Adaptive algorithms for low-rank approximation: subspace point of view and goal-oriented approximation

Anthony Nouy

Ecole Centrale Nantes / GeM
Nantes, France

Collaborators: Marie Billaud-Friess, Loic Giraldi, Gregory Legrain, Olivier Zahm

Low-rank approximation of high-dimensional functions

- Objective: approximation of a multivariate function

$$u(x_1, \dots, x_d)$$

seen as an element of a tensor space

$$V_{\|\cdot\|} = \overline{V_1 \otimes \dots \otimes V_d}^{\|\cdot\|}.$$

Low-rank approximation of high-dimensional functions

- Objective: approximation of a multivariate function

$$u(x_1, \dots, x_d)$$

seen as an element of a tensor space

$$V_{\|\cdot\|} = \overline{V_1 \otimes \dots \otimes V_d}^{\|\cdot\|}.$$

- Low-rank approximation

$$u_m(x_1, \dots, x_d) = \sum_{i=1}^{m} (v_i^1 \otimes \dots \otimes v_i^d)(x_1, \dots, x_d) = \sum_{i=1}^{m} v_i^1(x_1) \dots v_i^d(x_d)$$

- Greedy construction of the approximation (canonical decomposition):

$$\mathcal{E}(u, u_m) = \min_{v_m^1, \dots, v_m^d} \mathcal{E}(u, u_{m-1} + v_m^1 \otimes \dots \otimes v_m^d)$$

where $\mathcal{E}(u, \cdot)$ is a certain distance to u.

Low-rank approximation of high-dimensional functions

- Greedy construction of the approximation (canonical decomposition):

$$\mathcal{E}(u, u_m) = \min_{v_m^1, \dots, v_m^d} \mathcal{E}(u, u_{m-1} + v_m^1 \otimes \dots \otimes v_m^d)$$

where $\mathcal{E}(u, \cdot)$ is a certain distance to u.

  - Convergence may be slow

Low-rank approximation of high-dimensional functions

- Greedy construction of the approximation (canonical decomposition):

$$\mathcal{E}(u, u_m) = \min_{v_m^1,\dots,v_m^d} \mathcal{E}(u, u_{m-1} + v_m^1 \otimes \dots \otimes v_m^d)$$

  where $\mathcal{E}(u, \cdot)$ is a certain distance to u.
  - Convergence may be slow
  - Does not exploit the tensor structure

- Greedy construction of the approximation (canonical decomposition):

$$\mathcal{E}(u, u_m) = \min_{v_m^1, \ldots, v_m^d} \mathcal{E}(u, u_{m-1} + v_m^1 \otimes \ldots \otimes v_m^d)$$

where $\mathcal{E}(u, \cdot)$ is a certain distance to u.

  - Convergence may be slow
  - Does not exploit the tensor structure
  - Redundant representation : linear dependencies in $\{v_1^k(x_k), \ldots, v_m^k(x_k)\}$?

Outline

Outline

Best approximation in Tucker format: a subspace point of view

- Tucker tensors with bounded multilinear rank:

$$\mathcal{T}_r = \{v : rank_\mu(v) \le r_\mu, \forall \mu \in \{1, \dots, d\}\}$$
$$= \left\{ v = \sum_{i_1=1}^{r_1} \dots \sum_{i_d=1}^{r_d} a_{i_1 \dots i_d} v^1_{i_1} \otimes \dots \otimes v^d_{i_d} : v^\mu_k \in V^\mu, a \in \mathbb{R}^{r_1 \times \dots \times r_d} \right\}$$

Best approximation in Tucker format: a subspace point of view

- Tucker tensors with bounded multilinear rank:

$$\mathcal{T}_r = \{v : rank_\mu(v) \leq r_\mu, \forall \mu \in \{1, \ldots, d\}\}$$
$$= \left\{ v = \sum_{i_1=1}^{r_1} \ldots \sum_{i_d=1}^{r_d} a_{i_1 \ldots i_d} v^1_{i_1} \otimes \ldots \otimes v^d_{i_d} : v^\mu_k \in V^\mu, a \in \mathbb{R}^{r_1 \times \ldots \times r_d} \right\}$$

