LABORATOIRE MODÉLISATION ET SIMULATION MULTI ÉCHELLE

Sous la co-tutelle de :
CNRS
UNIVERSITÉ GUSTAVE EIFFEL
UPEC • UNIVERSITÉ PARIS-EST CRÉTEIL

Recalage d'un modèle de substitution en dynamique stochastique non linéaire avec données partielles : application aux tuyères aérospatiales

Evangéline Capiez-Lernout, Olivier Ezvan, Christian Soize

Université Gustave Eiffel, MSME Laboratory UMR 8208 CNRS

17-18 octobre 2024

Journées 2024 du GDR EX-MODELI

General context

Nozzle structures

- Updating **the nonlinear stochastic dynamics of a nozzle with uncertainties** in a context of **partial observation and incomplete data set**
- Aerospatial and aeronautics applications

High-Fidelity Computational Model

- Geometry of the nozzle structure : **structural complexity** involving an optimized shape for gas expansion
- ☞ **a high fidelity computational model with a large number of dof**
- Thin structures that can involve **large nonlinear displacements**
- **Complex materials**, such as ceramic matrix or carbon fiber composites, are represented by **elastic homogeneized materials with slight random anisotropy** driven by **uncontrolled uncertain parameters U**.
- **Stochastic excitation** for representing **the pressure jet** driven by **control parameter w**.
- Under-observed system
- ☞ **1 computation** corresponding to a given value for controlled and uncontrolled parameters is **highly costly**.

General context

Difficulties related to the updating context

- **Under-observed system**
- **Incomplete available measurements dataset** induced by the limited number of sensors located at the surface of the nozzle output.

☞ Large-dimensional parameterized HFCM with a high-dimensional parameter space in a context of partial observation and incomplete target data set.

- The updating procedure requires the exploration of a large dimension of the parameter space.

☞ **Practically, it is not possible to solve this optimization problem with usual optimization algorithms because of the large number of required calls.**

Strategies

- **The training set** is constructed using the parameterized HFCM yielding a small training data set.
- **The target set** is issued from another parameterization of the HFCM
- **The learning set** is constructed from PLoM under the target constraints

Definition of the parameterized HFCM

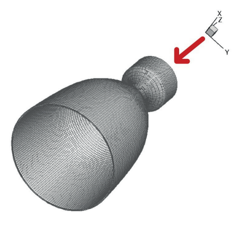
Finite element model

- Axisymmetric nozzle geometry, clamped at the combustion chamber represented by a FEM ($n_y = 216\,720$ dof) and requiring $n_t = 5000$ time sampling points.
- $\mathbf{F}(t;\mathbf{W})$ from a Gaussian stationary stochastic process as a function of control parameters $\mathbf{W}$ (standard deviation and frequency peak of its psd).
- Uncontrolled parameter $\mathbf{U}$ through full random Hooke matrix
- Stochastic observations
  - Displacement vector $\mathbf{Y}(t;\mathbf{W})$, $(n_y \times n_t)$
  - **Under-observed System** $(\overline{n}_y = 168)$
    $\mathbb{Y}(t;\mathbf{W}) = [B_{\text{UO}}]\mathbf{Y}(t;\mathbf{W})$, $\overline{n}_y \ll n_y$
  - **Observation** $\mathbb{O}(\nu;\mathbf{W}) = \log\left(|\widehat{\mathbb{Y}}(2\pi\nu;\mathbf{W})|\right)$, $N_o = 168 \times 5000 = 840\,000$
  - **Identification observation** $(n_{\text{exp}} = 8)$
    $\mathbb{O}_{\text{id}}(\nu;\mathbf{W}) = [B_{\text{id}}]\,\mathbb{O}(\nu;\mathbf{W})$, $n_o = 8 \times 5000 = 40000$
- High Fidelity Stochastic Uncertain Nonlinear Computational Model

$$[M]\ddot{\mathbf{Y}}(t;\mathbf{W}) + [D]\dot{\mathbf{Y}}(t;\mathbf{W}) + [K(\mathbf{U})]\mathbf{Y}(t;\mathbf{W}) + \mathbf{f}^{\text{NL}}(\mathbf{Y}(t;\mathbf{W}),\mathbf{U}) = \mathbf{F}(t;\mathbf{W})\,, \forall t \in T,$$

