Reduced-Order Modeling Of Hidden Dynamics

Patrick Héas & **Cédric Herzet**

INRIA, RENNES

SIAM UQ - 2016

High-Dimensional Model

- System of equations

$$\begin{cases} x_t = f_t(x_{t-1}, \theta_{t-1}), \\ x_1 = \theta_1, \end{cases}$$

  with

  - state variable $x_t \in \mathbb{R}^n$,
  - $f_t : \mathbb{R}^n \times \mathbb{R}^{p_t} \to \mathbb{R}^n$,
  - some parameters $\theta_t \in \mathbb{R}^{p_t}$.

$\Rightarrow$ unacceptable computational burdens.

Reduced-Order Models (ROMs)

- Reasonable ROM approximation for a set of *operating regimes*

$$\mathcal{X} \triangleq \{\boldsymbol{x} \triangleq (x_1 \cdots x_T) : \boldsymbol{x} \text{ trajectory with } \{\theta_t\}_{t=1}^T \in \mathcal{R}\ \},$$

where $\mathcal{R}$ is a set of admissible parameters.

ROM Building Scenarios

- Scenarios:
  - i) representative set of trajectories ($\mathcal{X}$ perfectly known)
  - ii) no representative trajectories because:
    - uncertainty on $\mathcal{R}$, *i.e.* on $\mathcal{X}$,
    - intractable computation of $\mathcal{X}$.
- *i)* is standard, but how to manage *ii*)?
  - substitute trajectories in $\mathcal{X}$ with observations ?
  - will ignore noise, incompleteness of observations.

Our Contributions

We propose a methodology for model-reduction which

- accounts for the uncertainties in the system to reduce;
- exploits observations in the reduction process while taking into account their imperfect nature.

Idea: recast model reduction as a Bayesian inference problem.

Ingredients

- Unknown regimes $\boldsymbol{x} = (x_1 \cdots x_T)$:
  - θ_t realization of a random variable Θ_t (unknown support $\mathcal{R}$),
  - *i.e.* , x_t realization of a random variable X_t (unknown support $\mathcal{X}$).
- Set of M observations, say matrix $\mathbf{Y} = (Y_1^1 \cdots Y_T^M)$ satisfying

$$Y_t^i = h_t(X_t) + W_t,$$

where $h_t : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$ and W_t's are mutually independent noises.

Bayesian Framework

- Uncertainty on state $\boldsymbol{x}$ given observation $\mathbf{Y} = \mathbf{y}$ is quantified by the posterior, say $\mu(d\boldsymbol{x}, \mathbf{y})$.
- Final goal is to include this information in the model reduction process.

Surrogate Prior

- We assume that trajectories in $\mathcal{X}$ must have a non-zero probability
- Surrogate models are commonly used in data assimilation with PDEs:
  - linear sub-space (*e.g.* Galerkin projection) + bounded error,
  - probabilistic structure (*e.g.* random fields, Markovian dynamics).

ROM A Posteriori Inference

- ROM parametrization given by

$$\hat{\mathbf{u}} \in \underset{\mathbf{u} \in \mathcal{U}}{\mathit{arg\ min}} \, \langle \mu(d\boldsymbol{x}), \|\boldsymbol{x} - \tilde{\boldsymbol{x}}(\mathbf{u})\| \rangle,$$

  where the ROM is defined by
  - an approximation $\tilde{\boldsymbol{x}}(\mathbf{u})$
  - a set $\mathcal{U}$.
  - an error norm $\|\boldsymbol{x} - \tilde{\boldsymbol{x}}(\mathbf{u})\|$,

Particularization 1: Galerkin Projections

- Models the subspace spanned by the columns of $\mathbf{u} \in \mathbb{R}^{n\times k}$, $k < n$: $\tilde{x}_t = \mathbf{u} z_t$ with k-dimensional recursion

$$\begin{cases} z_t = \mathbf{u}^* f_t(\mathbf{u} z_{t-1}, \theta_{t-1}), \\ z_1 = \mathbf{u}^* \theta_1. \end{cases}$$

- Subspace may be chosen as

$$\mathbf{u} \in \underset{\mathbf{v}\in\mathcal{U}}{arg\ min} \langle \mu(d\boldsymbol{x}), \| \boldsymbol{x}_{1:T} - \tilde{\boldsymbol{x}}_{1:T}(\mathbf{v}) \|_F^2 \rangle,$$

with $\mathcal{U} = \{\mathbf{v} \in \mathbb{R}^{n\times k} | \mathbf{v}^*\mathbf{v} = \mathbf{i}_k\}$.

