Reduced Basis ANOVA for PDEs with High-Dimensional Random Inputs

Qifeng Liao[+] and Guang Lin[*]

[+]School of Information Science and Technology, ShanghaiTech University

[*]Department of Mathematics, Purdue University

Outline

Outline

Partial Differential Equations with Uncertain Coefficients

Let $\xi \in I^M$ be a random vector. We find a random function $u(x, \xi)$:

$$\mathcal{L}(x, \xi; u(x, \xi)) = f(x, \xi), \qquad (x, \xi) \in D \times I^M,$$

$$\mathfrak{b}(x, \xi; u(x, \xi)) = g(x, \xi), \qquad (x, \xi) \in \partial D \times I^M.$$

- $\mathcal{L}$: a partial differential operator.
- $\mathfrak{b}$: a boundary operator.
- Both of $\mathcal{L}$ and $\mathfrak{b}$ can have random coefficients.
- The random source ξ is high dimensional.

ANOVA decomposition (Cao, Chen, Gunzburger, Gao, Hesthaven, ...)

$$\mathcal{L}\left(x,\xi;u\left(x,\xi\right)\right)=f(x,\xi),\qquad (x,\xi)\in D\times I^M,$$

$$\mathfrak{b}\left(x,\xi;u\left(x,\xi\right)\right)=g(x,\xi),\qquad (x,\xi)\in \partial D\times I^M.$$

Decompose the (global) random solution $u(x,\xi)$ w.r.t ξ:

$$u(x,\xi)=u_\emptyset(x)+u_1(x,\xi_1)+\ldots+u_{1,2}(x,\xi_{1,2})+\ldots=\sum_{t\in\mathcal{P}}u_t(x,\xi_t).$$

Given anchor point $c=(c_1,\ldots,c_M)\in I^M$

Define index set $\mathcal{P}:=\{\mathcal{P}_0,\mathcal{P}_1,\ldots,\mathcal{P}_M\}$ $\rightarrow$

$\mathcal{P}_0$: $\{\emptyset\}$
$\mathcal{P}_1$: $\{1,\ldots,M\}$
$\mathcal{P}_2$: $\{(1,2),\ (1,3),\ \ldots,\ (2,3),\ \ldots\}$
$\ldots$
$\mathcal{P}_M$: $\{(1,2,\ldots,M)\}$

- $u_\emptyset(x):=u(x,c)$
- $u_1(x,\xi_1):=u\Big(x,(\xi_1,c_2,\ldots,c_M)\Big)-u_\emptyset(x)$
- Define a local solution for $t\in\mathcal{P}$: $u(x,c,\xi_t):=u\Big(x,(c_1,..,\xi_{t_1},...\xi_{t_2},...)\Big)$
- $u_t(x,\xi_t):=u(x,c,\xi_t)-\sum_{s\subset t}u_s(x,\xi_s)$

Stochastic collocation for each ANOVA term

$$u(x,\xi) = \sum_{t\in\mathcal{P}} u_t(x,\xi_t), \qquad u_t(x,\xi_t) := u(x,c,\xi_t) - \sum_{s\subset t} u_s(x,\xi_s),$$
$$u(x,c,\xi_t) := u\Big(x,(c_1,..,\xi_{t_1},...\xi_{t_2},...)\Big).$$

$$u(x,c,\xi_t) \text{ satisfies: } \begin{cases} \mathcal{L}\left(x,\xi_t;u\left(x,c,\xi_t\right)\right) = f(x), & (x,\xi_t)\in D\times I^{|t|}, \\ \mathfrak{b}\left(x,\xi_t;u\left(x,c,\xi_t\right)\right) = g(x), & (x,\xi_t)\in \partial D\times I^{|t|}. \end{cases}$$

- $|t|$ (dimension of t) is expected to be $\ll M$.
- Approximate $u(x,c,\xi_t)$ using stochastic collocation:

$$u^q\left(x,c,\xi_t\right) := \sum_{\xi_t^{(k)}\in\Theta_q^{|t|}} u\left(x,c,\xi_t^{(k)}\right)\Phi_{\xi_t^{(k)}}\left(\xi_t\right) \qquad \approx u(x,c,\xi_t).$$

- Overall approximation: $u(x,\xi) \approx u^q(x,\xi) := \sum_{t\in\mathcal{P}} u_t^q(x,\xi_t)$,

$$u_t^q(x,\xi_t) := u^q(x,c,\xi_t) - \sum_{s\subset t} u_s(x,\xi_s).$$

- Stochastic collocation: Xiu, Hesthaven, Babuška, Nobile, Tempone, Webster ...
- ANOVA-Collocation: Ma, Zabaras, Yang, Lin, Karniadakis ...