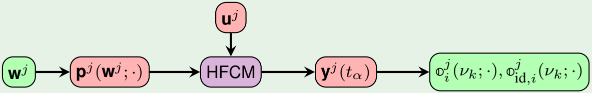
$$\mathbf{Y}(t_0;\mathbf{W}) = \mathbf{Y}_0(\mathbf{W},\mathbf{U})\,,$$

$$\dot{\mathbf{Y}}(t_0;\mathbf{W}) = \mathbf{Y}_1(\mathbf{W},\mathbf{U})\,.$$



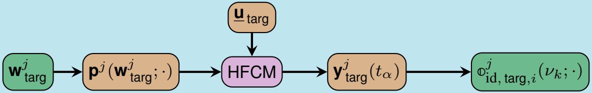
Definition of the training set and of the target data set

Organization of the training data set



- time/frequency discretization $n_\nu = n_t = 5000$
- partial observation $\overline{n}_y = 168 \Rightarrow \phi_i^j(\nu_k;\mathbf{w}^j,\mathbf{u}^j,\mathbf{p}^j) \quad (840000 \times 1)$
- identification observation $n_{\text{exp}} = 8 \Rightarrow \phi_{\text{id},i}^j(\nu_k;\mathbf{w}^j,\mathbf{u}^j,\mathbf{p}^j) \quad (40000 \times 1)$
- control parameter $n_w = 2 \Rightarrow \mathbf{w}^j \quad (2 \times 1)$
- ☞ $n_d = 400$ training realizations $\Rightarrow$ data matrix $[\varkappa^d] \quad (880002 \times 400)$ (a storage of $2.6GB$)

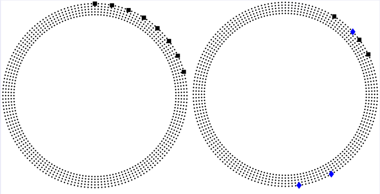
Organization of the target data set representing the experimental data



- identification observation $n_{\text{exp}} = 8 \Rightarrow \phi_{\text{id, targ},i}^j(\nu_k;\mathbf{w}_{\text{targ}}^j,\underline{\mathbf{u}}_{\text{targ}},\mathbf{p}^j) \quad (40000 \times 1)$
- control parameter $n_w = 2 \Rightarrow \mathbf{w}_{\text{targ}}^j \quad (2 \times 1)$
- ☞ $n_{\text{targ}} = 40$ target realizations $\Rightarrow$ data matrix $[\boldsymbol{\xi}_{\text{targ}}] \quad (40002 \times 40)$
- $\mathbf{w}_{\text{targ}}^j \neq \mathbf{w}^j$ and $\underline{\mathbf{u}}_{\text{targ}} = \alpha\mu_{\mathbf{u}} \neq \mathbf{u}^j$ deterministic
- **Target data set cannot be reproduced by the Training procedure.**

Visualisation

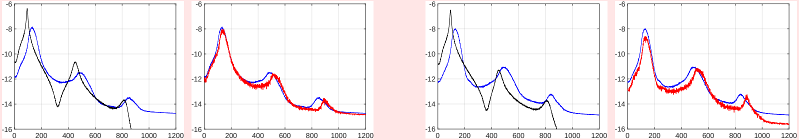
Location of the observations (normal displacements)



(*a*) Training and target. (*b*) Learning post-processing.

Data visualization

Identification observation — Validation observation



- mean linear (black) and nonlinear (blue) response of the training data set
- mean nonlinear response of the target data set (red)

☞ **How to update the parameterized HFCM with respect to the target ?**

Constructing the learning data set with PLoM

Actual context

- **Small training data set**
- **Small target data set**

☞ A lack of data prevents the use of deep learning techniques

Probabilistic learning strategy

- Utilisation de PLoM ”Probabilistic Learning on Manifolds”
  - C. Soize, R. Ghanem, JPC, 2016
  - C. Soize, R. Ghanem, CMAME, 2021

☞ adapted to **small training data set**

☞ capability of preserving the concentration of the **learned probability density function** in the vinicity of a random manifold.

