Sparse grid and reduced basis approximation for Bayesian inverse problems

Peng Chen [1]

joint work with

Christoph Schwab [1]

Acknowledgement: Alfio Quarteroni [2] Gianluigi Rozza [3]

[1]SAM - D-MATH - ETH, Zurich

[2]CMCS - MATHICSE - EPFL, Lausanne

[3]MathLab - SISSA - International School for Advanced Studies, Trieste, Italy

Pro∗Doc Retreat Disentis, August 13 – 15, 2014

Outline

Bayesian inverse problems [Stuart, 2010]

Problem: given noisy observation data for system output, to calibrate unknown input.

- X – (separable Banach) space for unknown input ;
- Y – (separable Banach) space for system output.

Given a forward operator (e.g. PDEs, system of ODEs, etc.)

$$G : X \to Y,$$

and a observation operator (a set of sensors, e.g. pointwise data, Gaussian average)

$$\mathcal{O} : Y \to \mathbb{R}^K,$$

with $K \in \mathbb{N}$. We define the map from unknown input to finite data

$$\mathcal{G} := \mathcal{O} \circ G : X \to \mathbb{R}^K.$$

The inverse problem: find $u \in X$ given the noisy observation

$$\delta = \mathcal{G}(u) + \eta,$$

where $\eta \in \mathbb{R}^K$ represents the noise, e.g. drawn from the Gaussian measure $\mathcal{N}(0, \Gamma)$.

Bayesian inverse problems [Stuart, 2010]

Problem: given noisy observation data for system output, to calibrate unknown input.

- X – (separable Banach) space for unknown input ;
- Y – (separable Banach) space for system output.

Given a forward operator (e.g. PDEs, system of ODEs, etc.)

$$G : X \to Y,$$

and a observation operator (a set of sensors, e.g. pointwise data, Gaussian average)

$$\mathcal{O} : Y \to \mathbb{R}^K,$$

with $K \in \mathbb{N}$. We define the map from unknown input to finite data

$$\mathcal{G} := \mathcal{O} \circ G : X \to \mathbb{R}^K.$$

The inverse problem: find $u \in X$ given the noisy observation

$$\delta = \mathcal{G}(u) + \eta,$$

where $\eta \in \mathbb{R}^K$ represents the noise, e.g. drawn from the Gaussian measure $\mathcal{N}(0, \Gamma)$.

Bayesian inverse problems [Stuart, 2010]

Problem: given noisy observation data for system output, to calibrate unknown input.

- X – (separable Banach) space for unknown input ;
- Y – (separable Banach) space for system output.

Given a forward operator (e.g. PDEs, system of ODEs, etc.)

$$G : X \to Y,$$

and a observation operator (a set of sensors, e.g. pointwise data, Gaussian average)

$$\mathcal{O} : Y \to \mathbb{R}^K,$$

with $K \in \mathbb{N}$. We define the map from unknown input to finite data

$$\mathcal{G} := \mathcal{O} \circ G : X \to \mathbb{R}^K.$$

The inverse problem: find $u \in X$ given the noisy observation

$$\delta = \mathcal{G}(u) + \eta,$$

where $\eta \in \mathbb{R}^K$ represents the noise, e.g. drawn from the Gaussian measure $\mathcal{N}(0, \Gamma)$.

Bayesian inverse problems [Stuart, 2010]

Bayesian approach: given data δ, to update the distribution of the unknown input u.

- Let u be a random variable with Lebesgue density $\rho_0(u)$;
- Assume the noise η is independent of u with Lebesgue density $\rho(\eta)$;
- So (u, δ) is a random variable with Lebesgue density $\rho(\delta - \mathcal{G}(u))\rho_0(u)$.

Bayes' theorem

Assume that the probability of δ is positive, i.e.

$$Z := \int_X \rho(\delta - \mathcal{G}(u))\rho_0(u)du > 0,$$

Then $u|\delta$ is a random variable with Lebesgue density ρ^δ given by

$$\underbrace{\rho^\delta(u)}_{\text{posterior density}} = \frac{1}{Z}\underbrace{\rho(\delta - \mathcal{G}(u))}_{\text{likelihood}}\underbrace{\rho_0(u)}_{\text{prior density}}. \qquad (1)$$

Given data δ and the prior density $\rho_0(u)$, to determine the posterior density $\rho^\delta(u)$.

Bayesian inverse problems [Stuart, 2010]

Bayesian approach: given data δ, to update the distribution of the unknown input u.

- Let u be a random variable with Lebesgue density $\rho_0(u)$;
- Assume the noise η is independent of u with Lebesgue density $\rho(\eta)$;
- So (u, δ) is a random variable with Lebesgue density $\rho(\delta - \mathcal{G}(u))\rho_0(u)$.

Bayes' theorem

Assume that the probability of δ is positive, i.e.

$$Z := \int_X \rho(\delta - \mathcal{G}(u))\rho_0(u)du > 0,$$

Then $u|\delta$ is a random variable with Lebesgue density ρ^δ given by

$$\underbrace{\rho^\delta(u)}_{\text{posterior density}} = \frac{1}{Z} \underbrace{\rho(\delta - \mathcal{G}(u))}_{\text{likelihood}} \underbrace{\rho_0(u)}_{\text{prior density}} . \quad (1)$$

Given data δ and the prior density $\rho_0(u)$, to determine the posterior density $\rho^\delta(u)$.

Bayesian inverse problems [Stuart, 2010]

Bayesian approach: given data δ, to update the distribution of the unknown input u.

- Let u be a random variable with Lebesgue density $\rho_0(u)$;
- Assume the noise η is independent of u with Lebesgue density $\rho(\eta)$;
- So (u, δ) is a random variable with Lebesgue density $\rho(\delta - \mathcal{G}(u))\rho_0(u)$.

Bayes' theorem

Assume that the probability of δ is positive, i.e.

$$Z := \int_X \rho(\delta - \mathcal{G}(u))\rho_0(u)du > 0,$$

Then $u|\delta$ is a random variable with Lebesgue density ρ^δ given by

$$\underbrace{\rho^\delta(u)}_{\text{posterior density}} = \frac{1}{Z}\underbrace{\rho(\delta - \mathcal{G}(u))}_{\text{likelihood}}\underbrace{\rho_0(u)}_{\text{prior density}}. \quad (1)$$

Given data δ and the prior density $\rho_0(u)$, to determine the posterior density $\rho^\delta(u)$.

Parametric representation of the unknown input u.

The input u admits parametric representation, e.g. with an affine structure

$$u(y) = \psi_0 + \sum_{j\in\mathbb{J}} y_j\psi_j, \quad \psi_0, \psi_j \in X, \quad y_j \sim \mathcal{U}(-1,1),$$

being $\mathbb{J}$ a finite or countably infinite set, i.e. $\mathbb{J} = \{1, \dots, J\}$ with $J \in \mathbb{N}$, or $\mathbb{J} = \mathbb{N}$.

Parametric problem: let $\mathcal{X}$ and $\mathcal{Y}$ be two reflexive Banach spaces with duals $\mathcal{X}', \mathcal{Y}'$; let $A : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ denote a bilinear form and $F : \mathcal{Y} \to \mathbb{R}$ a linear functional; we consider

$$\text{find } p(y) \in \mathcal{X} \text{ such that} \quad A(p(y), v; y) = F(v) \quad \forall v \in \mathcal{Y}, \tag{2}$$

where we assume that the bilinear form admits the affine structure

$$\textbf{Affine parametrization:} \quad A(w, v; y) = A_0(w, v) + \sum_{j\in\mathbb{J}} y_j A_j(w, v). \tag{3}$$

$$\textbf{inf-sup condition:} \quad \inf_{0\neq w\in\mathcal{X}} \sup_{0\neq v\in\mathcal{Y}} \frac{|A(w, v; y)|}{||w||_{\mathcal{X}} ||v||_{\mathcal{Y}}} = \beta(y).$$

diffusion problem, Stokes flow, linear elasticity, acoustic problem, electromagnetics, etc.

$$\textbf{Example:} \quad A_j(w, v) = \int_D \psi_j(x)\nabla w(x) \cdot \nabla v(x) d(x) \quad \forall w, v \in H_0^1(D), \ j \in \{0\} \cup \mathbb{J}.$$

Bayesian inverse problems: parametrization [Schwab and Stuart, 2012]

Parametric representation of the unknown input u.

The input u admits parametric representation, e.g. with an affine structure

$$u(y) = \psi_0 + \sum_{j \in \mathbb{J}} y_j \psi_j, \quad \psi_0, \psi_j \in X, \quad y_j \sim \mathcal{U}(-1,1),$$

being $\mathbb{J}$ a finite or countably infinite set, i.e. $\mathbb{J} = \{1, \dots, J\}$ with $J \in \mathbb{N}$, or $\mathbb{J} = \mathbb{N}$.

Parametric problem: let $\mathcal{X}$ and $\mathcal{Y}$ be two reflexive Banach spaces with duals $\mathcal{X}', \mathcal{Y}'$; let $A : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ denote a bilinear form and $F : \mathcal{Y} \to \mathbb{R}$ a linear functional; we consider

$$\text{find } p(y) \in \mathcal{X} \text{ such that } \quad A(p(y), v; y) = F(v) \quad \forall v \in \mathcal{Y}, \tag{2}$$

where we assume that the bilinear form admits the affine structure

$$\textbf{Affine parametrization:} \quad A(w, v; y) = A_0(w, v) + \sum_{j \in \mathbb{J}} y_j A_j(w, v). \tag{3}$$

$$\textbf{inf-sup condition:} \quad \inf_{0 \neq w \in \mathcal{X}} \sup_{0 \neq v \in \mathcal{Y}} \frac{|A(w, v; y)|}{||w||_{\mathcal{X}} ||v||_{\mathcal{Y}}} = \beta(y).$$

diffusion problem, Stokes flow, linear elasticity, acoustic problem, electromagnetics, etc.

Example: $A_j(w, v) = \int_D \psi_j(x) \nabla w(x) \cdot \nabla v(x) d(x) \quad \forall w, v \in H_0^1(D), \ j \in \{0\} \cup \mathbb{J}.$

Parametric representation of the unknown input u.

The input u admits parametric representation, e.g. with an affine structure

$$u(y) = \psi_0 + \sum_{j \in \mathbb{J}} y_j \psi_j, \quad \psi_0, \psi_j \in X, \quad y_j \sim \mathcal{U}(-1, 1),$$

being $\mathbb{J}$ a finite or countably infinite set, i.e. $\mathbb{J} = \{1, \dots, J\}$ with $J \in \mathbb{N}$, or $\mathbb{J} = \mathbb{N}$.

Parametric problem: let $\mathcal{X}$ and $\mathcal{Y}$ be two reflexive Banach spaces with duals $\mathcal{X}', \mathcal{Y}'$; let $A : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ denote a bilinear form and $F : \mathcal{Y} \to \mathbb{R}$ a linear functional; we consider

$$\text{find } p(y) \in \mathcal{X} \text{ such that} \quad A(p(y), v; y) = F(v) \quad \forall v \in \mathcal{Y}, \tag{2}$$

where we assume that the bilinear form admits the affine structure

Affine parametrization: $$A(w, v; y) = A_0(w, v) + \sum_{j \in \mathbb{J}} y_j A_j(w, v). \tag{3}$$

inf-sup condition: $$\inf_{0 \neq w \in \mathcal{X}} \sup_{0 \neq v \in \mathcal{Y}} \frac{|A(w, v; y)|}{||w||_{\mathcal{X}} ||v||_{\mathcal{Y}}} = \beta(y).$$

diffusion problem, Stokes flow, linear elasticity, acoustic problem, electromagnetics, etc.

Example: $$A_j(w, v) = \int_D \psi_j(x) \nabla w(x) \cdot \nabla v(x) d(x) \quad \forall w, v \in H_0^1(D), \ j \in \{0\} \cup \mathbb{J}.$$

Bayesian inverse problems: parametrization [Schwab and Stuart, 2012]

Let $U = [-1,1]^{\mathbb{J}}$ and $\mathcal{B}$ be the σ-algebra on U. We equip $(U,\mathcal{B})$ with the prior measure

$$\mu_0(dy) = \bigotimes_{j\in\mathbb{J}} \frac{dy_j}{2}. \quad (4)$$

By Radon–Nikodym theorem, the posterior measure is given by

$$\frac{d\mu^\delta}{d\mu_0}(y) = \frac{1}{Z}\Theta(y), \quad (5)$$

where

$$\Theta(y) := \rho(\delta - \mathcal{O}(p(y))) \text{ and } Z := \mathbb{E}[\Theta] = \int_U \Theta(y)\mu_0(dy). \quad (6)$$

In the case $\eta \sim \mathcal{N}(0,\Gamma)$, we have

$$\Theta(y) = \frac{1}{\sqrt{(2\pi)^K|\Gamma|}} \exp\left(-\frac{1}{2}(\delta - \mathcal{O}(p(y)))^\top \Gamma^{-1}(\delta - \mathcal{O}(p(y)))\right).$$

Given the prior measure μ_0 and the data δ, to determine the posterior measure μ^δ.