Particularization 1: Galerkin Projections

- Models the subspace spanned by the columns of $\mathbf{u} \in \mathbb{R}^{n \times k}$, $k < n$:

  $\tilde{x}_t = \mathbf{u} z_t$ with k-dimensional recursion

$$\begin{cases} z_t = \mathbf{u}^* f_t(\mathbf{u} z_{t-1}, \theta_{t-1}), \\ z_1 = \mathbf{u}^* \theta_1. \end{cases}$$

- Subspace may be chosen as

$$\mathbf{u} \in \underset{\mathbf{v} \in \mathcal{U}}{arg\ min} \langle \mu(d\boldsymbol{x}), \| \boldsymbol{x}_{1:T} - \tilde{\boldsymbol{x}}_{1:T}(\mathbf{v}) \|_F^2 \rangle,$$

  with $\mathcal{U} = \{\mathbf{v} \in \mathbb{R}^{n \times k} | \mathbf{v}^* \mathbf{v} = \mathbf{i}_k\}$.

- POD: tractable error bound

$$\| \boldsymbol{x}_{1:T} - \tilde{\boldsymbol{x}}_{1:T}(\mathbf{u}) \|_F^2 \leq cte \| \boldsymbol{x}_{1:T} - \mathbf{u}\mathbf{u}^* \boldsymbol{x}_{1:T} \|_F^2 .$$

Particularization 2: POP Approximation

- Krylov subspaces approximation

$$\tilde{x}_t \approx \mathbf{p}\mathbf{q}^* x_{t-1}$$

with $\mathbf{p}, \mathbf{q} \in \mathbb{R}^{n \times k}$, $k < n$.

$\Leftrightarrow$ Equivalent to $\tilde{x}_t = \mathbf{p} z_t$ with k-dimensional recursion

$$\begin{cases} z_t = \mathbf{q}^* \mathbf{p} z_{t-1}, \\ z_1 = \mathbf{p}^\dagger \theta_1. \end{cases}$$

- Krylov subspaces chosen as

$$(\mathbf{p}, \mathbf{q}) \in \underset{\mathbf{p},\mathbf{q} \in \mathbb{R}^{n \times k}}{arg\ min} \langle \mu(d\boldsymbol{x}), \| \boldsymbol{x}_{2:T} - \tilde{\boldsymbol{x}}_{2:T} \|_F^2 \rangle.$$

Particularization 2: POP Approximation

- Krylov subspaces approximation

$$\tilde{x}_t \approx \mathbf{pq}^* x_{t-1}$$

with $\mathbf{p}, \mathbf{q} \in \mathbb{R}^{n \times k}$, $k < n$.

$\Leftrightarrow$ Equivalent to $\tilde{x}_t = \mathbf{p} z_t$ with k-dimensional recursion

$$\begin{cases} z_t = \mathbf{q}^* \mathbf{p} z_{t-1}, \\ z_1 = \mathbf{p}^\dagger \theta_1. \end{cases}$$

- Krylov subspaces chosen as

$$(\mathbf{p}, \mathbf{q}) \in \underset{\mathbf{p},\mathbf{q} \in \mathbb{R}^{n \times k}}{arg\ min} \langle \mu(d\boldsymbol{x}), \| \boldsymbol{x}_{2:T} - \tilde{\boldsymbol{x}}_{2:T} \|_F^2 \rangle.$$

- Tractable error bound $\| \boldsymbol{x}_{2:T} - \tilde{\boldsymbol{x}}_{2:T} \|_F^2 \leq cte \| \boldsymbol{x}_{2:T} - \mathbf{pq}^* \boldsymbol{x}_{1:T-1} \|_F^2$.

