KoopSTD: Reliable Similarity Analysis between Dynamical Systems via Approximating **Koop**man **S**pectrum with **T**imescale **D**ecoupling

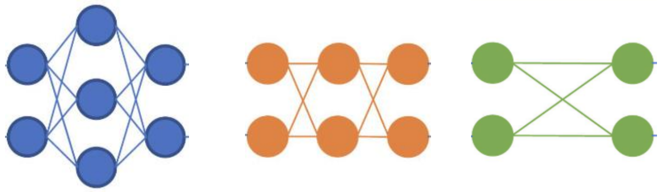
Shimin Zhang*, Ziyuan Ye*, Yinsong Yan, Zeyang Song, Yujie Wu, Jibin Wu†

Background

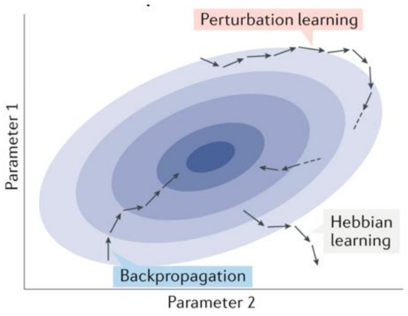
Neural representation similarity analysis: a family of computational and statistical methods designed to **quantify the similarity between representations of neural activity** across experimental conditions, time points, or between brain regions and computational models.

Why so important?

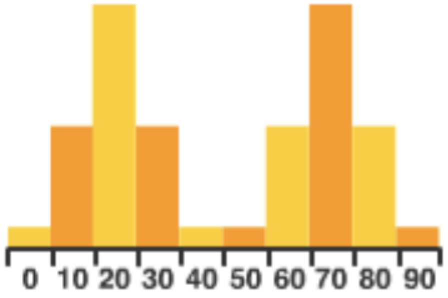
To machine learning, it reveals how internal representations of deep neural networks are influenced by:



Network Architectures

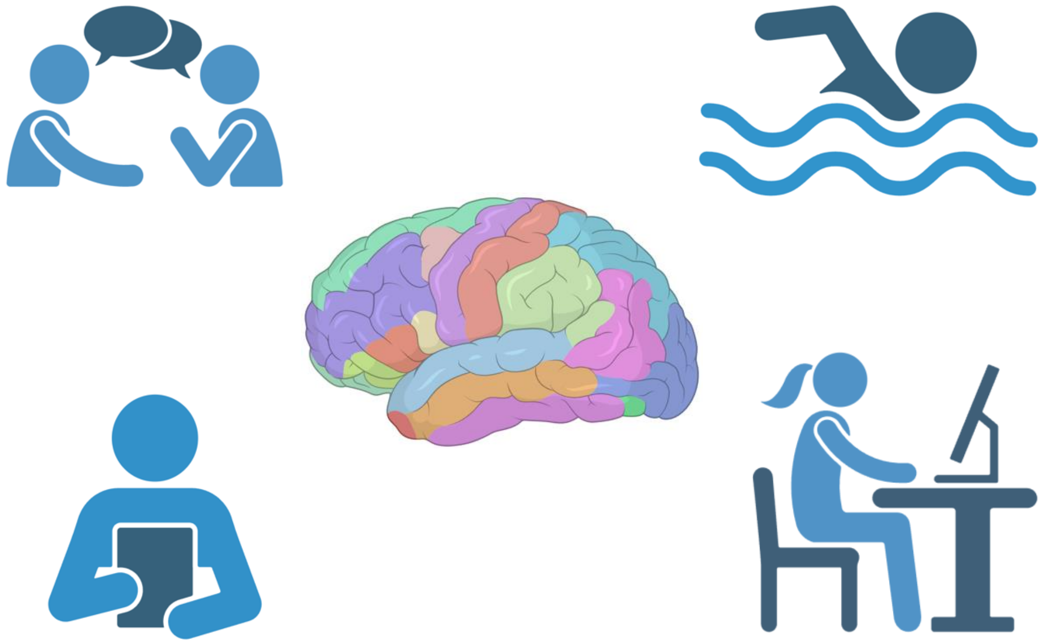


Training Methods



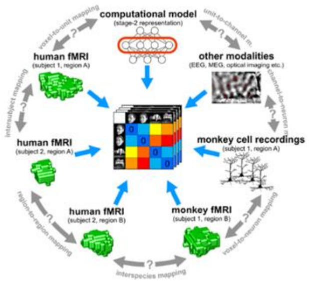
Data Distributions

To neuroscience, It reveals how brain regions encode information, offering insights into their roles in perception, cognition, and functional specialization:

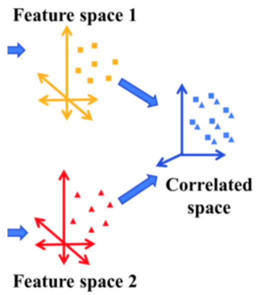


Related Work

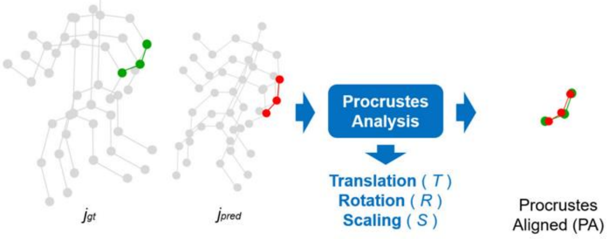
- **Representation-based Similarity Metric:**



Representation Similarity Analysis

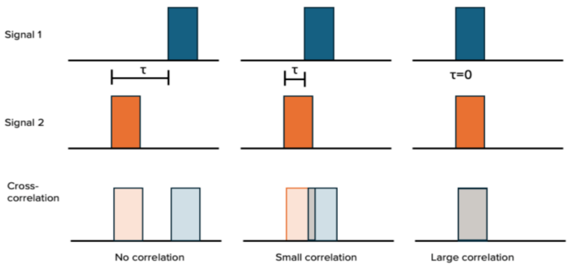


Canonical Correlation Analysis

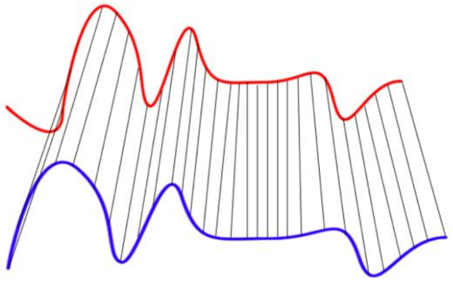


Procrustes Analysis (general shape metric)