Computational aspects of ANOVA-Collocation approximation

ANOVA-Collocation: $u(x,\xi) \approx u^q(x,\xi) := \sum_{t\in\mathcal{P}} u_t^q(x,\xi_t)$,

$$u_t^q(x,\xi_t) := u^q(x,c,\xi_t) - \sum_{s\subset t} u_s(x,\xi_s),$$

$$u^q\left(x,c,\xi_t\right) := \sum_{\xi_t^{(k)}\in\Theta_q^{|t|}} u\left(x,c,\xi_t^{(k)}\right)\Phi_{\xi_t^{(k)}}\left(\xi_t\right).$$

$\mathcal{P} := \{\mathcal{P}_0, \mathcal{P}_1, \dots, \mathcal{P}_M\}$

$\mathcal{P}_0$: $\{\emptyset\}$

$\mathcal{P}_1$: $\{1,\dots,M\}$

$\mathcal{P}_2$: $\{(1,2),\ (1,3),\ \dots,(2,3),\ \dots\}$

$\cdots$

$\mathcal{P}_M$: $\{(1,2,\dots,M)\}$

Computation challenges

- Many ANOVA terms ($|\mathcal{P}|$ is large)

  Adaptive ANOVA (selecting important terms)

  - Ma, Zabaras (2010); Yang, et al. (2012)

- Spatial $d.o.f$ can be very large (computing each collocation coefficient $u\left(x,c,\xi_t^{(k)}\right)$ is expensive).

  Reduced basis collocation:

  - Elman, Liao (2013)

Adaptive ANOVA–selecting important terms (indices)

ANOVA-Collocation: $u(x,\xi) \approx u^q(x,\xi) := \sum_{t\in\mathcal{P}} u_t^q(x,\xi_t)$.

$\mathcal{P}_0$:	$\{\emptyset\}$
$\mathcal{P}_1$:	$\{1, \ 2, \ 3, \ 4, \ 5\}$
$\mathcal{P}_2$:	$\{(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)\}$
$\mathcal{P}_3$:	$\{(1,3,5)\}$
$\mathcal{P}_4$:	No $4th$ order terms.

Figure: A example of adaptive index selection.

Selecting criterion:

- Relative mean value $\rightarrow$

$$\text{relative-mean}_t := \frac{\left\|\mathbb{E}(u_t^q)\right\|}{\left\|\mathbb{E}\left(\sum_{s\in\mathcal{P},|s|\le|t|-1} u_s^q\right)\right\|}$$

■ Adaptive ANOVA: Ma and Zabaras (2010); Yang, et al. (2012).

Outline

$u^q(x,\xi) := \sum_{t\in\mathcal{P}} u_t^q(x,\xi_t), \quad u_t^q(x,\xi_t) := u^q(x,c,\xi_t) - \sum_{s\subset t} u_s(x,\xi_s),$

$$u^q\left(x,c,\xi_t\right) := \sum_{\xi_t^{(k)}\in\Theta_q^{|t|}} u\left(x,c,\xi_t^{(k)}\right)\Phi_{\xi_t^{(k)}}\left(\xi_t\right).$$

Finite element methods

- Let $\mathfrak{B}_{\xi_t}(\cdot,\cdot) = l(\cdot)$ denote a weak form, and X^h a FEM space.
- Seek $u_h(\cdot,c,\xi_t) \in X^h \rightarrow \quad \mathfrak{B}_\xi\left(u_h\left(\cdot,c,\xi_t\right),v\right) = l(v), \quad \forall v \in X^h.$

Each FEM solution $u_h(\cdot,c,\xi_t)$ is called a *snapshot*.