Bayesian inverse problems: parametrization [Schwab and Stuart, 2012]

Let $U = [-1,1]^{\mathbb{J}}$ and $\mathcal{B}$ be the σ-algebra on U. We equip $(U, \mathcal{B})$ with the prior measure

$$\mu_0(dy) = \bigotimes_{j\in\mathbb{J}} \frac{dy_j}{2}. \tag{4}$$

By Radon–Nikodym theorem, the posterior measure is given by

$$\frac{d\mu^\delta}{d\mu_0}(y) = \frac{1}{Z}\Theta(y), \tag{5}$$

where

$$\Theta(y) := \rho(\delta - \mathcal{O}(p(y))) \text{ and } Z := \mathbb{E}[\Theta] = \int_U \Theta(y)\mu_0(dy). \tag{6}$$

In the case $\eta \sim \mathcal{N}(0,\Gamma)$, we have

$$\Theta(y) = \frac{1}{\sqrt{(2\pi)^K|\Gamma|}} \exp\left(-\frac{1}{2}(\delta - \mathcal{O}(p(y)))^\top \Gamma^{-1}(\delta - \mathcal{O}(p(y)))\right).$$

Given the prior measure μ_0 and the data δ, to determine the posterior measure μ^δ.

Bayesian inverse problems: parametrization [Schwab and Stuart, 2012]

Let $U = [-1, 1]^{\mathbb{J}}$ and $\mathcal{B}$ be the σ-algebra on U. We equip $(U, \mathcal{B})$ with the prior measure

$$\mu_0(dy) = \bigotimes_{j \in \mathbb{J}} \frac{dy_j}{2}. \tag{4}$$

By Radon–Nikodym theorem, the posterior measure is given by

$$\frac{d\mu^\delta}{d\mu_0}(y) = \frac{1}{Z}\Theta(y), \tag{5}$$

where

$$\Theta(y) := \rho(\delta - \mathcal{O}(p(y))) \text{ and } Z := \mathbb{E}[\Theta] = \int_U \Theta(y)\mu_0(dy). \tag{6}$$

In the case $\eta \sim \mathcal{N}(0, \Gamma)$, we have

$$\Theta(y) = \frac{1}{\sqrt{(2\pi)^K |\Gamma|}} \exp\left(-\frac{1}{2}(\delta - \mathcal{O}(p(y)))^\top \Gamma^{-1} (\delta - \mathcal{O}(p(y)))\right).$$

Given the prior measure μ_0 and the data δ, to determine the posterior measure μ^δ.

Computational quantities of interests (QoIs): 1. pointwise $\Theta(y)$ and 2. integration Z.

Computational requests

1. Given any $y \in U$, solve the parametric problem (2), and evaluate $\Theta(y)$ through (6).
2. Evaluate Z by some integration scheme, e.g. Monte Carlo, Gauss quadrature rule.

Computational challenges

- Curse-of-dimensionality: when the dimension $|\mathbb{J}|$ of the parameter space becomes **very high** or **infinite**, too many (millions or more) solutions are needed, e.g. MC.
- Large-scale computation: **one solution** is very expensive (taking hours by the fastest supercomputers), so only a few tens or hundreds of them are affordable.

Computational opportunities

- Sparsity: the dimensions are anisotropic and/or only have low mutual interaction.
- Reducibility: the solution/QoIs live in an intrinsically low-dimensional manifold.

Computational quantities of interests (QoIs): 1. pointwise $\Theta(y)$ and 2. integration Z.

Computational requests

1. Given any $y \in U$, solve the parametric problem (2), and evaluate $\Theta(y)$ through (6).
2. Evaluate Z by some integration scheme, e.g. Monte Carlo, Gauss quadrature rule.

Computational challenges

- Curse-of-dimensionality: when the dimension $|\mathbb{J}|$ of the parameter space becomes **very high** or **infinite**, too many (millions or more) solutions are needed, e.g. MC.
- Large-scale computation: **one solution** is very expensive (taking hours by the fastest supercomputers), so only a few tens or hundreds of them are affordable.

Computational opportunities

- Sparsity: the dimensions are anisotropic and/or only have low mutual interaction.
- Reducibility: the solution/QoIs live in an intrinsically low-dimensional manifold.

Bayesian inverse problems: computational aspects

Computational quantities of interests (QoIs): 1. pointwise $\Theta(y)$ and 2. integration Z.

Computational requests

1. Given any $y \in U$, solve the parametric problem (2), and evaluate $\Theta(y)$ through (6).
2. Evaluate Z by some integration scheme, e.g. Monte Carlo, Gauss quadrature rule.

Computational challenges

- Curse-of-dimensionality: when the dimension $|\mathbb{J}|$ of the parameter space becomes **very high** or **infinite**, too many (millions or more) solutions are needed, e.g. MC.
- Large-scale computation: **one solution** is very expensive (taking hours by the fastest supercomputers), so only a few tens or hundreds of them are affordable.

Computational opportunities

- Sparsity: the dimensions are anisotropic and/or only have low mutual interaction.
- Reducibility: the solution/QoIs live in an intrinsically low-dimensional manifold.

Bayesian inverse problems: computational aspects

Computational quantities of interests (QoIs): 1. pointwise $\Theta(y)$ and 2. integration Z.

Computational requests

1. Given any $y \in U$, solve the parametric problem (2), and evaluate $\Theta(y)$ through (6).
2. Evaluate Z by some integration scheme, e.g. Monte Carlo, Gauss quadrature rule.

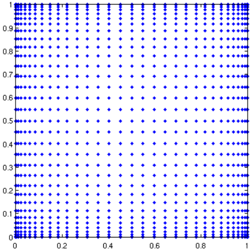
Computational challenges

- Curse-of-dimensionality: when the dimension $|\mathbb{J}|$ of the parameter space becomes **very high** or **infinite**, too many (millions or more) solutions are needed, e.g. MC.
- Large-scale computation: **one solution** is very expensive (taking hours by the fastest supercomputers), so only a few tens or hundreds of them are affordable.

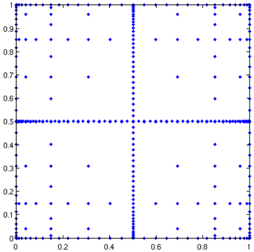
Computational opportunities

- Sparsity: the dimensions are anisotropic and/or only have low mutual interaction.
- Reducibility: the solution/QoIs live in an intrinsically low-dimensional manifold.

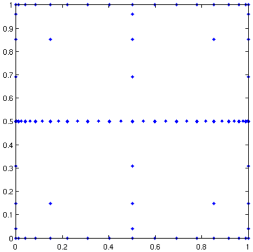
Sparsity: low mutual dimensional interaction and/or anisotropic property



tensor grid



sparse grid



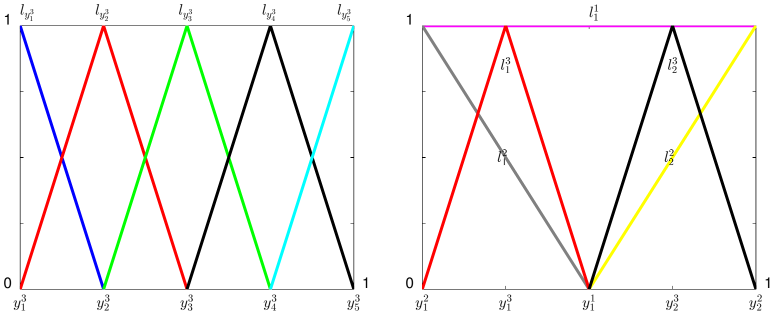
anisotropic sparse grid

Let $\mathcal{I}_q$ denote a univariate interpolation operator given by

$$\mathcal{I}_q g = \sum_{k=1}^{m(q)} g(y_k^q) l_{y_k^q}(y) \quad \text{vs} \quad \mathcal{I}_q g = \sum_{i=1}^{q} \triangle^i g \equiv \sum_{i=1}^{q} (\mathcal{I}_i - \mathcal{I}_{i-1}) g, \tag{7}$$

where q is grid level, $m(q)$ is $\#$ nodes, $m_\triangle^i$ is index set for additional nodes at level i.

$$\mathcal{I}_q g = \sum_{i=1}^{q} (\mathcal{I}_i g - \underbrace{\mathcal{I}_i \circ \mathcal{I}_{i-1}}_{\mathcal{I}_{i-1}} g) = \sum_{i=1}^{q} \sum_{k \in m_\triangle^i} \underbrace{(g(y_k^i) - \mathcal{I}_{i-1} g(y_k^i))}_{s_k^i} l_k^i(y), \tag{8}$$

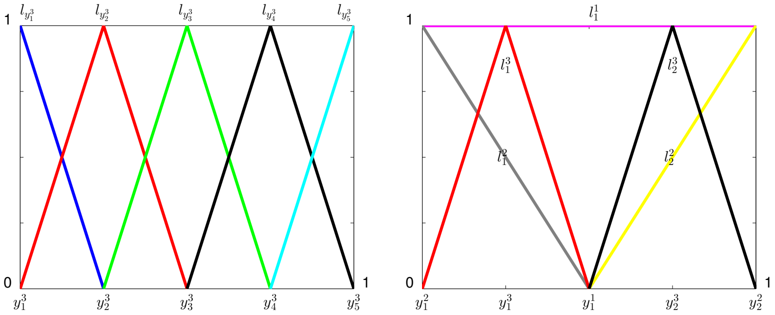


Let $\mathcal{I}_q$ denote a univariate interpolation operator given by

$$\mathcal{I}_q g = \sum_{k=1}^{m(q)} g(y_k^q) l_{y_k^q}(y) \quad \text{vs} \quad \mathcal{I}_q g = \sum_{i=1}^{q} \triangle^i g \equiv \sum_{i=1}^{q} (\mathcal{I}_i - \mathcal{I}_{i-1}) g, \tag{7}$$

where q is grid level, $m(q)$ is $\#$ nodes, $m_\triangle^i$ is index set for additional nodes at level i.

$$\mathcal{I}_q g = \sum_{i=1}^{q} (\mathcal{I}_i g - \underbrace{\mathcal{I}_i \circ \mathcal{I}_{i-1}}_{\mathcal{I}_{i-1}} g) = \sum_{i=1}^{q} \sum_{k \in m_\triangle^i} \underbrace{(g(y_k^i) - \mathcal{I}_{i-1} g(y_k^i))}_{s_k^i} l_k^i(y), \tag{8}$$



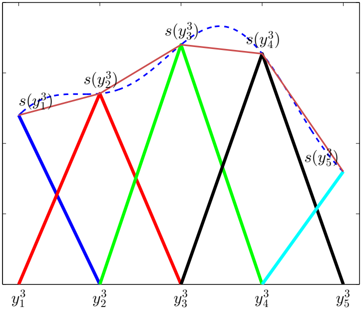
Sparse grid approximation: univariate hierarchical construction

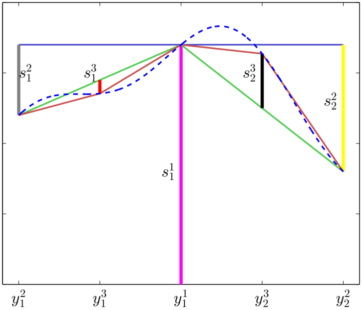
Let $\mathcal{I}_q$ denote a univariate interpolation operator given by

$$\mathcal{I}_q g = \sum_{k=1}^{m(q)} g(y_k^q) l_{y_k^q}(y) \quad \text{vs} \quad \mathcal{I}_q g = \sum_{i=1}^{q} \triangle^i g \equiv \sum_{i=1}^{q} (\mathcal{I}_i - \mathcal{I}_{i-1}) g, \tag{7}$$

where q is grid level, $m(q)$ is $\#$ nodes, $m_\triangle^i$ is index set for additional nodes at level i.

$$\mathcal{I}_q g = \sum_{i=1}^{q} (\mathcal{I}_i g - \underbrace{\mathcal{I}_i \circ \mathcal{I}_{i-1}}_{\mathcal{I}_{i-1}} g) = \sum_{i=1}^{q} \sum_{k \in m_\triangle^i} \underbrace{(g(y_k^i) - \mathcal{I}_{i-1} g(y_k^i))}_{s_k^i} l_k^i(y), \tag{8}$$

