A Dual Koopman Approach to Low-order Modeling of Ensemble Weather Simulations

Zhicheng Zhang and Yoshihiko Susuki

Department of Electrical Engineering, Kyoto University

Summary

- Methodologies for **low-order modeling** and **anomaly detection** of **ensemble weather simulations** are developed.
- A **dual Koopman approach** to **deviation dynamics** is introduced for our development as a new mathematical framework.
- An **explicit formula** characterizing the **accuracy** of the low-order modeling is derived, which is computationally tractable using DMD methods (e.g., kernel eDMD).

Motivation

Given: Ensemble (a collection of trajectories from M initial conditions $x_0^1, x_0^2, \ldots, x_0^M$)

Goal: To establish methodologies and tools to
(1) *develop a low-order model representing the given ensembles*, and
(2) *detect anomalous members (possibly related to bad weather conditions, e.g., heavy rainfall).*

$$\begin{bmatrix} y_0^1 & y_1^1 & y_2^1 & \ldots & y_{N-1}^1 \\ y_0^2 & y_1^2 & y_2^2 & \ldots & y_{N-1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_0^M & y_1^M & y_2^M & \ldots & y_{N-1}^M \end{bmatrix}$$
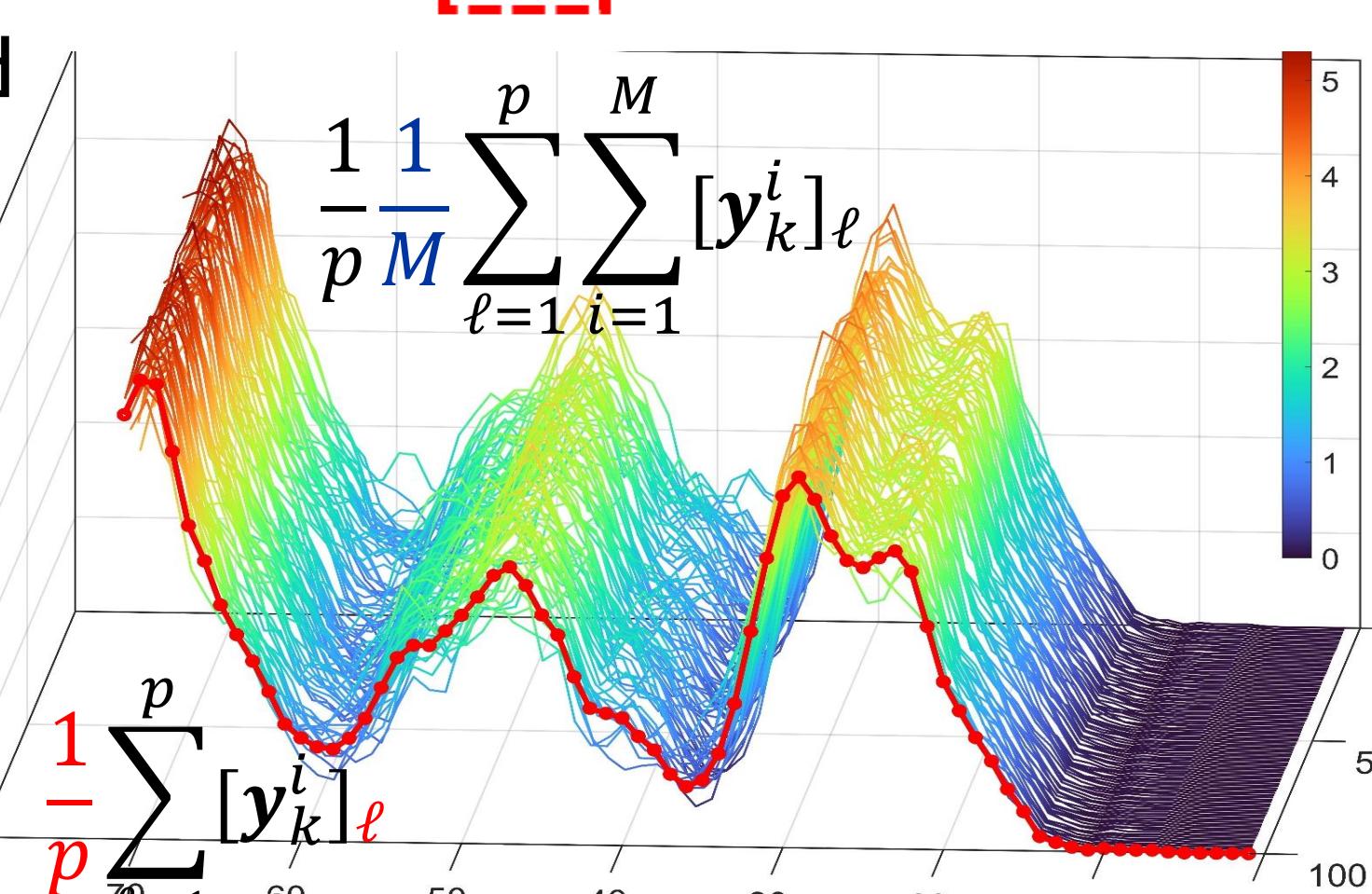
Key Idea: *Focus on "deviation dynamics" between an ensemble and their ensemble mean!*