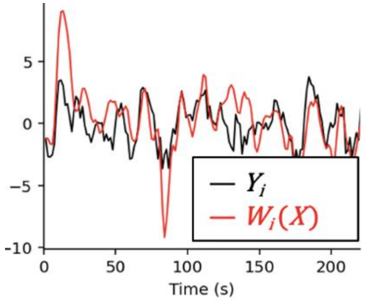
- **Dynamics-based Similarity Metric:**



Cross Correlation



Dynamic Time Warping



$$R = \mathrm{corr}(Y, W(X))$$

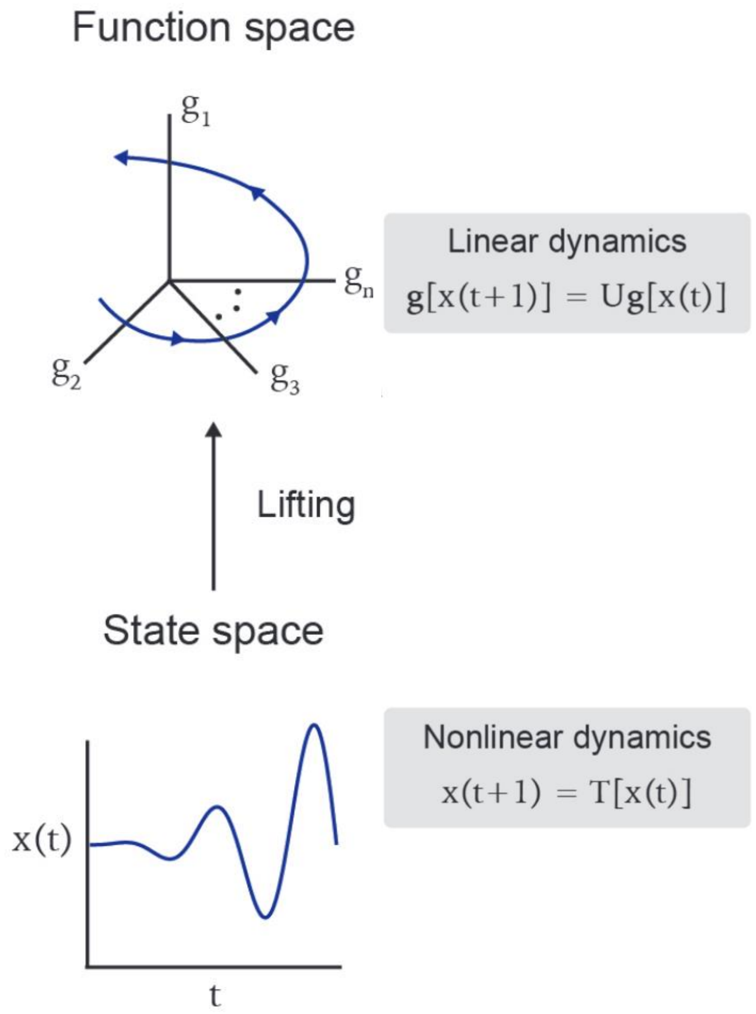
Temporal Response Function

Challenge:

- Static metrics **overlook the temporal dynamics** inherent in many real-world systems.
- Dynamic metrics **fail to capture the nonlinear, and complex temporal patterns** observed in biological and artificial systems.

Preliminary: Koopman Operator Theory

The Koopman operator theoretically **embeds nonlinear systems into infinite-dimensional Hilbert space**, which permits an **exact and globally linear description** of the dynamics.



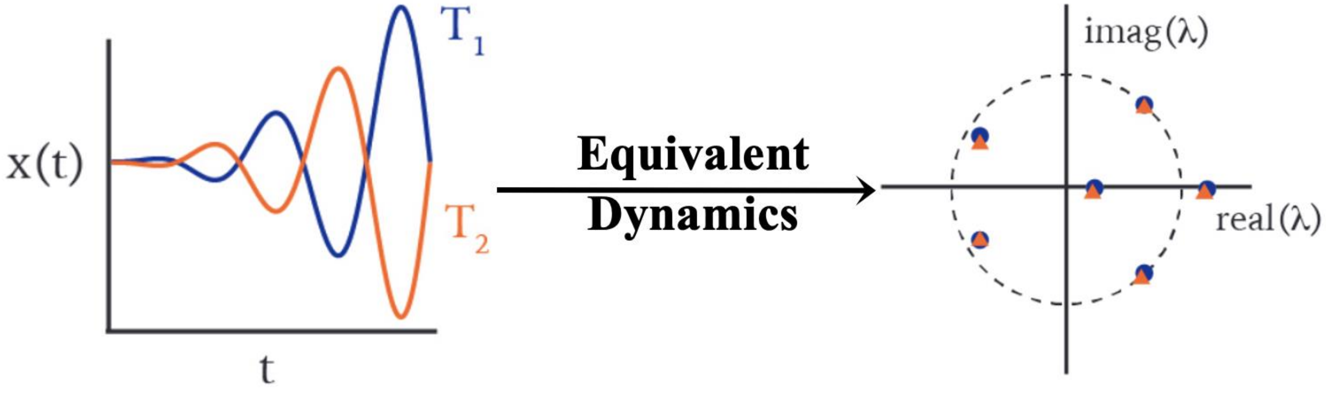
Finite approximation: **Dynamic Mode Decomposition (DMD)**

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}(t_1) & \mathbf{x}(t_2) & \cdots & \mathbf{x}(t_m) \\ | & | & & | \end{bmatrix} \qquad \mathbf{X}' \approx \mathbf{A}\mathbf{X}.$$

$$\mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathbf{x}(t_1') & \mathbf{x}(t_2') & \cdots & \mathbf{x}(t_m') \\ | & | & & | \end{bmatrix} \qquad \mathbf{A} = \underset{\mathbf{A}}{\operatorname{argmin}} \|\mathbf{X}' - \mathbf{A}\mathbf{X}\|_F = \mathbf{X}'\mathbf{X}^\dagger$$