- Subspace-based parametrization

$$\mathcal{T}_r = \left\{ v \in \bigotimes_{\mu=1}^{d} U_\mu \ : \ dim(U_\mu) \leq r_\mu, \forall \mu \right\}$$

Best approximation in Tucker format: a subspace point of view

- Tucker tensors with bounded multilinear rank:

$$\mathcal{T}_r = \{v : rank_\mu(v) \leq r_\mu, \forall \mu \in \{1, \ldots, d\}\}$$
$$= \left\{ v = \sum_{i_1=1}^{r_1} \ldots \sum_{i_d=1}^{r_d} a_{i_1 \ldots i_d} v_{i_1}^1 \otimes \ldots \otimes v_{i_d}^d : v_k^\mu \in V^\mu, a \in \mathbb{R}^{r_1 \times \ldots \times r_d} \right\}$$

- Subspace-based parametrization

$$\mathcal{T}_r = \left\{ v \in \bigotimes_{\mu=1}^{d} U_\mu \ : \ dim(U_\mu) \leq r_\mu, \forall \mu \right\}$$

- Best approximation — a subspace point of view:

$$\min_{v \in \mathcal{T}_r} \mathcal{E}(u, v) = \min_{dim(U_1)=r_1} \ldots \min_{dim(U_d)=r_d} \min_{v \in U_1 \otimes \ldots \otimes U_d} \mathcal{E}(u, v)$$

This yields sequences of optimal but non necessarily nested subspaces $\{U_\mu^{r_\mu} : r_\mu \geq 1\}$.

Adaptive construction of subspaces for higher-order tensors

- Greedy construction of subspaces with nestedness property

$$\mathcal{E}(u, u_m) = \min_{U_1^m \supset U_1^{m-1}} \dots \min_{U_d^m \supset U_d^{m-1}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

Adaptive construction of subspaces for higher-order tensors

- Greedy construction of subspaces with nestedness property

$$\mathcal{E}(u, u_m) = \min_{U_1^m \supset U_1^{m-1}} \dots \min_{U_d^m \supset U_d^{m-1}} \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

- Suboptimal greedy construction of subspaces with nestedness property (isotropic enrichment)

$$\mathcal{E}(u, u_{m-1} + \otimes_{\mu=1}^d w_m^{(\mu)}) = \min_{w \in \mathcal{R}_1} \mathcal{E}(u, u_{m-1} + w), \quad U_\mu^m = U_\mu^{m-1} + span\{w_m^{(\mu)}\}$$

$$\mathcal{E}(u, u_m) \leq (1 + \epsilon) \min_{v \in U_1^m \otimes \dots \otimes U_d^m} \mathcal{E}(u, v)$$

Simple Benchmark: Poisson equation

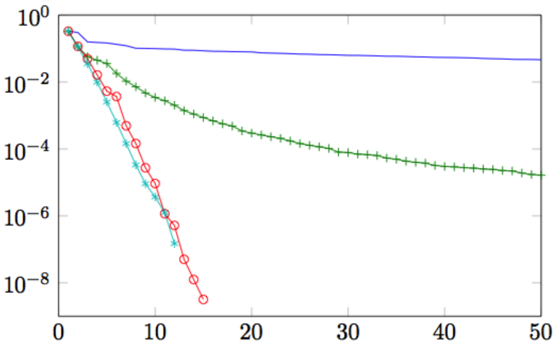
$$-\Delta u = 1 \quad \text{on} \quad \Omega = (0,1)^d$$

$$\mathcal{E}(u,w) = \mathcal{J}(w) - \mathcal{J}(u), \quad \mathcal{J}(w) = \int_\Omega \nabla w \cdot \nabla w - 2\int_\Omega w$$

$$\mathcal{E}(u,w) = \|w-u\|^2_{H^1_0}$$

Simple Benchmark: Poisson equation in dimension $d = 8$

Error (norm of algebraic residual, fixed finite element discretization) with respect to the rank for different tensor formats and algorithms



