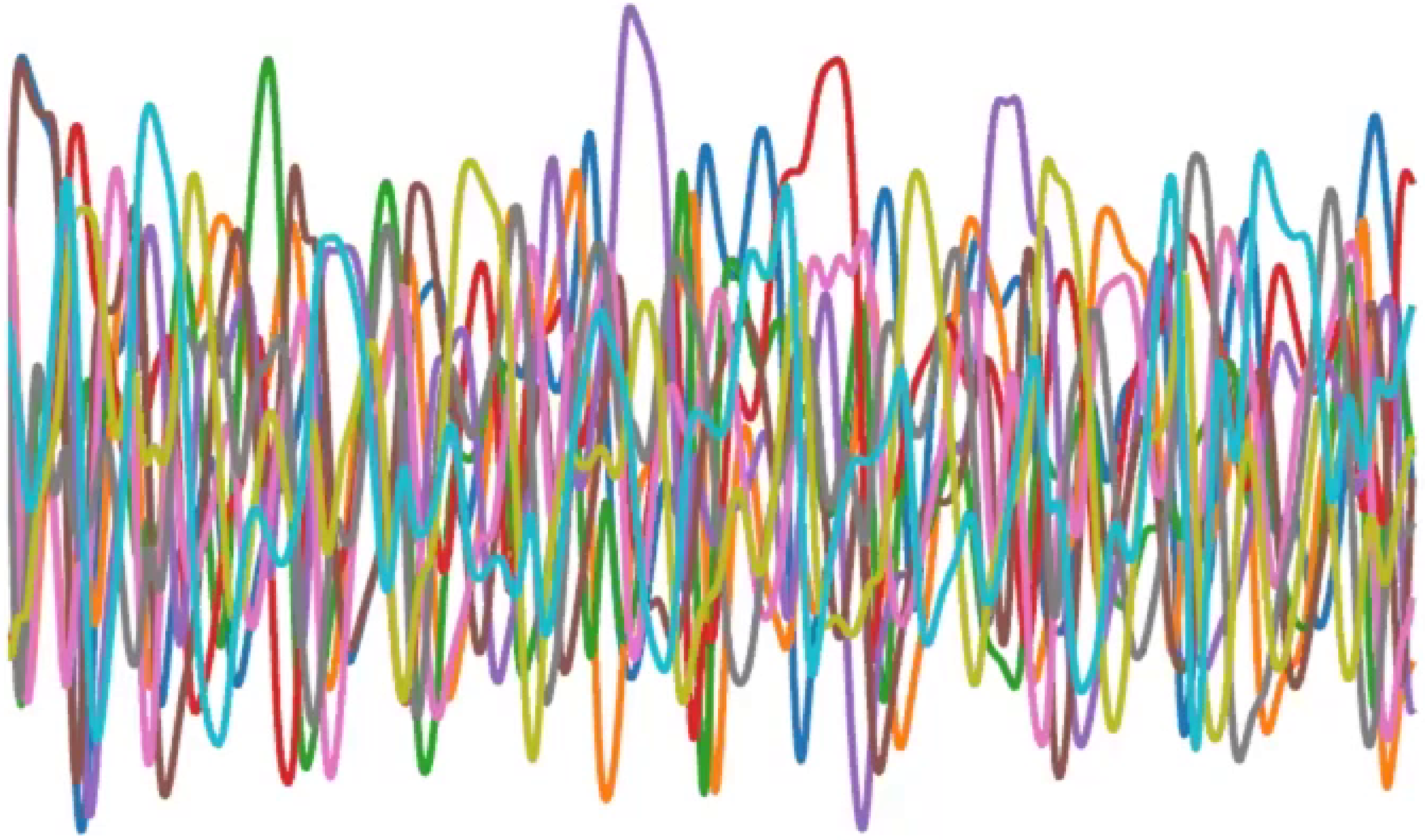
$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = x_1(\rho - x_3) - x_2$$