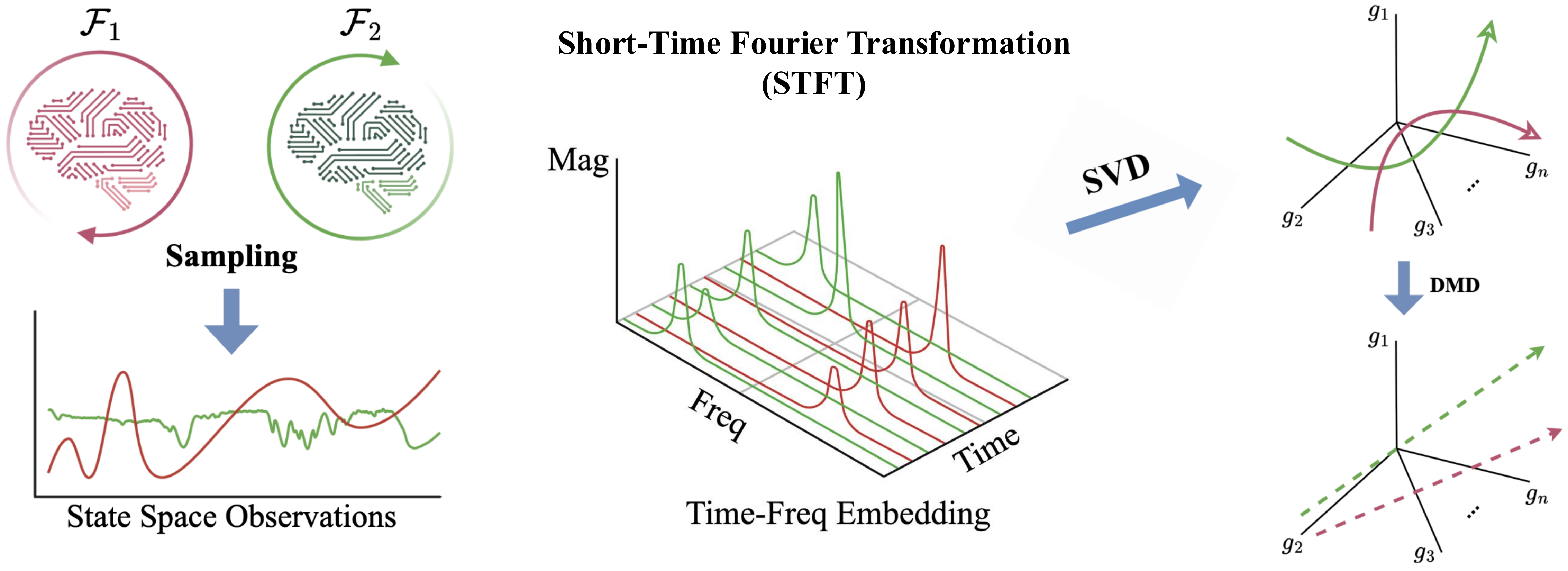
Temporal snapshots **Ordinary Least Square (OLS)**

How can it relate to the dynamic similarity?



Previous attempt: *Fujii et al., 2017; Ishikawa et al., 2018; Ostrow et al., 2024*

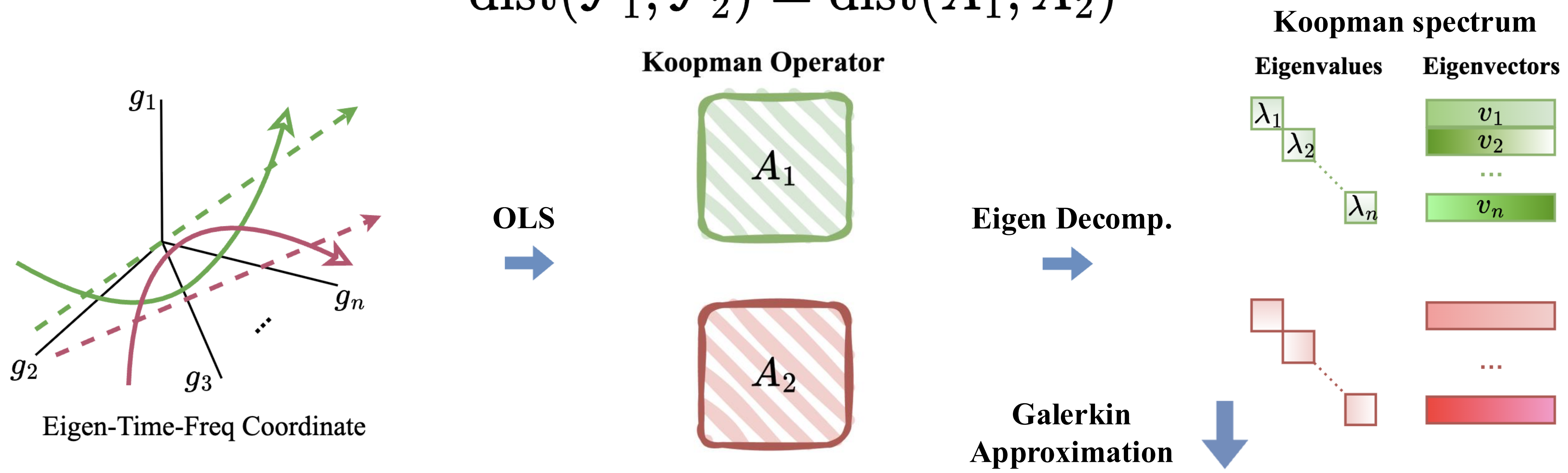
Method



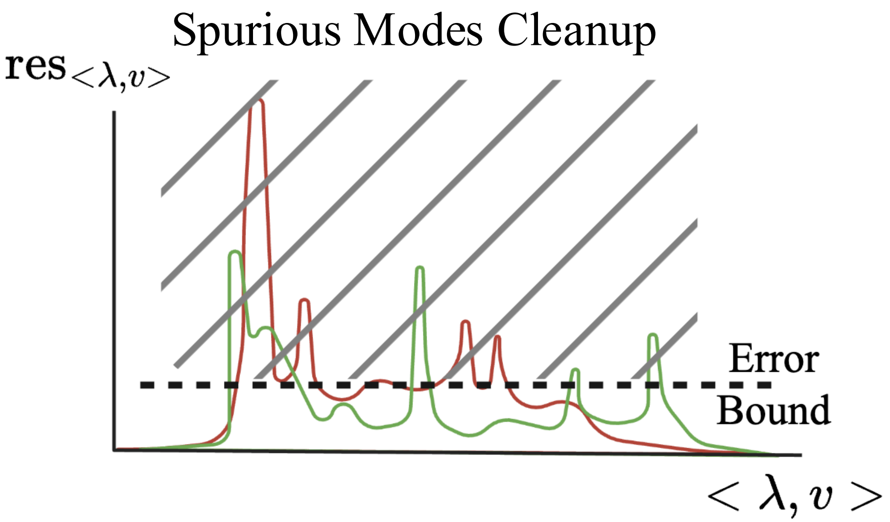
Nonlinear systems are hard to analyze directly in the time domain due to the **intricate interactions across multiple timescales.**

