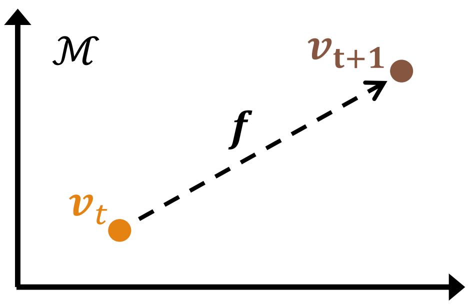
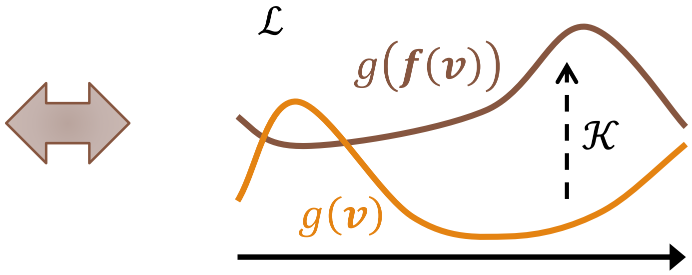
Factorially Switching Dynamic Mode Decomposition for Koopman Analysis of Time-Variant Systems

Naoya Takeishi (RIKEN)
Takehisa Yairi (The University of Tokyo)
Yoshinobu Kawahara (Osaka University / RIKEN)

Koopman Operator

[Koopman 31; Mezić 05]





$$\boldsymbol{v}_{t+1} = \boldsymbol{f}(\boldsymbol{v}_t)$$
$$\boldsymbol{v} \in \mathcal{M},\ \boldsymbol{f}\colon \mathcal{M} \to \mathcal{M}$$

$\boldsymbol{f}$ may be nonlinear
usually, $\dim(\mathcal{M}) < \infty$

$$g(\boldsymbol{f}(\boldsymbol{v})) = \mathcal{K} g(\boldsymbol{v})$$
$$g \in \mathcal{L}\colon \mathcal{M} \to \mathbb{C},\ \mathcal{K}\colon \mathcal{L} \to \mathcal{L}$$

$\mathcal{K}$ is a **linear** operator
in general, $\dim(\mathcal{L}) = \infty$

Modal Decomposition Based on Koopman Operator [Mezić 05]

For simplicity, suppose $\mathcal{K}$ has only discrete spectra (eigenvalues)

$$\mathcal{K}\varphi_j(\boldsymbol{v}) = \lambda_j \varphi_j(\boldsymbol{v}) \qquad (\lambda_j \in \mathbb{C},\ \varphi: \mathcal{M} \to \mathbb{C},\ \text{and } j \in \mathbb{N})$$

Assume that observable g is in the span of the eigenfunctions:

$$g(\boldsymbol{v}) = \sum_j w_j \varphi_j(\boldsymbol{v})$$

With these assumptions, because $\varphi_j(\boldsymbol{f}(\boldsymbol{v})) = \mathcal{K}\varphi_j(\boldsymbol{v}) = \lambda\varphi_j(\boldsymbol{v})$,

$$g(\boldsymbol{f}^t(\boldsymbol{v})) = \sum_j (\lambda_j)^t w_j \varphi_j(\boldsymbol{v})$$

frequency / decay rate

mode

Dynamic Mode Decomposition

[Rowley+ 09; Schmid 10]

Dynamic mode decomposition (DMD) can compute the Koopman-based modal decomposition under some conditions.