$$\frac{\mathrm{d}x_3}{\mathrm{d}t} = x_1 x_2 - \beta x_3$$

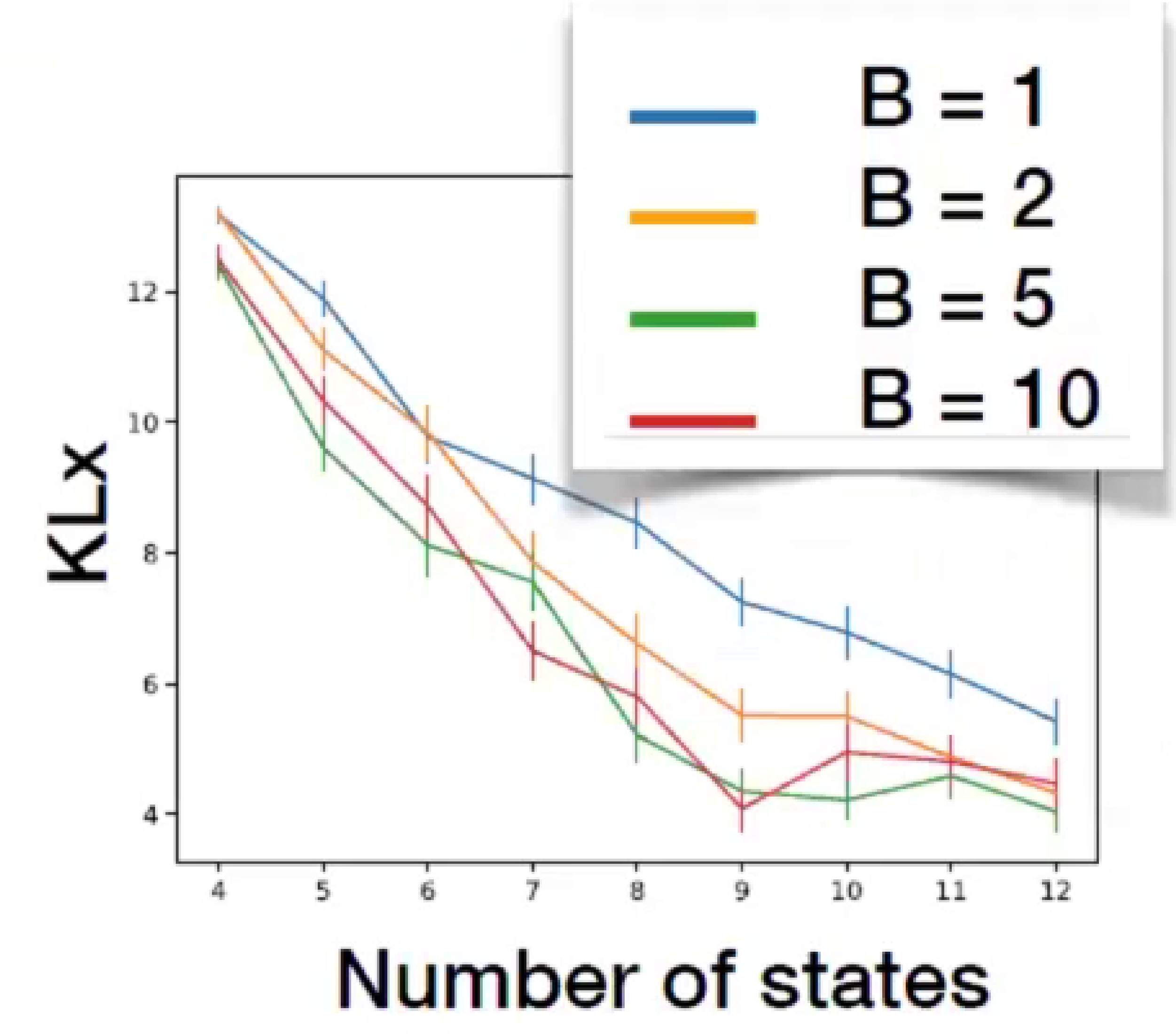


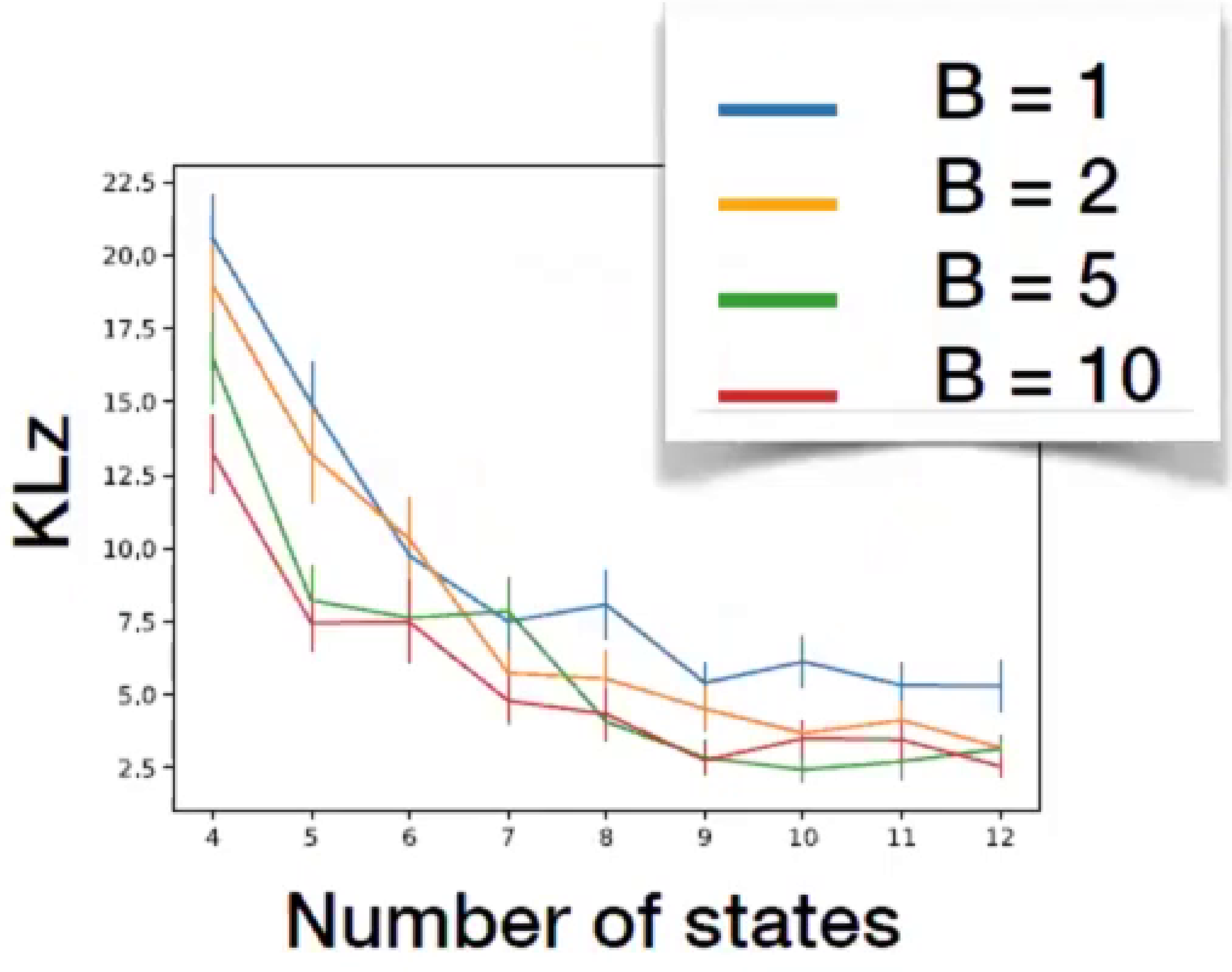
Lorenz-96 System

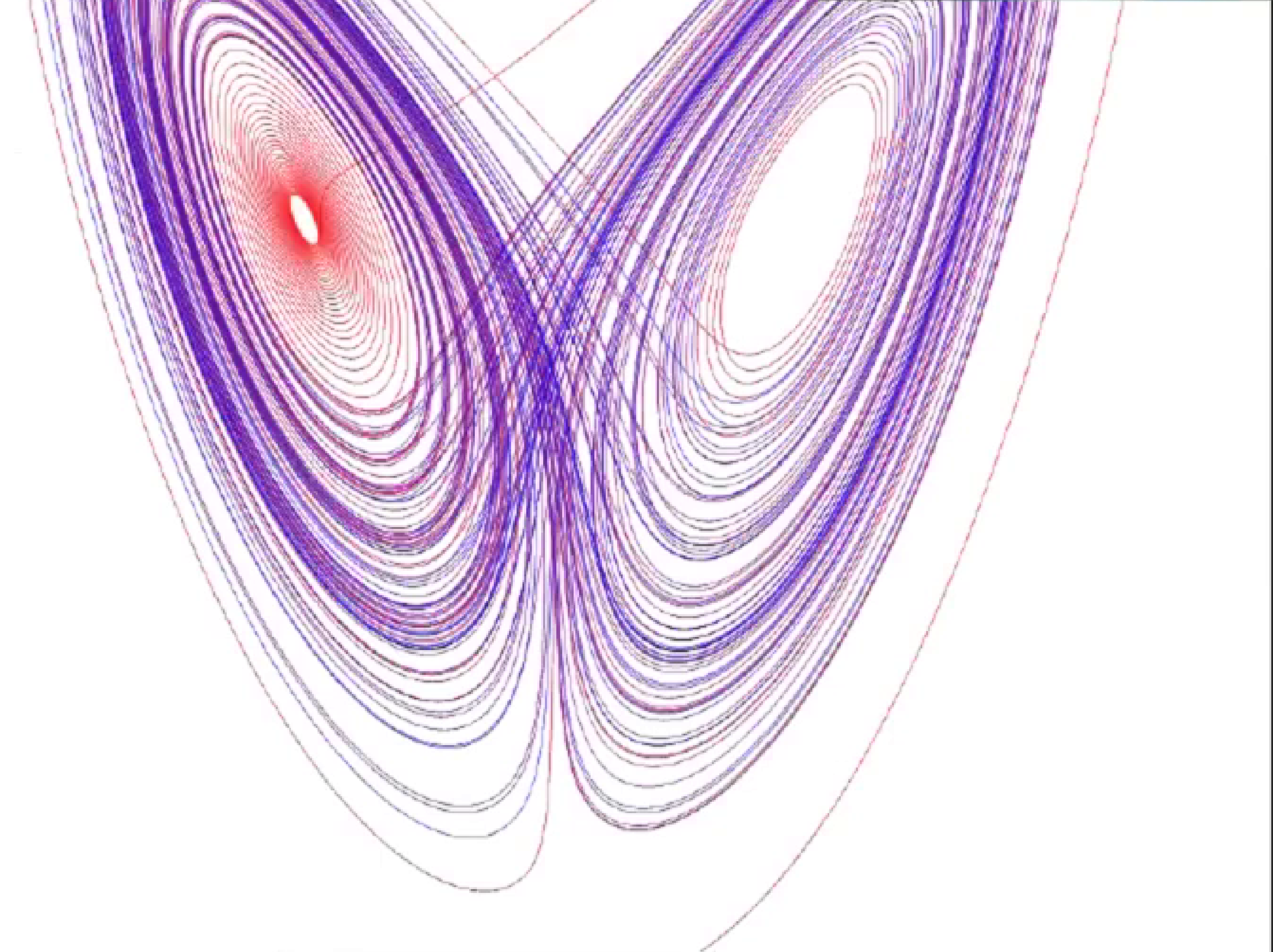
$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F$$