Method

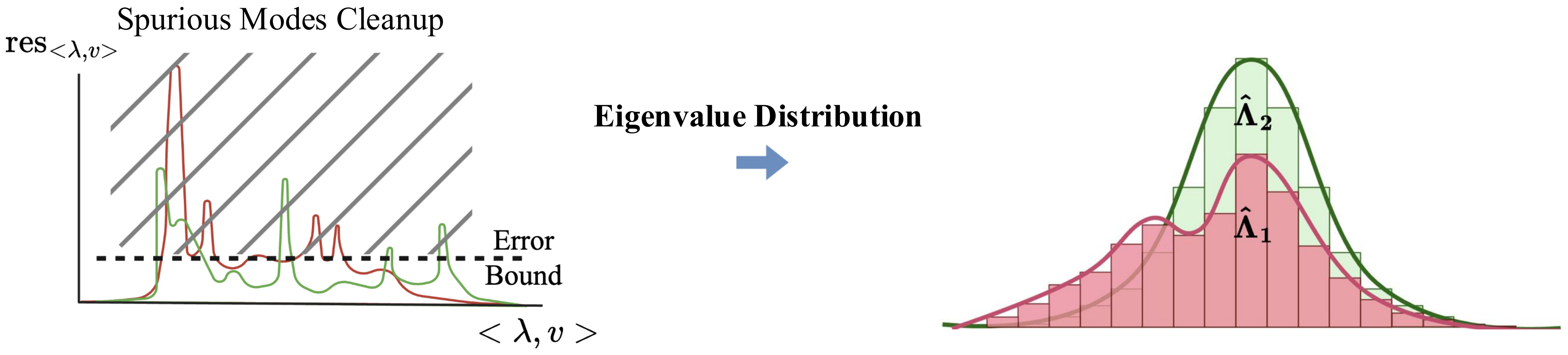
$$\mathrm{dist}(\mathcal{F}_1, \mathcal{F}_2) \triangleq \mathrm{dist}(A_1, A_2)$$



- The **right singular vector** extracts the dominant modes and captures **multi-scale temporal dynamics**.
- Question: How can we guarantee the DMD (OLS) **convergence**?
- **Spectral pollution**: Spectral discretization introduces spurious modes.

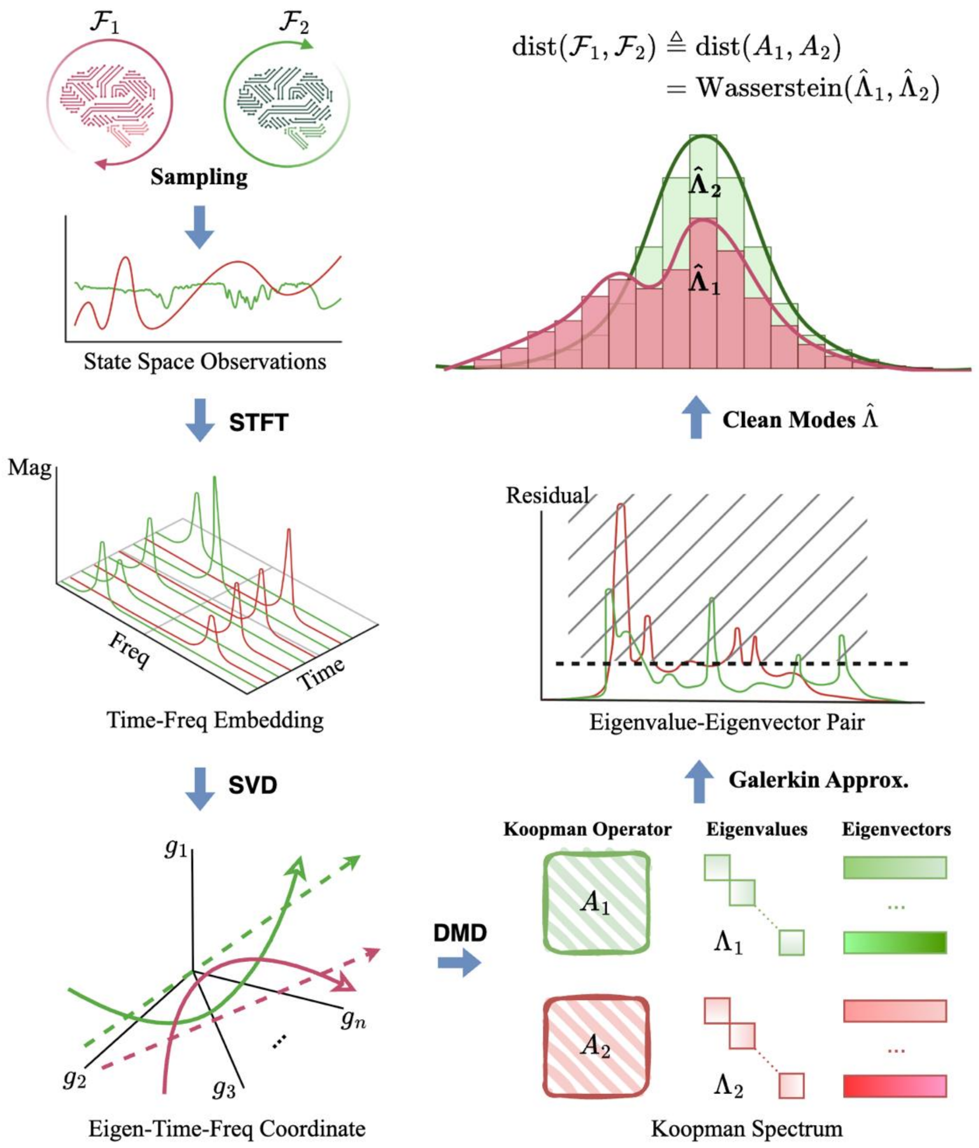


Method



$$\begin{aligned}\text{dist}(\mathcal{F}_1, \mathcal{F}_2) &\triangleq \text{dist}(A_1, A_2) \\ &= \text{Wasserstein}(\hat{\Lambda}_1, \hat{\Lambda}_2)\end{aligned}$$

Method



KoopSTD overview.