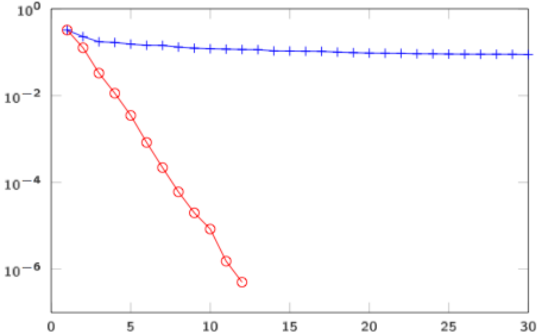
$+$ Greedy approximation in $\mathcal{R}_r$ (rank-one updates)

$+$ "Optimal approximation" $\mathcal{R}_r$

$\star$ "Optimal approximation in $\mathcal{H}_r(V_1 \otimes \ldots \otimes V_d)$"

$\circ$ Suboptimal greedy construction of subspaces U^r_μ (approximation in $\mathcal{H}_r(U^r_1 \otimes \ldots \otimes U^r_d)$)

Error with respect to the rank for different tensor formats and algorithms



$+$ Greedy approximation in $\mathcal{R}_r$ (rank-one updates)

$\circ$ Suboptimal greedy construction of subspaces U^r_μ (approximation in $\mathcal{H}_r(U^r_1 \otimes \ldots \otimes U^r_d)$)

Outline

Eigenvalue problems

Virginie Ehrlacher and Loic Giraldi

Computing the lowest eigenvalue of

$$Au = \lambda u, \quad u \in V = \overline{V_1 \otimes \ldots \otimes V_d}$$

with greedy construction of approximation or subspaces.

Algorithm (symmetric operator):
Start from $u_0 = 0$. At step m:

- Compute a rank-one correction $w_m = \otimes_{\mu=1}^d w_m^{(\mu)}$ of u_{m-1}, e.g.

$$\min_{w_m \in \mathcal{R}_1} \frac{\langle A(u_{m-1} + w_m), u_{m-1} + w_m \rangle}{\langle u_{m-1} + w_m, u_{m-1} + w_m \rangle}$$

- Update of subspaces $U_\mu^m = U_\mu^{m-1} + span\{w_m^{(\mu)}\}$, $U^m = U_1^m \otimes \ldots \otimes U_d^m$
- Approximate solution of the eigenproblem projected on U_m:

$$u_m \in U^m = U_1^m \otimes \ldots \otimes U_d^m, \quad \langle v, Au_m \rangle \approx \lambda \langle v, u_m \rangle \quad \forall v \in U^m$$

$$\min_{u_m \in \mathcal{M} \subset U^m} \frac{\langle Au_m, u_m \rangle}{\langle u_m, u_m \rangle}$$

Illustration: PDE eigenvalue problem

Example considered in [Kressner & Tobler 2011]

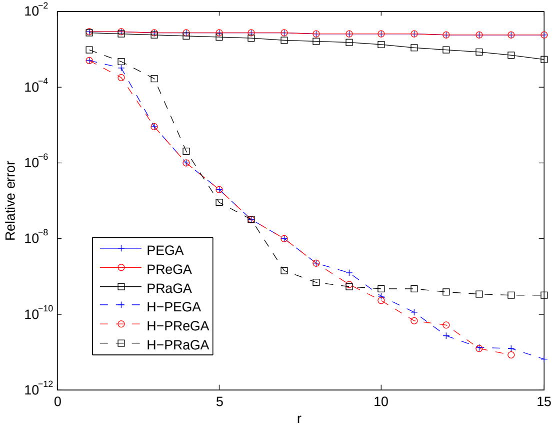
$$\begin{cases} -\Delta u(x) + V(x)u(x) = \lambda u(x), & x \in \Omega \\ u(x) = 0, & x \in \partial\Omega \end{cases}$$