Closed-form solutions

- We obtain solutions for POD-Galerkin and POP in terms of eigen-decomposition of matrices

$$\langle \mu(d\boldsymbol{x}), \boldsymbol{x}_{1+\ell:T} \boldsymbol{x}^*_{1:T-\ell} \rangle, \quad \ell \in \{0, 1\}$$

Comparison With State-Of-The-Art

$$\langle \mu(d\boldsymbol{x}), \boldsymbol{x}_{1+\ell:T}\boldsymbol{x}^*_{1:T-\ell} \rangle = \sum_{t=1+\ell}^{T} \underbrace{\mathbf{cov}(x_{t-\ell}, x_t)}_{covariance} + \underbrace{\bar{x}_{t-\ell}(\mathbf{y})}_{mean} \bar{x}^*_t(\mathbf{y}).$$

- In standard methods, unknown regimes are substituted by snapshots's posterior mean, or by some other estimate, yielding the approximation

$$\langle \mu(d\boldsymbol{x}), \boldsymbol{x}_{1+\ell:T}\boldsymbol{x}^*_{1:T-\ell} \rangle \approx \sum_{t=1+\ell}^{T} \bar{x}_{t-\ell}(\mathbf{y})\bar{x}^*_t(\mathbf{y}).$$

Comparison With State-Of-The-Art

$$\langle \mu(d\boldsymbol{x}), \boldsymbol{x}_{1+\ell:T} \boldsymbol{x}^*_{1:T-\ell} \rangle = \sum_{t=1+\ell}^{T} \underbrace{\mathbf{cov}(x_{t-\ell}, x_t)}_{covariance} + \underbrace{\bar{x}_{t-\ell}(\mathbf{y})}_{mean} \bar{x}^*_t(\mathbf{y}).$$

- In standard methods, unknown regimes are substituted by snapshots's posterior mean, or by some other estimate, yielding the approximation

$$\langle \mu(d\boldsymbol{x}), \boldsymbol{x}_{1+\ell:T} \boldsymbol{x}^*_{1:T-\ell} \rangle \approx \sum_{t=1+\ell}^{T} \bar{x}_{t-\ell}(\mathbf{y}) \bar{x}^*_t(\mathbf{y}).$$

$\Rightarrow$ Uncertainty (or covariance) on regimes is neglected!

Simulations

Rayleigh-Bénard convection (generalization of Navier-Stokes advection):

- **quadratic dynamics** f_t w.r.t. velocities $\{x_t\}_{t=1}^{50}$ where $x_t \in \mathbb{R}^n$ satisfy

$$\begin{cases} x_{t+1} = f_{t+1}(x_t, \theta_t), \\ x_1 = \theta_1, \end{cases}$$

  with some initial condition θ_1 and forcing θ_t;

- **linear observations** $\{y_t\}_{t=1}^{50}$ where $y_t \in \mathbb{R}^m$ satisfy

$$y_t = \mathbf{h}_t x_t + w_t, \quad W_t \sim \mathcal{N}(0, \sigma^2 \mathbf{i}_m).$$

Simulations: a simple case

- Assume a Gaussian time-uncorrelated model:
  - $n = 2^{15}$, $m = 2^{14}$,
  - 1 observation ($M = 1$),
  - 1 unknown regime to reproduce (Dirac on $\mathcal{R}$),
  - self-similar surrogate prior on x_t's[1]: $\mathcal{N}(0, \mathbf{q}_t)$.

[1] see references in *P. Héas, F. Lavancier, S. Kadri Harouna. Self-similar prior and wavelet bases for hidden incompressible turbulent motion. SIAM Journal on Imaging Sciences, Volume 7, Issue 2, pp. 1171-1209, 2014.*

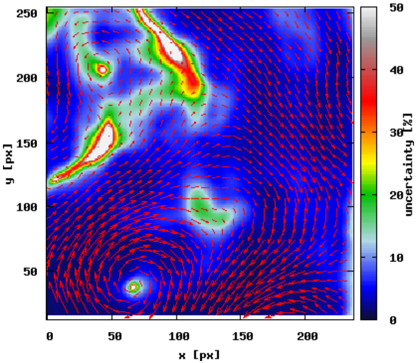
Simulations: a simple case

- Assume a Gaussian time-uncorrelated model:
  - $n = 2^{15}$, $m = 2^{14}$,
  - 1 observation ($M = 1$),
  - 1 unknown regime to reproduce (Dirac on $\mathcal{R}$),
  - self-similar surrogate prior on x_t's[1]: $\mathcal{N}(0, \mathbf{q}_t)$.
- in this linear Gaussian setting, we obtain a closed-form posterior distribution μ

