Dynamical low rank approximation of time dependent random PDEs

Eleonora Musharbash†, Fabio Nobile†, Eva Vidličková†

†MATH-CSQI, École Polytechnique Fédérale de Lausanne, Switzerland

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Motivation

Problem setting:

Physical space $D \subset \mathbb{R}^d$, $d = 1, 2, 3$; time interval $\mathcal{T} = [0, T]$; random variables $\xi = (\xi_1, \dots, \xi_N) \in \Gamma$ with a joint density function $\rho \in L^\infty(\Gamma)$

$$\begin{aligned} \frac{\partial u(x,t,\xi)}{\partial t} &= \mathcal{L}(u(x,t,\xi),\xi), && x \in D,\ t \in \mathcal{T},\ \xi \in \Gamma \\ u(x,0,\xi) &= u_0(x,\xi), && x \in D,\ \xi \in \Gamma \\ u(x,t,\xi) &= h(x,t), && x \in \partial D,\ t \in \mathcal{T},\ \xi \in \Gamma \end{aligned}$$

Goal: Low rank approximation

- **POD** (spatial basis fixed in time)

$$u_R(x,t,\xi) = \bar{u}_R(x,t) + \sum_{i=1}^{R} Y_i(t,\xi) U_i(x),$$

where $\{U_i(x)\}_{i=1}^R$ orthogonal functions in $L^2(D)$

- **Issues**

solution manifold changes over time → R too high

- **PCE** (stochastic basis fixed in time)

$$u_R(x,t,\xi) = \bar{u}_R(x,t) + \sum_{i=1}^{R} Y_i(\xi) U_i(x,t),$$

where $\{Y_i(\xi)\}_{i=1}^R$ zero mean stochastic processes orthonormal in $L^2_\rho(\Gamma)$

- **Issues**

"curse of dimensionality", long time integration

Introduction to DLR

Goal: Dynamical low rank approximation

$$u_R(x,t,\xi) = \bar{u}_R(x,t) + \sum_{i=1}^{R} Y_i(t,\xi)\, U_i(x,t), \tag{1}$$

where $\bar{u}_R(x,t) = \mathbb{E}[u_R(x,t,\xi)]$, $\{Y_i(t,\xi)\}_{i=1}^R$ are zero mean stochastic processes in $L^2_\rho(\Gamma)$, $\{U_i(x,t)\}_{i=1}^R$ are orthonormal functions in $L^2(D)$.

DO condition:

$$\left\langle \frac{\partial U_i(\cdot,t)}{\partial t}, U_j(\cdot,t) \right\rangle_{L^2(D)} = 0, \qquad \forall i,j = 1,\dots,R,\ \forall t \in \mathcal{T}$$

$\Longrightarrow$ unique representation of (1) up to rotation of the basis $\{U_i\}_{i=1}^R$.

Dynamically orthogonal approximation:

$$\frac{\partial \bar{u}_R(x,t)}{\partial t} = \mathbb{E}[\mathcal{L}(u_R(x,t,\xi))] \tag{2}$$

$$\frac{\partial U_i(x,t)}{\partial t} = \mathbb{E}[\mathcal{L}(u_R(x,t,\xi)) Y_i(t,\xi)], \qquad \forall i = 1,\dots,R \tag{3}$$

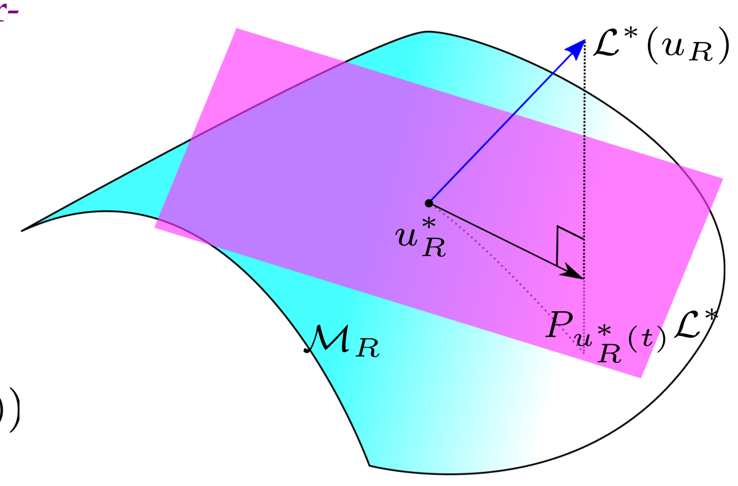
$$\sum_{j=1}^{R} \mathbf{M}_{ij} \frac{\partial Y_j(t,\xi)}{\partial t} = \Pi^\perp_{\mathcal{Y}} \langle \mathcal{L}^*(u_R(x,t,\xi),\xi), U_i(x,t) \rangle_{L^2(D)}, \qquad \forall i = 1,\dots,R \tag{4}$$