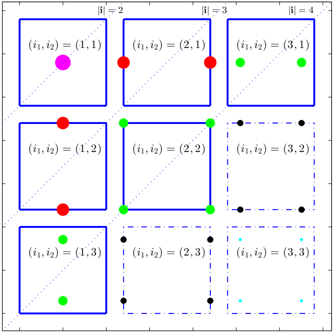


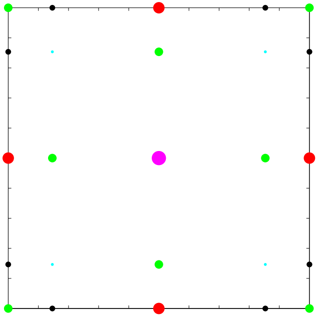


Sparse grid approximation: Smolyak sparse grid [Smolyak, 1963]

$$\mathcal{S}_q g = \sum_{|\boldsymbol{i}| \le q} \left(\triangle_1^{i_1} \otimes \cdots \otimes \triangle_J^{i_J} \right) g = \sum_{|\boldsymbol{i}|=J}^{q} \triangle \mathcal{S}_{|\boldsymbol{i}|} g(y),$$

$$\triangle \mathcal{S}_q g(y) = \sum_{|\boldsymbol{i}|=q} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} \underbrace{\left(g(y_{k_1}^{i_1}, \dots, y_{k_J}^{i_J}) - \mathcal{S}_{q-1} g(y_{k_1}^{i_1}, \dots, y_{k_J}^{i_J}) \right)}_{s_{\boldsymbol{k}}^{\boldsymbol{i}}} \underbrace{\left(l_{k_1}^{i_1}(y_1) \otimes \cdots \otimes l_{k_J}^{i_J}(y_J) \right)}_{l_{\boldsymbol{k}}^{\boldsymbol{i}}}.$$

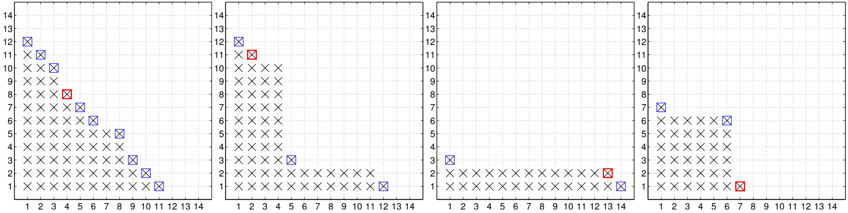




Hierarchical construction of Smolyak sparse grid.

$$\Lambda_s = \{\boldsymbol{i} \in \mathbb{N}_+^K : |\boldsymbol{i}| \leq q\} \to \Lambda_M = \{\boldsymbol{i} \in \mathbb{N}_+^K : \boldsymbol{i} \text{ admissible } \}.$$

Admissible: if $\boldsymbol{i} \in \Lambda_M$ then $\boldsymbol{i} - e_j \in \Lambda_M$ for any $j \in \mathbb{J}$.



Admissible set of indices for dimension adaptive sparse grid construction. Colored square: active index set $\mathscr{A}$; red square: the index to process in next step.

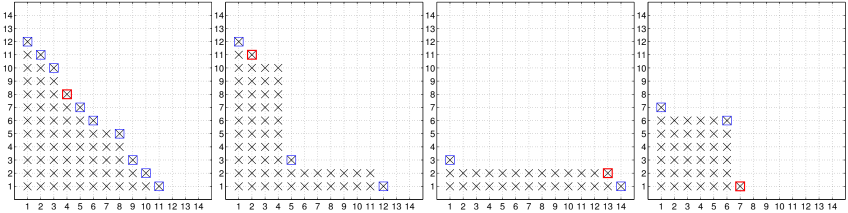
$$\mathcal{S}_{\Lambda_M} g(y) = \sum_{i \in \Lambda_M} \sum_{k \in m_\triangle^i} \underbrace{\left(g(y_k^i) - \mathcal{S}_{\Lambda_M \setminus \{i\}} g(y_k^i)\right)}_{s_k^i} l_k^i(y). \tag{9}$$

$$\mathbb{E}[g] \approx \mathbb{E}[\mathcal{S}_{\Lambda_M} g] = \sum_{i \in \Lambda_M} \sum_{k \in m_\triangle^i} s_k^i \mathbb{E}[l_k^i] = \sum_{i \in \Lambda_M} \sum_{k \in m_\triangle^i} s_k^i w_k^i. \tag{10}$$

Sparse grid approximation: adaptive SG [Gerstner and Griebel, 2003]

$$\Lambda_s = \{\boldsymbol{i} \in \mathbb{N}_+^K : |\boldsymbol{i}| \leq q\} \to \Lambda_M = \{\boldsymbol{i} \in \mathbb{N}_+^K : \boldsymbol{i} \text{ admissible } \}.$$

Admissible: if $\boldsymbol{i} \in \Lambda_M$ then $\boldsymbol{i} - e_j \in \Lambda_M$ for any $j \in \mathbb{J}$.



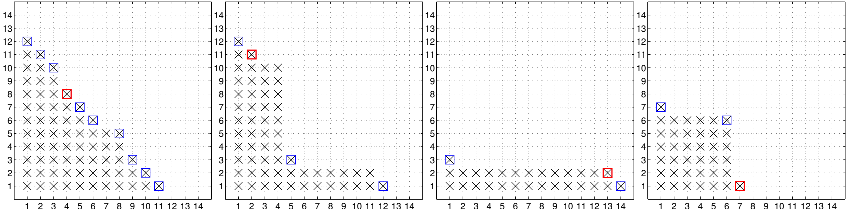
Admissible set of indices for dimension adaptive sparse grid construction.
Colored square: active index set $\mathscr{A}$; red square: the index to process in next step.

$$\mathcal{S}_{\Lambda_M} g(y) = \sum_{\boldsymbol{i} \in \Lambda_M} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} \underbrace{\left(g(y_{\boldsymbol{k}}^{\boldsymbol{i}}) - \mathcal{S}_{\Lambda_M \setminus \{\boldsymbol{i}\}} g(y_{\boldsymbol{k}}^{\boldsymbol{i}}) \right)}_{s_{\boldsymbol{k}}^{\boldsymbol{i}}} l_{\boldsymbol{k}}^{\boldsymbol{i}}(y). \quad (9)$$

$$\mathbb{E}[g] \approx \mathbb{E}[\mathcal{S}_{\Lambda_M} g] = \sum_{\boldsymbol{i} \in \Lambda_M} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} \mathbb{E}[l_{\boldsymbol{k}}^{\boldsymbol{i}}] = \sum_{\boldsymbol{i} \in \Lambda_M} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} w_{\boldsymbol{k}}^{\boldsymbol{i}}. \quad (10)$$

$$\Lambda_s = \{\boldsymbol{i} \in \mathbb{N}_+^K : |\boldsymbol{i}| \leq q\} \to \Lambda_M = \{\boldsymbol{i} \in \mathbb{N}_+^K : \boldsymbol{i} \text{ admissible } \}.$$

Admissible: if $\boldsymbol{i} \in \Lambda_M$ then $\boldsymbol{i} - e_j \in \Lambda_M$ for any $j \in \mathbb{J}$.



Admissible set of indices for dimension adaptive sparse grid construction.
Colored square: active index set $\mathscr{A}$; red square: the index to process in next step.

$$\mathcal{S}_{\Lambda_M} g(y) = \sum_{\boldsymbol{i} \in \Lambda_M} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} \underbrace{\left(g(y_{\boldsymbol{k}}^{\boldsymbol{i}}) - \mathcal{S}_{\Lambda_M \setminus \{\boldsymbol{i}\}} g(y_{\boldsymbol{k}}^{\boldsymbol{i}}) \right)}_{s_{\boldsymbol{k}}^{\boldsymbol{i}}} l_{\boldsymbol{k}}^{\boldsymbol{i}}(y). \quad (9)$$

$$\mathbb{E}[g] \approx \mathbb{E}[\mathcal{S}_{\Lambda_M} g] = \sum_{\boldsymbol{i} \in \Lambda_M} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} \mathbb{E}[l_{\boldsymbol{k}}^{\boldsymbol{i}}] = \sum_{\boldsymbol{i} \in \Lambda_M} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} w_{\boldsymbol{k}}^{\boldsymbol{i}}. \quad (10)$$

Sparse grid approximation: error indicators and estimators

Interpolation error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_i(\boldsymbol{i}'), \text{ with } \mathcal{E}_i(\boldsymbol{i}') = \frac{1}{|m_\triangle^{\boldsymbol{i}'}|} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}'}} |s_{\boldsymbol{k}}^{\boldsymbol{i}'}|.$$

Integration error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_e(\boldsymbol{i}'), \text{ with } \mathcal{E}_e(\boldsymbol{i}') = \frac{1}{|m_\triangle^{\boldsymbol{i}'}|} \left| \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}'}} s_{\boldsymbol{k}}^{\boldsymbol{i}'} w_{\boldsymbol{k}}^{\boldsymbol{i}'} \right|.$$

Interpolation and integration error estimators

$$\mathcal{E}_i(\mathscr{A}) = \max_{\boldsymbol{i} \in \mathscr{A}} \max_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} |s_{\boldsymbol{k}}^{\boldsymbol{i}}| \text{ and } \mathcal{E}_e(\mathscr{A}) = \left| \sum_{\boldsymbol{i} \in \mathscr{A}} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} w_{\boldsymbol{k}}^{\boldsymbol{i}} \right|.$$

Verification algorithm for stagnation problem [Chen and Quarteroni, 2014].

Sparse grid approximation: error indicators and estimators

Interpolation error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_i(\boldsymbol{i}'), \text{ with } \mathcal{E}_i(\boldsymbol{i}') = \frac{1}{|m_\triangle^{\boldsymbol{i}'}|} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}'}} |s_{\boldsymbol{k}}^{\boldsymbol{i}'}|.$$

Integration error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_e(\boldsymbol{i}'), \text{ with } \mathcal{E}_e(\boldsymbol{i}') = \frac{1}{|m_\triangle^{\boldsymbol{i}'}|} \left| \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}'}} s_{\boldsymbol{k}}^{\boldsymbol{i}'} w_{\boldsymbol{k}}^{\boldsymbol{i}'} \right|.$$

Interpolation and integration error estimators

$$\mathcal{E}_i(\mathscr{A}) = \max_{\boldsymbol{i} \in \mathscr{A}} \max_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} |s_{\boldsymbol{k}}^{\boldsymbol{i}}| \text{ and } \mathcal{E}_e(\mathscr{A}) = \left| \sum_{\boldsymbol{i} \in \mathscr{A}} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} w_{\boldsymbol{k}}^{\boldsymbol{i}} \right|.$$

Verification algorithm for stagnation problem [Chen and Quarteroni, 2014].

Sparse grid approximation: error indicators and estimators

Interpolation error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_i(\boldsymbol{i}'), \text{ with } \mathcal{E}_i(\boldsymbol{i}') = \frac{1}{|m_\triangle^{\boldsymbol{i}'}|} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}'}} |s_{\boldsymbol{k}}^{\boldsymbol{i}'}|.$$

Integration error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_e(\boldsymbol{i}'), \text{ with } \mathcal{E}_e(\boldsymbol{i}') = \frac{1}{|m_\triangle^{\boldsymbol{i}'}|} \left| \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}'}} s_{\boldsymbol{k}}^{\boldsymbol{i}'} w_{\boldsymbol{k}}^{\boldsymbol{i}'} \right|.$$

Interpolation and integration error estimators

$$\mathcal{E}_i(\mathscr{A}) = \max_{\boldsymbol{i} \in \mathscr{A}} \max_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} |s_{\boldsymbol{k}}^{\boldsymbol{i}}| \text{ and } \mathcal{E}_e(\mathscr{A}) = \left| \sum_{\boldsymbol{i} \in \mathscr{A}} \sum_{\boldsymbol{k} \in m_\triangle^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} w_{\boldsymbol{k}}^{\boldsymbol{i}} \right|.$$

Verification algorithm for stagnation problem [Chen and Quarteroni, 2014].

Sparse grid approximation: error indicators and estimators

Interpolation error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_i(\boldsymbol{i}'), \text{ with } \mathcal{E}_i(\boldsymbol{i}') = \frac{1}{|m_{\triangle}^{\boldsymbol{i}'}|} \sum_{\boldsymbol{k} \in m_{\triangle}^{\boldsymbol{i}'}} |s_{\boldsymbol{k}}^{\boldsymbol{i}'}|.$$

Integration error indicator

$$\boldsymbol{i} = \operatorname*{argmax}_{\boldsymbol{i}' \in \mathscr{A}} \mathcal{E}_e(\boldsymbol{i}'), \text{ with } \mathcal{E}_e(\boldsymbol{i}') = \frac{1}{|m_{\triangle}^{\boldsymbol{i}'}|} \left| \sum_{\boldsymbol{k} \in m_{\triangle}^{\boldsymbol{i}'}} s_{\boldsymbol{k}}^{\boldsymbol{i}'} w_{\boldsymbol{k}}^{\boldsymbol{i}'} \right|.$$

Interpolation and integration error estimators

$$\mathcal{E}_i(\mathscr{A}) = \max_{\boldsymbol{i} \in \mathscr{A}} \max_{\boldsymbol{k} \in m_{\triangle}^{\boldsymbol{i}}} |s_{\boldsymbol{k}}^{\boldsymbol{i}}| \text{ and } \mathcal{E}_e(\mathscr{A}) = \left| \sum_{\boldsymbol{i} \in \mathscr{A}} \sum_{\boldsymbol{k} \in m_{\triangle}^{\boldsymbol{i}}} s_{\boldsymbol{k}}^{\boldsymbol{i}} w_{\boldsymbol{k}}^{\boldsymbol{i}} \right|.$$

Verification algorithm for stagnation problem [Chen and Quarteroni, 2014].