☞ capability of constructing fastly a large number of **learned realizations** that (1) capture the statistical characteristics of the training data and (2) take into account the **target constraint**.

☞ **The learned probability distribution function** allows for estimating the conditional statistics of the model output observations given the input control parameters.

Constructing the learning data set with PLoM

A Very Brief Summary of usual PLoM

- **Small training data set** is collected into $[\mathbb{x}^d]$
- A PCA statistical reduction of the data, preserving its statistical properties
- ☞ from $[\mathbb{x}^d]$ with dim. (840002×400) to data matrix $[\eta^d]$ with dim. (307×400)
- Estimating the pdf of random matrix $[\mathbf{H}_d]$ by Multi-Dimensional Gaussian KDE
- Constructing a MCMC generator obtained by projection of an Itô SDE onto a diffusion map basis for avoiding data scatter.
- Generating **the learned realizations** $[\mathbf{H}_{\text{learned}}]$

Including target data set constraint

- Constructing an updated statistical surrogate model with PLoM, from **the training data with partial observability** constrained by **an available incomplete target data set**
- Projection of **the incomplete target data set** on the PCA model representation
- **Learned data set under constraint** $[\mathbf{H}_{\text{ud}}]$ $(307 \times 50\,000)$
  - Admissible set $\mathcal{C}_{\text{ad},p} = \{\boldsymbol{\eta} \mapsto p(\boldsymbol{\eta}) : \mathbb{R}^{\nu_p} \to \mathbb{R}^+, \int_{\mathbb{R}^{\nu_p}} p(\boldsymbol{\eta})\, d\boldsymbol{\eta} = 1, \int_{\mathbb{R}^{\nu_p}} \mathbf{h}^c(\boldsymbol{\eta})\, p(\boldsymbol{\eta})\, d\boldsymbol{\eta} = \mathbf{b}^c\}$
  - Minimization of the Kullback-Leibler divergence in order to get the posterior pdf as closest to the prior pdf within $\mathcal{C}_{\text{ad},p}$. $\quad p_{\mathbf{H}_{\text{ud}}} = \arg\min_{p \in \mathcal{C}_{\text{ad},p}} \int_{\mathbb{R}^{\nu_p}} p(\boldsymbol{\eta}) \log\left(\frac{p(\boldsymbol{\eta})}{p_{\mathbf{H}}(\boldsymbol{\eta})}\right) d\boldsymbol{\eta}.$
- Constructing the surrogate model by estimating the conditional statistics of **the learned observations** knowing the input control parameters

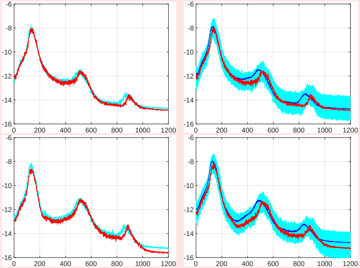
Results

Graphical observations

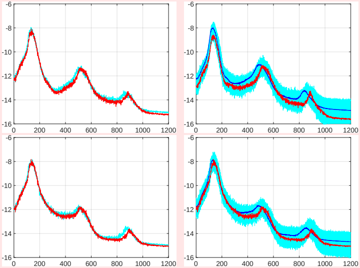
- mean nonlinear response of the target data set (red) / of the training data set (blue)
- mean conditional expectation of the learned data set (given **w**) (cyan)
- conditional confidence region of the learned data set (given **w**) (cyan area)

Graphical representations

Identification observation



Validation observation



Results

- Good results for observed and unobserved outputs
- A good capture of the mean statistics
- A target centered in the C.R., a C.R. that follows the target peaks, a C.R. with a reasonable width

Conclusion

- Context of the updating of a parameterized HFCM with geometrical non-linearity, uncontrolled uncertainty and random controlled parameter, under stochastic excitation, that is underobserved.

☞ Complexity yields computational costs yielding a small training data set.

- An available incomplete target data set
- An updating by constructing a statistical surrogate model that maps the control parameters to the observations using the "PLoM under constraints" methodology.

Some references

- C. Soize, R. Ghanem, JPC, 2016
- C. Soize, R. Ghanem, CMAME, 2023
- E. Capiez-Lernout, O. Ezvan, C. Soize, JCISE, 2024