Algorithm 1 KoopSTD Pseudocode

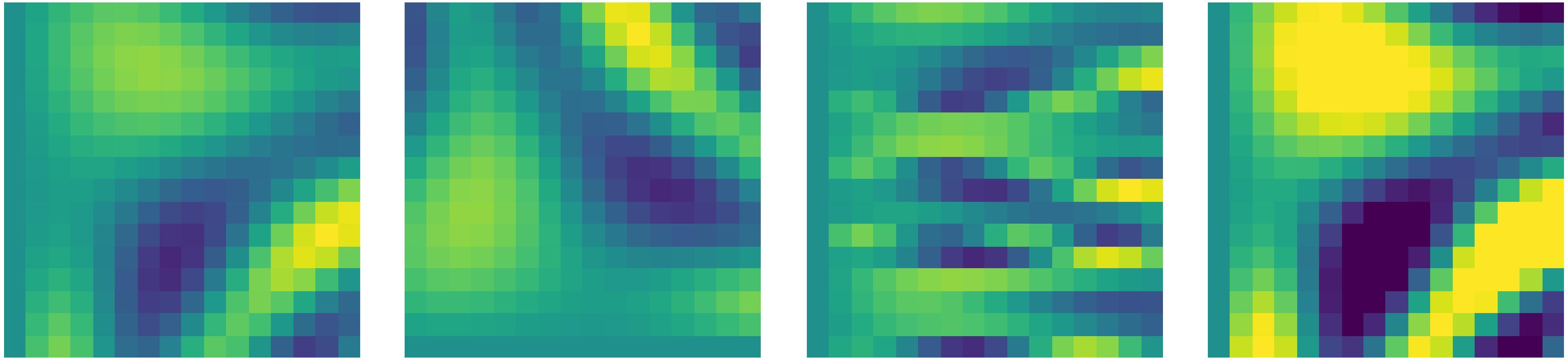
Input: two time series, $\mathbf{X}_1 \in \mathbb{R}^{T_1 \times N_{d_1}}$ and $\mathbf{X}_2 \in \mathbb{R}^{T_2 \times N_{d_2}}$; STFT window size, $l \in \mathbb{Z}^+$; STFT hop size, $s \in \mathbb{Z}^+$; number of preserved modes, $r \in \mathbb{Z}^+$
Output: Dynamics dissimilarity d between $\mathbf{X}_1$ and $\mathbf{X}_2$
Procedure $\mathrm{DMD}_{STFT}(\mathbf{X}, l, s)$
  $\mathbf{Z} = \mathrm{STFT}(\mathbf{X}, l, s)$
  Solve $\mathbf{Z} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$
  Approximate $\mathbf{A}_{tf}$ by Eq. (6)
  Return $\mathbf{A}_{tf}$
End Procedure
Procedure ResControl$(\mathbf{A}_{tf}, r)$
  Solve $\mathbf{A}_{tf}\Phi = \Phi\Lambda$ for eigenpairs $\{\hat{\lambda}_j, \hat{v}_j\}_{j=1}^{N_f}$
  for $j = 1$ to N_f **do**
    Compute the residual of $\{\hat{\lambda}_j, \hat{v}_j\}$ by Eq. (7)
  end for
  Top r accurate eigenvalues $\Lambda = \mathrm{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_r)$
  Return Λ
End Procedure
$\mathbf{A_{tf,1}} \leftarrow \mathrm{DMD}_{STFT}(\mathbf{X}_1, l, s)$
$\mathbf{A_{tf,2}} \leftarrow \mathrm{DMD}_{STFT}(\mathbf{X}_2, l, s)$
$\Lambda_1 \leftarrow$ ResControl$(\mathbf{A}_{tf,1}, r)$
$\Lambda_2 \leftarrow$ ResControl$(\mathbf{A}_{tf,2}, r)$
Compute the dynamics dissimilarity d by Eq. (8)

Transformation-Invariant Property

Let $\mathbf{X}_1[t+1] = \mathcal{F}_1(\mathbf{X}_1[t])$ and $\mathbf{X}_2[t+1] = \mathcal{F}_2(\mathbf{X}_2[t])$ be two time-discrete dynamical systems with state variables $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{N_d}$, and they are governed by mappings $\mathcal{F}_1, \mathcal{F}_2 : \mathbb{R}^{N_d} \to \mathbb{R}^{N_d}$. Now we prove that the distance $d(\mathcal{F}_1, \mathcal{F}_2)$ between two systems calculated by KoopSTD remains invariant under invertible linear transformations $\mathcal{T}$, such that:

$$d(\mathcal{T}(\mathcal{F}_1, \mathcal{F}_2)) = d(\mathcal{F}_1, \mathcal{F}_2), \tag{16}$$

where $\mathcal{T} = \{\mathbf{X} \mapsto \mathbf{XQ} : \mathbf{Q} \in GL(N_d, \mathbb{R})\}$. $GL(N_d, \mathbb{R})$ denotes the general linear group of all invertible matrices $\mathbf{Q} \in \mathbb{R}^{N_d \times N_d}$.

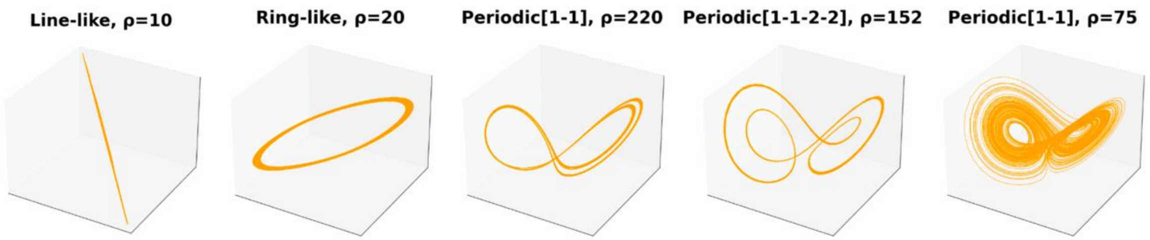


Original Pattern | Rotation | Permutation | Isotropic Scaling

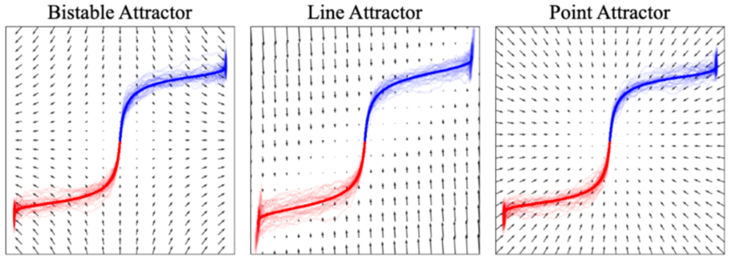
This theoretical groundedness **ensures robustness to common transformations in the representation space**, highlighting its potential for broad applicability in challenging scenarios.