Let $\boldsymbol{g}\colon \mathcal{M} \to \mathbb{C}^m$ or $\mathbb{R}^m$ (vector-valued observable).
Suppose we have time-series data from time t_0 to t_n ($t_i = t_0 + i\Delta t$).

$$\boldsymbol{g}(\boldsymbol{v}_{t_0}), \boldsymbol{g}(\boldsymbol{v}_{t_1}), \dots, \boldsymbol{g}(\boldsymbol{v}_{t_i}), \dots, \boldsymbol{g}(\boldsymbol{v}_{t_{n-1}}), \boldsymbol{g}(\boldsymbol{v}_{t_n})$$

$$\boldsymbol{X} = \left[\boldsymbol{g}(\boldsymbol{v}_{t_0}) \quad \cdots \quad \boldsymbol{g}(\boldsymbol{v}_{t_{n-1}})\right] \quad \text{and} \quad \boldsymbol{Y} = \left[\boldsymbol{g}(\boldsymbol{v}_{t_1}) \quad \cdots \quad \boldsymbol{g}(\boldsymbol{v}_{t_n})\right]$$

DMD computes eigenvalues λ & eigenvectors $\boldsymbol{w}$ of $\boldsymbol{A} = \boldsymbol{Y}\boldsymbol{X}^{+}$.

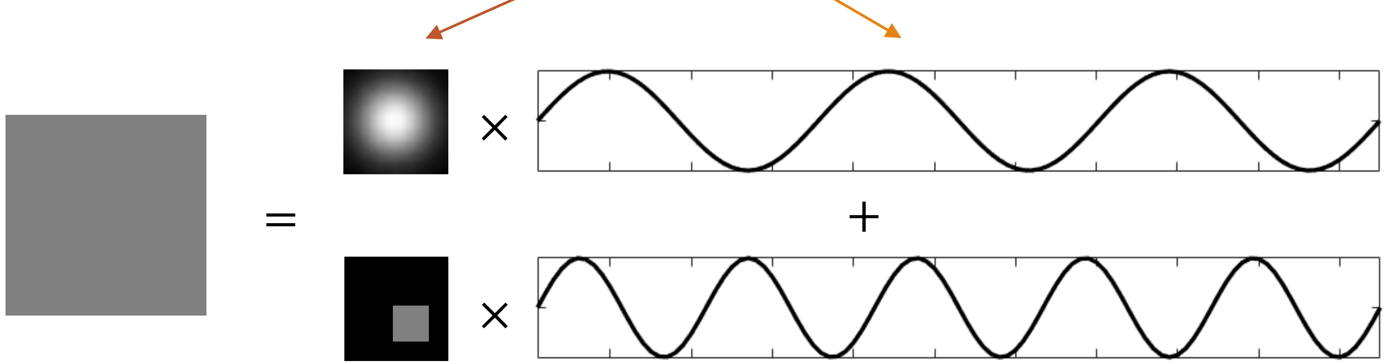
Under some conditions, these yield $\displaystyle \boldsymbol{g}(\boldsymbol{v}_{t_i}) = \sum_j (\lambda_j)^i \boldsymbol{w}_j \left(\boldsymbol{z}_j^* \boldsymbol{g}(\boldsymbol{v}_{t_0})\right)$

Dynamic Mode Decomposition (cont'd)

[Rowley+ 09; Schmid 10]

data — frequency / decay rate — coherent mode

$$\boldsymbol{g}(\boldsymbol{v}_{t_i}) = \sum_j (\lambda_j)^i \boldsymbol{w}_j \left(\boldsymbol{z}_j^* \boldsymbol{g}(\boldsymbol{v}_{t_0})\right)$$



Limitation of Standard DMDs

Within the dataset at hand, system is assumed to be time-invariant, and only a single set of dynamic modes is computed for the dataset.