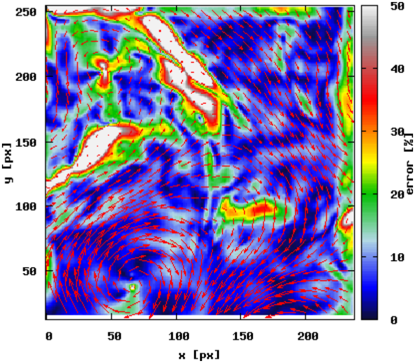
$$\begin{cases} \bar{x}_t(\mathbf{y}) = \sigma^{-2}\mathbf{p}_t\mathbf{h}_t^*(y_t - \xi_t), \\ \mathbf{cov}(x_t, x_t) = \sigma^2(\mathbf{h}_t^*\mathbf{h}_t + \sigma^2\mathbf{q}_t)^{-1}. \end{cases}$$

[1] see references in *P. Héas, F. Lavancier, S. Kadri Harouna. Self-similar prior and wavelet bases for hidden incompressible turbulent motion. SIAM Journal on Imaging Sciences, Volume 7, Issue 2, pp. 1171-1209, 2014.*

Simulations: A Simple Case



$\mu(dx_t, y_t)$



$\|\bar{x}_t - x_t^{true}\|$

Simulations: A More Involved Scenario

- Assume a general hidden Markov model:
  - $n = 2^{10}$, $m = 2^9$,
  - $M = 2^5$ observations,
  - unknown uniform prior on $\mathcal{R}$,
  - Markovian surrogate prior,

$$\begin{cases} X_t = f_{t+1}(X_t, \Theta_t), \\ X_1 \sim \mu_1(d\theta_1), \end{cases}$$

where $\{\Theta_t\}_t$ is uniformly distributed of support including $\mathcal{R}$, but 2 times larger.

Simulations: A More Involved Scenario

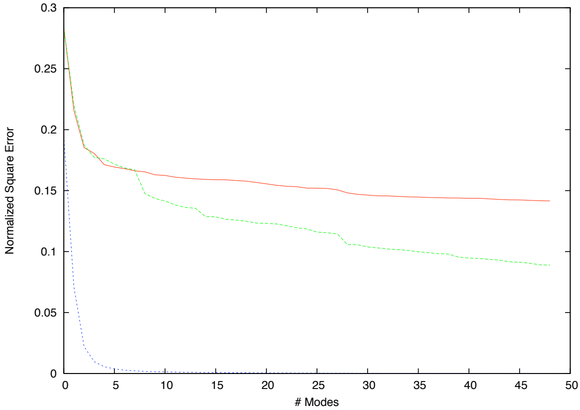
- Assume a general hidden Markov model:
  - $n = 2^{10}$, $m = 2^9$,
  - $M = 2^5$ observations,
  - unknown uniform prior on $\mathcal{R}$,
  - Markovian surrogate prior,

$$\begin{cases} X_t = f_{t+1}(X_t, \Theta_t), \\ X_1 \sim \mu_1(d\theta_1), \end{cases}$$

  where $\{\Theta_t\}_t$ is uniformly distributed of support including $\mathcal{R}$, but 2 times larger.

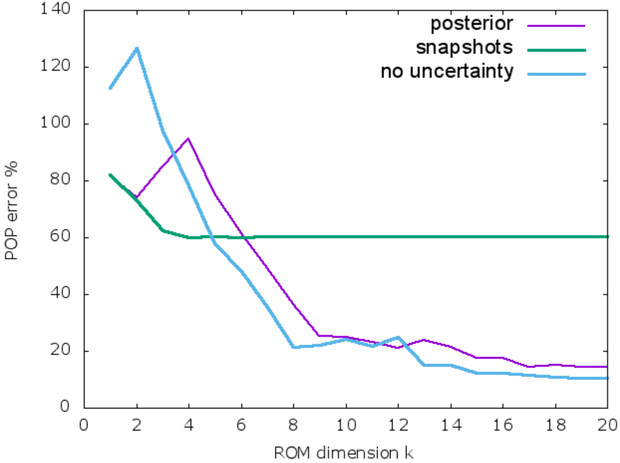
- The posterior μ is approximated by particle filtering.

