Identification and Prediction of Multiscale Dynamical Systems Using Bayesian Models

P.S. Koutsourelakis
Cornell University
pk285@cornell.edu

SIAM - CSE’09
March 6 2009
Deterministic (Stochastic) Dynamical Systems

\[
\frac{d\mathbf{y}_t}{dt} = f(\mathbf{y}_t) \\
\left( \frac{d\mathbf{y}_t}{dt} = f(\mathbf{y}_t; \omega) \right)
\]

- \( \mathbf{y}_t \in \mathbb{R}^n, n >> 1 \)
- Data/Observables:
  - a single time series (e.g. weather)
  - multiple time series (e.g. *inexpensive* computational model)
  - multiple time series but at an increasing cost (e.g. *expensive* computational model)
Motivation

Deterministic (Stochastic) Dynamical Systems

\[
\frac{dy_t}{dt} = f(y_t) \quad \left( \frac{dy_t}{dt} = f(y_t; \omega) \right)
\]

- In several cases, we can simulate the evolution of this high-dimensional system but on much shorter time ranges than what we are interested.

- Example - Molecular Dynamics: We can have a huge number of atoms \((n >> 1)\) and the typical time step is \(10^{-12} - 10^{-15} \text{s}\).
**Motivation**

### Deterministic (Stochastic) Dynamical Systems

\[
\frac{dy_t}{dt} = f(y_t) \quad \left( \frac{dy_t}{dt} = f(y_t; \omega) \right)
\]

- In several cases, we can simulate the evolution of this high-dimensional system but on much shorter time ranges than what we are interested.

- **Example - Molecular Dynamics**: We can have a huge number of atoms \((n >> 1)\) and the typical time step is \(10^{-12} - 10^{-15} s\).
Motivation

Deterministic (Stochastic) Dynamical Systems

\[
\frac{dy_t}{dt} = f(y_t) \quad \left( \frac{dy_t}{dt} = f(y_t; \omega) \right)
\]

- There is hope!
- In many cases the dynamics of such systems evolve in a much lower dimensional manifolds and at time scales that are much slower.
Objectives

- Can we use data from *short bursts of simulation* to identify those essential reaction coordinates?
- Can we learn how the dynamics on these lower-dimensional manifolds look like?
- Can we evolve the lower-dimensional system forward in time and still predict where our original system will be in the long-run?
- Can we quantify associated uncertainties that are inherent or simply arise from this reduced representation?

Desiderata:
- Model-Dimensionality Reduction
- Identification
- Prediction

pk285@cornell.edu
Cornell University
Identification and Prediction of Dyn. Systems
Objectives

- Can we use data from *short bursts of simulation* to identify those essential reaction coordinates?
- Can we learn how the dynamics on these lower-dimensional manifolds look like?
- Can we evolve the lower-dimensional system forward in time and still predict where our original system will be in the long-run?
- Can we quantify associated uncertainties that are inherent or simply arise from this reduced representation?

**Desiderata:**
- Model-Dimensionality Reduction
- Identification
- Prediction
Existing Approaches - PCA or POD

- Assume $y_t$ is stationary
- Evaluate sample covariance $C = Y Y^T$, diagonalize $C = P \Lambda P^T$
- Truncate after first few eigenvectors and approximate $y_t = \hat{P} x_t$. ($x_t \in \mathbb{R}^m$, $m << n$)
- Determine dynamics in $x$-space:

$$\frac{dx_t}{dt} = \hat{P}^T f(\hat{P} x)$$

Figure: Observables

$Y = y_{t_i} = (y_{t_i}^{(1)}, y_{t_i}^{(2)})$

pk285@cornell.edu
Cornell University

Identification and Prediction of Dyn. Systems
Existing Approaches - Nonlinear PCA

- Assume $y_t$ is stationary
- Learn nonlinear mapping $x = \hat{p}^{-1}(y)$.
- Determine dynamics in the reduced space $x$ using data and fitting appropriate models (e.g. diffusion maps, spectral clustering).