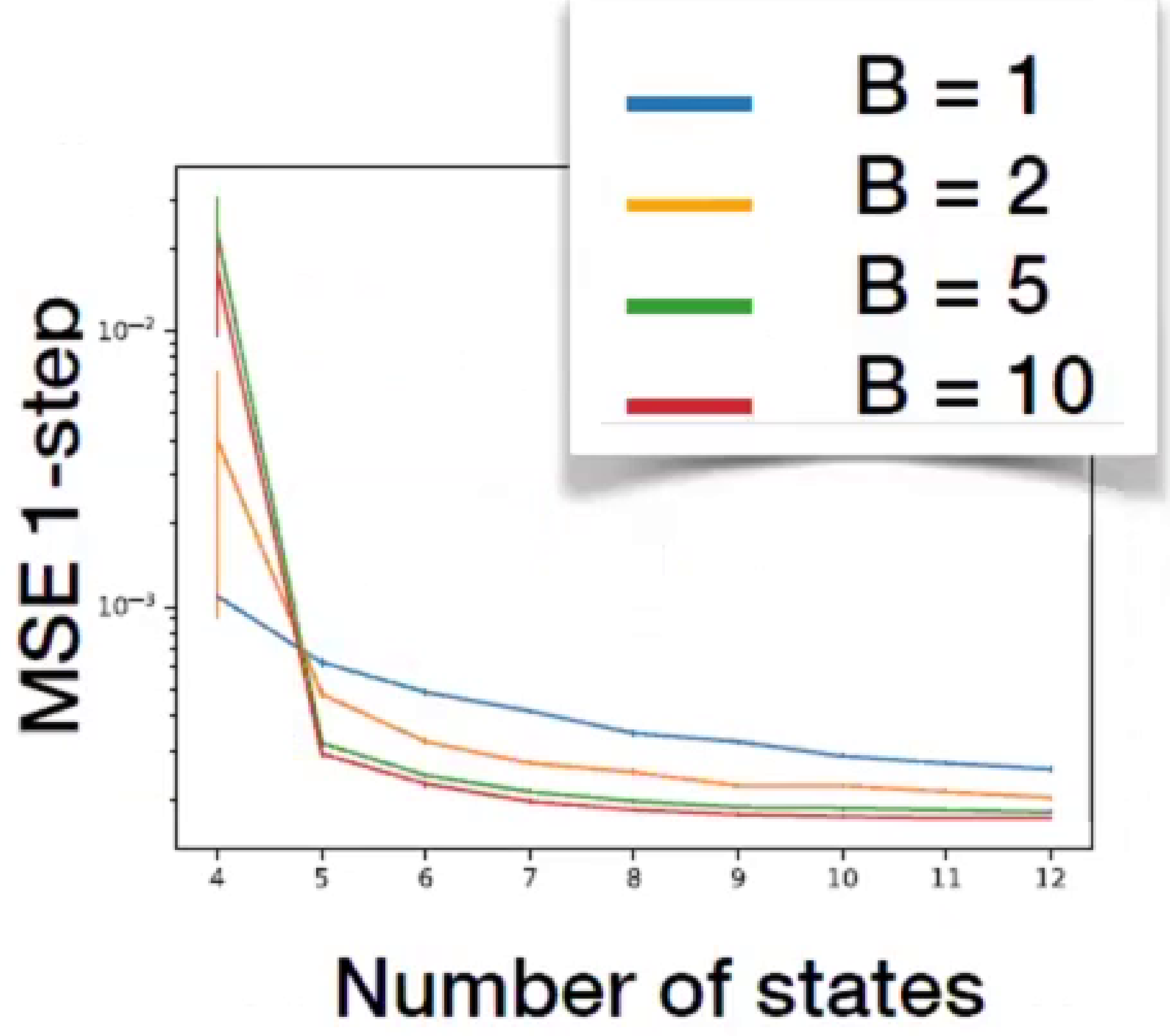
Results - Lorenz System

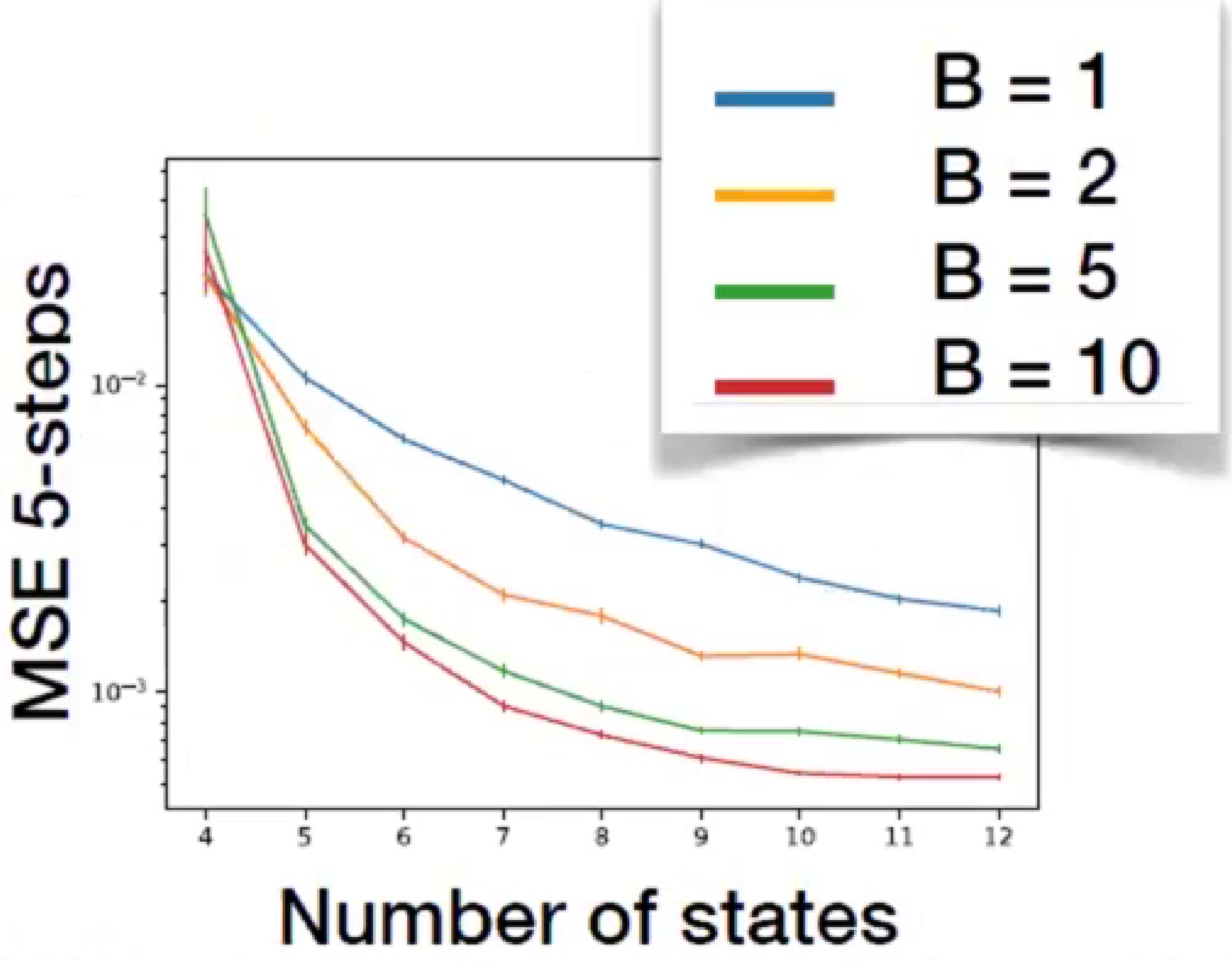


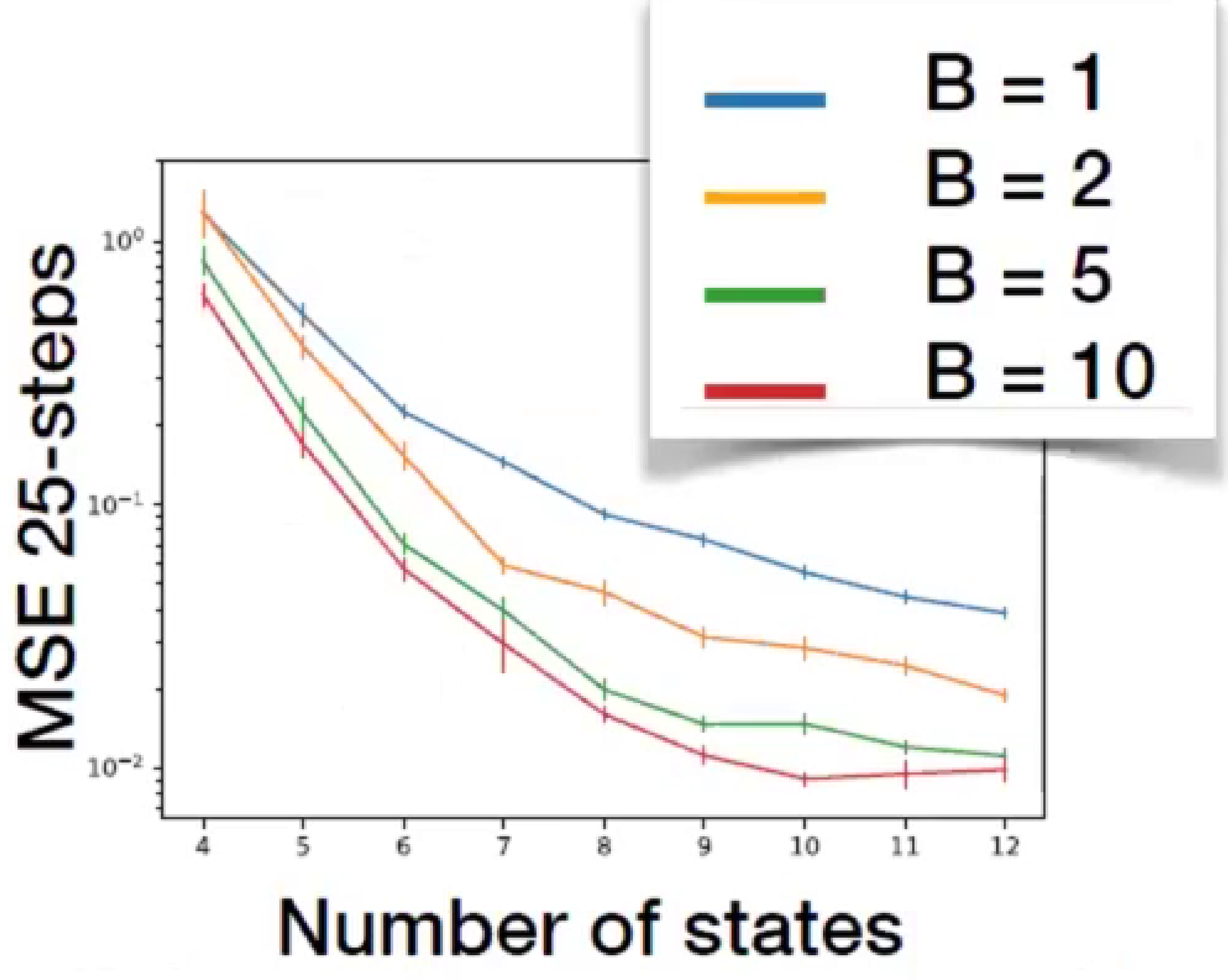