- Henon-Heiles potential

$$V(x_1, \ldots, x_d) = \frac{1}{2} \sum_{i=1}^{d} \sigma_i x_i^2 + \sum_{i=1}^{d-1} \left(\sigma_\star (x_i x_{i+1}^2 - \frac{1}{3} x_i^3) + \frac{\sigma_\star^2}{16} (x_i^2 + x_{i+1}^2)^2 \right)$$

  with $\sigma_j = 1$ and $\sigma_\star = 0.2$
- $\Omega = (-10, 2)^d$, $d = 20$
- Finite element approximation with 128 nodes per dimension

Error with respect to the rank for different tensor formats and algorithms



- Solid lines : Greedy approximation in $\mathcal{R}_r$ (rank-one updates)
- Dashed lines : Suboptimal greedy construction of subspaces U_μ^r (approximation in $\mathcal{H}_r(U_1^r \otimes \ldots \otimes U_d^r)$)

$$u_r = \sum_{i_1=1}^{r} \ldots \sum_{i_d=1}^{r} a^r_{i_1 \ldots i_d} w^{(1)}_{i_1} \otimes \ldots w^{(d)}_{i_d} \quad \text{with low-rank approximation of } a^r$$

Outline

Direct approximation

If

$$\alpha\|u - w\| \leq \mathcal{E}(u, w) \leq \beta\|u - w\|$$

for all $w \in \mathcal{M}_{\leq r}$, then

$$u_r = \arg\min_{w \in \mathcal{M}_{\leq r}} \mathcal{E}(u, w)$$

is such that

$$\|u - u_r\| \leq \frac{\beta}{\alpha} \min_{w \in \mathcal{M}_{\leq r}} \|u - w\|$$

Interest of working with well conditioned formulations,
i.e. such that $\beta/\alpha \approx 1$

(Explicit) Preconditioning in tensor format

- Given an operator $A \in W = \bigotimes_{\mu=1}^{d} W_\mu$, $W_\mu \simeq \mathbb{R}^{n_\mu \times n_\mu}$, we want to construct an approximation P of the inverse A^{-1} using low-rank format

$$P = \sum_{i \in I_1 \times \ldots \times I_d} a_i \bigotimes_{\mu=1}^{d} P^{\mu}_{i_\mu} \quad \text{(structured } a\text{)}$$

(Explicit) Preconditioning in tensor format

- Given an operator $A \in W = \bigotimes_{\mu=1}^{d} W_\mu$, $W_\mu \simeq \mathbb{R}^{n_\mu \times n_\mu}$, we want to construct an approximation P of the inverse A^{-1} using low-rank format

$$P = \sum_{i \in I_1 \times \ldots \times I_d} a_i \bigotimes_{\mu=1}^{d} P^{\mu}_{i_\mu} \quad \text{(structured } a\text{)}$$

- Given a low rank subset $\mathcal{M}$ of operators in W, we would like

$$\min_{P \in \mathcal{M}} \|P - A^{-1}\|_\star$$

for a norm $\|\cdot\|_\star$ which makes computable the approximate inverse.

(Explicit) Preconditioning in tensor format

- Given an operator $A \in W = \bigotimes_{\mu=1}^{d} W_\mu$, $W_\mu \simeq \mathbb{R}^{n_\mu \times n_\mu}$, we want to construct an approximation P of the inverse A^{-1} using low-rank format

$$P = \sum_{i \in I_1 \times \ldots \times I_d} a_i \bigotimes_{\mu=1}^{d} P^\mu_{i_\mu} \quad \text{(structured } a\text{)}$$