**Figure:** Observables
$Y = y_t = (y^{(1)}_{t_i}, y^{(2)}_{t_i})$

pk285@cornell.edu
Cornell University
Existing Approaches - HMM-PCA

- Hidden State Process:
  \[ S_t = 1, \ldots, K \]

- \( K \) Hidden evolution in \textit{projected} space, e.g.:
  \[
  d\mathbf{x}_t^{(k)} = -\nabla U^{(k)}(\mathbf{x}_t^{(k)}) + d\mathbf{W}_t^{(k)}, \quad k = 1, \ldots, K
  \]

- Mixture model:
  \[
  y_t \mid S_t = k \sim \mathcal{N}(P^{(k)}x_t^{(k)}, \Sigma^{(k)})
  \]

Figure: Observables
\[ Y = y_{t_i} = (y_{t_i}^{(1)}, y_{t_i}^{(2)}) \]
Existing Approaches - HMM-PCA

Figure: Double-well potential for OU process

Figure: Time series for OU process with double-well potential
Existing Approaches - HMM-PCA

Figure: Approximation with two OU processes with harmonic potential

Two hidden states $K = 2$

Figure: Time series for OU process with double-well potential
Dynamic Mixture Model

- Hidden assignment process:
  \[ Pr[S_t = k \mid S_{t-1} = j] = Q_{k,j} \]

- Hidden evolution at reduced-dimension spaces:
  \[ dX^{(k)}_t = -D^{(k)} (X^{(k)}_t - \mu^{(k)}) + \sigma^{(k)} dW^{(k)}_t, \quad k = 1, \ldots, K \]

- Likelihood:
  \[ Y_t \mid S_t = k \sim N \left( \mathcal{P}^{(k)} X^{(k)}_t, \Sigma \right) \]
Dynamic Mixture Model

- **Hidden assignment process:**

\[ Pr[S_t = k \mid S_{t-1} = j] = Q_{k,j} \]

- **Hidden evolution at reduced-dimension spaces:**

\[ dX_t^{(k)} = -D^{(k)} \left( X_t^{(k)} - \mu^{(k)} \right) + \sigma^{(k)} dW_t^{(k)}, \quad k = 1, \ldots, K \]

- **Likelihood:**

\[ Y_t \mid S_t = k \sim N \left( P^{(k)} X_t^{(k)}, \Sigma \right) \]

\[
p(y_t \mid \Theta_t) = \sum_{k=1}^{K} \pi_{k,t} p_k(y_t \mid \theta_{k,t})
\]

\[
\Theta_t = \left\{ S_t, \left\{ X_t^{(k)} \right\}_{k=1}^{K}, Q, \left\{ P^{(k)}, D^{(k)}, \mu^{(k)}, \sigma^{(k)} \right\}_{k=1}^{K}, \Sigma \right\}
\]
Dynamic Mixture Models

**Figure:** Notably non-harmonic potential

**Figure:** Multimodal potential
Dynamic Mixture Models

- How many hidden states $K$?
- Infinite HMM (see Caron et al. 2007, and HDP of Teh et al. 2006) could provide an interesting extension.
- This could still be expensive computationally if the number of clusters needs to be large.

    Basic Assumption: At any point in time a single cluster (potential well) dictates the behavior of the system.

    Furthermore, there is no chance of indentifying such a potential well unless it’s visited by the observables.
Dynamic Mixture Models

- How many hidden states $K$?
- Infinite HMM (see Caron et al. 2007, and HDP of Teh et al. 2006) could provide an interesting extension.
- This could still be expensive computationally if the number of clusters needs to be large.
- Basic Assumption: At any point in time a single cluster (potential well) dictates the behavior of the system.
- Furthermore, there is no chance of indentifying such a potential well unless it’s visited by the observables.
Dynamic Partial-Membership Mixture Models

- (Dynamic) mixture model. For fixed $t$:

$$p(y | \Theta) = \sum_{k=1}^{K} \pi_k p_k(y | \theta_k)$$

- can also be written as:

$$p(y | \Theta) = \sum_z p(z) \prod_{k=1}^{K} p_k^z(y_t | \theta_{k,t})$$

where:

- $z = (z_1, \ldots, z_K)$, $z_k \in \{0, 1\}$ and $\sum_{k=1}^{K} z_k = 1$
- $p(z = (0, \ldots, 1, \ldots, 0)) = \pi_k$
Dynamic Partial-Membership Mixture Models

What if we dropped the requirement that $z_k$ are binary and allowed them to take values in $[0, 1]$ such that $\sum_{k=1}^{K} z_k = 1$ still (Heller et al. 2008).

$$p(y|\Theta) = \sum_z p(z) \prod_{k=1}^{K} p_{z_k}(y_t | \theta_{k,t})$$

Figure: Mixture of two Gaussians

pk285@cornell.edu
Cornell University

Identification and Prediction of Dyn. Systems
Dynamic *Partial-Membership* Mixture Models