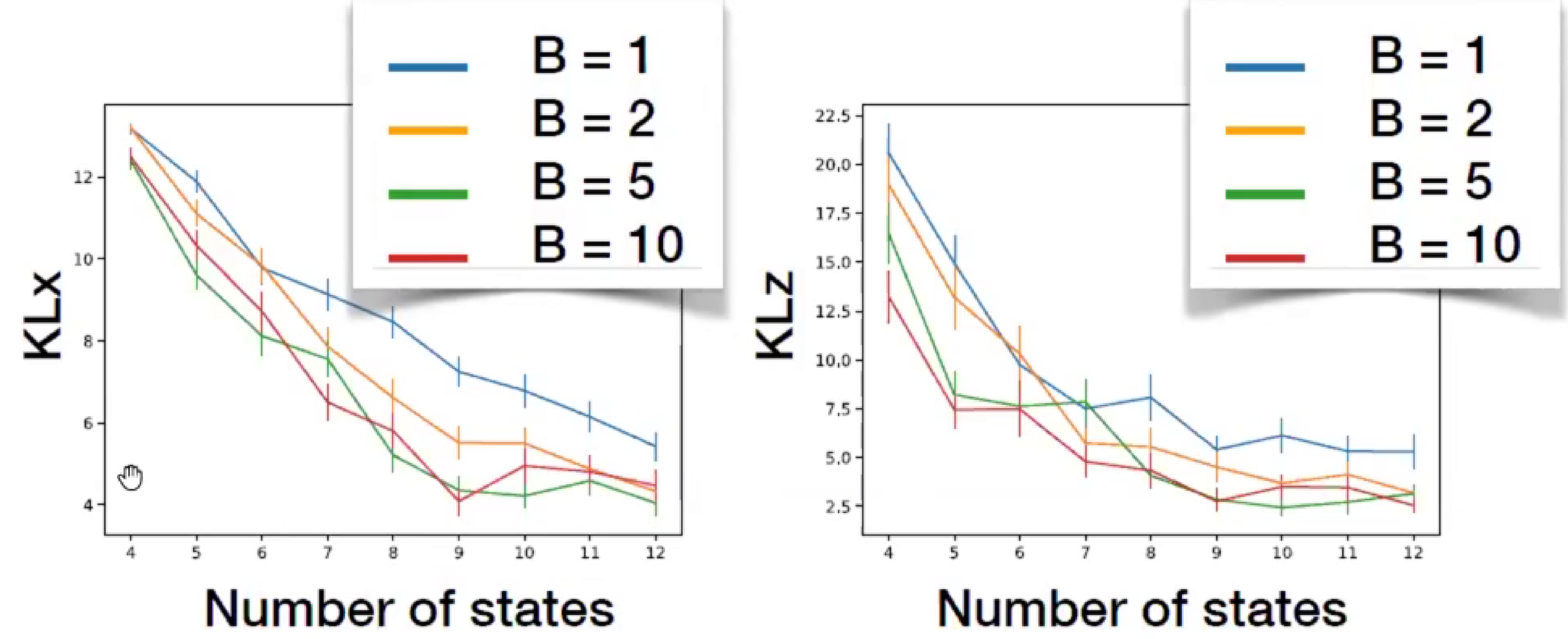

Manuel Brenner

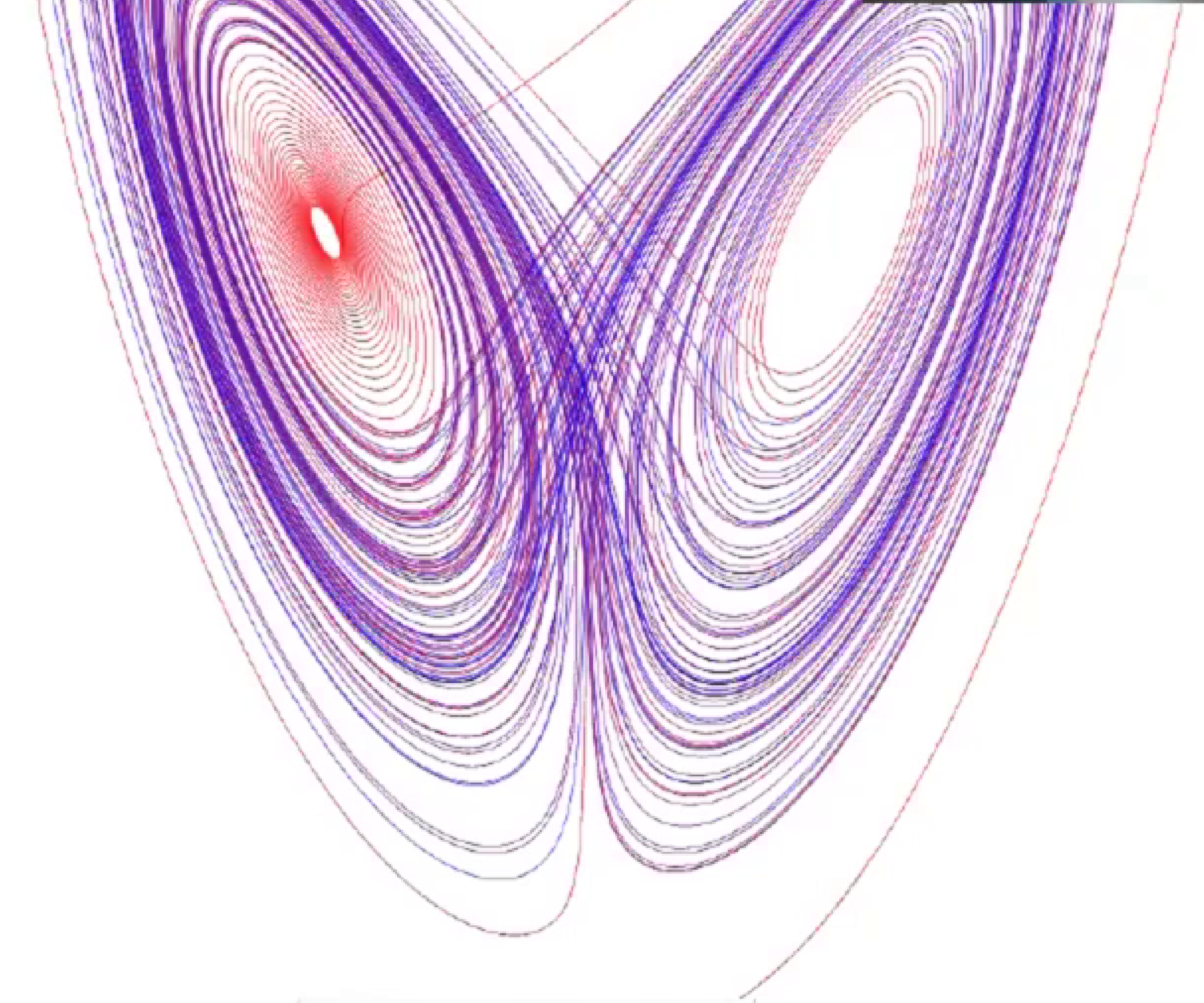


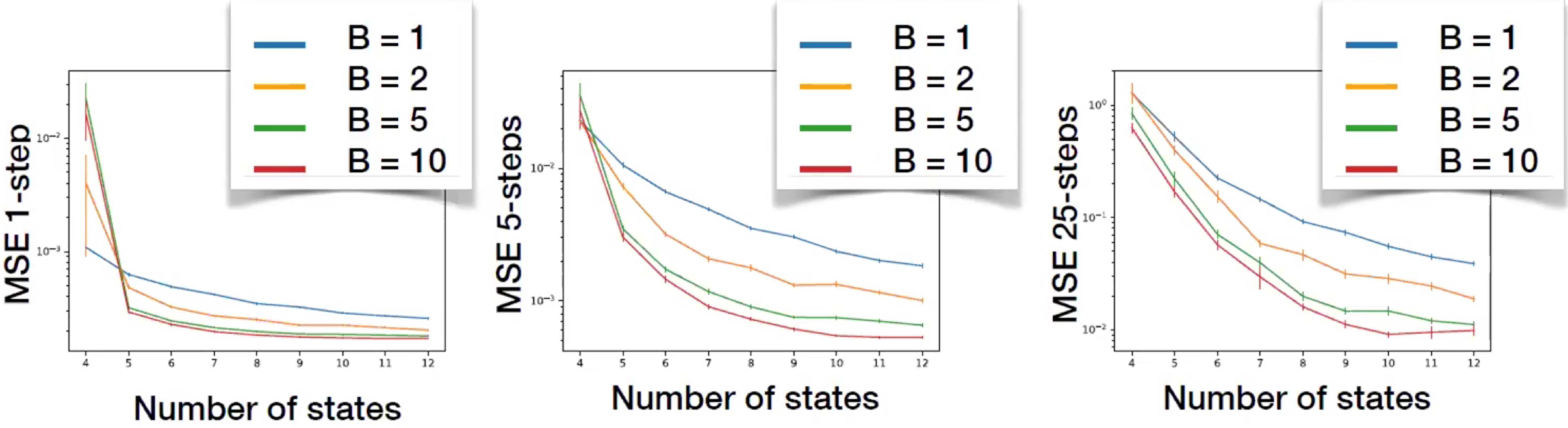




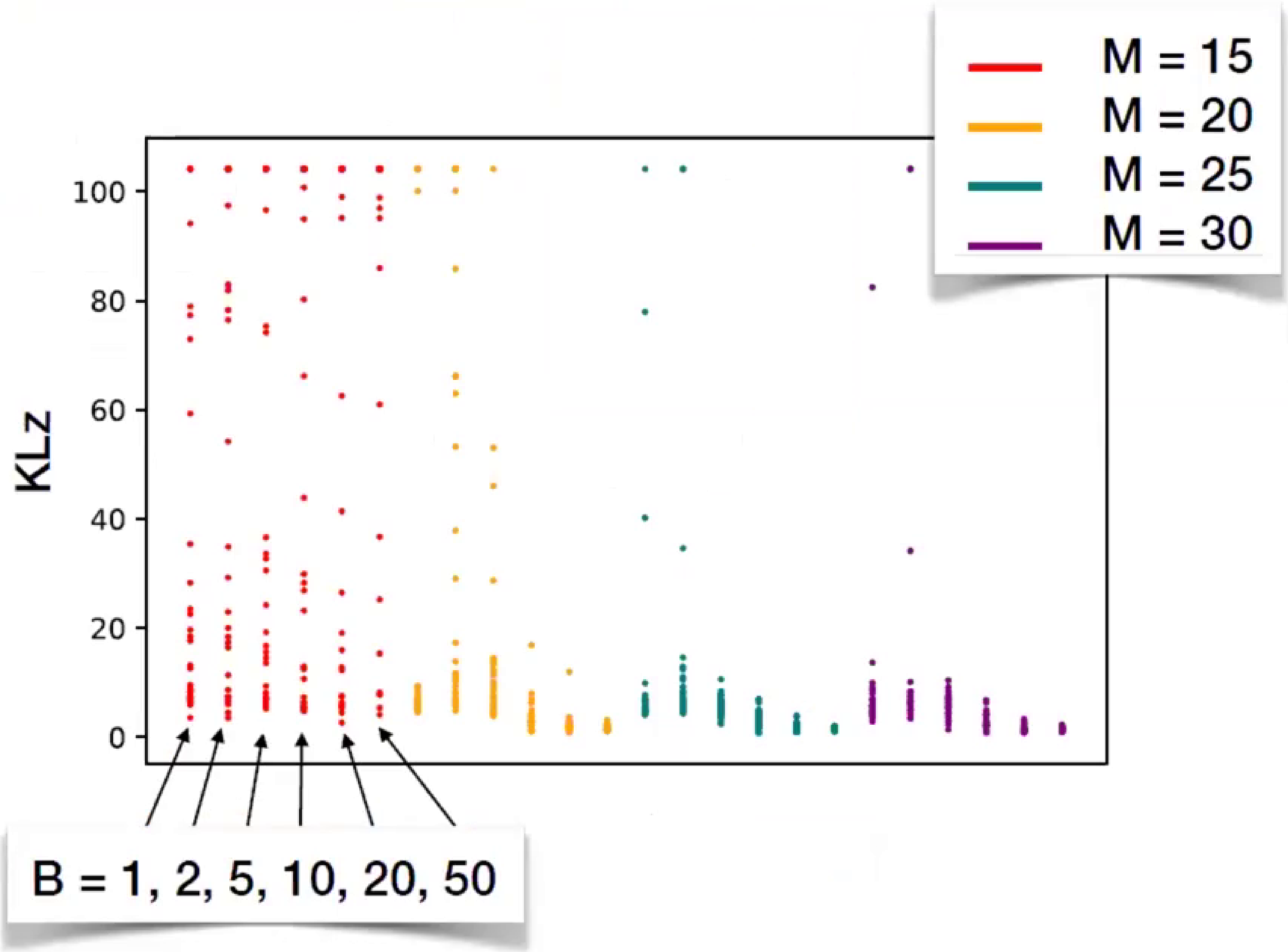