Experiments

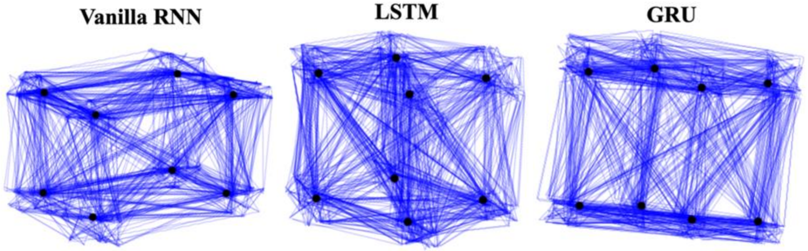
We construct **three synthetic datasets** derived from distinct physical and neural systems, each exhibiting different dynamic behaviors.



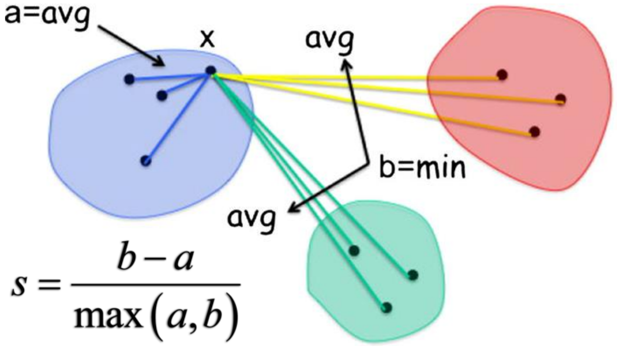
- **Dataset 1:** Trajectories of Lorenz63 system with different ρ.



- **Dataset 2:** Noisy 2D attractors for Perceptual Decision Making.



- **Dataset 3:** Hidden states of RNNs for solving the Flip-Flop task.

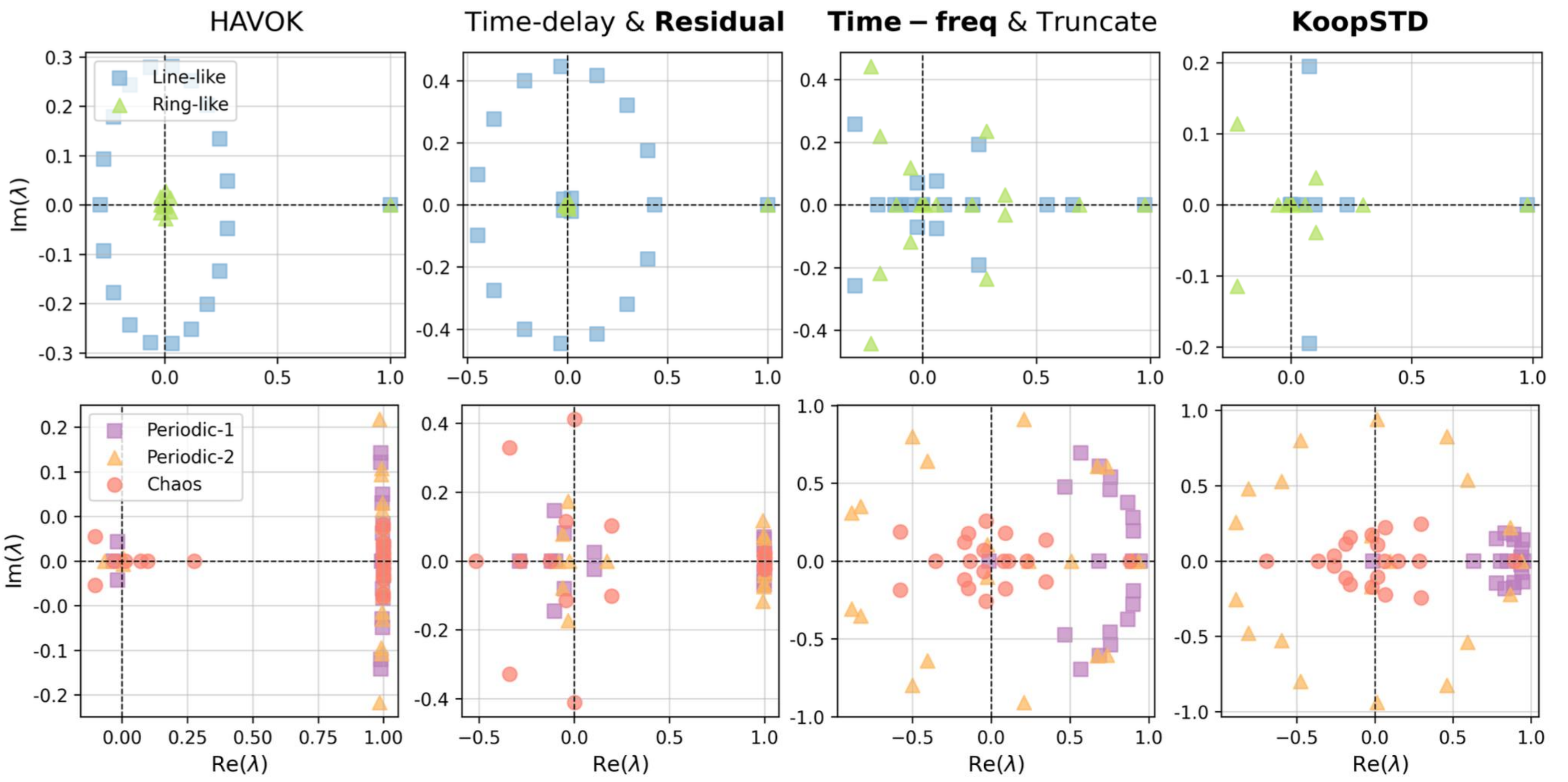


To evaluate effectiveness, we use the **Silhouette Coefficient** to quantify how well the metric distinguishes data according to their underlying dynamics.