High-fidelity approximation: large-scale computation

- High-fidelity approximation spaces: $\mathcal{X}_h \subset \mathcal{X}$ and $\mathcal{Y}_h \subset \mathcal{Y}$;
- Let $(w_h^n)_{n=1}^{\mathcal{N}}$ and $(v_h^n)_{n=1}^{\mathcal{N}}$ denote the bases of $\mathcal{X}_h$ and $\mathcal{Y}_h$;

The high-fidelity solution $p_h(y)$ can be expanded on the bases $(w_h^n)_{n=1}^{\mathcal{N}}$ as

$$p_h(y) = \sum_{n=1}^{\mathcal{N}} p_h^n(y) w_h^n, \quad (11)$$

with $\mathbf{p}_h(y) = (p_h^1(y), \dots, p_h^{\mathcal{N}}(y))^\top$. The high-fidelity (Petrov)-Galerkin approximation

$$\text{given any } y \in U, \text{ find } p_h(y) \in \mathcal{X}_h \text{ such that } A(p_h(y), v_h; y) = F(v_h) \quad \forall v_h \in \mathcal{Y}_h. \quad (12)$$

Let $(\mathbb{A}_h^j)_{nn'} := A_j(w_h^n, v_h^{n'})$ $1 \le n, n' \le \mathcal{N}$, $\mathbf{f}_h = (F(v_h^1), \dots, F(v_h^{\mathcal{N}}))^\top$, then

$$\text{given any } y \in U, \text{ find } \mathbf{p}_h(y) \in \mathbb{R}^{\mathcal{N}} \text{ such that } \left(\mathbb{A}_h^0 + \sum_{j \in \mathbb{J}} y_j \mathbb{A}_h^j \right) \mathbf{p}_h(y) = \mathbf{f}_h, \quad (13)$$

which is a $\mathcal{N} \times \mathcal{N}$ system, requiring large-scale computation when $\mathcal{N}$ is very large.

High-fidelity approximation: large-scale computation

- High-fidelity approximation spaces: $\mathcal{X}_h \subset \mathcal{X}$ and $\mathcal{Y}_h \subset \mathcal{Y}$;
- Let $(w_h^n)_{n=1}^{\mathcal{N}}$ and $(v_h^n)_{n=1}^{\mathcal{N}}$ denote the bases of $\mathcal{X}_h$ and $\mathcal{Y}_h$;

The high-fidelity solution $p_h(y)$ can be expanded on the bases $(w_h^n)_{n=1}^{\mathcal{N}}$ as

$$p_h(y) = \sum_{n=1}^{\mathcal{N}} p_h^n(y) w_h^n, \tag{11}$$

with $\mathbf{p}_h(y) = (p_h^1(y), \dots, p_h^{\mathcal{N}}(y))^\top$. The high-fidelity (Petrov)-Galerkin approximation

$$\text{given any } y \in U, \text{ find } p_h(y) \in \mathcal{X}_h \text{ such that } A(p_h(y), v_h; y) = F(v_h) \quad \forall v_h \in \mathcal{Y}_h. \tag{12}$$

Let $(\mathbb{A}_h^j)_{nn'} := A_j(w_h^n, v_h^{n'})$ $1 \leq n, n' \leq \mathcal{N}$, $\mathbf{f}_h = (F(v_h^1), \dots, F(v_h^{\mathcal{N}}))^\top$, then

$$\text{given any } y \in U, \text{ find } \mathbf{p}_h(y) \in \mathbb{R}^{\mathcal{N}} \text{ such that } \left(\mathbb{A}_h^0 + \sum_{j \in \mathbb{J}} y_j \mathbb{A}_h^j\right) \mathbf{p}_h(y) = \mathbf{f}_h, \tag{13}$$

which is a $\mathcal{N} \times \mathcal{N}$ system, requiring large-scale computation when $\mathcal{N}$ is very large.

High-fidelity approximation: large-scale computation

- High-fidelity approximation spaces: $\mathcal{X}_h \subset \mathcal{X}$ and $\mathcal{Y}_h \subset \mathcal{Y}$;
- Let $(w_h^n)_{n=1}^{\mathcal{N}}$ and $(v_h^n)_{n=1}^{\mathcal{N}}$ denote the bases of $\mathcal{X}_h$ and $\mathcal{Y}_h$;

The high-fidelity solution $p_h(y)$ can be expanded on the bases $(w_h^n)_{n=1}^{\mathcal{N}}$ as

$$p_h(y) = \sum_{n=1}^{\mathcal{N}} p_h^n(y) w_h^n, \quad (11)$$

with $\mathbf{p}_h(y) = (p_h^1(y), \dots, p_h^{\mathcal{N}}(y))^\top$. The high-fidelity (Petrov)-Galerkin approximation

$$\text{given any } y \in U, \text{ find } p_h(y) \in \mathcal{X}_h \text{ such that } A(p_h(y), v_h; y) = F(v_h) \quad \forall v_h \in \mathcal{Y}_h. \quad (12)$$

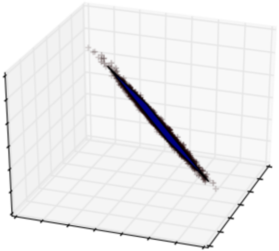
Let $(\mathbb{A}_h^j)_{nn'} := A_j(w_h^n, v_h^{n'})$ $1 \leq n, n' \leq \mathcal{N}$, $\mathbf{f}_h = (F(v_h^1), \dots, F(v_h^{\mathcal{N}}))^\top$, then

$$\text{given any } y \in U, \text{ find } \mathbf{p}_h(y) \in \mathbb{R}^{\mathcal{N}} \text{ such that } \left(\mathbb{A}_h^0 + \sum_{j \in \mathbb{J}} y_j \mathbb{A}_h^j\right) \mathbf{p}_h(y) = \mathbf{f}_h, \quad (13)$$

which is a $\mathcal{N} \times \mathcal{N}$ system, requiring large-scale computation when $\mathcal{N}$ is very large.

Reduced basis approximation: low-dimensional manifold

Reducibility: the solution manifold $\mathcal{M} = \{p_h(y) \in \mathcal{X}_h, y \in U\}$ is low-dimensional.



Mathematically, the best approximation error decays very fast

Kolmogorov N-width: $d_N(\mathcal{X}_h, \mathcal{M}) := \inf_{\mathcal{Z}_N \subset \mathcal{X}_h} \sup_{v \in \mathcal{M}} \inf_{w \in \mathcal{Z}_N} ||v - w||_{\mathcal{X}} \equiv \inf_{\mathcal{Z}_N \subset \mathcal{X}_h} \text{dist}(\mathcal{Z}_N, \mathcal{M}).$

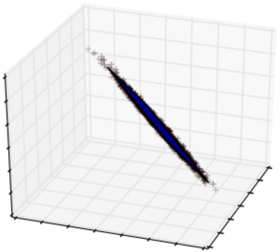
Look for a low-dimensional reduced basis space $\mathcal{X}_N \subset \mathcal{M}$ such that

reduced basis error: $\sigma_N(\mathcal{X}_N, \mathcal{M}) := \sup_{v \in \mathcal{M}} \inf_{w \in \mathcal{X}_N} ||v - w||_{\mathcal{X}} \equiv \text{dist}(\mathcal{X}_N, \mathcal{M}),$

converges with rate not far from (ideally achieves) that of the best approximation error.

Reduced basis approximation: low-dimensional manifold

Reducibility: the solution manifold $\mathcal{M} = \{p_h(y) \in \mathcal{X}_h, y \in U\}$ is low-dimensional.



Mathematically, the best approximation error decays very fast

Kolmogorov N-width: $d_N(\mathcal{X}_h, \mathcal{M}) := \inf_{\mathcal{Z}_N \subset \mathcal{X}_h} \sup_{v \in \mathcal{M}} \inf_{w \in \mathcal{Z}_N} ||v - w||_{\mathcal{X}} \equiv \inf_{\mathcal{Z}_N \subset \mathcal{X}_h} \text{dist}(\mathcal{Z}_N, \mathcal{M}).$

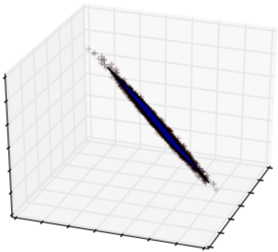
Look for a low-dimensional reduced basis space $\mathcal{X}_N \subset \mathcal{M}$ such that

reduced basis error: $\sigma_N(\mathcal{X}_N, \mathcal{M}) := \sup_{v \in \mathcal{M}} \inf_{w \in \mathcal{X}_N} ||v - w||_{\mathcal{X}} \equiv \text{dist}(\mathcal{X}_N, \mathcal{M}),$

converges with rate not far from (ideally achieves) that of the best approximation error.

Reduced basis approximation: low-dimensional manifold

Reducibility: the solution manifold $\mathcal{M} = \{p_h(y) \in \mathcal{X}_h, y \in U\}$ is low-dimensional.



Mathematically, the best approximation error decays very fast

Kolmogorov N-width: $d_N(\mathcal{X}_h, \mathcal{M}) := \inf_{\mathcal{Z}_N \subset \mathcal{X}_h} \sup_{v \in \mathcal{M}} \inf_{w \in \mathcal{Z}_N} ||v - w||_{\mathcal{X}} \equiv \inf_{\mathcal{Z}_N \subset \mathcal{X}_h} \text{dist}(\mathcal{Z}_N, \mathcal{M}).$

Look for a low-dimensional reduced basis space $\mathcal{X}_N \subset \mathcal{M}$ such that

reduced basis error: $\sigma_N(\mathcal{X}_N, \mathcal{M}) := \sup_{v \in \mathcal{M}} \inf_{w \in \mathcal{X}_N} ||v - w||_{\mathcal{X}} \equiv \text{dist}(\mathcal{X}_N, \mathcal{M}),$

converges with rate not far from (ideally achieves) that of the best approximation error.

Reduced basis approximation: reduction [Patera and Rozza, 2007]

- Reduced basis approximation spaces: $\mathcal{X}_N \subset \mathcal{X}_h$ and $\mathcal{Y}_N \subset \mathcal{Y}_h$;
- Let $(w_N^n)_{n=1}^N$ and $(v_N^n)_{n=1}^N$ denote the bases of $\mathcal{X}_N$ and $\mathcal{Y}_N$;

The reduced solution $p_h(y)$ can be expanded on the bases $(w_N^n)_{n=1}^N$ as

$$p_N(y) = \sum_{n=1}^{N} p_N^n(y) w_N^n, \tag{14}$$

with $\mathbf{p}_N(y) = (p_N^1(y), \dots, p_N^N(y))^\top$. The reduced basis (Petrov)-Galerkin approximation

$$\text{given any } y \in U, \text{ find } p_N(y) \in \mathcal{X}_N \text{ such that } A(p_N(y), v_N; y) = F(v_N) \quad \forall v_N \in \mathcal{Y}_N. \tag{15}$$

Let $\mathbb{W} = (\mathbf{w}_N^1, \dots, \mathbf{w}_N^N)$ and $\mathbb{V} = (\mathbf{v}_N^1, \dots, \mathbf{v}_N^N)$, $\mathbb{A}_N^j = \mathbb{V}^\top \mathbb{A}_h^j \mathbb{W}$, $j \in \{0\} \cup \mathbb{J}$; $\mathbf{f}_N = \mathbb{V}^\top \mathbf{f}_h$.

$$\text{given any } y \in U, \text{ find } \mathbf{p}_N(y) \in \mathbb{R}^N \text{ such that } \left(\mathbb{A}_N^0 + \sum_{j \in \mathbb{J}} y_j \mathbb{A}_N^j \right) \mathbf{p}_N(y) = \mathbf{f}_N. \tag{16}$$

which is a $N \times N$ system, needs small-scale computation as $N \ll \mathcal{N}$, e.g. $(10 \sim 100)$.

