Inverse Reduced-Order Modeling

Patrick Héas & **Cédric Herzet**

Inria, Rennes

ROM-2015

How to study high-dimensional differential equations ?

- *e.g. wave equation, Navier-Stokes and Maxwell's equations*.
- discretization yields high-dimensional systems

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

  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 $\theta_t \in \mathbb{R}^{p_t}$.
- computing a trajectory $\{x_t\}_t$ given some parameters $\{\theta_t\}_t$ may lead to unacceptable computational burdens.

Reduced-Order Models (ROMs)

- ROM is a system
  - involving a small number of degrees of freedoms,
  - providing a reasonable approximation for a set of operating regimes.
- Set of operating regimes

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

  where $\mathcal{R}$ is a set of admissible parameters.
- *e.g.* Galerkin projections, Hankel norm approximations, principal oscillating patterns (POPs) or principal interacting patterns (PIPs).

Example of POD-Galerkin projection

- *Galerkin* projection is a low-rank approximation.
- obtained by projecting x_t's onto a subspace spanned by the columns of some matrix $\mathbf{u} \in \mathbb{R}^{n\times k}$ where $k < n$.
- yields tractable k-dimensional recursion

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

- approximation of x_t obtained by matrix-vector multiplication $\mathbf{u} z_t$.
- *POD-Galerkin* approximation chooses matrix $\mathbf{u}$ solution of

$$\underset{\mathbf{u}\in\mathcal{U}}{arg\ min} \ \| \mathbf{x} - \mathbf{u}\mathbf{u}^*\mathbf{x} \|_F^2,$$

with $\mathcal{U}$ the set of unitary matrices in $\mathbb{R}^{n\times k}$, given some $\mathbf{x} \in \mathcal{X}$.

Ingredients to build a ROM

- ROM construction scenarios:
  i) representative set of trajectories because $\mathcal{X}$ perfectly known
  ii) no representative trajectories because:
    - uncertainty on $\mathcal{R}$, *i.e.* on $\mathcal{X}$,
    - intractable computation of $\mathcal{X}$.

- Common approach substitutes trajectories in $\mathcal{X}$ by their incomplete observation (*e.g.* in geophysics, satellite images provide partial observation of the Ocean state).
  $\Rightarrow$ ignores noise, incompleteness of observations.

Our contribution

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 an a posteriori inference problem

Surrogate prior

- Regimes uncertainties
  - θ_t realization of a random variable Θ_t (probability measure of support $\mathcal{R}$).
  - *i.e.* , $\mathbf{x}_t$ seen as the realization of a random variable X_t (probability measure of support $\mathcal{X}$)
- Surrogate prior η_t
  - dominating measure
    i.e. trajectories in $\mathcal{X}$ will have a non-zero probability
  - assume probabilistic model of the form:

    $$\begin{cases} X_t = b_t(X_{t-1}) + V_t, \\ X_1 \sim \eta_1(dx_1), \end{cases}$$

    where $b_t : \mathbb{R}^n \to \mathbb{R}^n$ and where the V_t's are mutually independent random variables.
  - often exist probabilistic models for deterministic chaotic systems.

Observation model

- set of M observations, say matrix $\mathbf{Y} = (Y_1^1 \cdots Y_T^M)$.
- assume matrix $\mathbf{X} = (X_1 \cdots X_T)$ and $\mathbf{Y}$ satisfy

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

  where $h_t : \mathbb{R}^n \to \mathbb{R}^m$ and W_t's are mutually independent noises.
- Given the observations, one can hope to remove certain uncertainties on the system regime, the final goal being to include this information in the model reduction process.

Posterior model

- Bayesian estimators relying on the joint posterior measure of $\mathbf{X}$ given some observation $\mathbf{Y} = \mathbf{y}$, say μ.
- Hidden Markov model (HMM) factorization
- Helpful to compute expectation for some integrable function φ

$$\langle \mu(\cdot, \mathbf{y}), \varphi \rangle \triangleq \int_{\mathbb{R}^{n \times T}} \mu(d\mathbf{x}, \mathbf{y}) \varphi(d\mathbf{x}).$$

  - Gaussian linear case or finite state HMMs $\Rightarrow$ closed-form expectation given by *Kalman* recursions or by the *Baum-Welsh* re-estimation formulae
  - General case, *sequential Monte-Carlo methods* provide asymptotically consistent estimators

ROM a posteriori inference

- The uncertainty on the state $\mathbf{x}$ given some observation $\mathbf{Y} = \mathbf{y}$ is quantified by the posterior μ.
- ROM parametrization $\mathbf{u}$ inferred as the solution of

$$\mathop{arg\ min}_{\mathbf{u}\in\mathcal{U}} \langle \mu(d\mathbf{x}, \mathbf{y}), \phi(\mathbf{x}, \mathbf{u}) \rangle.$$

  where the set $\mathcal{U}$ and cost $\phi(\mathbf{x}, \mathbf{u})$ are specific to the ROM
- May extend the construction of ROMs such as POD-Galerkin, POPs, POPs, etc.