where $\mathcal{Y} = span\{Y_1(t,\xi),\dots,Y_R(t,\xi)\}$, $\Pi^\perp_{\mathcal{Y}}$ is the orthogonal projection operator in $L^2_\rho(\Gamma)$ to the complement of $\mathcal{Y}$, $\mathcal{L}^*(\cdot) = \mathcal{L}(\cdot) - \mathbb{E}[\mathcal{L}(\cdot)]$ and $\mathbf{M}_{ij} = \langle U_i(\cdot,t), U_j(\cdot,t) \rangle_{L^2(D)}$.

Geometrical interpretation

- The set of all zero mean R-rank random fields form a differential manifold:

$$\mathcal{M}_R = \left\{ u^*_R = \sum_{i=1}^{R} U_i Y_i,\ \ U_i, Y_i \text{ as in (1)} \right\}.$$



- u_R is the DO solution $\iff$

$$\frac{\partial u_R(x,t,\xi)}{\partial t} = \mathbb{E}[\mathcal{L}(u_R(x,t,\cdot))] + P_{u^*_R(t)}(\mathcal{L}^*(u_R(x,t,\xi)))$$

where $P_{u^*_R(t)}$ is the orthogonal projection onto the tangent space of $\mathcal{M}_R$ at $u^*_R = u_R - \mathbb{E}[u_R]$ [1, 3].

Discretization

- FEM for spatial discretization, semi-implicit Euler for time discretization of $\bar{u}_R(\cdot,t),\ U_i(\cdot,t)$ and explicit Euler for $Y_i(\cdot,t)$
- sparse grid for stochastic discretization:
  - $m(i)$ - no. of collocation points at level i
  - $knots$ - type of collocation points
  - I - set of multiindices

  $\Big\}\ \{y_k\}_{k=1}^{N_c(I)},\ \{w_k\}_{k=1}^{N_c(I)}$ - colloc. points and weights
- sparse grid interpolant $S_I : L^2_\rho(\Gamma) \cap C^0(\Gamma) \to \bigoplus_{i \in I} \mathbb{P}_{m(i)-1}(\Gamma)$

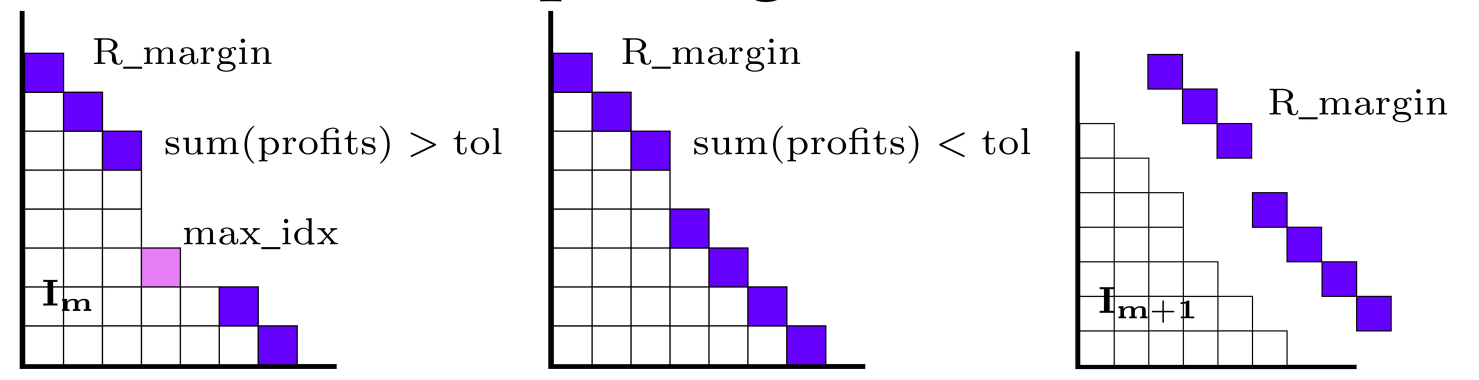
$$S_I[f](\xi) = \sum_{i \in I} \triangle^{m(i)}[f](\xi) = \sum_{k=1}^{N_c(I)} f(y_k)\, L_k(\xi)$$