Reduced basis approximation: reduction [Patera and Rozza, 2007]

- Reduced basis approximation spaces: $\mathcal{X}_N \subset \mathcal{X}_h$ and $\mathcal{Y}_N \subset \mathcal{Y}_h$;
- Let $(w_N^n)_{n=1}^N$ and $(v_N^n)_{n=1}^N$ denote the bases of $\mathcal{X}_N$ and $\mathcal{Y}_N$;

The reduced solution $p_h(y)$ can be expanded on the bases $(w_N^n)_{n=1}^N$ as

$$p_N(y) = \sum_{n=1}^{N} p_N^n(y) w_N^n, \tag{14}$$

with $\mathbf{p}_N(y) = (p_N^1(y), \dots, p_N^N(y))^\top$. The reduced basis (Petrov)-Galerkin approximation

$$\text{given any } y \in U, \text{ find } p_N(y) \in \mathcal{X}_N \text{ such that } A(p_N(y), v_N; y) = F(v_N) \quad \forall v_N \in \mathcal{Y}_N. \tag{15}$$

Let $\mathbb{W} = (\mathbf{w}_N^1, \dots, \mathbf{w}_N^N)$ and $\mathbb{V} = (\mathbf{v}_N^1, \dots, \mathbf{v}_N^N)$, $\mathbb{A}_N^j = \mathbb{V}^\top \mathbb{A}_h^j \mathbb{W}$, $j \in \{0\} \cup \mathbb{J}$; $\mathbf{f}_N = \mathbb{V}^\top \mathbf{f}_h$.

$$\text{given any } y \in U, \text{ find } \mathbf{p}_N(y) \in \mathbb{R}^N \text{ such that } \left(\mathbb{A}_N^0 + \sum_{j \in \mathbb{J}} y_j \mathbb{A}_N^j \right) \mathbf{p}_N(y) = \mathbf{f}_N. \tag{16}$$

which is a $N \times N$ system, needs small-scale computation as $N \ll \mathcal{N}$, e.g. $(10 \sim 100)$.

Reduced basis approximation: reduction [Patera and Rozza, 2007]

- Reduced basis approximation spaces: $\mathcal{X}_N \subset \mathcal{X}_h$ and $\mathcal{Y}_N \subset \mathcal{Y}_h$;
- Let $(w_N^n)_{n=1}^N$ and $(v_N^n)_{n=1}^N$ denote the bases of $\mathcal{X}_N$ and $\mathcal{Y}_N$;

The reduced solution $p_h(y)$ can be expanded on the bases $(w_N^n)_{n=1}^N$ as

$$p_N(y) = \sum_{n=1}^{N} p_N^n(y) w_N^n, \tag{14}$$

with $\mathbf{p}_N(y) = (p_N^1(y), \dots, p_N^N(y))^\top$. The reduced basis (Petrov)-Galerkin approximation

$$\text{given any } y \in U, \text{ find } p_N(y) \in \mathcal{X}_N \text{ such that } A(p_N(y), v_N; y) = F(v_N) \quad \forall v_N \in \mathcal{Y}_N. \tag{15}$$

Let $\mathbb{W} = (\mathbf{w}_N^1, \dots, \mathbf{w}_N^N)$ and $\mathbb{V} = (\mathbf{v}_N^1, \dots, \mathbf{v}_N^N)$, $\mathbb{A}_N^j = \mathbb{V}^\top \mathbb{A}_h^j \mathbb{W}$, $j \in \{0\} \cup \mathbb{J}$; $\mathbf{f}_N = \mathbb{V}^\top \mathbf{f}_h$.

$$\text{given any } y \in U, \text{ find } \mathbf{p}_N(y) \in \mathbb{R}^N \text{ such that } \left(\mathbb{A}_N^0 + \sum_{j \in \mathbb{J}} y_j \mathbb{A}_N^j \right) \mathbf{p}_N(y) = \mathbf{f}_N. \tag{16}$$

which is a $N \times N$ system, needs small-scale computation as $N \ll \mathcal{N}$, e.g. $(10 \sim 100)$.

Reduced basis approximation: construction of reduced spaces

Greedy algorithm [Patera and Rozza, 2007]

Initialize $\mathcal{X}_1 = \text{span}\{p_h(y^{(1)})\}$ at some random sample $y^{(1)}$, then for $N = 2, 3, \dots,$

$$y^{(N)} = \underset{y \in U}{\text{argsup}} \, ||p_h(y) - p_{N-1}(y)||_{\mathcal{X}} \quad \text{or} \quad y^{(N)} = \underset{y \in U}{\text{argsup}} \, |\Theta_h(y) - \Theta_{N-1}(y)| \tag{17}$$

and the reduced space $\mathcal{X}_N$ can be constructed by the snapshots

$$\mathcal{X}_N = \mathcal{X}_{N-1} \oplus \text{span}\{p_h(y^{(N)})\}. \tag{18}$$

Gram–Schmidt process → orthnormal bases $(w_N^n)_{n=1}^N$ of $\mathcal{X}_N$ for better stability of $\mathbb{A}_N(y)$.

- In case of symmetric coercive A, we can directly take $\mathcal{Y}_N = \mathcal{X}_N$;
- otherwise, we solve a 'supremizer' problem (to guarantee the inf-sup condition)

$$\text{given } y \in U, \text{find } v_N^n(y) \in \mathcal{Y}_h \text{ such that } (v_N^n(y), v_h)_{\mathcal{Y}_h} = A(w_N^n, v_h; y) \quad \forall v_h \in \mathcal{Y}_h, \tag{19}$$

Let $\mathbb{A}_N^{j,j'} = (\mathbb{A}_h^j \mathbb{W})^\top \mathbb{M}_h^{-1} \mathbb{A}_h^{j'} \mathbb{W}$, and $\mathbf{f}_N^j = (\mathbb{A}_h^j \mathbb{W})^\top \mathbb{M}_h^{-1} \mathbf{f}_h$; we solve the $N \times N$ system

$$\left(\mathbb{A}_N^{0,0} + 2 \sum_{j \in \mathbb{J}} y_j \mathbb{A}_N^{0,j} + \sum_{j \in \mathbb{J}} \sum_{j' \in \mathbb{J}} y_j y_{j'} \mathbb{A}_N^{j,j'} \right) \mathbf{p}_N(y) = \mathbf{f}_N^0 + \sum_{j \in \mathbb{J}} y_j \mathbf{f}_N^j. \tag{20}$$

Greedy algorithm [Patera and Rozza, 2007]

Initialize $\mathcal{X}_1 = \text{span}\{p_h(y^{(1)})\}$ at some random sample $y^{(1)}$, then for $N = 2, 3, \dots,$

$$y^{(N)} = \operatorname*{argsup}_{y \in U} ||p_h(y) - p_{N-1}(y)||_{\mathcal{X}} \quad \text{or} \quad y^{(N)} = \operatorname*{argsup}_{y \in U} |\Theta_h(y) - \Theta_{N-1}(y)| \tag{17}$$

and the reduced space $\mathcal{X}_N$ can be constructed by the snapshots

$$\mathcal{X}_N = \mathcal{X}_{N-1} \oplus \text{span}\{p_h(y^{(N)})\}. \tag{18}$$

Gram–Schmidt process → orthnormal bases $(w_N^n)_{n=1}^N$ of $\mathcal{X}_N$ for better stability of $\mathbb{A}_N(y)$.

- In case of symmetric coercive A, we can directly take $\mathcal{Y}_N = \mathcal{X}_N$;
- otherwise, we solve a 'supremizer' problem (to guarantee the inf-sup condition)

$$\text{given } y \in U, \text{find } v_N^n(y) \in \mathcal{Y}_h \text{ such that } (v_N^n(y), v_h)_{\mathcal{Y}_h} = A(w_N^n, v_h; y) \quad \forall v_h \in \mathcal{Y}_h, \tag{19}$$

Let $\mathbb{A}_N^{j,j'} = (\mathbb{A}_h^j \mathbb{W})^\top \mathbb{M}_h^{-1} \mathbb{A}_h^{j'} \mathbb{W}$, and $\mathbf{f}_N^j = (\mathbb{A}_h^j \mathbb{W})^\top \mathbb{M}_h^{-1} \mathbf{f}_h$; we solve the $N \times N$ system

$$\left(\mathbb{A}_N^{0,0} + 2 \sum_{j \in \mathbb{J}} y_j \mathbb{A}_N^{0,j} + \sum_{j \in \mathbb{J}} \sum_{j' \in \mathbb{J}} y_j y_{j'} \mathbb{A}_N^{j,j'} \right) \mathbf{p}_N(y) = \mathbf{f}_N^0 + \sum_{j \in \mathbb{J}} y_j \mathbf{f}_N^j. \tag{20}$$

Reduced basis approximation: construction of reduced spaces

Greedy algorithm [Patera and Rozza, 2007]

Initialize $\mathcal{X}_1 = \text{span}\{p_h(y^{(1)})\}$ at some random sample $y^{(1)}$, then for $N = 2, 3, \dots,$

$$y^{(N)} = \underset{y \in U}{\text{argsup}} \, ||p_h(y) - p_{N-1}(y)||_{\mathcal{X}} \quad \text{or} \quad y^{(N)} = \underset{y \in U}{\text{argsup}} \, |\Theta_h(y) - \Theta_{N-1}(y)| \tag{17}$$

and the reduced space $\mathcal{X}_N$ can be constructed by the snapshots

$$\mathcal{X}_N = \mathcal{X}_{N-1} \oplus \text{span}\{p_h(y^{(N)})\}. \tag{18}$$

Gram–Schmidt process → orthnormal bases $(w_N^n)_{n=1}^N$ of $\mathcal{X}_N$ for better stability of $\mathbb{A}_N(y)$.

- In case of symmetric coercive A, we can directly take $\mathcal{Y}_N = \mathcal{X}_N$;
- otherwise, we solve a 'supremizer' problem (to guarantee the inf-sup condition)

$$\text{given } y \in U, \text{find } v_N^n(y) \in \mathcal{Y}_h \text{ such that } (v_N^n(y), v_h)_{\mathcal{Y}_h} = A(w_N^n, v_h; y) \quad \forall v_h \in \mathcal{Y}_h, \tag{19}$$

Let $\mathbb{A}_N^{j,j'} = (\mathbb{A}_h^j \mathbb{W})^\top \mathbb{M}_h^{-1} \mathbb{A}_h^{j'} \mathbb{W}$, and $\mathbf{f}_N^j = (\mathbb{A}_h^j \mathbb{W})^\top \mathbb{M}_h^{-1} \mathbf{f}_h$; we solve the $N \times N$ system

$$\left(\mathbb{A}_N^{0,0} + 2 \sum_{j \in \mathbb{J}} y_j \mathbb{A}_N^{0,j} + \sum_{j \in \mathbb{J}} \sum_{j' \in \mathbb{J}} y_j y_{j'} \mathbb{A}_N^{j,j'} \right) \mathbf{p}_N(y) = \mathbf{f}_N^0 + \sum_{j \in \mathbb{J}} y_j \mathbf{f}_N^j. \tag{20}$$

We consider the error estimator for the nonlinear, nonaffine QoI. Recall by definition

$$\Theta_h(y) = \frac{1}{\sqrt{(2\pi)^K|\Gamma|}} \exp\left(-\frac{1}{2}(\delta - \mathcal{O}(p_h(y)))^\top \Gamma^{-1}(\delta - \mathcal{O}(p_h(y)))\right),$$

which can be expanded as

$$\Theta_h(y) = \Theta_N(y) + \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)} (p_h(y) - p_N(y)) + O(||p_h(y) - p_N(y)||^2_{\mathcal{X}}). \tag{21}$$

We can estimate the error by

$$|\Theta_h(y) - \Theta_N(y)| \approx \left|\frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)} (p_h(y) - p_N(y))\right| \leq \left|\left|\frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}\right|\right|_{\mathcal{X}'} ||p_h(y) - p_N(y)||_{\mathcal{X}} =: \triangle_N^{(1)}(y).$$

Here, the reduced solution error can be bounded by

$$||p_h(y) - p_N(y)||_{\mathcal{X}} \leq \frac{||R_h(\cdot; y)||_{\mathcal{Y}'}}{\beta_h(y)} =: \triangle_N^p(y),$$

where the residual $R_h(\cdot; y) \in \mathcal{Y}'$, defined as

$$R_h(v_h; y) = F(v_h) - A(p_N(y), v_h; y) \quad \forall v_h \in \mathcal{Y}_h,$$

and the inf-sup constant $\beta_h(y)$ is uniformly bounded from below by β_h^{LB}.

Reduced basis approximation: a posteriori error estimators I

We consider the error estimator for the nonlinear, nonaffine QoI. Recall by definition

$$\Theta_h(y) = \frac{1}{\sqrt{(2\pi)^K|\Gamma|}} \exp\left(-\frac{1}{2}(\delta - \mathcal{O}(p_h(y)))^\top \Gamma^{-1}(\delta - \mathcal{O}(p_h(y)))\right),$$

which can be expanded as

$$\Theta_h(y) = \Theta_N(y) + \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)} (p_h(y) - p_N(y)) + O(||p_h(y) - p_N(y)||^2_{\mathcal{X}}). \quad (21)$$

We can estimate the error by

$$|\Theta_h(y) - \Theta_N(y)| \approx \left|\frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)} (p_h(y) - p_N(y))\right| \leq \left|\left|\frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}\right|\right|_{\mathcal{X}'} ||p_h(y) - p_N(y)||_{\mathcal{X}} =: \triangle_N^{(1)}(y).$$

Here, the reduced solution error can be bounded by

$$||p_h(y) - p_N(y)||_{\mathcal{X}} \leq \frac{||R_h(\cdot; y)||_{\mathcal{Y}'}}{\beta_h(y)} =: \triangle_N^p(y),$$

where the residual $R_h(\cdot; y) \in \mathcal{Y}'$, defined as

$$R_h(v_h; y) = F(v_h) - A(p_N(y), v_h; y) \quad \forall v_h \in \mathcal{Y}_h,$$

and the inf-sup constant $\beta_h(y)$ is uniformly bounded from below by β_h^{LB}.