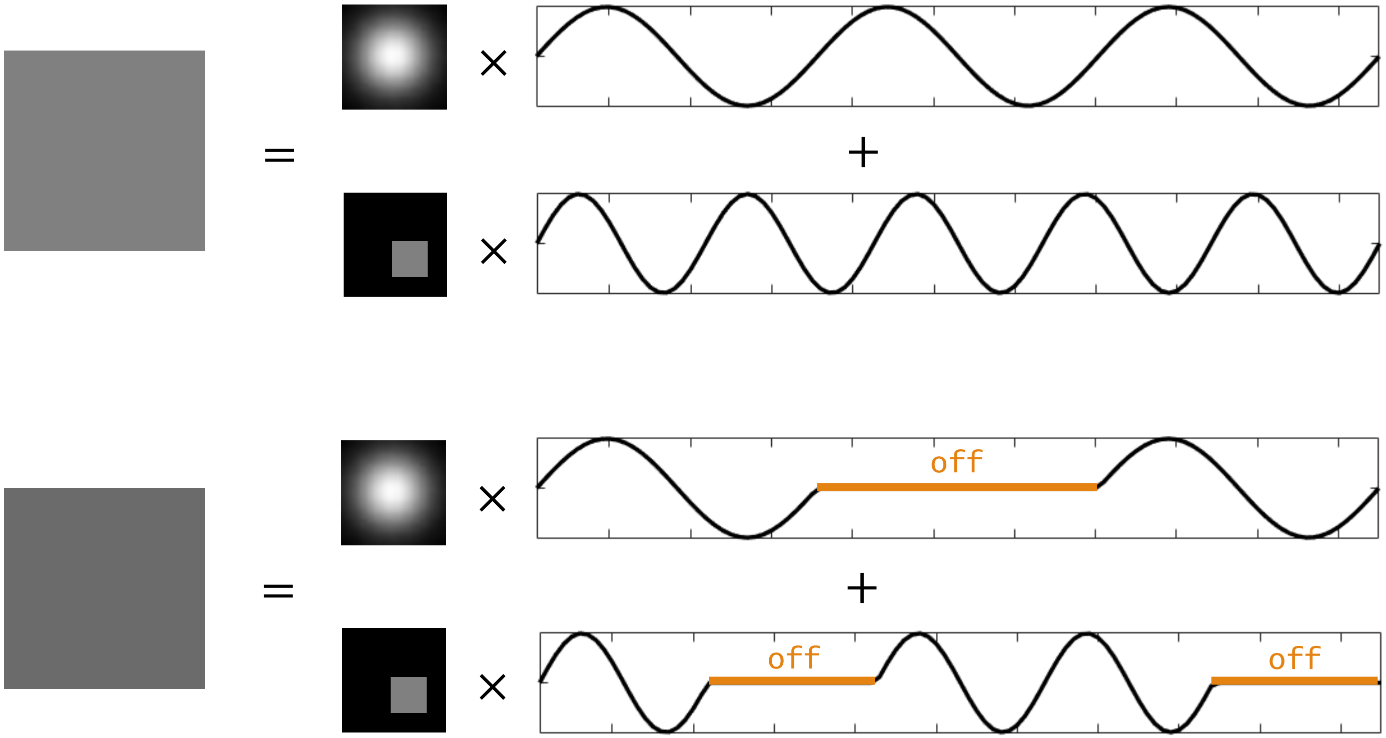
In practice, however,

- This assumption may not hold. (e.g., switching systems)
- Even if f is time-invariant, within finite-data regime,
  dynamic modes adequate for different periods of data may vary with time.
  (e.g., transient phenomena)

Existing approaches:

- Manual separation as preprocessing
- Multi-resolution DMD [Kutz+ 16]

Core Idea: Introducing "On-off Switching" to Dynamic Modes



→ Implement this idea via probabilistic formulation.

Preliminary: Probabilistic DMD

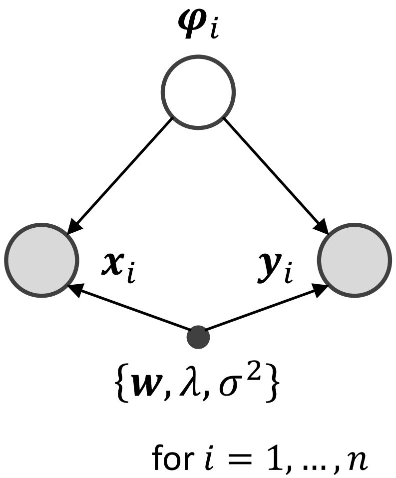
[Takeishi+ 17]