$$\frac{1}{p}\frac{1}{M}\sum_{\ell=1}^{p}\sum_{i=1}^{M}[y_k^i]_\ell$$

Method: *Dual Koopman acting on kernel sections decodes ensemble members via $\kappa_{x_0^i}$.*
(i.e., the *RKHS feature representation of the point x_0^i*)

$$\frac{1}{p}\sum_{\ell=1}^{p}[y_k^i]_\ell$$

Preliminaries and Problem Formulation

- **Ensemble dynamics with different initial realizations (i.e., nonlinear evolution)**

$$U^k f = f \circ T^k$$

$$x_{k+1}^i = T(x_k^i) := T^k(x_0^i), \quad x_k \in \mathcal{X}, i = 1, \ldots, M$$

$$y_k^i = f(x_k^i), \qquad f: \mathcal{X} \to \mathbb{R}^p, \qquad f \in \mathcal{F},$$

Distinguishable? VS.

$$(U^k)^* \kappa_{x_0^i} = \kappa_{T^k(x_0^i)}$$

- **Ensemble Mean Dynamics** (i.e., ensembles M)

$$\langle y_k \rangle := \frac{1}{M}\sum_{i=1}^{M} y_k^i = \frac{1}{M}\sum_{i=1}^{M} U^k f(x_0^i), \implies (U^k)^* \bar{\kappa}_{x_0} := \frac{1}{M}\sum_{i=1}^{M}(U^k)^*\kappa_{x_0^i} = (U^k)^*\left(\frac{1}{M}\sum_{i=1}^{M}\kappa_{x_0^i}\right)$$

Linear evolution!

Problem 1 (Ensemble Trajectories vs. Ensemble Mean Trajectory): Can we provide an explicit formula to quantify the deviation from the vector-valued ensemble mean $y_k^i - \langle y_k \rangle$ in the ensemble, that is, multiple trajectories.

Koopman and its Dual Koopman Approach

- **Koopman and its Dual Koopman Mode Decomposition on RKHS [1]**

Standard Koopman on RKHS:

$$U^k f = \sum_{j=1}^{\infty} \lambda_j^k \psi_j \underbrace{\langle \phi_j, f \rangle}_{v_j(f)}, f \in \mathcal{H}$$

Koopman (X): In-Distinguish the ensemble member between any x_0^i using the **same observable f**.

$$\langle \phi_i, \psi_i \rangle = \delta_{ij}$$
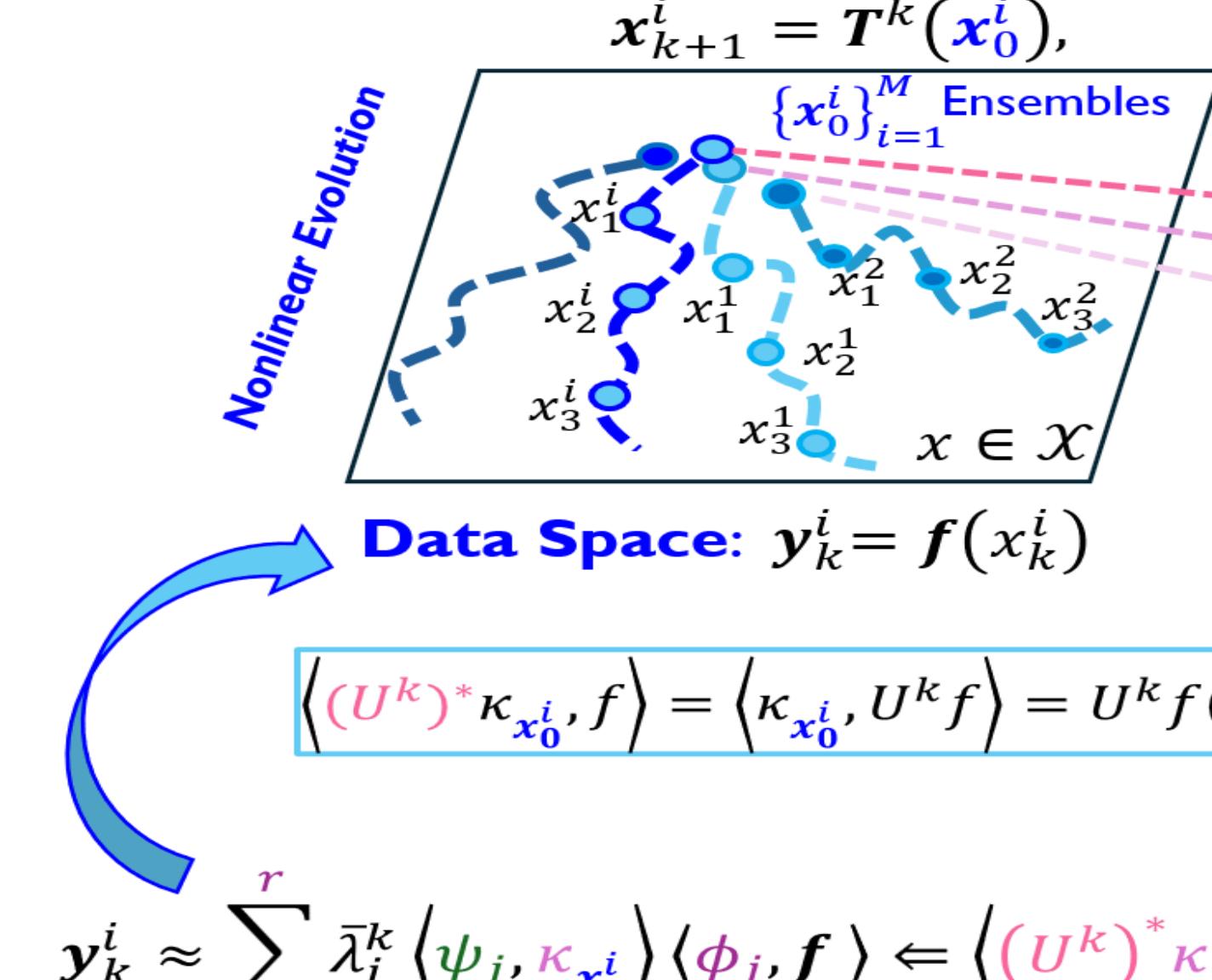
Dual Koopman on RKHS:

$$(U^k)^* g = \sum_{j=1}^{\infty} \bar{\lambda}_j^k \phi_j \langle \psi_j, g \rangle, g \in \mathcal{H}$$

Relationship

Koopman & its Dual:

$$\langle f, (U^k)^* g \rangle_{\mathcal{H}} = \langle U^k f, g \rangle_{\mathcal{H}}$$
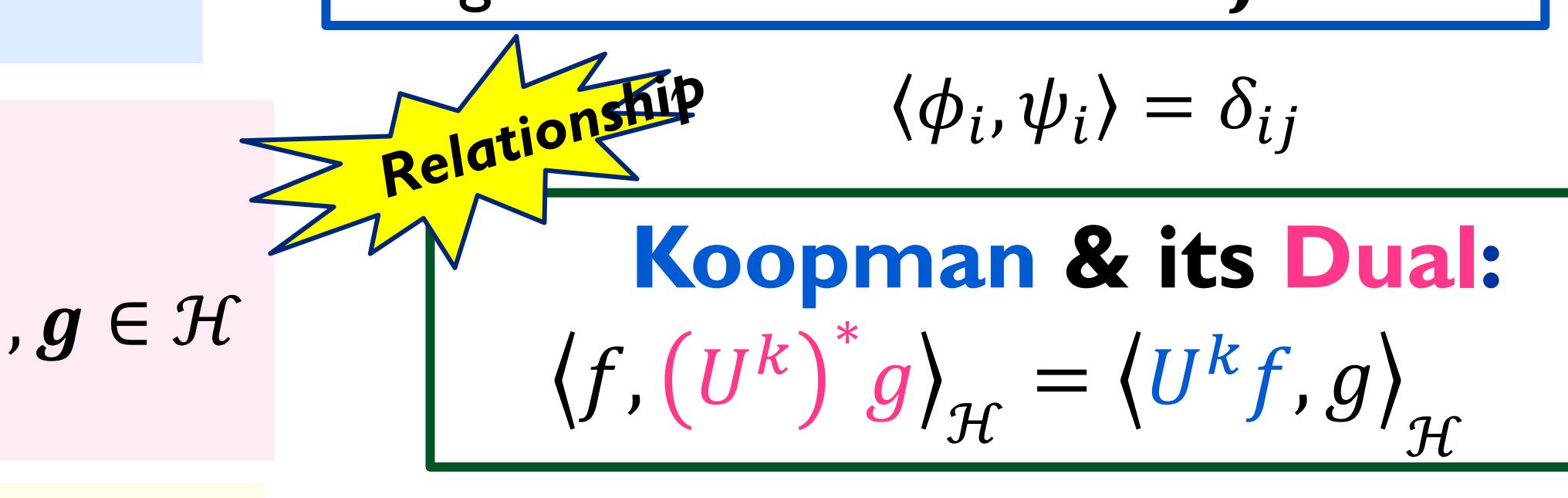
Let $g = \kappa_{x_0^i} \in \mathcal{H}$ (RKHS)

$$(U^k)^* \kappa_{x_0^i} = \sum_{j=1}^{\infty} \bar{\lambda}_j^k \phi_j \underbrace{\langle \psi_j, \kappa_{x_0^i} \rangle}_{\psi_j(x_0^i)},$$

Dual Koopman ($\checkmark$): Distinguish the ensemble member between any x_0^i using the **kernel sections $\kappa_{x_0^i}$**.