Reduced basis approximation: a posteriori error estimators II

We consider a dual problem corresponding to the primal problem (12) reads as

$$\text{given any } y \in U, \text{ find } \psi_h(y) \in \mathcal{Y}_h \text{ such that } A(w_h, \psi_h; y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(w_h) \quad \forall w_h \in \mathcal{X}_h.$$

We may approximate this high-fidelity solution with a reduced dual solution by solving

$$\text{find } \psi_{N_{du}}(y) \in \mathcal{Y}_{N_{du}} \text{ such that } A(w^{du}_{N_{du}}, \psi_{N_{du}}; y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(w^{du}_{N_{du}}) \quad \forall w^{du}_{N_{du}} \in \mathcal{X}_{N_{du}}. \quad (22)$$

The second error estimator (dual-weighted residual) is simply defined as

$$\triangle_N^{(2)}(y) := R(\psi_{N_{du}}(y); y) = \mathbf{f}_h^\top \mathbb{W}_{du} \boldsymbol{\psi}_{N_{du}}(y) - \sum_{j \in \{0\} \cup \mathbb{J}} y_j(\mathbf{p}_N(y))^\top \mathbb{W}^\top \mathbb{A}_h^j \mathbb{W}_{du} \boldsymbol{\psi}_{N_{du}}(y) \quad (23)$$

A closer look at the residual (by Galerkin orthogonality):

$$R(\psi_h(y); y) = A(p_h(y) - p_N(y), \psi_h(y); y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(p_h(y) - p_N(y)), \quad (24)$$

which is nothing but the first term in the expansion. Moreover,

$$R(\psi_h(y); y) - \triangle_N^{(2)}(y) = R(e_h^{du}(y); y) = A(e_h(y), e_h^{du}(y); y) \le \gamma_h(y)||e_h(y)||_{\mathcal{X}}||e_h^{du}(y)||_{\mathcal{Y}}, \quad (25)$$

where we denote the reduced errors $e_h(y) = p_h(y) - p_N(y)$ and $e_h^{du}(y) = \psi_h(y) - \psi_{N_{du}}(y)$.

Reduced basis approximation: a posteriori error estimators II

We consider a dual problem corresponding to the primal problem (12) reads as

$$\text{given any } y \in U, \text{ find } \psi_h(y) \in \mathcal{Y}_h \text{ such that } A(w_h, \psi_h; y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(w_h) \quad \forall w_h \in \mathcal{X}_h.$$

We may approximate this high-fidelity solution with a reduced dual solution by solving

$$\text{find } \psi_{N_{du}}(y) \in \mathcal{Y}_{N_{du}} \text{ such that } A(w^{du}_{N_{du}}, \psi_{N_{du}}; y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(w^{du}_{N_{du}}) \quad \forall w^{du}_{N_{du}} \in \mathcal{X}_{N_{du}}. \tag{22}$$

The second error estimator (dual-weighted residual) is simply defined as

$$\triangle_N^{(2)}(y) := R(\psi_{N_{du}}(y); y) = \mathbf{f}_h^\top \mathbb{W}_{du} \boldsymbol{\psi}_{N_{du}}(y) - \sum_{j \in \{0\} \cup \mathbb{J}} y_j (\mathbf{p}_N(y))^\top \mathbb{W}^\top \mathbb{A}_h^j \mathbb{W}_{du} \boldsymbol{\psi}_{N_{du}}(y) \tag{23}$$

A closer look at the residual (by Galerkin orthogonality):

$$R(\psi_h(y); y) = A(p_h(y) - p_N(y), \psi_h(y); y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(p_h(y) - p_N(y)), \tag{24}$$

which is nothing but the first term in the expansion. Moreover,

$$R(\psi_h(y); y) - \triangle_N^{(2)}(y) = R(e_h^{du}(y); y) = A(e_h(y), e_h^{du}(y); y) \le \gamma_h(y) ||e_h(y)||_{\mathcal{X}} ||e_h^{du}(y)||_{\mathcal{Y}}, \tag{25}$$

where we denote the reduced errors $e_h(y) = p_h(y) - p_N(y)$ and $e_h^{du}(y) = \psi_h(y) - \psi_{N_{du}}(y)$.

Reduced basis approximation: a posteriori error estimators II

We consider a dual problem corresponding to the primal problem (12) reads as

$$\text{given any } y \in U, \text{ find } \psi_h(y) \in \mathcal{Y}_h \text{ such that } A(w_h, \psi_h; y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(w_h) \quad \forall w_h \in \mathcal{X}_h.$$

We may approximate this high-fidelity solution with a reduced dual solution by solving

$$\text{find } \psi_{N_{du}}(y) \in \mathcal{Y}_{N_{du}} \text{ such that } A(w^{du}_{N_{du}}, \psi_{N_{du}}; y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(w^{du}_{N_{du}}) \quad \forall w^{du}_{N_{du}} \in \mathcal{X}_{N_{du}}. \quad (22)$$

The second error estimator (dual-weighted residual) is simply defined as

$$\triangle_N^{(2)}(y) := R(\psi_{N_{du}}(y); y) = \mathbf{f}_h^\top \mathbb{W}_{du} \boldsymbol{\psi}_{N_{du}}(y) - \sum_{j \in \{0\} \cup \mathbb{J}} y_j (\mathbf{p}_N(y))^\top \mathbb{W}^\top \mathbb{A}_h^j \mathbb{W}_{du} \boldsymbol{\psi}_{N_{du}}(y) \quad (23)$$

A closer look at the residual (by Galerkin orthogonality):

$$R(\psi_h(y); y) = A(p_h(y) - p_N(y), \psi_h(y); y) = \frac{\partial \Theta_h}{\partial p_h}\Big|_{p_N(y)}(p_h(y) - p_N(y)), \quad (24)$$

which is nothing but the first term in the expansion. Moreover,

$$R(\psi_h(y); y) - \triangle_N^{(2)}(y) = R(e_h^{du}(y); y) = A(e_h(y), e_h^{du}(y); y) \le \gamma_h(y) ||e_h(y)||_{\mathcal{X}} ||e_h^{du}(y)||_{\mathcal{Y}}, \quad (25)$$

where we denote the reduced errors $e_h(y) = p_h(y) - p_N(y)$ and $e_h^{du}(y) = \psi_h(y) - \psi_{N_{du}}(y)$.

Reduced basis approximation: a posteriori error estimators III

We may propose the use of a (improved/corrected) reduced output

$$\Theta_N^c(y) = \Theta_N(y) + \triangle_N^{(2)}(y). \tag{26}$$

The last term in the expansion can be further expanded as

$$O(||p_h(y) - p_N(y)||^2_{\mathcal{X}}) = \frac{1}{2}\frac{\partial^2 \Theta_h}{\partial p_h^2}\Big|_{p_N(y)}(p_h(y) - p_N(y), p_h(y) - p_N(y)) + O(||p_h(y) - p_N(y)||^3_{\mathcal{X}}).$$

So that the third a posteriori error estimator can be given by

$$\triangle_N^{(3)}(y) := \max\left\{\triangle_N^{(4)}(y), \triangle_N^{(5)}(y)\right\}, \tag{27}$$

where (note $|\Theta_h(y) - \Theta_N^c(y)| \approx$ (first term $- \triangle_N^{(2)}$) $+$ second term)

$$\triangle_N^{(4)}(y) := \gamma_h(y)||e_h(y)||_{\mathcal{X}}||e_h^{du}(y)||_{\mathcal{Y}}, \text{ and } \triangle_N^{(5)}(y) := \gamma_h'(y)||e_h(y)||^2_{\mathcal{X}}. \tag{28}$$

Note that $\triangle_N^{(4)}$ and $\triangle_N^{(5)}$ exhibit a **quadratic** dependence on the reduced solution error.

The second (dual-weighted residual) error estimator is the cheapest to evaluate.

Reduced basis approximation: a posteriori error estimators III

We may propose the use of a (improved/corrected) reduced output

$$\Theta_N^c(y) = \Theta_N(y) + \triangle_N^{(2)}(y). \quad (26)$$

The last term in the expansion can be further expanded as

$$O(||p_h(y) - p_N(y)||_{\mathcal{X}}^2) = \frac{1}{2}\frac{\partial^2\Theta_h}{\partial p_h^2}\Big|_{p_N(y)}(p_h(y) - p_N(y), p_h(y) - p_N(y)) + O(||p_h(y) - p_N(y)||_{\mathcal{X}}^3).$$

So that the third a posteriori error estimator can be given by

$$\triangle_N^{(3)}(y) := \max\left\{\triangle_N^{(4)}(y), \triangle_N^{(5)}(y)\right\}, \quad (27)$$

where (note $|\Theta_h(y) - \Theta_N^c(y)| \approx$ (first term $-\triangle_N^{(2)}$) + second term)

$$\triangle_N^{(4)}(y) := \gamma_h(y)||e_h(y)||_{\mathcal{X}}||e_h^{du}(y)||_{\mathcal{Y}}, \text{ and } \triangle_N^{(5)}(y) := \gamma_h'(y)||e_h(y)||_{\mathcal{X}}^2. \quad (28)$$

Note that $\triangle_N^{(4)}$ and $\triangle_N^{(5)}$ exhibit a **quadratic** dependence on the reduced solution error.

The second (dual-weighted residual) error estimator is the cheapest to evaluate.

Reduced basis approximation: a posteriori error estimators III

We may propose the use of a (improved/corrected) reduced output

$$\Theta_N^c(y) = \Theta_N(y) + \triangle_N^{(2)}(y). \tag{26}$$

The last term in the expansion can be further expanded as

$$O(||p_h(y) - p_N(y)||_{\mathcal{X}}^2) = \frac{1}{2}\frac{\partial^2 \Theta_h}{\partial p_h^2}\Big|_{p_N(y)} (p_h(y) - p_N(y), p_h(y) - p_N(y)) + O(||p_h(y) - p_N(y)||_{\mathcal{X}}^3).$$

So that the third a posteriori error estimator can be given by

$$\triangle_N^{(3)}(y) := \max\left\{\triangle_N^{(4)}(y), \triangle_N^{(5)}(y)\right\}, \tag{27}$$

where (note $|\Theta_h(y) - \Theta_N^c(y)| \approx$ (first term $- \triangle_N^{(2)}$) + second term)

$$\triangle_N^{(4)}(y) := \gamma_h(y)||e_h(y)||_{\mathcal{X}}||e_h^{du}(y)||_{\mathcal{Y}}, \text{ and } \triangle_N^{(5)}(y) := \gamma_h'(y)||e_h(y)||_{\mathcal{X}}^2. \tag{28}$$

Note that $\triangle_N^{(4)}$ and $\triangle_N^{(5)}$ exhibit a **quadratic** dependence on the reduced solution error.

The second (dual-weighted residual) error estimator is the cheapest to evaluate.

Reduced basis approximation: a posteriori error estimators III

We may propose the use of a (improved/corrected) reduced output

$$\Theta_N^c(y) = \Theta_N(y) + \triangle_N^{(2)}(y). \tag{26}$$

The last term in the expansion can be further expanded as

$$O(||p_h(y) - p_N(y)||^2_{\mathcal{X}}) = \frac{1}{2}\frac{\partial^2 \Theta_h}{\partial p_h^2}\Big|_{p_N(y)} (p_h(y) - p_N(y), p_h(y) - p_N(y)) + O(||p_h(y) - p_N(y)||^3_{\mathcal{X}}).$$

So that the third a posteriori error estimator can be given by

$$\triangle_N^{(3)}(y) := \max\left\{\triangle_N^{(4)}(y), \triangle_N^{(5)}(y)\right\}, \tag{27}$$

where (note $|\Theta_h(y) - \Theta_N^c(y)| \approx$ (first term $- \triangle_N^{(2)}$) $+$ second term)

$$\triangle_N^{(4)}(y) := \gamma_h(y)||e_h(y)||_{\mathcal{X}}||e_h^{du}(y)||_{\mathcal{Y}}, \text{ and } \triangle_N^{(5)}(y) := \gamma_h'(y)||e_h(y)||^2_{\mathcal{X}}. \tag{28}$$

Note that $\triangle_N^{(4)}$ and $\triangle_N^{(5)}$ exhibit a **quadratic** dependence on the reduced solution error.

The second (dual-weighted residual) error estimator is the cheapest to evaluate.

Reduced basis approximation: an adaptive greedy algorithm

Adaptively construct the reduced bases using sparse grid nodes as training samples.