Dataset:

$$\boldsymbol{X} = [\boldsymbol{g}(\boldsymbol{v}_{t_0}) \quad \cdots \quad \boldsymbol{g}(\boldsymbol{v}_{t_{n-1}})] = [\boldsymbol{x}_1 \quad \cdots \quad \boldsymbol{x}_n] \quad \text{and} \quad \boldsymbol{Y} = [\boldsymbol{g}(\boldsymbol{v}_{t_1}) \quad \cdots \quad \boldsymbol{g}(\boldsymbol{v}_{t_n})] = [\boldsymbol{y}_1 \quad \cdots \quad \boldsymbol{y}_n]$$

Likelihood (observation model):

$$p(\boldsymbol{x}_i \mid \boldsymbol{\varphi}_i) = \mathcal{CN}_{\boldsymbol{x}_i}\left(\textstyle\sum_j \boldsymbol{w}_j \varphi_{i,j}\,,\ \sigma^2 \boldsymbol{I}\right)$$
$$p(\boldsymbol{y}_i \mid \boldsymbol{\varphi}_i) = \mathcal{CN}_{\boldsymbol{y}_i}\left(\textstyle\sum_j \lambda_j \boldsymbol{w}_j \varphi_{i,j}\,,\ \sigma^2 \boldsymbol{I}\right)$$



Prior:

$$p(\varphi_{i,j}) = \mathcal{CN}_{\varphi_{i,j}}(0,\ 1)$$

→ MLE in $\sigma^2 \to 0$ coincides with TLS-DMD

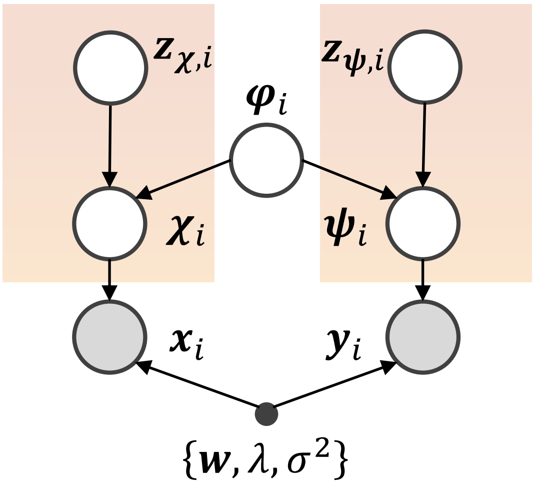
Proposed Model: Factorially-Switching DMD

Likelihood (observation model):

$$p(\boldsymbol{x}_i \mid \boldsymbol{\chi}_i) = \mathcal{CN}_{\boldsymbol{x}_i}\left(\textstyle\sum_j \boldsymbol{w}_j \chi_{i,j}\,,\ \sigma^2 \boldsymbol{I}\right)$$
$$p(\boldsymbol{y}_i \mid \boldsymbol{\psi}_i) = \mathcal{CN}_{\boldsymbol{y}_i}\left(\textstyle\sum_j \lambda_j \boldsymbol{w}_j \psi_{i,j}\,,\ \sigma^2 \boldsymbol{I}\right)$$

Priors:

$$p(\chi_{i,j}) = \delta(\chi_{i,j})^{(1-z_{\chi,i,j})}\,\delta(\varphi_{i,j} - \chi_{i,j})^{z_{\chi,i,j}}$$
$$p(\psi_{i,j}) = \delta(\psi_{i,j})^{(1-z_{\psi,i,j})}\,\delta(\varphi_{i,j} - \psi_{i,j})^{z_{\psi,i,j}}$$
$$p(\varphi_{i,j}) = \mathcal{CN}_{\varphi_{i,j}}(0,\ 1)$$