State Space

$$x_{k+1}^i = T^k(x_0^i),$$

$\{x_0^i\}_{i=1}^M$ Ensembles

Reproducing Kernel Hilbert Space

$$\langle \kappa_x, f \rangle = f(x) \qquad \kappa_x \in \mathcal{H}$$

$$U\psi_j = \lambda_j \psi_j$$
$$U^*\phi_j = \bar{\lambda}_j \phi_j$$

Dual Koopman $\langle \phi_i, \psi_j \rangle = \delta_{ij}$

$$(U^k)^* \kappa_{x_0^i} = \kappa_{T^k(x_0^i)} = \kappa_{x_k}$$

Data Space: $y_k^i = f(x_k^i)$

$$\langle (U^k)^* \kappa_{x_0^i}, f \rangle = \langle \kappa_{x_0^i}, U^k f \rangle = U^k f(x_0^i)$$

$$y_k^i \approx \sum_{j=1}^{r} \bar{\lambda}_j^k \underbrace{\langle \psi_j, \kappa_{x_0^i} \rangle}_{\psi_j(x_0^i)} \underbrace{\langle \phi_j, f \rangle}_{c_j} \Leftarrow \langle (U^k)^* \kappa_{x_0^i}, f \rangle$$

$$\mathcal{H}_r = \text{span}\{\phi_1, \ldots, \phi_r\}$$
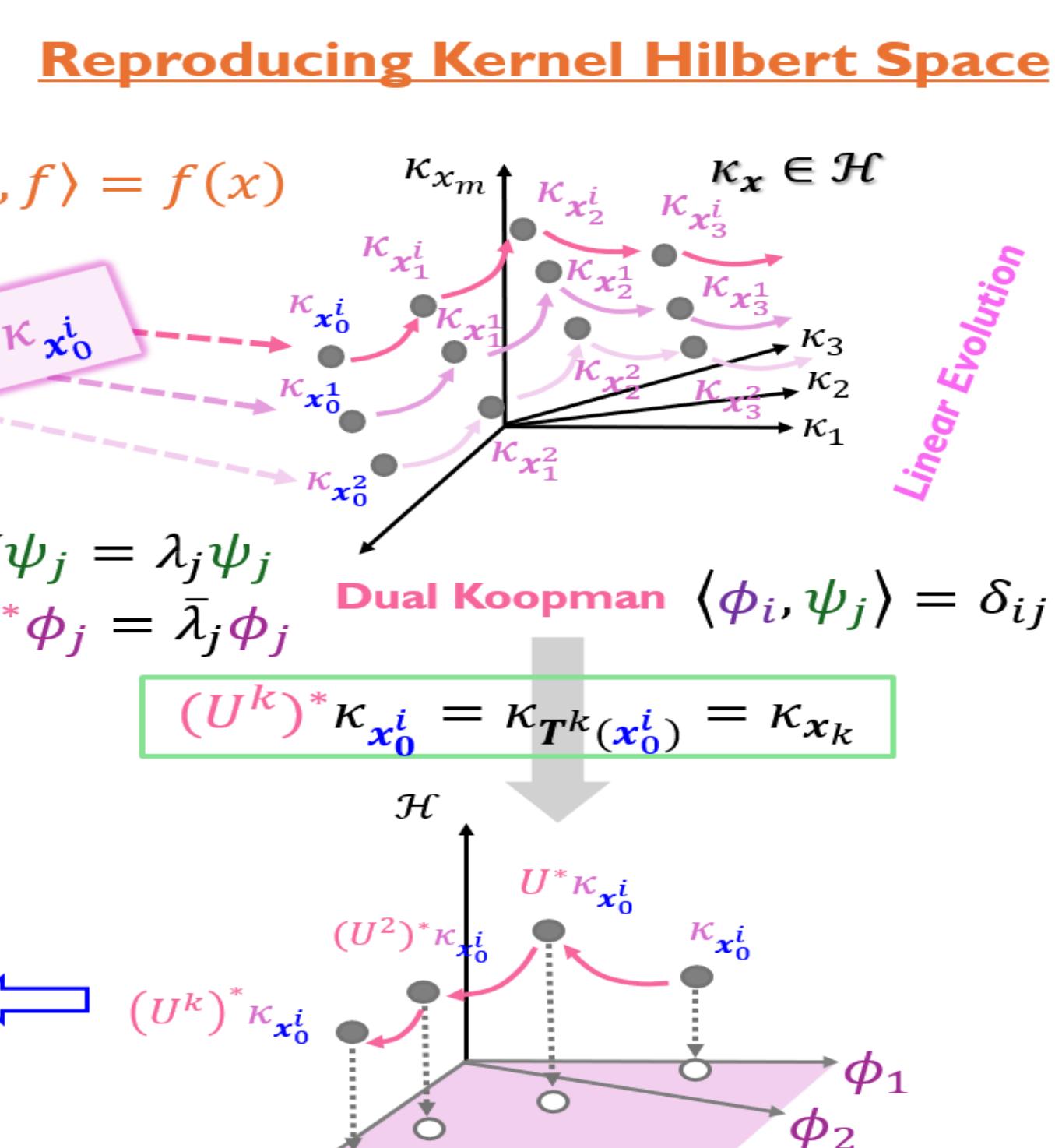
Sketch of Dual Koopman

Dual Koopman Theory to Deviation Dynamics

Assumption: All ensemble members (trajectories) converge to a **common attractor.**

Assumption (Invariant subspace) Let evaluation functional (or kernel section) $\kappa_{x_0^i} \in \mathcal{H}$ on RKHS. Assume there exists a finite dimensional Koopman-invariant subspace, spanned by the leading dual Koopman eigenfunctions $\mathcal{H}_r = \text{span}\{\phi_1, \ldots, \phi_r\}$.