Reduced basis approximation

- Introduce a reduced basis Q with a small size, $span(Q) \subset X^h$.
- Seek $u_r\left(\cdot,c,\xi_t\right) \in span(Q) \rightarrow \mathfrak{B}_\xi\left(u_r\left(\cdot,c,\xi_t\right),v\right) = l(v), \forall v \in span(Q).$

Each $u_r\left(\cdot,c,\xi_t\right)$ is called a *reduced solution*.

What information should Q contain, and how large is it?

- Ideally, $span(Q) \supset \{u_h\left(\cdot,c,\xi_t\right), \xi_t \in I^{|t|}\}$, (*the full snapshot set*).
- Size of Q = rank of $\{u_h\left(\cdot,c,\xi_t\right), \xi_t \in I^{|t|}\} \ll N_h$? ($N_h$: FEM $d.o.f$).

Algebraic Issue and Error Indicator, Linear PDEs

Original finite element approximation: $\mathbf{A}_{\xi_t} \in \mathbb{R}^{N_h \times N_h} \rightarrow$

$$\mathbf{A}_{\xi_t}\mathbf{u}_h = \mathbf{f}.$$

Reduced basis approximation: $\mathbf{Q} \in \mathbb{R}^{N_h \times N_r}$ with $N_r \ll N_h \rightarrow$

$$\mathbf{Q}^T\mathbf{A}_{\xi_t}\mathbf{Q}\mathbf{u}_r = \mathbf{Q}^T\mathbf{f}.$$

Reduced basis approximation is a projection:

- projects a large $N_h \times N_h$ system to a small $N_r \times N_r$ system $\rightarrow$ very cheap to solve.

To estimate the error $e = \mathbf{u}_h - \mathbf{Q}\mathbf{u}_r$, we use the residual indicator:

$$\text{error-indicator}_{\xi_t} = \|\mathbf{A}_{\xi_t}\mathbf{Q}\mathbf{u}_r - \mathbf{f}\|.$$

- The cost of this residual indicator is $O(N_r^2)$, independent of N_h.

Greedy Algorithm (Patera, Boyaval, Bris, Lelièvre, Maday, Nguyen, ...)

Goal for reduced solution: $u_r \approx u_h$, $\leftrightarrow$ $span(Q) \approx \{u_h(\cdot, c, \xi_t), \xi_t \in I^{|t|}\}$.

- SVD approach: get Q from $SVD\{u_h(\cdot, c, \xi_t), \xi_t \in I^{|t|}\}$, but may expensive.
- Greedy approach: find most important samples $\rightarrow Q$.

Given: a set of candidate parameters $\chi = \{\xi_t\}$,

an initial choice $\xi_t^{(1)} \in \chi$, and compute the snapshot $u_h(\cdot, c, \xi_t^{(1)})$.

Initialize: $Q = \{u_h(\cdot, c, \xi_t^{(1)})\}$

for each $\xi_t \in \chi$

compute reduced solution $u_r(\cdot, c, \xi_t)$

compute error-indicator$_{\xi_t}$ (an error indicator for $\|u_h - u_r\|$)

If error-indicator$_\xi > tol$

compute $u_h(\cdot, c, \xi_t)$, and update $Q = \{Q, u_h(\cdot, c, \xi_t)\}$

endif

endfor

■ Greedy on sparse grids: Elman and Liao (2013); Chen et al. (2015)

Reduced bases for ANVOA-Collocation terms

ANOVA-Collocation: $u(x,\xi) \approx u^q(x,\xi) := \sum_{t\in\mathcal{P}} u_t^q(x,\xi_t),$

$$u_t^q(x,\xi_t) := u^q(x,c,\xi_t) - \sum_{s\subset t} u_s(x,\xi_s),$$

$$u^q(x,c,\xi_t) := \sum_{\xi_t^{(k)}\in\Theta_q^{|t|}} u\left(x,c,\xi_t^{(k)}\right)\Phi_{\xi_t^{(k)}}(\xi_t).$$

1. Use collocation points $\Theta_q^{|t|}$ as candidate set χ.
2. Use reduced solution $u_r \to u_c := u\left(x,c,\xi_t^{(k)}\right)$ whenever possible.
3. Different reduced basis Q_t for different t, but use them hierarchically $\rightarrow$

$\mathcal{P}_0$: $\{\emptyset\}$

$\mathcal{P}_1$: $\{1, \quad 2, \quad 3, \quad 4, \quad 5\}$

$\mathcal{P}_2$: $\{(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)\}$

$\mathcal{P}_3$: $\{(1,3,5)\}$

... ...