Adaptive greedy algorithm [Chen and Quarteroni, 2014]

Initialization: specify tolerance ϵ_t, set $N = 1$, solve the high-fidelity problem at y_1^1, the root node in the sparse grid, and construct the first reduced space $\mathcal{X}_1 = \text{span}\{p_h(y_1^1)\}$;
While sparse grid construction continues
  at each new index $\boldsymbol{i}$, update the training set $\Xi_{train} = \Xi_{\triangle}^{\boldsymbol{i}}$;
  While $\max_{y \in \Xi_{train}} \triangle_N(y) \geq \epsilon_t$
    update Ξ_{train} as $\Xi_{train} = \Xi_{train} \setminus \{y \in \Xi_{train} : \triangle_N(y) < \epsilon_t\}$;
    set $y^{(N+1)} = \text{argmax}_{y \in \Xi_{train}} \triangle_N(y)$;
    solve high-fidelity problem at $y^{(N+1)}$ to obtain $p_h(y^{(N+1)})$;
    update $\mathcal{X}_{N+1} = \mathcal{X}_N \oplus \text{span}\{p_h(y^{(N+1)})\}$;
    set $N = N + 1$.
  end
end

Replace high-fidelity solve by reduced basis solve at **almost all** sparse grid nodes.

Reduced basis approximation: an adaptive greedy algorithm

Adaptively construct the reduced bases using sparse grid nodes as training samples.

Adaptive greedy algorithm [Chen and Quarteroni, 2014]

Initialization: specify tolerance ϵ_t, set $N = 1$, solve the high-fidelity problem at y_1^1, the root node in the sparse grid, and construct the first reduced space $\mathcal{X}_1 = \text{span}\{p_h(y_1^1)\}$;

While sparse grid construction continues

at each new index i, update the training set $\Xi_{train} = \Xi_\triangle^i$;

While $\max_{y \in \Xi_{train}} \triangle_N(y) \geq \epsilon_t$

set $y^{(N+1)} = \text{argmax}_{y \in \Xi_{train}} \triangle_N(y)$;

solve high-fidelity problem at $y^{(N+1)}$ to obtain $p_h(y^{(N+1)})$;

update $\mathcal{X}_{N+1} = \mathcal{X}_N \oplus \text{span}\{p_h(y^{(N+1)})\}$;

set $N = N + 1$.

end

end

Replace high-fidelity solve by reduced basis solve at **almost all** sparse grid nodes.

Adaptively construct the reduced bases using sparse grid nodes as training samples.

Adaptive greedy algorithm [Chen and Quarteroni, 2014]

Initialization: specify tolerance ϵ_t, set $N = 1$, solve the high-fidelity problem at $y_{\mathbf{1}}^{\mathbf{1}}$, the root node in the sparse grid, and construct the first reduced space $\mathcal{X}_1 =$ span$\{p_h(y_{\mathbf{1}}^{\mathbf{1}})\}$;

While sparse grid construction continues

  at each new index $\boldsymbol{i}$, update the training set $\Xi_{train} = \Xi_{\triangle}^{\boldsymbol{i}}$;

  While $\max_{y \in \Xi_{train}} \triangle_N(y) \geq \epsilon_t$

    update Ξ_{train} as $\Xi_{train} = \Xi_{train} \setminus \{y \in \Xi_{train} : \triangle_N(y) < \epsilon_t\}$;

    set $y^{(N+1)} = \mathrm{argmax}_{y \in \Xi_{train}} \triangle_N(y)$;

    solve high-fidelity problem at $y^{(N+1)}$ to obtain $p_h(y^{(N+1)})$;

    update $\mathcal{X}_{N+1} = \mathcal{X}_N \oplus$ span$\{p_h(y^{(N+1)})\}$;

    set $N = N + 1$.

  end

end

Replace high-fidelity solve by reduced basis solve at **almost all** sparse grid nodes.

Reduced basis approximation: an adaptive greedy algorithm

Adaptively construct the reduced bases using sparse grid nodes as training samples.

Adaptive greedy algorithm [Chen and Quarteroni, 2014]

Initialization: specify tolerance ϵ_t, set $N = 1$, solve the high-fidelity problem at $y_{\mathbf{1}}^{\mathbf{1}}$, the root node in the sparse grid, and construct the first reduced space $\mathcal{X}_1 = \text{span}\{p_h(y_{\mathbf{1}}^{\mathbf{1}})\}$;

While sparse grid construction continues

at each new index $\boldsymbol{i}$, update the training set $\Xi_{train} = \Xi_{\triangle}^{\boldsymbol{i}}$;

While $\max_{y \in \Xi_{train}} \triangle_N(y) \geq \epsilon_t$

update Ξ_{train} as $\Xi_{train} = \Xi_{train} \setminus \{y \in \Xi_{train} : \triangle_N(y) < \epsilon_t\}$;

set $y^{(N+1)} = \text{argmax}_{y \in \Xi_{train}} \triangle_N(y)$;

solve high-fidelity problem at $y^{(N+1)}$ to obtain $p_h(y^{(N+1)})$;

update $\mathcal{X}_{N+1} = \mathcal{X}_N \oplus \text{span}\{p_h(y^{(N+1)})\}$;

set $N = N + 1$.

end

end

Replace high-fidelity solve by reduced basis solve at **almost all** sparse grid nodes.

Reduced basis approximation: an adaptive greedy algorithm

Adaptively construct the reduced bases using sparse grid nodes as training samples.

Adaptive greedy algorithm [Chen and Quarteroni, 2014]

Initialization: specify tolerance ϵ_t, set $N = 1$, solve the high-fidelity problem at $y_{\mathbf{1}}^{\mathbf{1}}$, the root node in the sparse grid, and construct the first reduced space $\mathcal{X}_1 = \text{span}\{p_h(y_{\mathbf{1}}^{\mathbf{1}})\}$;
While sparse grid construction continues
  at each new index $\boldsymbol{i}$, update the training set $\Xi_{train} = \Xi_{\triangle}^{\boldsymbol{i}}$;
  While $\max_{y \in \Xi_{train}} \triangle_N(y) \geq \epsilon_t$
    update Ξ_{train} as $\Xi_{train} = \Xi_{train} \setminus \{y \in \Xi_{train} : \triangle_N(y) < \epsilon_t\}$;
    set $y^{(N+1)} = \text{argmax}_{y \in \Xi_{train}} \triangle_N(y)$;
    solve high-fidelity problem at $y^{(N+1)}$ to obtain $p_h(y^{(N+1)})$;
    update $\mathcal{X}_{N+1} = \mathcal{X}_N \oplus \text{span}\{p_h(y^{(N+1)})\}$;
    set $N = N + 1$.
  end
end

Replace high-fidelity solve by reduced basis solve at **almost all** sparse grid nodes.

Reduced basis approximation: an adaptive greedy algorithm

Adaptively construct the reduced bases using sparse grid nodes as training samples.

Adaptive greedy algorithm [Chen and Quarteroni, 2014]

Initialization: specify tolerance ϵ_t, set $N = 1$, solve the high-fidelity problem at $y^{\mathbf{1}}_{\mathbf{1}}$, the root node in the sparse grid, and construct the first reduced space $\mathcal{X}_1 = \text{span}\{p_h(y^{\mathbf{1}}_{\mathbf{1}})\}$;
While sparse grid construction continues
  at each new index $\boldsymbol{i}$, update the training set $\Xi_{train} = \Xi^{\boldsymbol{i}}_{\triangle}$;
  While $\max_{y \in \Xi_{train}} \triangle_N(y) \geq \epsilon_t$
    update Ξ_{train} as $\Xi_{train} = \Xi_{train} \setminus \{y \in \Xi_{train} : \triangle_N(y) < \epsilon_t\}$;
    set $y^{(N+1)} = \text{argmax}_{y \in \Xi_{train}} \triangle_N(y)$;
    solve high-fidelity problem at $y^{(N+1)}$ to obtain $p_h(y^{(N+1)})$;
    update $\mathcal{X}_{N+1} = \mathcal{X}_N \oplus \text{span}\{p_h(y^{(N+1)})\}$;
    set $N = N + 1$.
  end
end

Replace high-fidelity solve by reduced basis solve at **almost all** sparse grid nodes.

A priori error estimates

Assumption [Schwab and Stuart, 2012]

There exist $0 < a_{\min} \leq a_{\max} < \infty$, such that $\forall z \in \mathcal{U} := \bigotimes_{j\in\mathbb{J}}\{z \in \mathbb{C}^{\mathbb{J}} : |z_j| \leq 1\}$

$$a_{\min} \leq \Re(u(x,z)) \leq |u(x,z)| \leq a_{\max}, \quad \forall x \in D \tag{29}$$

There exists a constant $0 < \alpha < 1$, such that (recall $u(y) = \psi_0 + \sum_{j\in\mathbb{J}} y_j\psi_j$)

$$\sum_{j\in\mathbb{J}} ||\psi_j||^{\alpha}_{L^\infty(D)} < \infty. \tag{30}$$

Global approximation

The QoI $\Theta(y)$ and Z are approximated by

$$\Theta(y) = \underbrace{\Theta(y) - \Theta_s(y)}_{\text{interpolation error}} + \underbrace{\Theta_s(y) - \Theta_{s,h}(y)}_{\text{high-fidelity error}} + \underbrace{\Theta_{s,h}(y) - \Theta_{s,h,r}(y)}_{\text{reduced basis error}} + \Theta_{s,h,r}(y),$$

and

$$Z = \underbrace{Z - Z_s}_{\text{quadrature error}} + \underbrace{Z_s - Z_{s,h}}_{\text{high-fidelity error}} + \underbrace{Z_{s,h} - Z_{s,h,r}}_{\text{reduced basis error}} + Z_{s,h,r}.$$

A priori error estimates

Assumption [Schwab and Stuart, 2012]

There exist $0 < a_{\min} \leq a_{\max} < \infty$, such that $\forall z \in \mathcal{U} := \bigotimes_{j\in\mathbb{J}}\{z \in \mathbb{C}^{\mathbb{J}} : |z_j| \leq 1\}$

$$a_{\min} \leq \Re(u(x,z)) \leq |u(x,z)| \leq a_{\max}, \quad \forall x \in D \tag{29}$$

There exists a constant $0 < \alpha < 1$, such that (recall $u(y) = \psi_0 + \sum_{j\in\mathbb{J}} y_j\psi_j$)

$$\sum_{j\in\mathbb{J}} ||\psi_j||^{\alpha}_{L^\infty(D)} < \infty. \tag{30}$$

Global approximation

The QoI $\Theta(y)$ and Z are approximated by

$$\Theta(y) = \underbrace{\Theta(y) - \Theta_s(y)}_{\text{interpolation error}} + \underbrace{\Theta_s(y) - \Theta_{s,h}(y)}_{\text{high-fidelity error}} + \underbrace{\Theta_{s,h}(y) - \Theta_{s,h,r}(y)}_{\text{reduced basis error}} + \Theta_{s,h,r}(y),$$

and

$$Z = \underbrace{Z - Z_s}_{\text{quadrature error}} + \underbrace{Z_s - Z_{s,h}}_{\text{high-fidelity error}} + \underbrace{Z_{s,h} - Z_{s,h,r}}_{\text{reduced basis error}} + Z_{s,h,r}.$$

A priori error estimates

Sparse grid approximation error [Schillings and Schwab, 2013]

$$|\Theta(y) - \Theta_s(y)| \leq CM^{-s} \text{ and } |Z - Z_s| \leq CM^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

High-fidelity approximation error

$$|\Theta_s(y) - \Theta_{s,h}(y)| \leq Ch^t \text{ and } |Z_s - Z_{s,h}| \leq Ch^t, \quad t = \min\{t_{polynomial}, t_{regularity}\}.$$

RB approximation error [Binev et al., 2011, Cohen and DeVore, 2014]

$$|\Theta_{s,h}(y) - \Theta_{s,h,r}(y)| \leq CN^{-s} \text{ and } |Z_{s,h} - Z_{s,h,r}| \leq CN^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

Global approximation error

$$|\Theta(y) - \Theta_{s,h,r}(y)| \leq C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

$$|Z - Z_{s,h,r}| \leq C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

A priori error estimates

Sparse grid approximation error [Schillings and Schwab, 2013]

$$|\Theta(y) - \Theta_s(y)| \leq CM^{-s} \text{ and } |Z - Z_s| \leq CM^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

High-fidelity approximation error

$$|\Theta_s(y) - \Theta_{s,h}(y)| \leq Ch^t \text{ and } |Z_s - Z_{s,h}| \leq Ch^t, \quad t = \min\{t_{polynomial}, t_{regularity}\}.$$

RB approximation error [Binev et al., 2011, Cohen and DeVore, 2014]

$$|\Theta_{s,h}(y) - \Theta_{s,h,r}(y)| \leq CN^{-s} \text{ and } |Z_{s,h} - Z_{s,h,r}| \leq CN^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

Global approximation error

$$|\Theta(y) - \Theta_{s,h,r}(y)| \leq C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

$$|Z - Z_{s,h,r}| \leq C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

A priori error estimates

Sparse grid approximation error [Schillings and Schwab, 2013]

$$|\Theta(y) - \Theta_s(y)| \leq CM^{-s} \text{ and } |Z - Z_s| \leq CM^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