Results - Lorenz System

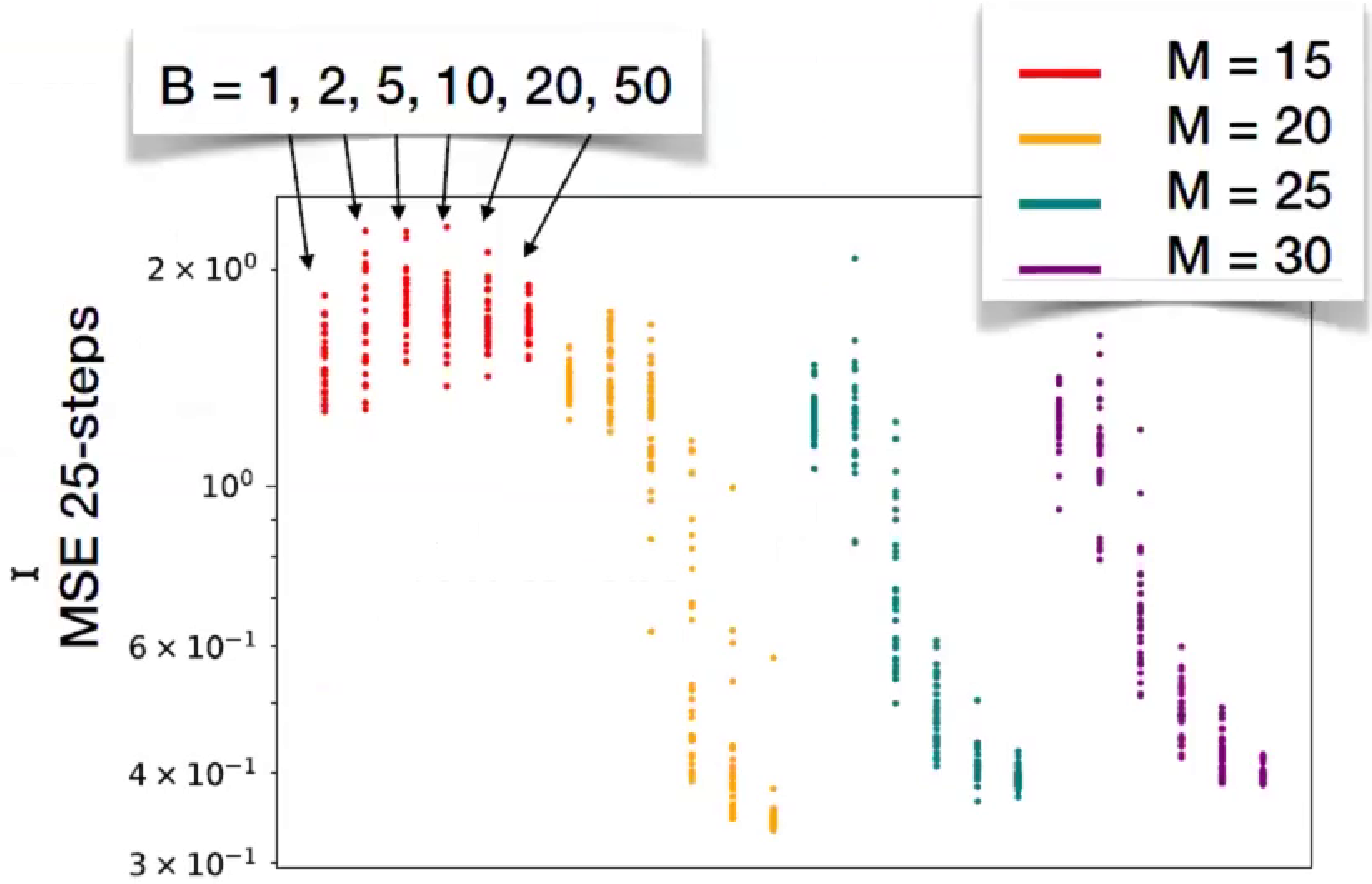


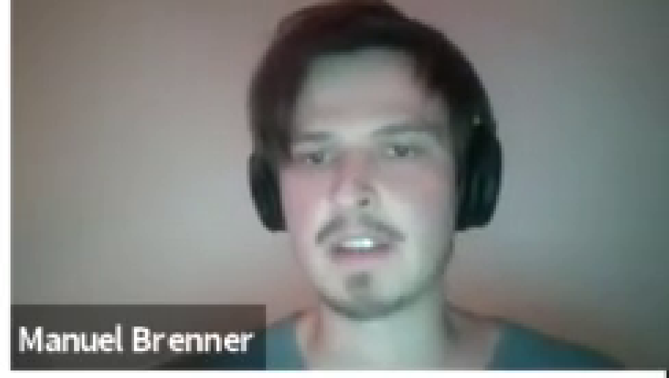




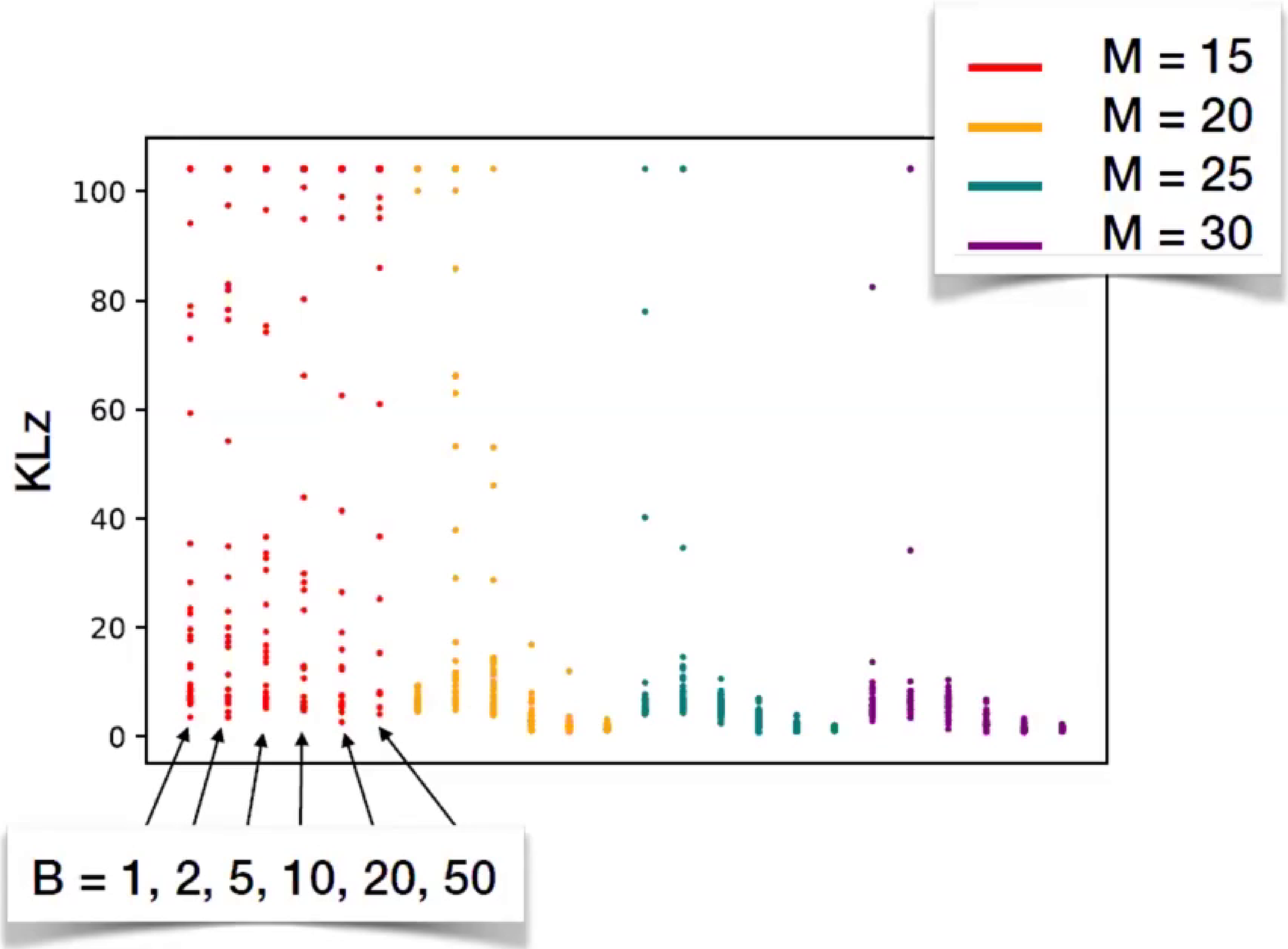