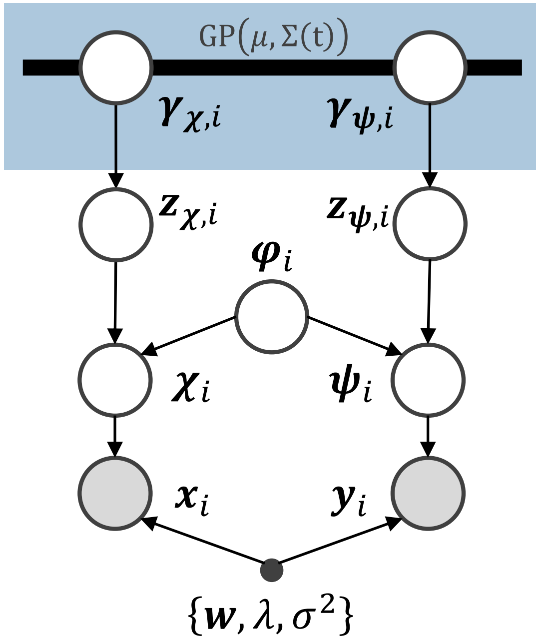
→ $z_{i,j}$ controls on-off of j-th mode at time i: $z_{i,j} = 1$ (on) / $z_{i,j} = 0$ (off)

Proposed Model: Factorially-Switching DMD (cont'd)

Likelihood (observation model):

$$p(\boldsymbol{x}_i \mid \boldsymbol{\chi}_i) = \mathcal{CN}_{\boldsymbol{x}_i}\left(\sum_j \boldsymbol{w}_j \chi_{i,j}\,,\ \sigma^2 \boldsymbol{I}\right)$$

$$p(\boldsymbol{y}_i \mid \boldsymbol{\psi}_i) = \mathcal{CN}_{\boldsymbol{y}_i}\left(\sum_j \lambda_j \boldsymbol{w}_j \psi_{i,j}\,,\ \sigma^2 \boldsymbol{I}\right)$$



Priors:

$$p(\chi_{i,j}) = \delta(\chi_{i,j})^{(1-z_{\chi,i,j})}\,\delta(\varphi_{i,j} - \chi_{i,j})^{z_{\chi,i,j}}$$

$$p(\psi_{i,j}) = \delta(\psi_{i,j})^{(1-z_{\psi,i,j})}\,\delta(\varphi_{i,j} - \psi_{i,j})^{z_{\psi,i,j}}$$

$$p(\varphi_{i,j}) = \mathcal{CN}_{\varphi_{i,j}}(0,\ 1)$$

$$p(z_{\chi/\psi,i,j} \mid \gamma_{\chi/\psi,i,j}) = \text{Bernoulli}\left(\Phi(\gamma_{\chi/\psi,i,j})\right)$$ (Φ: cdf of normal)

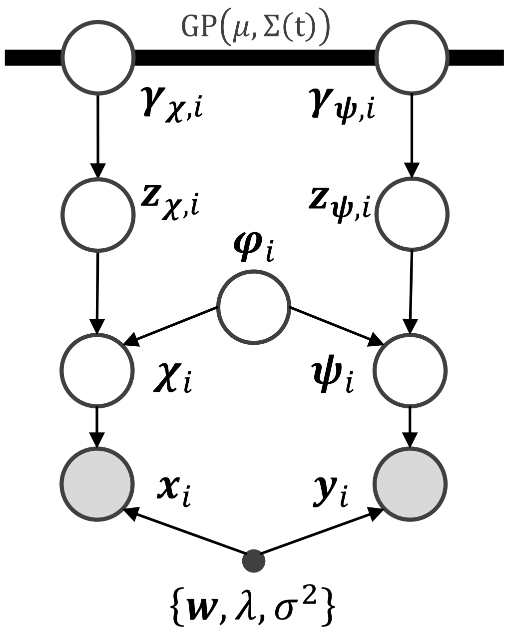
$$p(\boldsymbol{\gamma}_{\chi/\psi,:,j}) = \text{GaussianProcess}\left(\mu_{\chi/\psi,j}\mathbf{1},\ \boldsymbol{\Sigma}(t)\right)$$ → GP makes z smooth along time