Results: POD-Galerkin Projection Error (Simple Case)



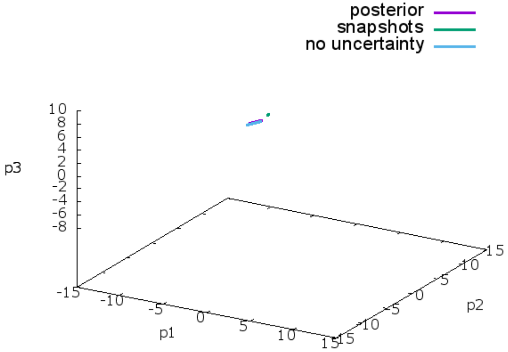
Square ℓ_2 error with respect to dimension k using state-of-the-art snapshot method (red solid line), the proposed method (green dashed line) or using directly the ground truth (blue dotted line).

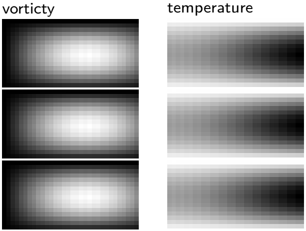
Results: POP Approximation Error (Involved Scenario)



Square ℓ_2 error with respect to dimension k using posterior-based method (in purple), state-of-the-art snapshot method (in green) or a construction with no regimes uncertainty (in turquoise).

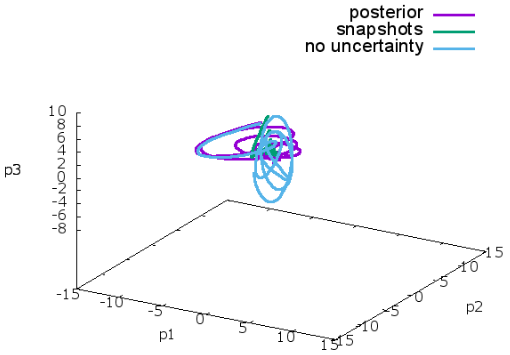
10-Dimensional POP Approximation Of Snapshot 1

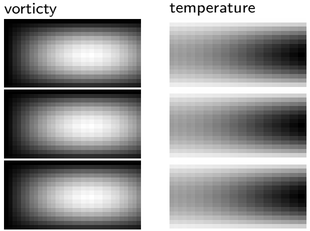




above: posterior-based,
middle: state-of-the-art snapshot method,
below: no regimes uncertainty.

10-Dimensional POP Prediction Of Snapshot 1





above: posterior-based,
middle: state-of-the-art snapshot method,
below: no regimes uncertainty.

Conclusions

- Integration of observations to reduce the regimes uncertainties.
- Bayesian inference context.
- Solutions for POD-Galerkin and POP given in term of eigendecomposition of

$$\sum_{t=1+\ell}^{T} \mathbf{cov}(x_{t-\ell}, x_t) + \bar{x}_{t-\ell}(\mathbf{y})\bar{x}_t^*(\mathbf{y}).$$

Perspectives

- Can we obtain a ROM expected error bound

$$\langle \mu(d\boldsymbol{x}), \|\boldsymbol{x} - \tilde{\boldsymbol{x}}(\hat{\mathbf{u}})\| \rangle \leq \textit{function of}(\textit{prior}, \textit{surrogate}, \text{likelihood})?$$

- Which surrogate will minimize this bound?

Perspectives

- Can we obtain a ROM expected error bound

$$\langle \mu(d\boldsymbol{x}), \|\boldsymbol{x} - \tilde{\boldsymbol{x}}(\hat{\mathbf{u}})\| \rangle \leq \textit{function of}(\textit{prior}, \textit{surrogate}, \text{likelihood})?$$

- Which surrogate will minimize this bound?

http://people.rennes.inria.fr/Patrick.Heas/