$$\triangle^{m(i)} = \bigotimes_{n=1}^{N} (\mathcal{U}_n^{m(i_n)} - \mathcal{U}_n^{m(i_n-1)}),$$

where $\{L_k(\xi)\}_{k=1}^{N_c(I)}$ - Lagrange polynomials, $\mathcal{U}_n^l$ - polynomial interpolant of order $l-1$ in the n-th variable ξ_n.

- sparse grid quadrature: $\mathbb{E}[S_I[f]] = \sum_{k=1}^{N_c(I)} f(y_k)\, w_k$

Adaptive choice of sparse grid



- time discretization $0 = t_0 < t_1 < \dots < t_M = T$

Goal: multiindex set I_{m+1} for sparse grid $S_{I_{m+1}}$ at time t_{m+1}

Data: multiindex set I_m from previous time step; so far algorithm works only with $I_m \subseteq I_{m+1}$, nested collocation points

Motivation: adapt w.r.t function **f**

Assume $\quad f(\xi) = \sum_{i \in \mathbb{N}^N_+} \triangle^{m(i)}[f](\xi),$

choose $I \supset I_m$ s.t.

$$\|f - S_I[f]\| \le \sum_{i \in \mathbb{N}^N_+ \setminus I} \|\triangle^{m(i)}[f]\| \approx \sum_{i \in \text{R\_margin}(I)} \|\triangle^{m(i)}[f]\| < \textbf{tol},$$

where R_margin$(I) = \{i \in \mathbb{N}^N_+ \setminus I : i - e_n \in I,$ for all $n = 1,\dots,N$ with $i_n > 1\}$ [2].

Profit function: profit_function$(I) = \{\|\triangle^{m(i)}[f]\|\}_{i \in I}$

Adapted_sparsegrid(I_m, R_margin, profit_function, tol)

```
I_{m+1} ← I_m;  profits = profit_function(R_margin);  max_idx ← argmax(profits)
while sum(profits) > tol do
    I_{m+1} ← I_m ∪ max_idx
    R_margin ← R_margin \ max_idx ∪ admiss_neighbours(max_idx)
    profits = profit_function(R_margin)
    max_idx ← argmax(profits)
end while
return I_{m+1}
```

Sparse grid adaptivity within DLR

Function **f** - unprojected right hand side of (4):

$$f(\xi) = (f_1(\xi),\dots,f_R(\xi))^T,$$

$$f_i(\xi) = \langle \mathcal{L}(u_R(x,t,\xi),\xi), U_i(x,t) \rangle_{L^2(D)}$$

Profit function:

$$\text{profit\_function}^2(i) = \|\triangle^{m(i)}[f]\|^2_{L^2_\rho} \approx \sum_{k=N_c(I_m)}^{N_c(I_m \cup i)} \|f(y_k) - S_{I_m}[f](y_k)\|^2_{l^2(\mathbb{R}^R)}\, \rho(y_k)$$

Numerical difficulties

To compute accurately the projections $\Pi^\perp_{\mathcal{Y}} \langle \mathcal{L}, U_i \rangle_{L^2(D)}$ in (4) and the right hand side in (3) one needs a good enough approximation of $\mathbb{E}[Y_j\, Y_k]$. However, unless Gaussian points are used:

$$\|\mathbb{E}[Y_j\, Y_k] - \mathbb{E}[S_I[Y_j\, Y_k]]\|_{Frob} \gg 0.$$

Solution: modal representation:

$$Y_j(\xi) = \sum_{l=1}^{K} a^l_j \psi_l(\xi), \qquad \psi_l\text{- ON polynomials in } L^2_\rho(\Gamma)$$

then

$$\mathbb{E}[Y_j\, Y_k] = \sum_{l=1}^{K} a^l_j\, a^l_k \qquad \to \qquad \text{compute } a^l_j \quad \forall j = 1,\dots,R;\, l = 1,\dots,K.$$