Parameter Estimation by Approx. EM

Input Number of modes, kernel function, & data matrices $\boldsymbol{X}, \boldsymbol{Y}$

1. Initialize quantities using DMD.
2. E-step
   Approximate posterior of $\varphi, \chi, \psi, \mathrm{z}, \gamma$ using expectation propagation [Minka 01].
3. M-step
   Maximize $\mathbb{E}[\mathcal{L}(\boldsymbol{w}, \lambda, \sigma^2, \mu)]$
   ($\mathbb{E}$ is wrt. distribution from E-step).
4. Repeat 2. and 3. until convergence.

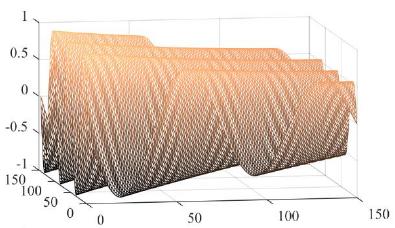
Output Posterior statistics of φ, z & estimated values of $\boldsymbol{w}, \lambda, \sigma^2, \mu$



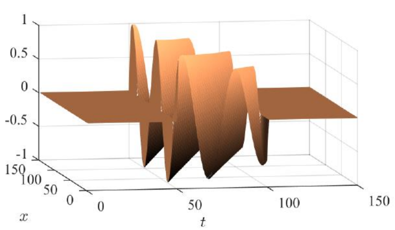
Toy Example (Local Traveling Wave)

DATA

=

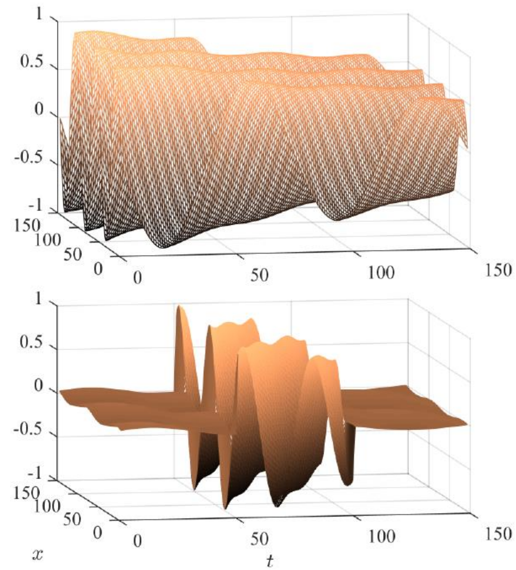


+



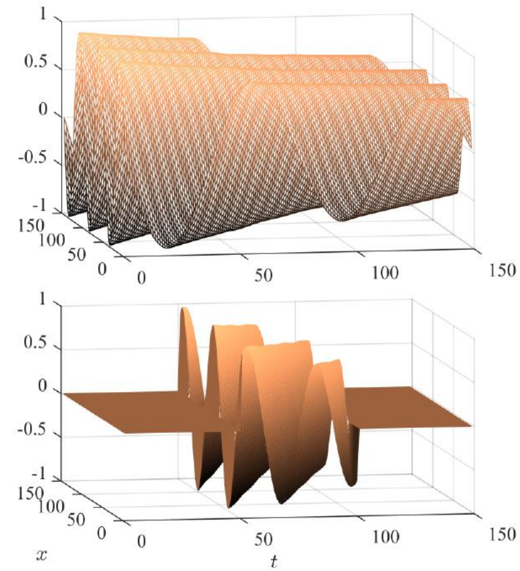
RESULTS

DMD



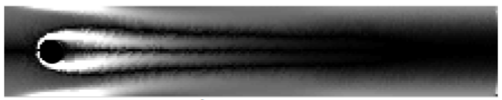
bumpy because of sudden changes

FSDMD (proposed)