- Given a low rank subset $\mathcal{M}$ of operators in W, we would like

$$\min_{P \in \mathcal{M}} \|P - A^{-1}\|_\star$$

  for a norm $\|\cdot\|_\star$ which makes computable the approximate inverse.
- Letting $\|\cdot\|$ denote the canonical inner product norm on W induced by Frobenius norms on the W_μ:
  - $\|P - A^{-1}\|_\star = \|I - AP\|$ (approximate right inverse)
  - $\|P - A^{-1}\|_\star = \|A^{-1/2} - A^{1/2}P\|$ (approximate right inverse for symmetric matrices)
  - $\|P - A^{-1}\|_\star = \|I - PA\|$ (approximate left inverse)
  - $\|P - A^{-1}\|_\star = \|A^{-1/2} - PA^{1/2}\|$ (approximate left inverse for symmetric matrices)

Greedy construction of operator subspaces

Let $P_0 = 0$. For $m = 1, 2, ...,$ do

- Correction step (possible additional constraints: symmetry, sparsity):

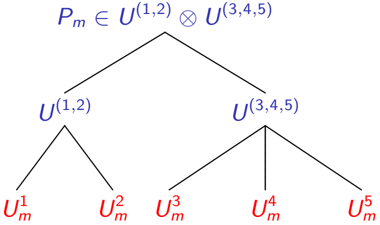
$$Q_m^1 \otimes \ldots \otimes Q_m^d \in \arg \min_{Q \in \mathcal{R}_1(W)} \|A^{-1} - P_{m-1} - Q\|_\star$$

- Update of operator subspaces:

$$U_m^\mu = span\{Q_1^\mu, \ldots, Q_m^\mu\}, \quad \mu \in \{1, \ldots, d\}, \quad \text{and} \quad U_m = U_m^1 \otimes \ldots \otimes U_m^d$$

- Approximate projection step in U_m using tree-based format :

$$P_m \in \arg \min_{v \in \mathcal{H}_m^T(U_m)} \|A^{-1} - P\|_\star$$



Poisson equation

- $d = 20$
- Construction of preconditioner using $\|A^{-1} - P\|_\star = \|A^{-1/2} - PA^{1/2}\|$
- Approximate PCG in $\mathcal{H}^T_{15}(V)$
- Symmetric approximation

Poisson equation

- $d = 20$
- Construction of preconditioner using $\|A^{-1} - P\|_\star = \|A^{-1/2} - PA^{1/2}\|$
- Approximate PCG in $\mathcal{H}^T_{15}(V)$
- Symmetric approximation

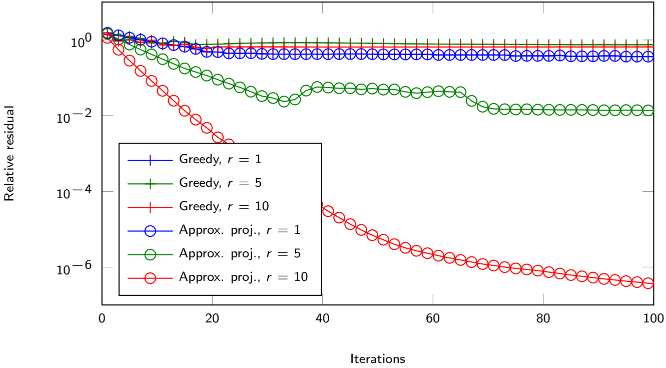


Figure: Convergence of the PCG for different preconditioners

Heat equation with uncertain parameters

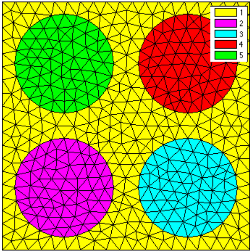


Figure: Geometry and mesh of Ω

$$-\nabla(K(\xi)\nabla u) = 1 \quad \text{on } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,$$

with a random diffusion field

$$K(\xi) = \sum_{i=1}^{5} I_{\omega_i} \xi_i$$

$$\xi_1 \sim U(1,2),\ \xi_2, \dots, \xi_5 \sim \log U(10,100)$$

- $u \in V = H_0^1(\Omega) \otimes L^2_{\mu_1}(\Xi_1) \otimes \ldots \otimes L^2_{\mu_5}(\Xi_5)$
- Finite element approximation in space
- Degree 10 Legendre polynomial expansions in parametric dimensions