Algorithm (Reduced Basis ANOVA)

1. Start with ANOVA level $i = 0$, initialize the index set $\mathcal{P}_0 = \{\emptyset\}$.
2. Set $Q_\emptyset := \{u_h(\cdot, c)\}$.
3. Set $\mathcal{P}_1 = \{1, \ldots, M\}$.
4. Update ANOVA level $i = i + 1$.
5. Loop over each $t \in \mathcal{P}_i$, i.e. $|t| = i$
   - Initialize local reduced basis: $Q_t := SVD\{q \mid q \in \cup_{s \subset t} Q_s\}$.
   - For each $\xi_t^{(k)} \in \Theta_q^{|t|}$ (collocation points),
     compute the reduced solution $u_r(\cdot, c, \xi_t^{(k)})$ and error-indicator$_{\xi^{(k)}}$.
     If error-indicator$_{\xi^{(k)}} < tol$, $u_c \leftarrow u_r(\cdot, c, \xi_t^{(k)})$.
     If error-indicator$_{\xi^{(k)}} \geq tol$, $u_c \leftarrow u_h(\cdot, c, \xi_t^{(k)})$ and $Q_t := \{Q_t, u_h\}$.
   - Compute $\mathrm{relative\text{-}mean}_t$.
   - If $\mathrm{relative\text{-}mean}_t < tol_{ANOVA}$, remove the index t: $\mathcal{P}_i = \mathcal{P}_i \setminus t$.
6. Generate $\mathcal{P}_{i+1}$ based on $\mathcal{P}_i$, and repeat step 5 for next level $i = i + 1$.

Qifeng Liao+ and Guang Lin* +ShanghaiTech University, *Purdue University

Outline

Test Problem

Diffusion equation: $-\nabla \cdot (a\nabla u) = f$ in $[0,1]^2$

The permeability coefficient a is a random field:

- mean function: $a_0(x) = 1$, standard deviation: $\sigma = 0.25$
- covariance function $C(x, y)$:

$$C(x,y) = \sigma^2 \exp\left(-\frac{|x_1 - y_1|}{c} - \frac{|x_2 - y_2|}{c}\right),$$

  where c is the correlation length.

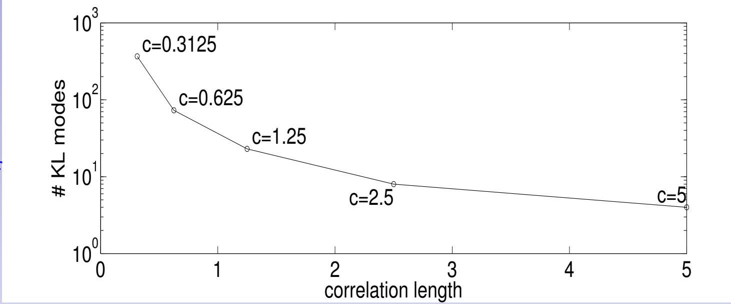
Parameterizing a using truncated KL expansion:

$$a(x,\xi) \approx a_0(x) + \sum_{k=1}^{M} \sqrt{\lambda_k} a_k(x)\xi_k,$$

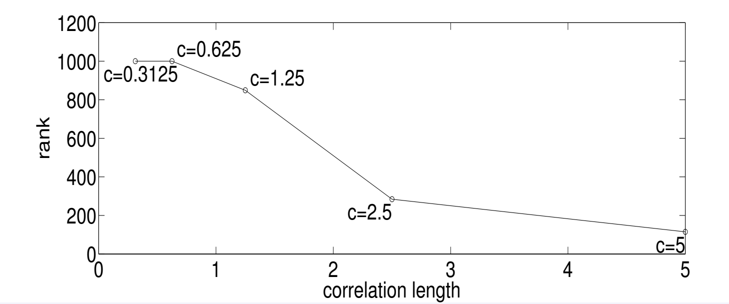
random vector $\xi = (\xi_1, \cdots, \xi_M)$ is uniformly distributed in $\Gamma = [-1,1]^M$.