High-fidelity approximation error

$$|\Theta_s(y) - \Theta_{s,h}(y)| \leq Ch^t \text{ and } |Z_s - Z_{s,h}| \leq Ch^t, \quad t = \min\{t_{polynomial}, t_{regularity}\}.$$

RB approximation error [Binev et al., 2011, Cohen and DeVore, 2014]

$$|\Theta_{s,h}(y) - \Theta_{s,h,r}(y)| \leq CN^{-s} \text{ and } |Z_{s,h} - Z_{s,h,r}| \leq CN^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

Global approximation error

$$|\Theta(y) - \Theta_{s,h,r}(y)| \leq C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

$$|Z - Z_{s,h,r}| \leq C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

A priori error estimates

Sparse grid approximation error [Schillings and Schwab, 2013]

$$|\Theta(y) - \Theta_s(y)| \le CM^{-s} \text{ and } |Z - Z_s| \le CM^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

High-fidelity approximation error

$$|\Theta_s(y) - \Theta_{s,h}(y)| \le Ch^t \text{ and } |Z_s - Z_{s,h}| \le Ch^t, \quad t = \min\{t_{polynomial}, t_{regularity}\}.$$

RB approximation error [Binev et al., 2011, Cohen and DeVore, 2014]

$$|\Theta_{s,h}(y) - \Theta_{s,h,r}(y)| \le CN^{-s} \text{ and } |Z_{s,h} - Z_{s,h,r}| \le CN^{-s}, \quad s = \frac{1}{\alpha} - 1.$$

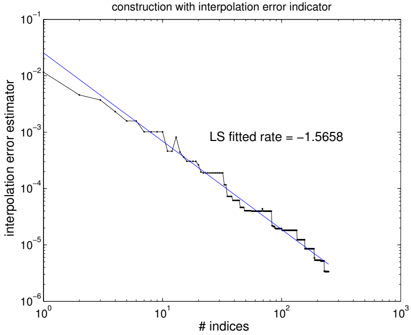
Global approximation error

$$|\Theta(y) - \Theta_{s,h,r}(y)| \le C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

$$|Z - Z_{s,h,r}| \le C_0 M^{-s} + C_1 h^t + C_2 N^{-s}.$$

Numerical experiments: sparse grid approximation error

We consider a diffusion problem with $K = 9$ observations. We take $\mathbb{J} = \{1, \dots, 64\}$ and $\psi_0 = 1$ and $\psi_j = 0.95 j^{-2} \chi_{D_j}$, $j \in \mathbb{J}$, so $\alpha > 1/2$ and the rate $-s = -(1/\alpha - 1) > -1$.



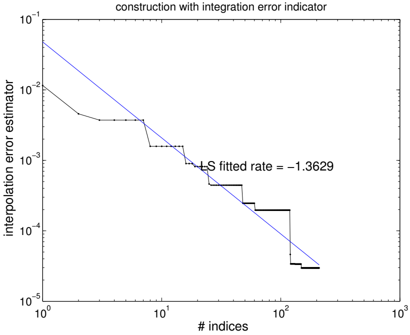


Figure: Interpolation error estimator of the dimension-adaptive sparse grid approximation constructed by the interpolation error indicator (left) and the integration error indicator (right).

Numerical experiments: sparse grid approximation error

We consider a diffusion problem with $K = 9$ observations. We take $\mathbb{J} = \{1, \dots, 64\}$ and $\psi_0 = 1$ and $\psi_j = 0.95 j^{-2} \chi_{D_j}$, $j \in \mathbb{J}$, so $\alpha > 1/2$ and the rate $-s = -(1/\alpha - 1) > -1$.

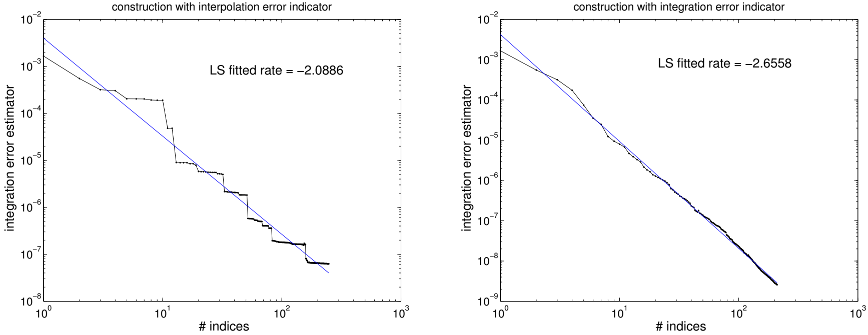


Figure: Integration error estimator of the dimension-adaptive sparse grid approximation constructed by the interpolation error indicator (left) and the integration error indicator (right).

Numerical experiments: high-fidelity approximation error

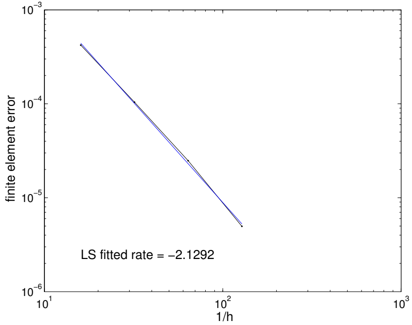


Figure: Left: decay of finite element error with respect to the mesh size $(1/h)$; right: change of the number of reduced bases (constructed with tolerance 10^{-7}) with respect to the mesh size $(1/h)$.

$$\text{effectivity} = \frac{\triangle_N}{|\Theta_h(y) - \Theta_N(y)|}$$

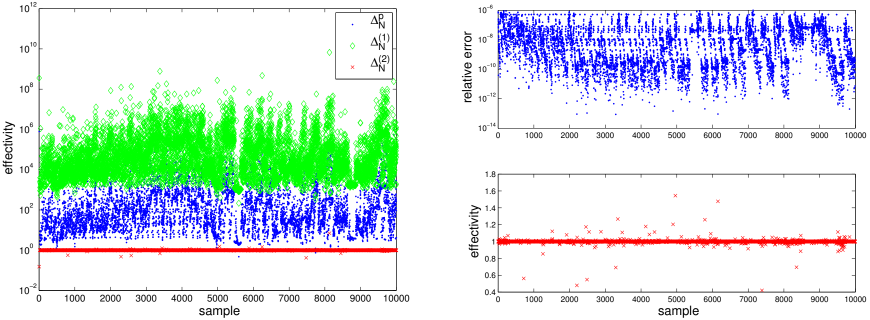
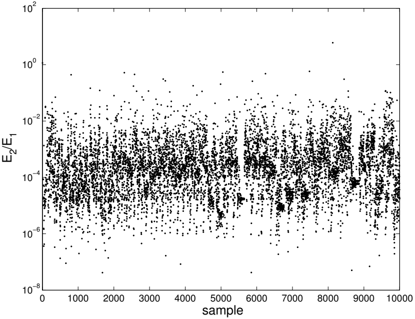


Figure: Left: effectivity of the three error estimators; right: the true reduced output error (truncated above 10^{-14}) and the effectivity of the dual-weighted residual with respect to this error.

Numerical experiments: effectivity of different error estimators

$$\text{effectivity} = \frac{\triangle_N}{|\Theta_h(y) - \Theta_N^c(y)|}$$



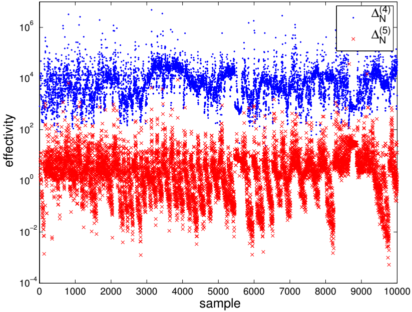


Figure: Left: effect of correction E_2/E_1 with $E_2 = |\Theta_h(y) - \Theta_N^c(y)|$ and $E_1 = |\Theta_h(y) - \Theta_N(y)|$; right: effectivity of $\triangle_N^{(4)}$ and $\triangle_N^{(5)}$ defined in (28) with respect to the corrected output error.

Relative output error without dual correction

$$\max_{y \in \Xi_{test}} \frac{|\Theta_h(y) - \Theta_N(y)|}{\Theta_h(y)}$$

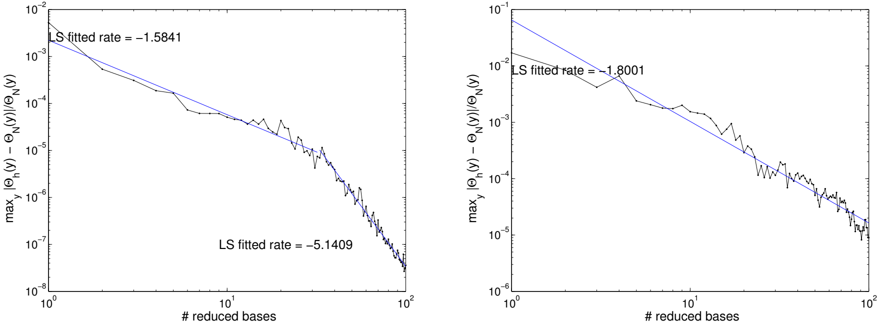


Figure: Decay of reduced basis approximation error with respect to the number reduced bases; left: 64 dimensions, fitted rates for the first 32 bases and the other 68 bases; right: 256 D.

Numerical experiments: reduced basis approximation error

Relative output error with dual correction

$$\max_{y \in \Xi_{test}} \frac{|\Theta_h(y) - \Theta_N^c(y)|}{\Theta_N^c(y)}$$

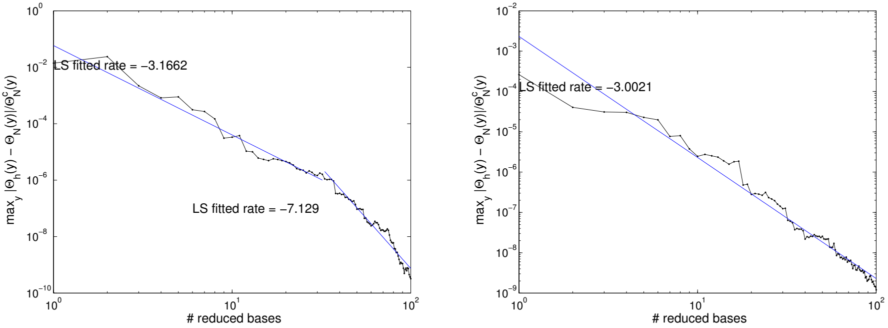


Figure: Decay of reduced basis approximation error with respect to the number reduced bases; left: 64 dimensions, fitted rates for the first 32 bases and the other 68 bases; right: 256 D.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

Conclusion and perspective

Conclusion

- Curse-of-dimensionality can be broken by sparsity – adaptive sparse grid.
- Large-scale computation can be harnessed by reducibility – reduced basis.
- Goal-oriented error estimator (dual-weighted residual) achieves **excellent effectivity** for the nonlinear and nonaffine output in Bayesian inverse problems.
- The adaptive sparse grid approximation error and **particularly the reduced basis** approximation error **converges faster** in practice than predicted by theory.

Perspective

- Work on the improvement of the theoretical estimate for faster convergence.
- Sparse grid reduced basis approximation for nonlinear and nonaffine problems.
- RB can be efficiently combined with any other quadrature rule, e.g. QMC.
- A global framework for adaptive approximation in balancing all the errors.

P. Chen and Ch. Schwab, Sparse grid and reduced basis approximation for Bayesian inverse problems, manuscript, 2014.

References I

Binev, P., Cohen, A., Dahmen, W., DeVore, R., Petrova, G., and Wojtaszczyk, P. (2011).
Convergence rates for greedy algorithms in reduced basis methods.
SIAM Journal of Mathematical Analysis, 43(3):1457–1472.

Chen, P. and Quarteroni, A. (2014).
A new algorithm for high-dimensional uncertainty problems based on dimension-adaptive and reduced basis methods.
EPFL, MATHICSE Report 09, submitted.

Cohen, A. and DeVore, R. (2014).
Kolmogorov widths under holomorphic mappings.

Gerstner, T. and Griebel, M. (2003).
Dimension–adaptive tensor–product quadrature.
Computing, 71(1):65–87.

Patera, A. and Rozza, G. (2007).
Reduced basis approximation and a posteriori error estimation for parametrized partial differential equations.
Copyright MIT, http://augustine.mit.edu.

Schillings, C. and Schwab, C. (2013).
Sparse, adaptive smolyak quadratures for bayesian inverse problems.
Inverse Problems, 29(6).

Schwab, C. and Stuart, A. (2012).
Sparse deterministic approximation of bayesian inverse problems.
Inverse Problems, 28(4):045003.

Smolyak, S. (1963).
Quadrature and interpolation formulas for tensor products of certain classes of functions.
In *Doklady Akademii Nauk SSSR*, volume 4, pages 240–243.

Stuart, A. (2010).
Inverse problems: a Bayesian perspective.
Acta Numerica, 19(1):451–559.

Thank you for your attention!