Results - Lorenz-96 System

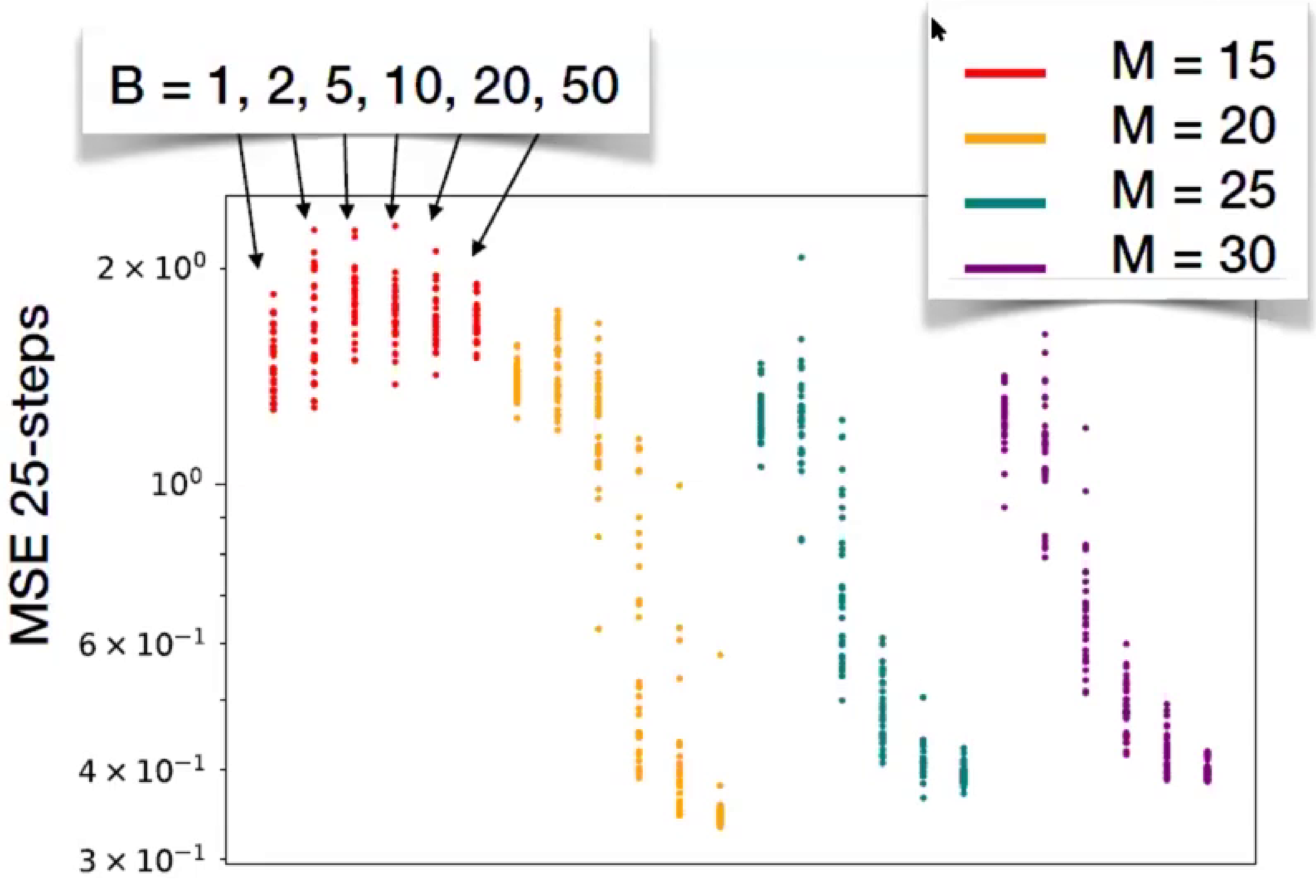






Results - Lorenz-96 System

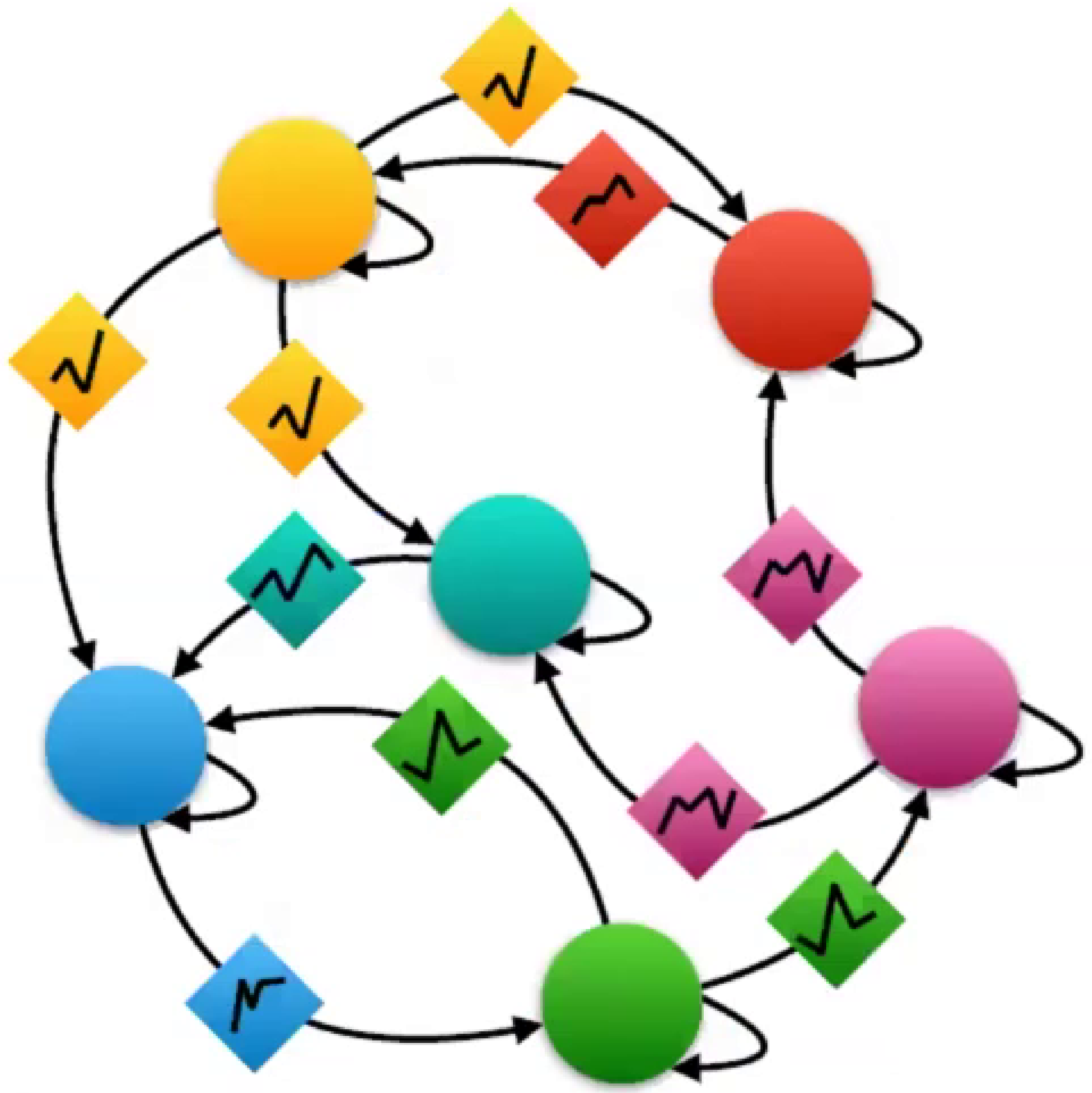




Summary

Basis expansion:

- Improves inference
- Encourages learning of more interesting dynamics
- Reduces dimensionality of the latent space

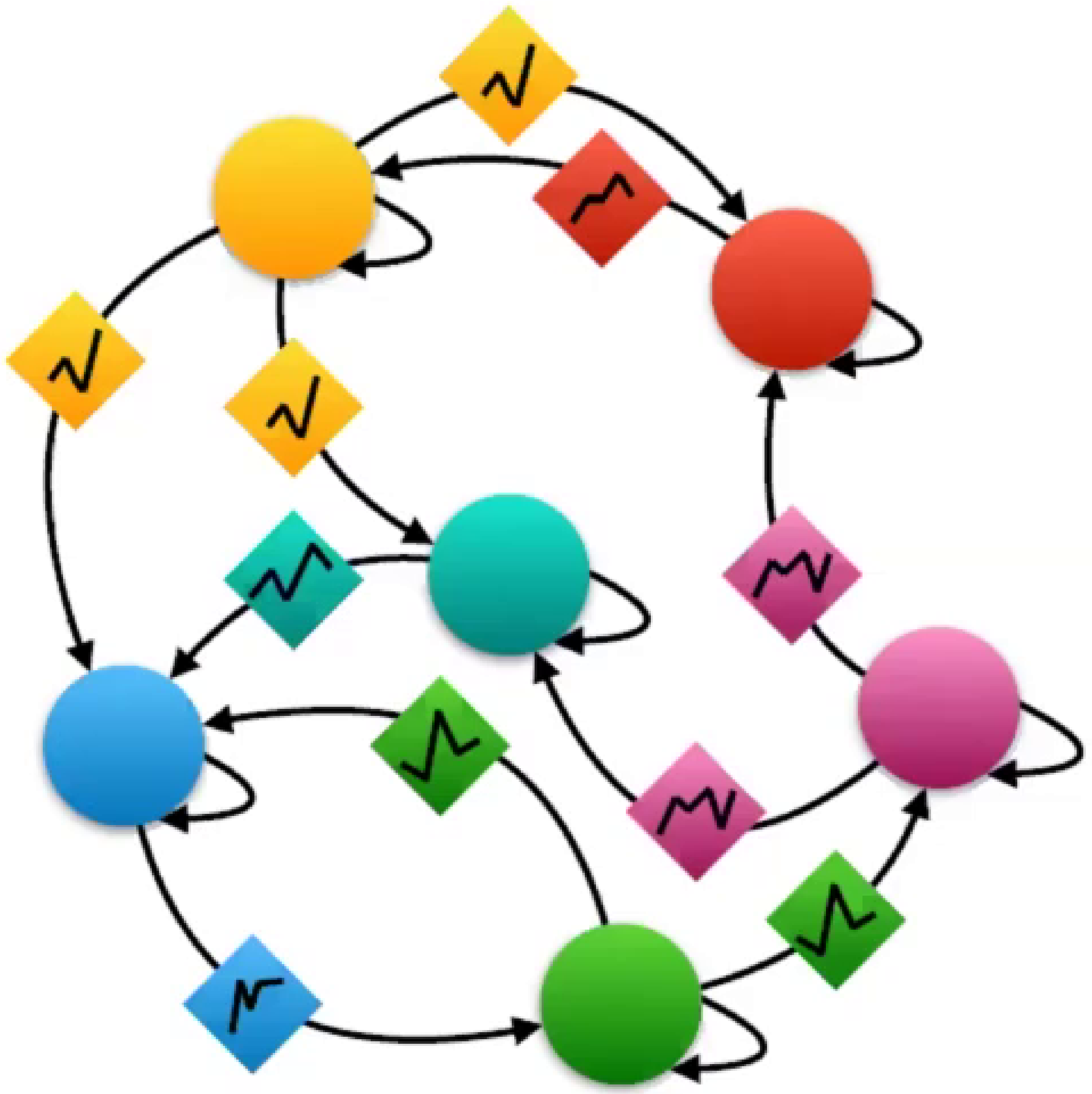


Acknowledgements: This work was funded by the **German Science Foundation** (DFG) through individual grant Du 354/10-1 to Daniel Durstewitz, and via the **Excellence Cluster *Structures* at Heidelberg University** (EXC-2181 – 390900948).

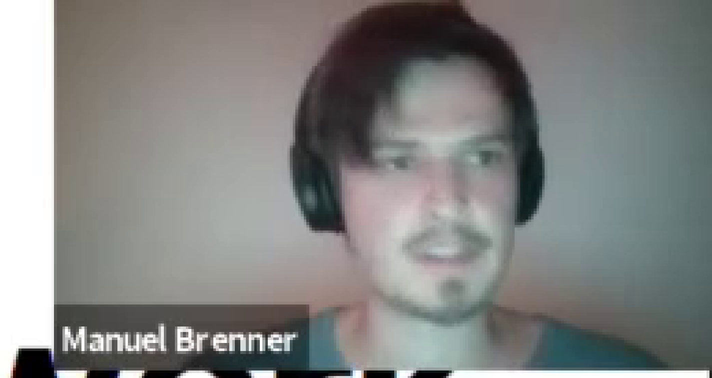
Summary

Basis expansion:

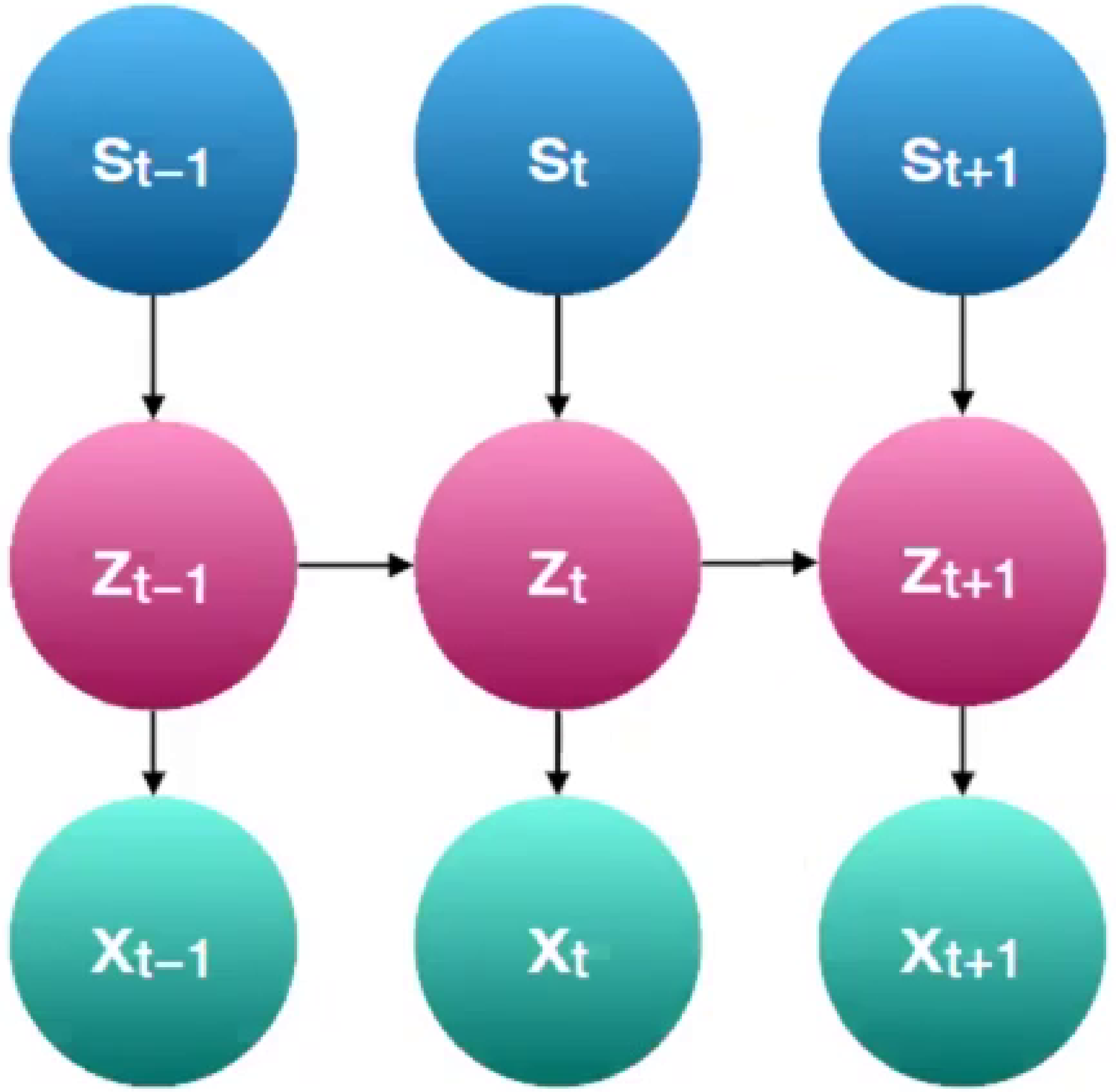
- Improves inference
- Encourages learning of more interesting dynamics
- Reduces dimensionality of the latent space



Acknowledgements: This work was funded by the **German Science Foundation** (DFG) through individual grant Du 354/10-1 to Daniel Durstewitz, and via the **Excellence Cluster *Structures* at Heidelberg University** (EXC-2181 – 390900948).