Proposition 1 (Dual KMD Viewpoint of Problem 1): *Dual KMD provides an explicit formula to quantify the deviation from the vector-valued* **ensemble trajectory** *and its* **ensemble mean trajectory,** $y_k^i - \langle y_k \rangle$, *in the ensemble members, which can be reformulated as the* underlined deviation of kernel sections (or evaluation functionals) in RKHS:

$$y_k^i - \langle y_k \rangle := C_f \mathcal{E}_k^i = \langle (U^k)^*(\kappa_{x_0^i} - \bar{\kappa}_{x_0}), f \rangle = \sum_{j=1}^{\infty} \bar{\lambda}_j^k \langle \psi_j, \kappa_{x_0^i} - \bar{\kappa}_{x_0} \rangle \underbrace{\langle \phi_j, f \rangle}_{c_j}$$

$$\|R_k^i\| \le \sum_{j \ge r+1} |\bar{\lambda}_j|^k |\langle \psi_j, \kappa_{x_0^i} - \bar{\kappa}_{x_0} \rangle| |C_j| \approx \sum_{j=1}^{r} \bar{\lambda}_j^k \langle \psi_j, \kappa_{x_0^i} - \bar{\kappa}_{x_0} \rangle C_j + R_k^i$$

- If for some ensemble index $i \in \{1, \ldots, M\}$, the residual R_k^i is small, then it gives the surrogate model $\mathcal{H}_r = \text{span}\{\phi_1, \ldots, \phi_r\}$ by the dominant dual Koopman KEFs, which well captures the **ensemble mean dynamics**.

Numerical Experiment: PREC Deviation Analysis

Return to Ensemble Data Space:

$$\langle (U^k)^* \kappa_{x_0^i}, f \rangle_{\mathcal{H}} = U^k f(x_0^i) = y_k^i$$

$$y_k^i - \langle y_k \rangle \approx \sum_{j=1}^{r} \bar{\lambda}_j^k \underbrace{\langle \kappa_{x_0^i} - \bar{\kappa}_{x_0}, \phi_j \rangle}_{\Delta \alpha_j^i(0)} \underbrace{\langle \phi_j, f \rangle}_{c_j} + R_k^i$$

Ensemble Precipitation Weather data:
- $Y^i \in \mathbb{R}^{20860 \times 71}$
- Weather simulation region: 127--132 E, 30--34 N,
- Snapshots N=71, p=20860, Ensembles M =100

- Dual Koopman amplitudes with time evolution:

$$\alpha_j^i(k) := \bar{\lambda}_j^k \psi_j(x_0^i) = \bar{\lambda}_j^k \langle \psi_j, \kappa_{x_0^i} \rangle,$$
$$\bar{\alpha}_j(k) = \bar{\lambda}_j^k \langle \psi_j, \bar{\kappa}_{x_0} \rangle$$
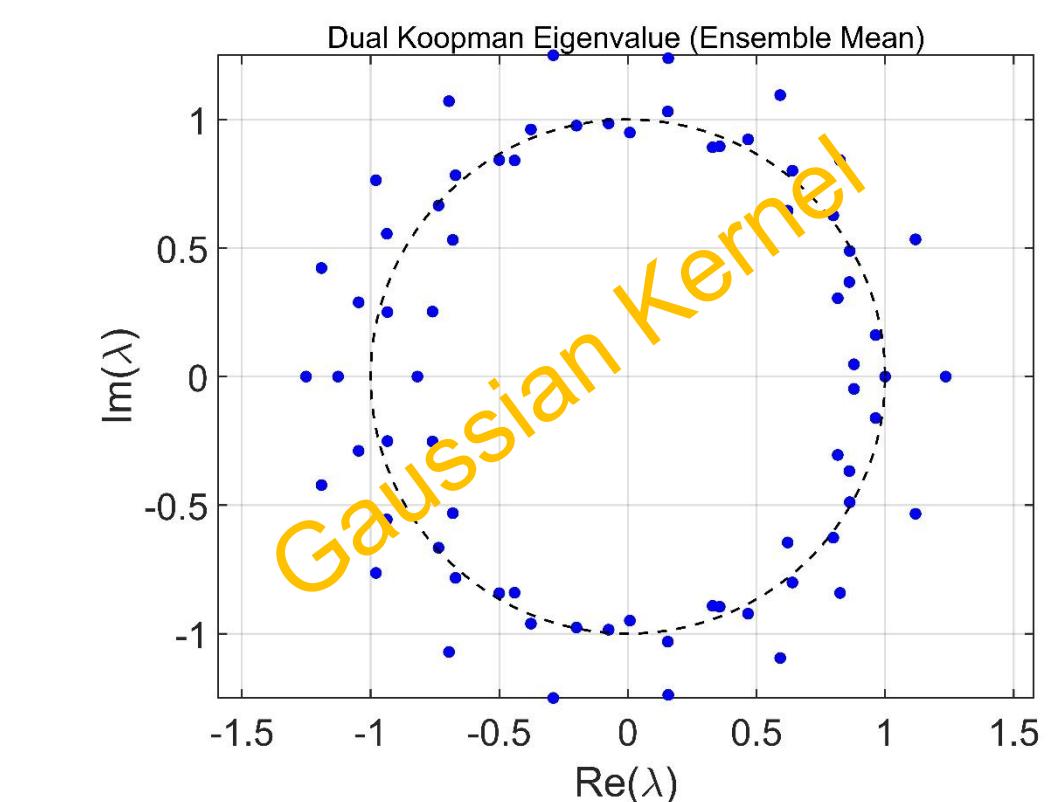
Fig. 1: Dual Koopman Eigenvalues and its absolute values
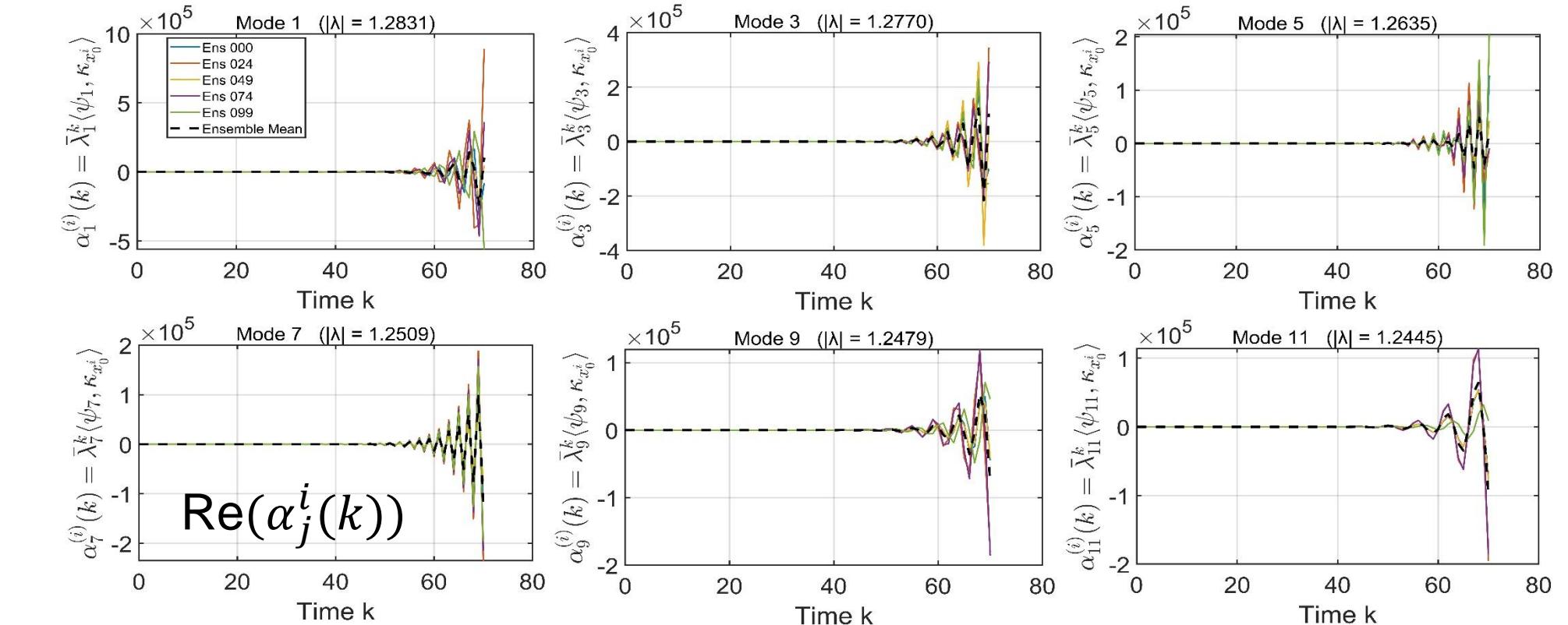