- **Likelihood.** Assume observations $y_t \in \mathbb{R}^n$ are available at discrete, equidistant time instants $t_i = i\tau$. Then:

  $$y_t \mid x_t, z_t \sim \mathcal{N}(\mu_{y,t}, \Sigma)$$

  where:
  - $\mu_{y,t} = \sum_{k=1}^{K} z_{k,t} P^{(k)} x^{(k)}_t$
  - $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$
Dynamic Partial-Membership Mixture Models

Prior Specification:

- Dynamic membership $z_t = (z_{1,t}, \ldots, z_{k,t})$:
  - Let $\hat{z}_t$ a Gaussian process defined such as:
    $$\hat{z}_t = \hat{z}_{t-1} - \tau D_z(\hat{z}_{t-1} - \mu_z) + \sigma_z \sqrt{\tau} \epsilon_{z,t}, \quad \epsilon_{z,t} \sim N(0, I_{K \times K})$$
  - and (logistic normal):
    $$z_{k,t} = \frac{e^{\hat{z}_{k,t}} + 1/K}{\sum_{j=1}^{K} e^{\hat{z}_{j,t}} + 1}$$

Note:

- If $\hat{z}_{k,t} \to -\infty, \forall k$, then $z_{k,t} \to 1/K$.
- If $\hat{z}_{k,t} \to +\infty, \forall k$, then $z_{k^*,t} = 1$ where $k^*$ corresponds to max. $\hat{z}_{k,t}$
Dynamic *Partial-Membership* Mixture Models

- **Prior Specification:**
  - Dynamic membership \( z_t = (z_{1,t}, \ldots, z_{k,t}) \):
    - Let \( \hat{z}_t \) a Gaussian process defined such as:
      \[
      \hat{z}_t = \hat{z}_{t-1} - \tau D_z(\hat{z}_{t-1} - \mu_z) + \sigma_z \sqrt{\tau} \epsilon_{z,t}, \quad \epsilon_{z,t} \sim N(0, I_{K \times K})
      \]
    - and (logistic normal):
      \[
      z_{k,t} = \frac{e^{\hat{z}_{k,t}} + 1/K}{\sum_{i=1}^{K} e^{\hat{z}_{i,t}} + 1}
      \]
  - Note:
    - If \( \hat{z}_{k,t} \rightarrow -\infty, \forall k \), then \( z_{k,t} \rightarrow \frac{1}{K} \).
    - If \( \hat{z}_{k,t} \rightarrow +\infty, \forall k \), then \( z_{k*,t} = 1 \) where \( k^* \) corresponds to max. \( \hat{z}_{k,t} \).
Dynamic Partial-Membership Mixture Models

Prior Specification:
- Dynamic evolution $X_t^{(k)} \in \mathbb{R}^m$ ($m << n$) of each model $k$:

$$x_t^{(k)} = x_{t-1}^{(k)} - \tau D_x^{(k)}(x_{t-1}^{(k)} - \mu_x^{(k)}) + \sigma_x^{(k)} \sqrt{\tau} \eta_t^{(k)}, \quad \eta_t^{(k)} \sim N(0, I_{m \times m})$$
Dynamic *Partial-Membership* Mixture Models

- Given $y_{1:T}$, the *model parameters* are:
  - **Dynamic** $q_{1:T}$:
    - $z_{1:T}, \hat{z}_{1:T}$
    - $\{x_{1:T}\}_{k=1}^K$
  - **Static** $\Theta$:
    - $D_z, \mu_z, \sigma_z$ (prior for $z_t$)
    - $\{D_{x_t}^{(k)}, \mu_{x_t}^{(k)}, \sigma_{x_t}^{(k)}\}_{k=1}^K$ (prior for $x_t^{(k)}$)
    - $\{P^{(k)}\}_{k=1}^K$, $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ (likelihood)

- Inference:
  - **Sequential**
    - Sampling from the posterior for dynamic parameters (fully Bayesian)
    - Point estimates (MLE or MAP) for static parameters
  - Online EM (e.g., Andrieu et al. 2005)

pk285@cornell.edu
Cornell University

Identification and Prediction of Dyn. Systems
Dynamic Partial-Membership Mixture Models

- Given $y_{1:T}$, the model parameters are:
  1. **Dynamic $q_{1:T}$:**
     - $z_{1:T},(\hat{z}_{1:T})$
     - $\{x_{1:T}^{(k)}\}_{k=1}^{K}$
  2. **Static $\Theta$:**
     - $D_z, \mu_z, \sigma_z$ (prior for $z_t$)
     - $\{D_x^{(k)}, \mu_x^{(k)}, \sigma_x^{(k)}\}_{k=1}^{K}$ (prior for $x_t^{(k)}$)
     - $\{P^{(k)}\}_{k=1}^{K}, \Sigma = diag(\sigma_1^2, \ldots, \sigma_n^2)$ (likelihood)