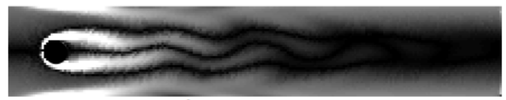


Transient Fluid Flow

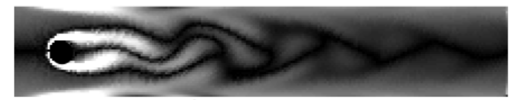
DATA



time = 1

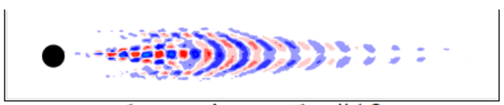


time = 200

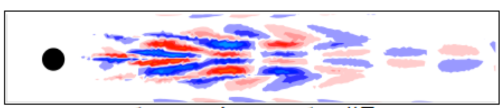


time = 400

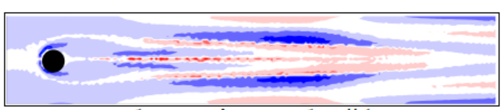
RESULTS



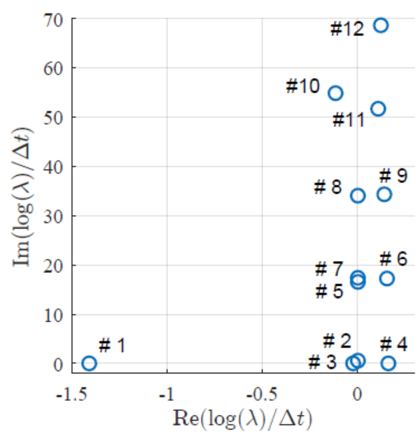
dynamic mode #12

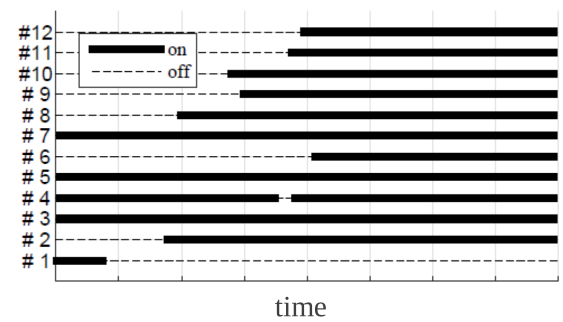


dynamic mode #7



dynamic mode #1



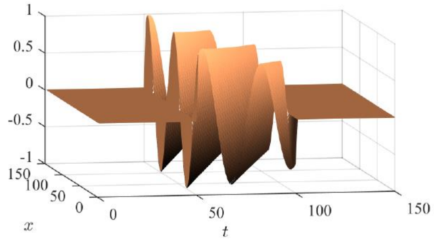


on/off states
of each dynamic mode

Summary

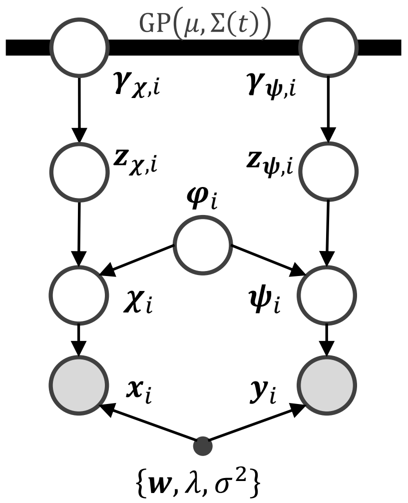
Objective

- Compute dynamic mode decomposition
  on time-varying systems & transient phenomena



Method

- Idea: Introducing on-off switching
  of each dynamic mode at each timestep
- Implemented it via probabilistic modeling/inference



Future Work

- Developing faster & more stable inference
- Considering interaction between dynamic modes