Particularization to POD-Galerkin projection

- In this case we have
  - $\mathcal{U}$ is the set of unitary matrices $\mathbf{u} \in \mathbb{R}^{n \times k}$,
  - $\phi(\mathbf{x}, \mathbf{u}) \triangleq \| \mathbf{x} - \mathbf{u}\mathbf{u}^* \mathbf{x} \|_F^2$.
- Closed-form solution $\hat{\mathbf{u}}$ by eigen-decomposition of

$$\langle \mu(d\mathbf{x}, \mathbf{y}), \mathbf{x}\mathbf{x}^* \rangle.$$

($\hat{\mathbf{u}}$ formed by the eigenvectors associated to the k largest eigenvalues)

Comparison with state-of-the-art ROM inference

- Expectation decomposition as

$$\langle \mu(d\mathbf{x},\mathbf{y}), \mathbf{x}\mathbf{x}^* \rangle = \sum_{t=1}^{T} \mathbf{p}_t + \bar{x}_t(\mathbf{y})\bar{x}_t^*(\mathbf{y}).$$

  where a posteriori mean $\bar{x}_t(\mathbf{y}) \in \mathbb{R}^n$ and covariance $\mathbf{p}_t \in \mathbb{R}^{n\times n}$.

- The standard *snapshot* method diagonalizes instead

$$\sum_{t=1}^{T} \hat{x}_t(\mathbf{y})\hat{x}_t^*(\mathbf{y}),$$

  where $\hat{x}_t(\mathbf{y})$ denotes some estimate of the state x_t given $\mathbf{y}$.

- Conclusion:
  - $\Rightarrow$ snapshot ignores directions where some uncertainty remains a posteriori. (*e.g.* for the MMSE estimator $\hat{x}_t(\mathbf{y}) = \bar{x}_t(\mathbf{y})$, neglects $\mathbf{p}_t$, *i.e.* directions where the covariance of the estimation error is large).

Numerical simulation: ROM of 2D turbulence

Goal: reduce the 2D Navier-Stokes equations from a video of scalar fields.

- linear observation model describes variation $y_t \in \mathbb{R}^{2^{14}}$ of the intensity of a scalar field conveyed by the flow.

$$h_t(x_t) = \mathbf{h}_t x_t + \xi_t,$$

  with a zero-mean uncorrelated Gaussian noise ξ_t of variance σ^2

- quadratic function with respect to velocity $x_t \in \mathbb{R}^{2^{16}}$

$$f_t(x_{t-1}, \theta_{t-1}) = \mathbf{c}^*(\mathbf{i}_n + \alpha\boldsymbol{\ell} - x_{t-1}^*\mathbf{r})\mathbf{c}x_{t-1} + \theta_{t-1},$$

  where variable $\theta_t \in \mathbb{R}^n$ accounts for some forcing.

$\Rightarrow$ sequence of 50 motion and scalar fields $\{x_t\}_{t=1}^{50}$ and $\{y_t\}_{t=1}^{50}$ for a given operating regime $\{\theta_t\}_{t=1}^{50} \in \mathcal{R}$.

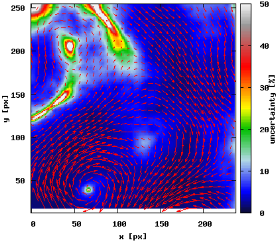
Numerical simulation: surrogate prior

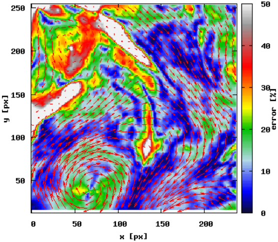
- standard or self-similar "optic-flow" priors[1] : time-uncorrelated Gaussians of covariance $\mathbf{q}_t$.
- In this linear Gaussian setting, we obtain the posterior mean and covariance

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

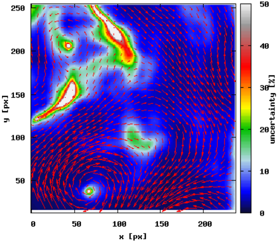
[1] see references in *P. Héas, F. Lavancier, S. Kadri Harouna. Self-similar prior and wavelet bases for hidden incompressible turbulent motion. SIAM Journal on Imaging Sciences, Volume 7, Issue 2, pp. 1171-1209, 2014.*

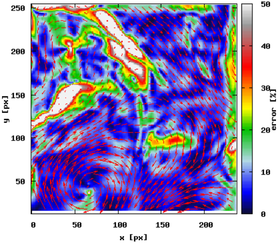
Results: uncertainty v.s. estimation error (t=25)





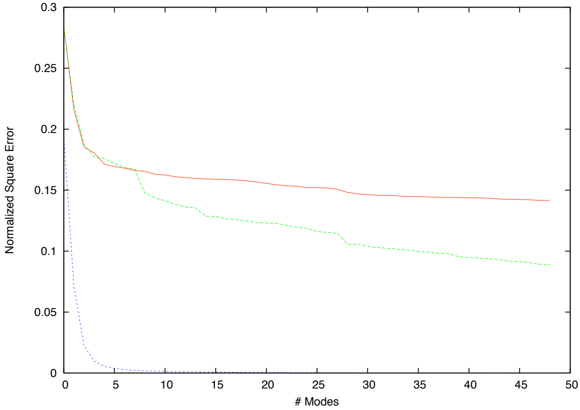
standard prior





self-similar prior

Results: reduction of POD-Galerkin projection error



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).

Any questions ?