- Approximate PCG in $\mathcal{H}^T_{15}(V)$
- Preconditioners constructed with $\|A^{-1} - P\|_\star = \|A^{-1/2} - PA^{1/2}\|$
- Symmetric approximation

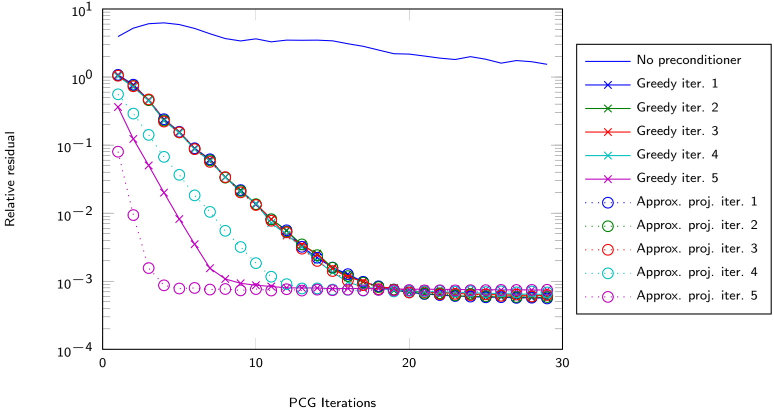


Figure: Convergence of the PCG for different preconditioners

Outline

- Greedy construction of subspaces with nestedness property (anisotropic enrichment)
  At iteration m, given dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \ldots \otimes U_d^m} \mathcal{E}(u, v)$$

Anisotropic construction of subspaces for higher-order tensors

- Greedy construction of subspaces with nestedness property (anisotropic enrichment)
  At iteration m, given dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \ldots \otimes U_d^m} \mathcal{E}(u, v)$$

- For the selection of the next subset of dimensions D_{m+1}:

Anisotropic construction of subspaces for higher-order tensors

- Greedy construction of subspaces with nestedness property (anisotropic enrichment)
  At iteration m, given dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \ldots \otimes U_d^m} \mathcal{E}(u, v)$$

- For the selection of the next subset of dimensions D_{m+1}:
  - $u_m \in U_\mu^m \otimes V_{[\mu]} \subset V_\mu \otimes V_{[\mu]}$, with $dim(U_\mu^m) = r_\mu^m$, admits the following SVD

$$u_m = \sum_{k=1}^{r_\mu^m} \sigma_k^\mu v_k^\mu \otimes w_k^{[\mu]}$$

Anisotropic construction of subspaces for higher-order tensors

- Greedy construction of subspaces with nestedness property (anisotropic enrichment)
  At iteration m, given dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \ldots \otimes U_d^m} \mathcal{E}(u, v)$$

- For the selection of the next subset of dimensions D_{m+1}:
  - $u_m \in U_\mu^m \otimes V_{[\mu]} \subset V_\mu \otimes V_{[\mu]}$, with $dim(U_\mu^m) = r_\mu^m$, admits the following SVD

$$u_m = \sum_{k=1}^{r_\mu^m} \sigma_k^\mu v_k^\mu \otimes w_k^{[\mu]}$$

  - If u_m is a good approximation of the truncated SVD of u, then

$$\|u - u_m\|_{\vee(V_\mu \otimes V_{[\mu]})} \approx \sigma^\mu_{r_\mu^m + 1} \leq \sigma^\mu_{r_\mu^m}$$

Anisotropic construction of subspaces for higher-order tensors

- Greedy construction of subspaces with nestedness property (anisotropic enrichment)
  At iteration m, given dimensions D_m for enrichment, let $U_\mu^m = U_\mu^{m-1}$ for $\mu \notin D_m$ and

$$\mathcal{E}(u, u_m) = \min_{\substack{U_\mu^m \supset U_\mu^{m-1} \\ \mu \in D_m}} \min_{v \in U_1^m \otimes \ldots \otimes U_d^m} \mathcal{E}(u, v)$$