- **Inference:**
  - Sequential sampling from the posterior for dynamic parameters (fully Bayesian)
  - Point estimates (MLE or MAP) for static parameters
  - Online EM (e.g., Andrieu et al. 2005)
Dynamic Partial-Membership Mixture Models

- Given $y_{1:T}$, the model parameters are:
  1. **Dynamic $q_{1:T}$**:
     - $z_{1:T}, (\hat{z}_{1:T})$
     - $\{x_{1:T}^{(k)}\}_{k=1}^{K}$
  2. **Static $\Theta$**:
     - $D_z, \mu_z, \sigma_z$ (prior for $z_t$)
     - $\{D_x^{(k)}, \mu_x^{(k)}, \sigma_x^{(k)}\}_{k=1}^{K}$ (prior for $x_t^{(k)}$)
     - $\{P^{(k)}\}_{k=1}^{K}, \Sigma = diag(\sigma_1^2, \ldots, \sigma_n^2)$ (likelihood)

- Inference:
  - **Sequential**
    - sampling from the posterior for dynamic parameters (fully Bayesian)
    - point estimates (MLE or MAP) for static parameters $\Theta$

Inference: Sequential

Online EM (e.g. Andrieu et al. 2005)
Dynamic Partial-Membership Mixture Models

- Given $y_{1:T}$, the model parameters are:
  1. **Dynamic** $q_{1:T}$:
     - $z_{1:T}, (\hat{z}_{1:T})$
     - $\{x^{(k)}_{1:T}\}_{k=1}^K$
  2. **Static** $\Theta$:
     - $D_z, \mu_z, \sigma_z$ (prior for $z_t$)
     - $\{D^{(k)}_x, \mu^{(k)}_x, \sigma^{(k)}_x\}_{k=1}^K$ (prior for $x^{(k)}_t$)
     - $\{P^{(k)}\}_{k=1}^K, \Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ (likelihood)

- Inference:
  - Sequential
  - sampling from the posterior for dynamic parameters (fully Bayesian)
  - point estimates (MLE or MAP) for static parameters $\Theta$

Inference: Sequential sampling from the posterior for dynamic parameters (fully Bayesian) point estimates (MLE or MAP) for static parameters $\Theta$
Dynamic Partial-Membership Mixture Models

Given $y_{1:T}$, the model parameters are:

1. **Dynamic $q_{1:T}$:**
   - $z_{1:T}, (\hat{z}_{1:T})$
   - $\{x_{1:T}^{(k)}\}_{k=1}^{K}$

2. **Static $\Theta$:**
   - $D_z, \mu_z, \sigma_z$ (prior for $z_t$)
   - $\{D_x^{(k)}, \mu_x^{(k)}, \sigma_x^{(k)}\}_{k=1}^{K}$ (prior for $x_t^{(k)}$)
   - $\{P^{(k)}\}_{k=1}^{K}, \Sigma = diag(\sigma_1^2, \ldots, \sigma_n^2)$ (likelihood)

**Inference:**
- Sequential
- sampling from the posterior for dynamic parameters (fully Bayesian)
- point estimates (MLE or MAP) for static parameters $\Theta$

Online EM (e.g. Andrieu et al. 2005)
Dynamic Partial-Membership Mixture Models

Expectation-Maximization

- E-step - Iteration $j$:
  
  \[ Q(\Theta, \Theta_j) = \int \log p_{\Theta}(q_{1:T}, y_{1:T}) p_{\Theta_j}(q_{1:T}, y_{1:T}) dq_{1:T} \]

- M-step:
  
  \[ \Theta_{j+1} = \arg\max_\Theta Q(\Theta, \Theta_j) \]

or with respect to sufficient statistics:

- E-step - Iteration $j$:
  
  \[ \Phi_{j+1} = E_{\Theta_j} [\Psi(q_{1:T}, y_{1:T})] \]

- M-step:
  
  \[ \Theta_{j+1} = \Lambda(\Phi_{j+1}) \text{ (i.e. invert the relation with } \Phi) \]
Dynamic *Partial-Membership* Mixture Models

**Expectation-Maximization**

- **E-step - Iteration $j$:**
  \[
  \Phi_j = E_{\Theta_j} \left[ \psi(q_1:T, y_1:T) \right]
  \]

- **M-step:**
  \[
  \Theta_{j+1} = \Lambda(\Phi_j) \text{ (i.e. invert the relation with } \Phi)\]

- A *recursive* (online) version is available which can employ SMC (particle filters) to perform the E-step.