- Small correlation length c leads to many KL terms.
- We consider small c situations (high-dimensional problems).

# of KL modes (M) for capturing 95% of the total variance



Rank of $\{u_h(\cdot,\xi), \xi \in I^{|M|}\}$



- FEM $d.o.f : N_h = 1089$
- Directly applying reduced basis methods may not be efficient for $c \leq 0.625$

Direct combination of MC and reduced basis (for comparison)

For each MC input sample $\xi^{(k)}$,

compute reduced solution $u_r(\cdot, \xi^{(k)})$ and error-indicator$_{\xi^{(k)}}$:

if error-indicator$_{\xi^{(k)}} < tol$, MC sample $\leftarrow u_r(\cdot, \xi^{(k)})$;

if error-indicator$_{\xi^{(k)}} \geq tol$, MC sample $\leftarrow u_h(\cdot, \xi^{(k)})$ and $Q := \{Q, u_h\}$.

Computational cost assessment model:

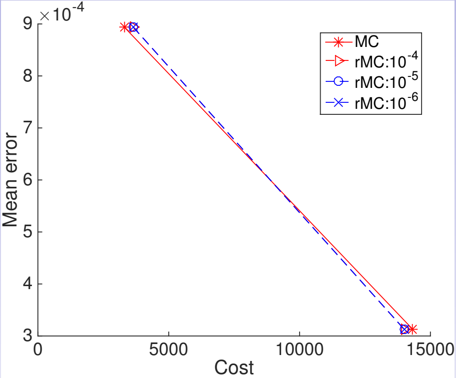
- Cost unit: 1 FEM system solve.
- Cost of a reduced system solve: N_r/N_h,

  (N_r: reduced basis size; N_h: FEM $d.o.f$).
- Cost of a full MC with N samples: N.
- Cost of a reduced basis MC with N samples and $\tilde{N}$ FEM solves:

$$\tilde{N} + \sum_{k=1}^{N} \frac{N_r(\xi^{(k)})}{N_h},$$

reduced basis size $N_r(\xi^{(k)})$ is dependent on $\xi^{(k)}$ in the greedy procedure.

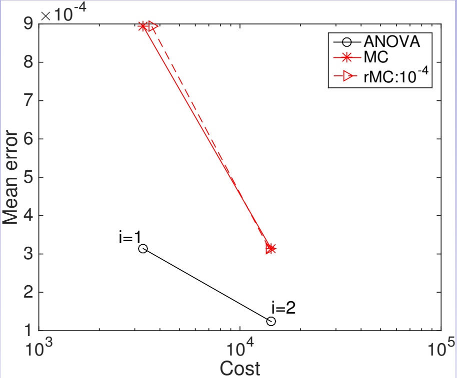
Direct reduced MC test, for $c = 0.3125$, $M = 367$; rank $\approx N_h = 1089$.



For this test, comparing MC and reduced basis MC (rMC),

- costs of the reduced basis MC are still large.

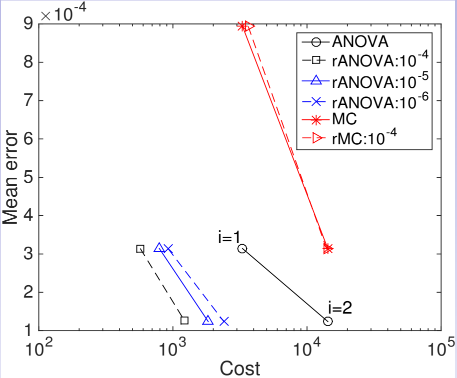
ANOVA vs MC, for $c = 0.3125$, $M = 367$; rank $\approx N_h = 1089$.



For this test,

- ANOVA has very small mean errors.

Reduced basis ANOVA, for $c = 0.3125$, $M = 367$; rank $\approx N_h = 1089$.



For this test,

- Reduced basis ANOVA (rANOVA) is very cheap.

Summary

- ANOVA methods have been designed to solve PDEs with high-dimensional random inputs.
- Many PDE solves can be involved for generating ANOVA-Collocation approximation.
- Our hierarchically-generated reduced bases can reduce the computational costs of ANOVA methods.