Piece-wise Linear Recurrent Neural Network



Latent Model $z_t \in \mathbb{R}^M$

$$z_t = \underset{\substack{\uparrow \\ \text{Diagonal}}}{A} z_{t-1} + \underset{\substack{\uparrow \\ \text{Off-diagonal}}}{W} \phi(z_{t-1}) + h_0 + C s_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma)$$

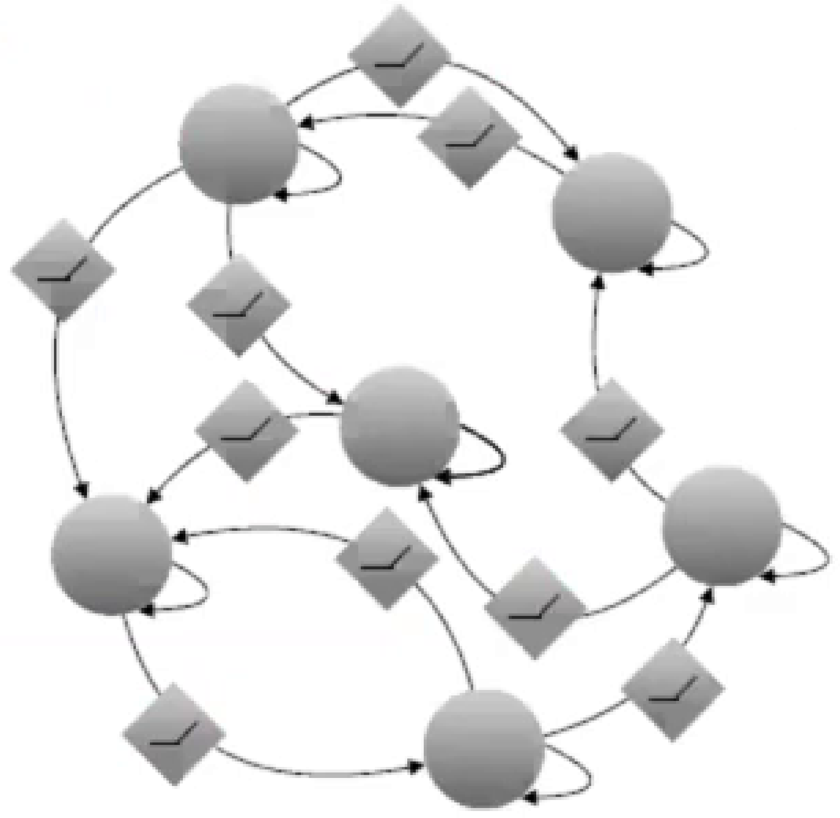
Observation Model

$$x_t = B z_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \Gamma)$$

Complete Likelihood

$$\begin{aligned} \log p_\theta(x, z) = &- \frac{1}{2}(z_1 - \mu_0 - s_1)^T \Sigma^{-1}(z_1 - \mu_0 - s_1) \\ &- \frac{1}{2} \sum_{t=2}^{T} (z_t - A z_{t-1} - W\phi(z_{t-1}) - s_t)^T \Sigma^{-1} (z_t - A z_{t-1} - W\phi(z_{t-1}) - s_t) \\ &- \frac{1}{2} \sum_{t=1}^{T} (x_t - B z_t)^T \Gamma^{-1} (x_t - B z_t) \end{aligned}$$

Activation Function



$$\phi(z_{t-1}) = \max(0, z_{t-1})$$

For **fMRI observation model** refer to: Koppe, G., Toutounji, H., Kirsch, P., Lis, S., & Durstewitz, D. (2019). Identifying nonlinear dynamical systems via generative recurrent neural networks with applications to fMRI. PLoS Computational Biology, 15.

Global vs. Local Metrics

Kullback-Leibler Divergence vs. Mean Squared Error

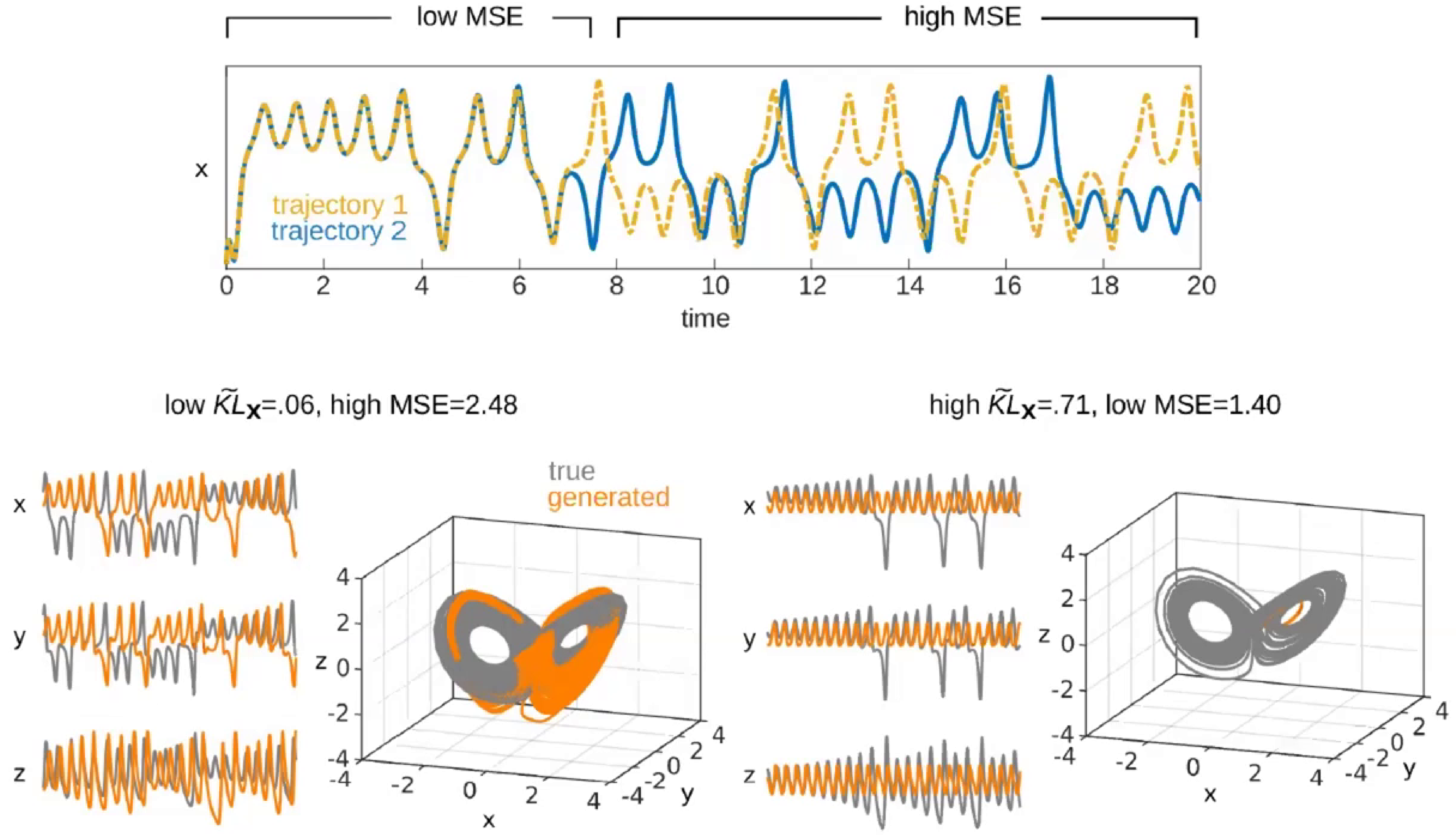


Figure from: Koppe, G., Toutounji, H., Kirsch, P., Lis, S., & Durstewitz, D. (2019).
Identifying nonlinear dynamical systems via generative recurrent neural networks with applications to fMRI. PLoS Computational Biology, 15.