- This essentially implies performing importance sampling in spaces of increasing dimension which generally fails due to error accumulation over time.
Dynamic Partial-Membership Mixture Models

Split-data likelihood or Pseudo-likelihood

If we break the data $y_{1:T}$ into $B$ blocks of length $L$ we obtain a pseudo-log-likelihood:

$$l_L(\Theta, y_{1:BL}) = \sum_{b=1}^{B} \log p_{\Theta}(y_{(b-1)L+1:bL})$$

which ignores the dependence between blocks.

- It can be shown that the $\Theta$ maximizing the split-data pseudo-likelihood includes the true value.
- Maximization requires expectation over finite horizon of length $L$. 
Dynamic Partial-Membership Mixture Models

Split-data likelihood or Pseudo-likelihood

If we break the data $y_{1:T}$ into $B$ blocks of length $L$ we obtain a pseudo-log-likelihood:

$$I_L(\Theta, y_{1:BL}) = \sum_{b=1}^{B} \log p_\Theta(y_{(b-1)L+1:bL})$$

which ignores the dependence between blocks.

- It can be shown that the $\Theta$ maximizing the split-data pseudo-likelihood includes the true value.
- Maximization requires expectation over finite horizon of length $L$. 

pk285@cornell.edu
Cornell University
Identification and Prediction of Dyn. Systems
Dynamic Partial-Membership Mixture Models

Split-data likelihood or Pseudo-likelihood

If we break the data $y_{1:T}$ into $B$ blocks of length $L$ we obtain a pseudo-log-likelihood:

$$I_L(\Theta, y_{1:BL}) = \sum_{b=1}^{B} \log p_{\Theta}(y_{(b-1)L+1:bL})$$

which ignores the dependence between blocks.

- It can be shown that the $\Theta$ maximizing the split-data pseudo-likelihood includes the true value.
- Maximization requires expectation over finite horizon of length $L$. 
Dynamic Partial-Membership Mixture Models

Online EM

Let $\gamma_j$ such that $\sum_j \gamma_j \to +\infty$, $\sum_j \gamma_j \to c < +\infty$. (e.g. $\gamma_j = 1/j$.

Then at each iteration $j$:

- **E-Step:**
  - Run SMC over horizon $L$ using stationary distribution and evaluate $E_{\Theta_j} [\psi(q_{(j-1)L+1:jL}, y_{(j-1)L+1:jL})]$
  - Update sufficient statistics:
    \[
    \phi_{j+1} = (1 - \gamma_j) + \gamma_j E_{\Theta_j} [\psi(q_{(j-1)L+1:jL}, y_{(j-1)L+1:jL})] \]

- **M-Step:**
  \[
  \theta_{j+1} = \Lambda(\phi_{j+1}) \text{ (i.e. invert the relation with } \phi) \]
Example 1

- Data $y_t$: Average daily temperatures at 50 locations in US states from 1/1/1995 to 1/13/2009.
Example 1

- Data $y_t$: Average daily temperatures at 50 locations in US states from 1/1/1995 to 1/13/2009.
Example 1

(a) Data

(b) Posterior $z_t$

Figure: Two hidden modes
Example 1

The image shows a graph with states marked along the x-axis: Alaska, Hawaii, Florida, Louisiana, Idaho, Nevada, Oregon, Washington, Texas, Arizona, South Dakota, and North Dakota. The y-axis represents temperature in degrees Fahrenheit. The graph plots temperature trends for each state over a period, with distinct lines for different states. The x-axis is labeled with numbers from 0 to 50, likely representing time or days. The temperature values range from 0 to 100 degrees Fahrenheit.
Example 2

- Data $y_t$: Exchange rates of 8 currencies w.r.t US dollar at closing from 1/1/1977 to 12/31/1998.

Figure: Normalized exchange rates
Example 2

Figure: Comparison of pairs of currencies

(a) NET-GER  
(b) CAN-JAP  
(c) AUS-SUI
Example 2

- Data $y_t$: Exchange rates of 8 currencies w.r.t US dollar at closing from 1/1/1977 to 12/31/1998.