Fig. 2: Time-evolved *dual* Koopman amplitudes act on kernel sections (**Gaussian**)
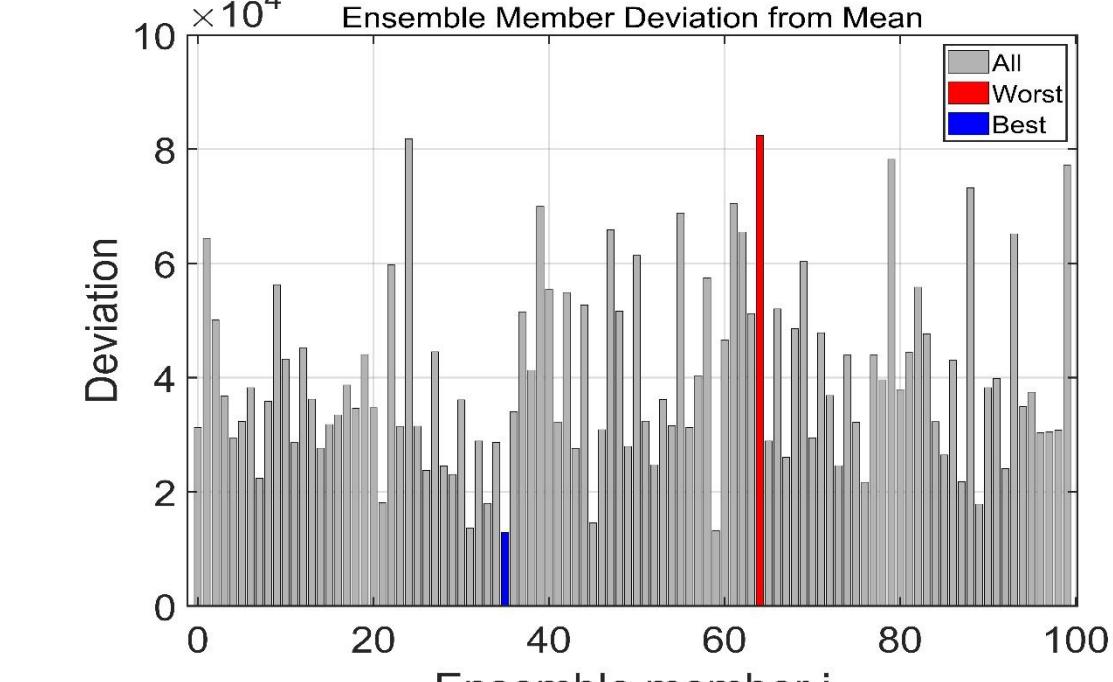
Fig. 3: Detect the worst/best case of ensembles to achieve Goal (2)
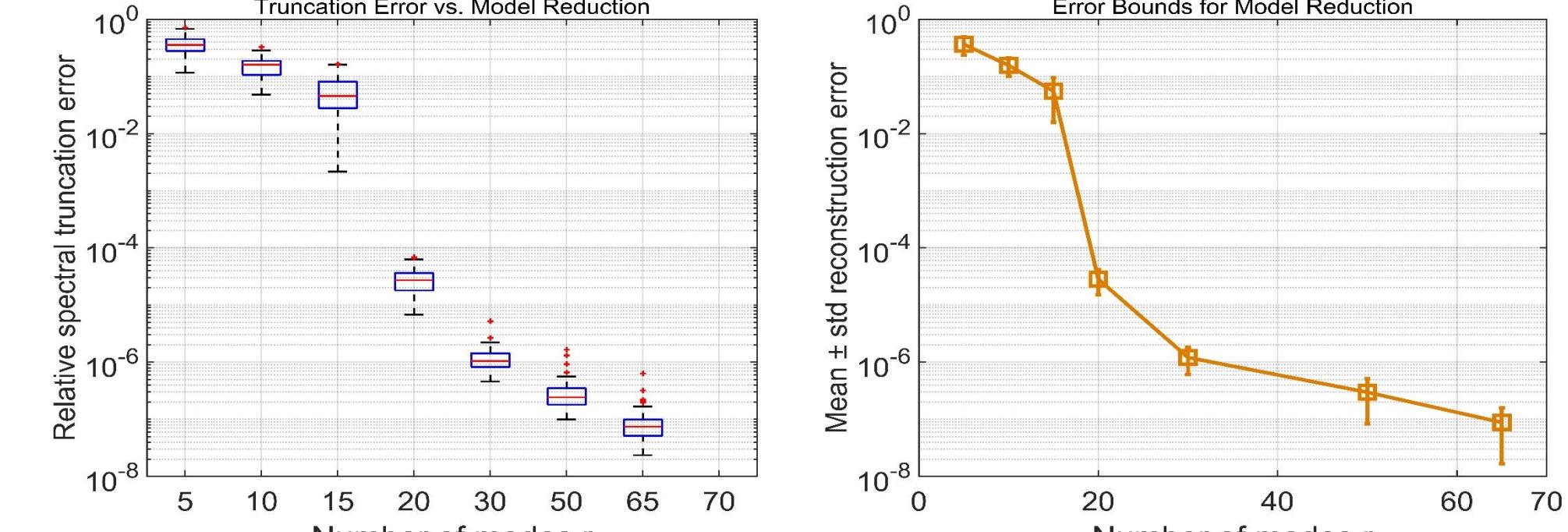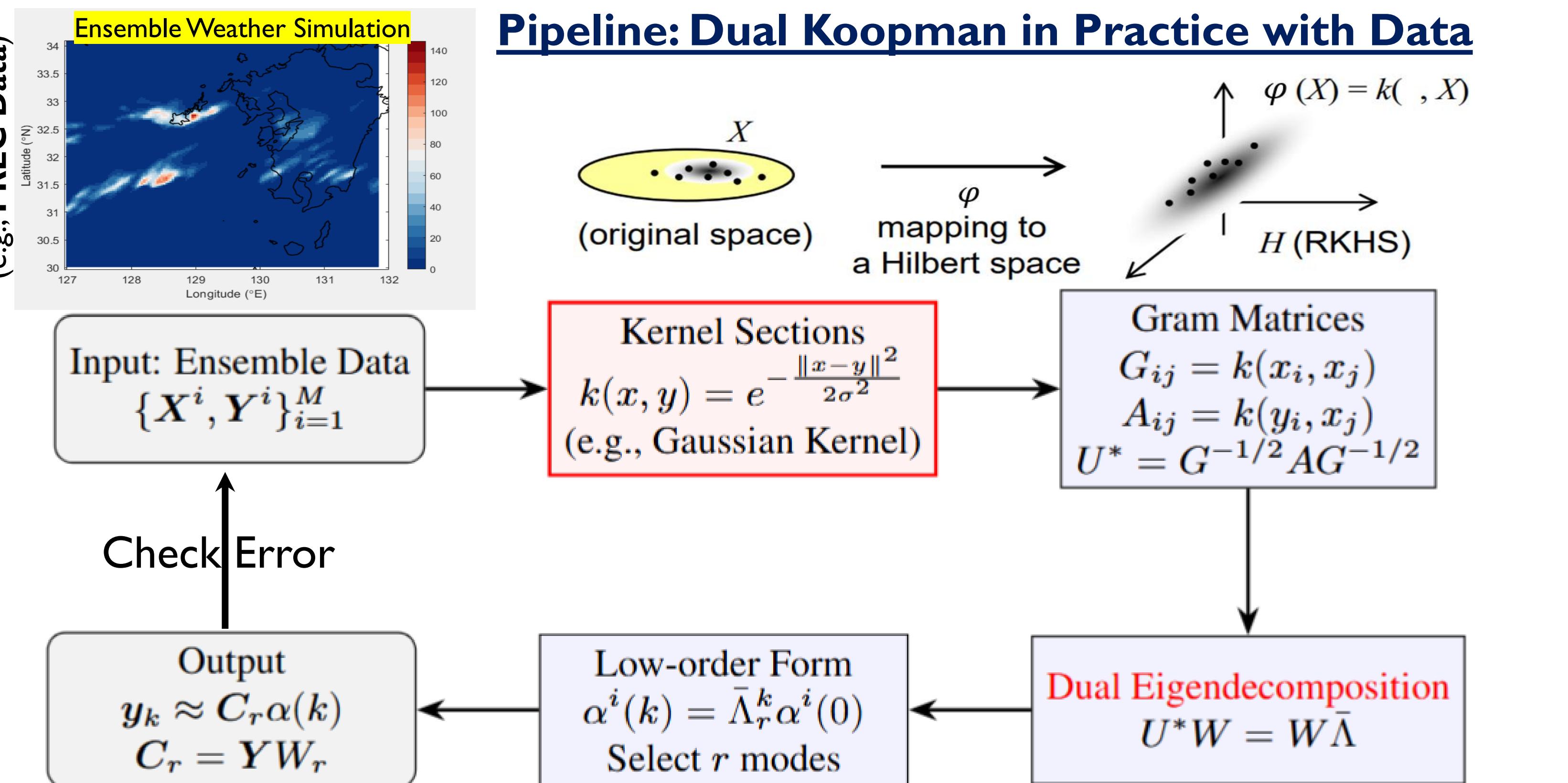
Fig. 4: Provide model-reduction and error bound to achieve Goal (1)