Metrics / Systems	Representational		Dynamical		
	CKA	Procrustes	CC	HAVOK	KoopSTD
Lorenz Systems	-0.05	-0.04	-0.27	0.47	**0.94**
PDM Attractors	-0.04	-0.02	-0.30	0.90	**0.99**
Flip-Flop RNNs	0.20	0.98	-0.16	0.10	**-0.04**

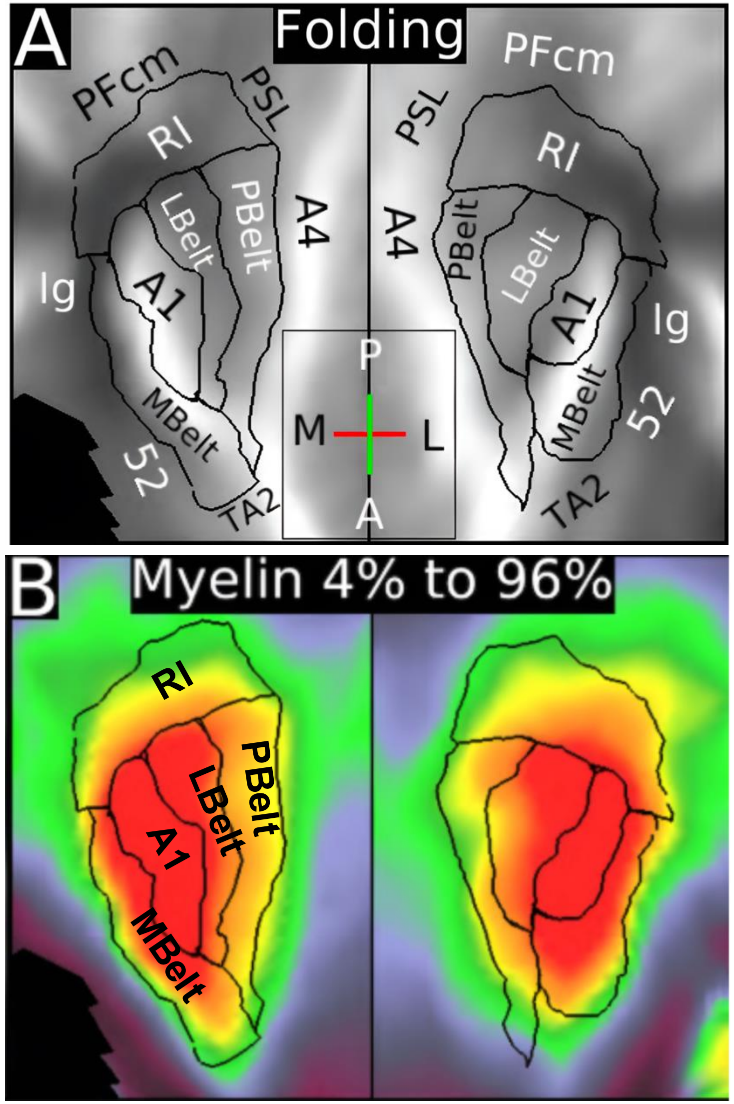
Ablation Study

We conduct an ablation study on the Lorenz63 system to separately examine the impact of **time-frequency representation** and **spectral residual control**.



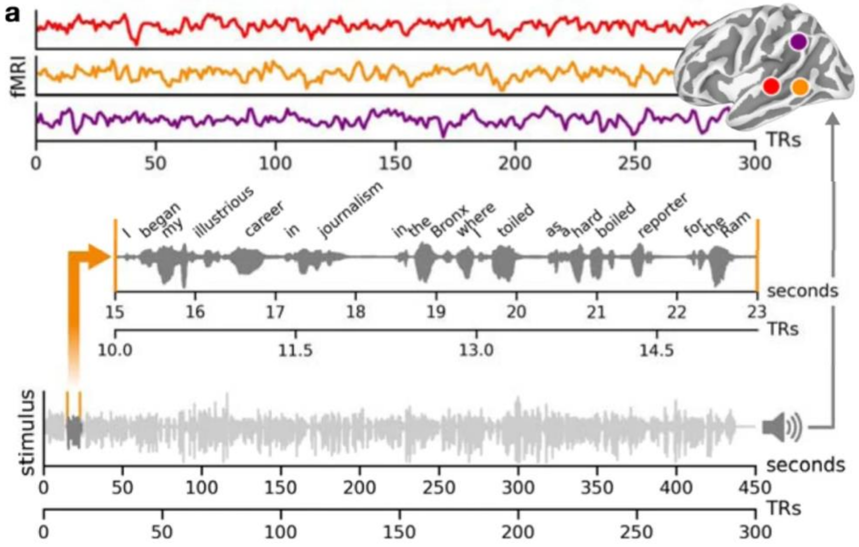
Discovery: Auditory Cortex Structural-Functional Relation