- For the selection of the next subset of dimensions D_{m+1}:
  - $u_m \in U_\mu^m \otimes V_{[\mu]} \subset V_\mu \otimes V_{[\mu]}$, with $dim(U_\mu^m) = r_\mu^m$, admits the following SVD

$$u_m = \sum_{k=1}^{r_\mu^m} \sigma_k^\mu v_k^\mu \otimes w_k^{[\mu]}$$

  - If u_m is a good approximation of the truncated SVD of u, then

$$\|u - u_m\|_{\vee(V_\mu \otimes V_{[\mu]})} \approx \sigma^\mu_{r_\mu^m + 1} \leq \sigma^\mu_{r_\mu^m}$$

  - Letting $0 \leq \theta \leq 1$, choose

$$D_{m+1} = \left\{ \mu \in \{1, \ldots, d\} : \sigma^\mu_{r_\mu^m} \geq \theta \max_{1 \leq \nu \leq d} \sigma^\nu_{r_\nu^m} \right\}$$

Illustration: PDE with random coefficients

$$-\nabla \cdot (K \nabla u) + \xi_2 u = 1 \quad on \quad \Omega, \quad u = 0 \quad on \quad \partial\Omega$$

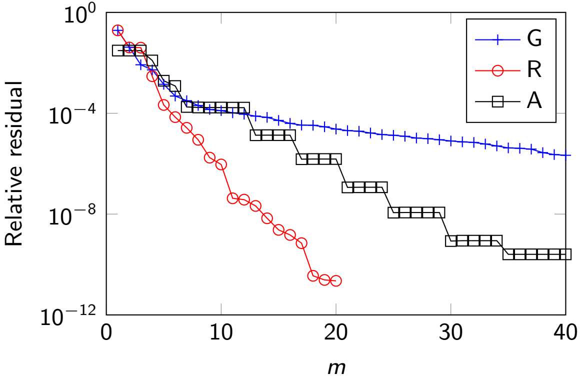
$$K = 1 + \xi_1 I_D(x)$$

$$\xi_1 \sim U(0, 10), \quad \xi_2 \sim U(0, 1)$$

$$u \in H_0^1(\Omega) \otimes L^2(\Xi_1) \otimes L^2(\Xi_2)$$

Illustration: PDE with random coefficients

Error with respect to iteration for different tensor formats and algorithms



$+$ Greedy approximation in $\mathcal{R}_r$ (rank-one updates)

$\circ$ Suboptimal greedy construction of subspaces based on rank-one corrections (isotropic)

$\square$ Suboptimal greedy construction of subspaces with anisotropic enrichment

Illustration: PDE with random coefficients

Ranks with respect to iteration m for anisotropic construction

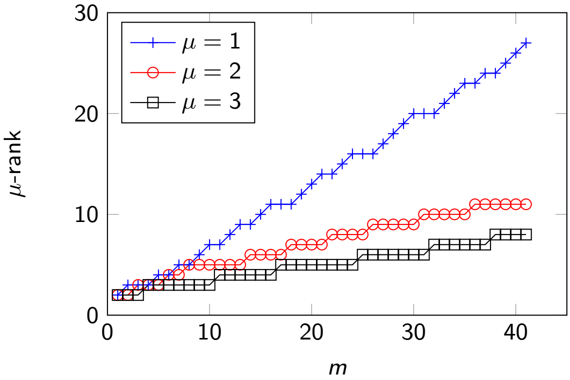
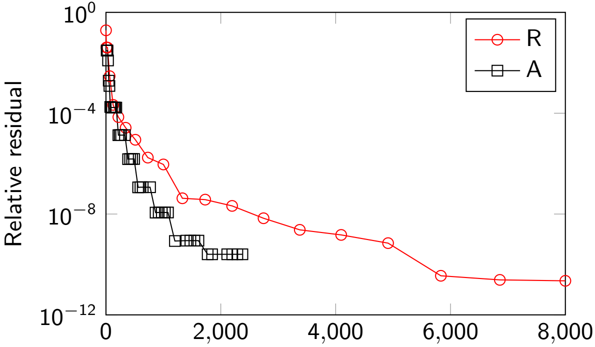