Numerical results

$$\begin{aligned} \frac{\partial u(x,t,\xi)}{\partial t} &= \nabla \cdot (a(x,\xi) \nabla u(x,t,\xi)), && x \in [0,\,1]^2,\ t \in [0,\,0.1],\ \xi \in \Gamma \\ u(x,0,\xi) &= 10 \sin(\pi x_1) \sin(\pi x_2) \xi_1, && x \in [0,\,1]^2,\ \xi \in \Gamma \\ u(x,t,\xi) &= 0, && x \in \partial[0,\,1]^2,\ t \in [0,\,0.1],\ \xi \in \Gamma \end{aligned}$$

where $a(x,\xi) = 0.5(\bar{a}(x) + \sum_{j=1}^{12} \sqrt{\lambda_j} X_j(x) \xi_j)$ the first 12 terms of the K-L expansion with $\xi_j \sim U[-1,1]$ and $X_j(x), \lambda_j$ given by the covariance function

$$\text{Cov}(x,y) = \exp(-\|x-y\|^2)/l^2, \qquad l = 0.5$$

- **adaptivity parameters:** Leja knots - nested, $m(i_n) = 2i_n - 1$ - adding 2 points in every adaptivity step
- **function f:**

$$f(\xi) = -\sum_{j=1}^{12} \sqrt{\lambda_j} \langle X_j\, \nabla \bar{u}_R, \nabla U \rangle_{L^2(D)} \xi_j - \sum_{i=1}^{R} \sum_{j=1}^{12} \sqrt{\lambda_j} \langle X_j \nabla U_i, \nabla U \rangle_{L^2(D)} \xi_j Y_i(\xi),$$

where $U = (U_1,\dots,U_R)^T,\ Y = (Y_1,\dots,Y_R)^T$.

type of grid	no. of collocation points	max no. of pts in each dim	error indicator
adapted sparse grid	3465	[7,5,5,5,3,3,3,5,3,3,5,5]	7.88e-2
tensor grid	23625	[7,5,5,5,3,3,3]	4.85e-1
tensor grid	354375	[7,5,5,5,3,3,3,5,3]	4.33e-1
isotropic sparse grid	6188	[6,6,6,6,6,6,6,6,6,6,6,6]	3.88e-1

Table 1: Discretization parameters: $\mathbb{P}_1$ FEM space (51×51 points), semi-implicit Euler scheme (10^3 time steps), rank $R = 13$, knots: Leja - nested for sparse grids, Gauss-Legendre for tensor grids. Multiindex set for isotropic sparse grid: $\sum_n i_n - 1 \le w$. **Reference solution:** $u_h(\cdot, 0.09, y_l),\ l = 1,\dots,9$ computed with a very small time step (10^5 steps) at 9 uniformly-randomly chosen points y_l. **Error indicator:** $\sum_l \|u_R(\cdot,0.09,y_l) - u_h(\cdot,0.09,y_l)\|_{L^2(D)} / \|u_h(\cdot,0.09,y_l)\|_{L^2(D)}$ averaged over $y_l,\ l = 1,\dots,9$.

References

[1] Musharbash, E., Nobile, F., Zhou, T., Error analysis of the dynamically orthogonal approximation of time dependent random PDEs, SIAM J. SCI. COMPUT.,37(2), A776-A810, 2015

[2] Nobile, F., Tamellini, L., Tesei, F., Tempone R, *An adaptive sparse grid algorithm for elliptic PDEs with lognormal diffusion coefficient*, Sparse Grids and Applications - Stuttgart, 2014

[3] Musharbash, E., Dynamical Low Rank approximation of PDEs with random parameters, PhD dissertation, no. 7813, École Polytechnique Fédérale de Lausanne, 2017