By wet experiment



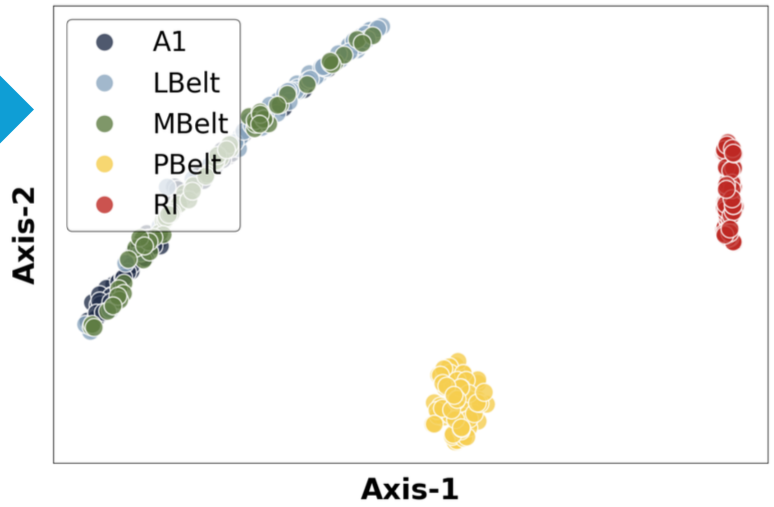
Glasser et al., 2016

Narratives fMRI dataset



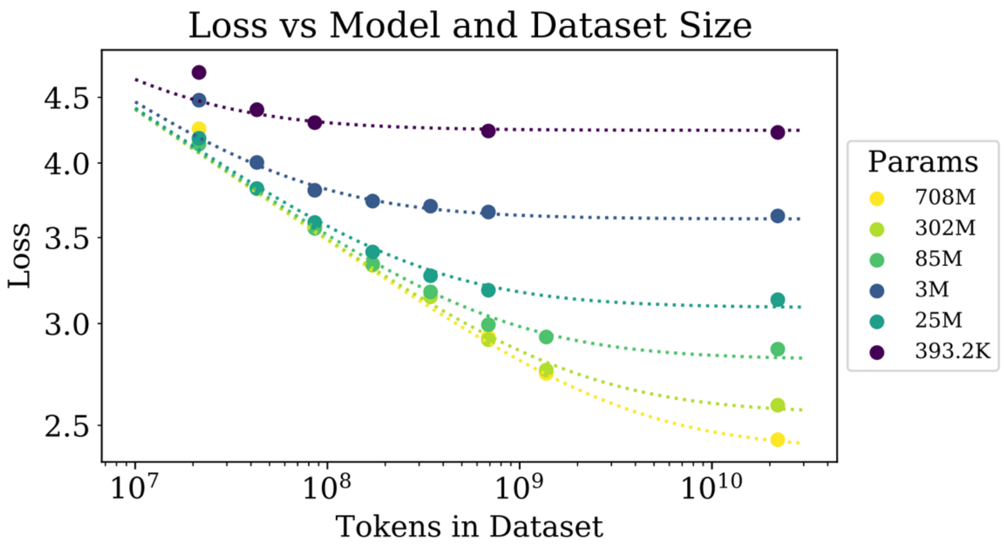
Nastase et al., 2021

By KoopSTD

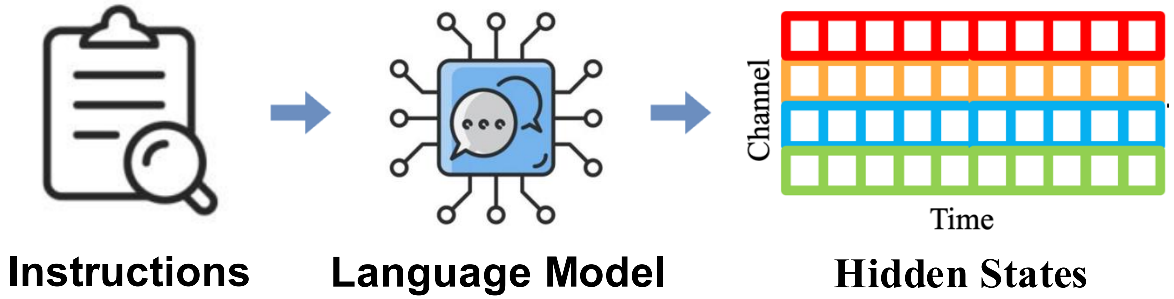


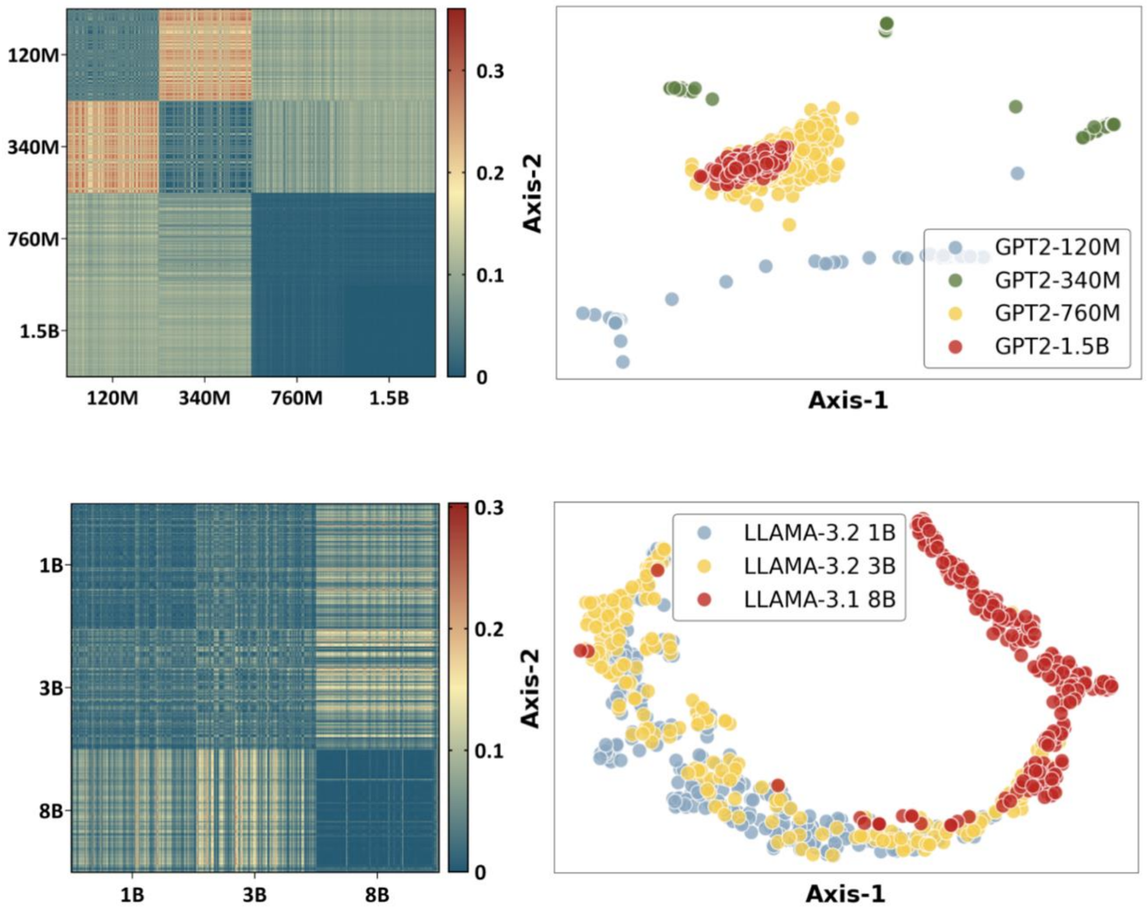
- The result from KoopSTD mirrors conclusion of myelination-based cortical parcellation.
- The potential of KoopSTD as a powerful tool for neuroscience research.

Discovery: LLMs Scaling Law



Kaplan et al., 2020





- Larger language models demonstrate **greater coherence in the dynamics of their hidden states**, whereas smaller models exhibit more divergent and unstable behaviors.
- This ***compactness* in the dynamical representation space** offers a novel perspective on the emergent capabilities of large language models.

Conclusion

- A novel similarity analysis framework **KoopSTD** for dynamical systems
- Theoretical soundness of **transformation-invariant property**
- Comprehensive experiments demonstrate **clear advantages over existing metrics**
- Great potential in **neuroscience research**
- A fresh lens on understanding the **LLM scaling law**