**Figure:** Normalized exchange rates
Example 2

Figure: Two hidden modes
Example 2-Prediction

(a) CAN
(b) FRA
(c) JAP

Figure: Sequential prediction over 100 time steps (days)
Example 3

- Multiscale PDEs:

\[
\begin{aligned}
&\quad \quad \quad u_t(x, t) = \frac{d}{dx} \left( a_\epsilon(x) \frac{du(x,t)}{dt} \right) \\
&\quad \quad \quad x \in [0, 1], u(0, t) = u(1, t) = 0
\end{aligned}
\]
Example 3

- Multiscale PDEs:

\[
\begin{align*}
  u_t(x, t) &= \frac{d}{dx} \left( a_\epsilon(x) \frac{du(x,t)}{dt} \right) \\
  x &\in [0, 1], \quad u(0, t) = u(1, t) = 0
\end{align*}
\]

- Discretization in space \( x \): \( \{u(x_i, t)\} = y_t \)

\[
\frac{dy_t}{dt} = A \ y_t
\]

- Discretization in time \( t \) with time step \( \tau = 10^{-4} \)
Example 3

- Temperature profile $u(x, T)$ at various $T$
Example 3

Temperature profile \( u(x, T) \) at various \( T \)
Example 3

Temperature profile $u(x, T)$ at various $T$
Example 3

Temperature profile \( u(x, T) \) at various \( T \)
Example 3

Temperature profile \( u(x, T) \) at various \( T \)
Example 3

Temperature profile $u(x, T)$ at various $T$
Example 3

Prediction using dynamic partial membership model (DPMM):

1. 2 hidden states/modes.
2. Full system is solved for 10 time steps i.e. $\delta t = 10^{-3}$
3. Inferred DPMM is used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$
4. Full system is re-initialized at posterior mean and run for another 10 time steps i.e. $\delta t = 10^{-3}$
5. DPMM is updated with new data and used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$
6. Goto step 3
Example 3

- Prediction using dynamic partial membership model (DPMM):
  1. 2 hidden states/modes.
  2. Full system is solved for 10 time steps i.e. $\delta t = 10^{-3}$
  3. Inferred DPMM is used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$
  4. Full system is re-initialized at posterior mean and run for another 10 time steps i.e. $\delta t = 10^{-3}$.
  5. DPMM is updated with new data and used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$.
  6. Goto step 3
Example 3

- Prediction using dynamic partial membership model (DPMM):
  1. 2 hidden states/modes.
  2. Full system is solved for 10 time steps i.e. $\delta t = 10^{-3}$
  3. Inferred DPMM is used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$
  4. Full system is re-initialized at posterior mean and run for another 10 time steps i.e. $\delta t = 10^{-3}$.
  5. DPMM is updated with new data and used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$.
  6. Goto step 3
Example 3

- **Prediction using dynamic partial membership model (DPMM):**

1. 2 hidden states/modes.
2. Full system is solved for 10 time steps i.e. $\delta t = 10^{-3}$
3. Inferred DPMM is used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$
4. Full system is re-initialized at *posterior mean* and run for another 10 time steps i.e. $\delta t = 10^{-3}$.
5. DPMM is updated with new data and used to “evolve” the system for 1,000 time steps, i.e. $\delta T = 10^{-1}$.
6. Goto step 3
Example 3

- Prediction using dynamic partial membership model (DPMM):
  1. 2 hidden states/modes.
  2. Full system is solved for 10 time steps i.e. \( \delta t = 10^{-3} \)
  3. Inferred DPMM is used to “evolve” the system for 1,000 time steps, i.e. \( \delta T = 10^{-1} \)
  4. Full system is re-initialized at posterior mean and run for another 10 time steps i.e. \( \delta t = 10^{-3} \).
  5. DPMM is updated with new data and used to “evolve” the system for 1,000 time steps, i.e. \( \delta T = 10^{-1} \).
  6. Goto step 3
Example 3

Figure: Exact vs. Predictive posterior obtained by DPMM at $x = 0.5$
Example 3

**Figure:** Exact vs. Predictive posterior obtained by DPMM at $x = 0.25$
Example 3

Figure: Exact vs. Predictive posterior obtained by DPMM at $x = 0.75$
A lightweight, structure prior model has been proposed for developing reduced-order models for high-dimensional dynamical systems.

It employs a SMC scheme for the dynamic hidden states and MAP (or MLE) for the static parameters.

Additional data can be ingested *sequentially*

The inferred model can be readily used to predict fast and accurately the evolution of the original system in time,

In multiscale computations, the predictive posterior can be used to adaptively determine when additional data from computations are needed.