Pipeline: Dual Koopman in Practice with Data

X (original space) $\xrightarrow{\varphi}$ mapping to a Hilbert space $\quad \varphi(X) = k(\cdot, X) \quad H$ (RKHS)

Input: Ensemble Data $\{X^i, Y^i\}_{i=1}^M$

Kernel Sections
$$k(x, y) = e^{-\frac{\|x-y\|^2}{2\sigma^2}}$$
(e.g., Gaussian Kernel)

Gram Matrices
$$G_{ij} = k(x_i, x_j)$$
$$A_{ij} = k(y_i, x_j)$$
$$U^* = G^{-1/2} A G^{-1/2}$$

Dual Eigendecomposition
$$U^* W = W \bar{\Lambda}$$

Low-order Form
$$\alpha^i(k) = \bar{\Lambda}_r^k \alpha^i(0)$$
Select r modes

Output
$$y_k \approx C_r \alpha(k)$$
$$C_r = Y W_r$$

Check Error

References

1. J. Mohet et. al., A dual Koopman approach to observer design for nonlinear systems, arXiv:2503.08345v1.
2. A. Mauroy, I. Mezic, and Y. Susuki, Koopman operator in systems and control. Springer, 2020.
3. Brunton and Kutz, Data-driven science and engineering: Machine learning, dynamical systems, and control, 2022.
4. R. Mohr et. al, Koopman reduced-order modeling with confidence bounds, *SIADS*, 2025
5. A. Navarra, J. Tribbia, S. Klus, and P. L.Sanchez, Variability of SST through Koopman modes, *J. Climate*, 2024.