Illustration: PDE with random coefficients

Error with respect to the dimension of the reduced space U^m for greedy constructions of subspaces.



$$dim(U_m) = \prod_{\mu=1}^{3} dim(U_\mu^m)$$

◦ Suboptimal greedy construction of subspaces based on rank-one corrections (isotropic)

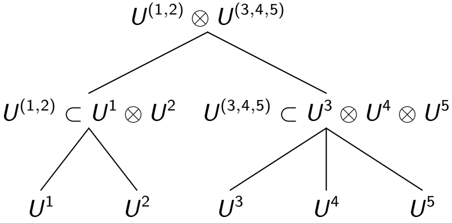
□ Suboptimal greedy construction of subspaces with anisotropic enrichment

Construction of subspaces for tree-based formats

- Tensors with bounded tree-based rank

$$\mathcal{H}_r^T = \{v \ : \ v \in U_\alpha \otimes U_{\alpha^c}, dim(U_\alpha) \leq r_\alpha, \alpha \in T\}$$

s.t. the set of subspaces $\{U_\alpha\}_{\alpha \in T}$ has a hierarchical structure

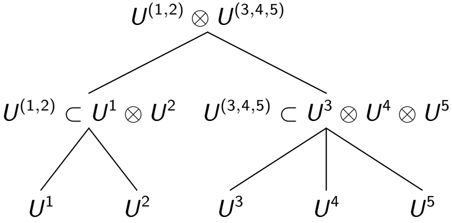


Construction of subspaces for tree-based formats

- Tensors with bounded tree-based rank

$$\mathcal{H}_r^T = \{v \ : \ v \in U_\alpha \otimes U_{\alpha^c}, dim(U_\alpha) \leq r_\alpha, \alpha \in T\}$$

s.t. the set of subspaces $\{U_\alpha\}_{\alpha \in T}$ has a hierarchical structure



- Best approximation problems - a subspace point of view

$$\min_{v \in \mathcal{H}_r^T} \|u - v\| = \min_{\substack{U_\mu \subset V_\mu \\ dim(U_\mu) = r_\mu \\ \mu \in \{1,\dots d\}}} \ \min_{\substack{U_\alpha \subset \bigotimes_{\beta \in S(\alpha)} U_\beta \\ dim(U_\alpha) = r_\alpha \\ \alpha \in I(T)}} \ \min_{v \in \bigotimes_{\beta \in S(D)} U_\beta} \|u - v\|$$

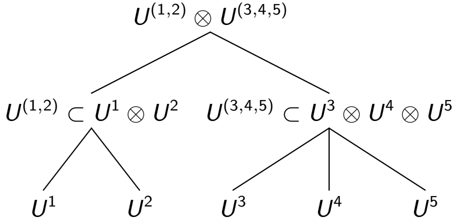
define sequences of optimal and non necessarily nested subspaces $\{U_\alpha^{r_\alpha}; r_\alpha \geq 1\}$.

Construction of subspaces for tree-based formats

- Tensors with bounded tree-based rank

$$\mathcal{H}_r^T = \{v \;:\; v \in U_\alpha \otimes U_{\alpha^c}, dim(U_\alpha) \leq r_\alpha, \alpha \in T\}$$

s.t. the set of subspaces $\{U_\alpha\}_{\alpha \in T}$ has a hierarchical structure



- Best approximation problems - a subspace point of view

$$\min_{v \in \mathcal{H}_r^T} \|u - v\| = \min_{\substack{U_\mu \subset V_\mu \\ dim(U_\mu) = r_\mu \\ \mu \in \{1,\ldots d\}}} \min_{\substack{U_\alpha \subset \bigotimes_{\beta \in S(\alpha)} U_\beta \\ dim(U_\alpha) = r_\alpha \\ \alpha \in I(T)}} \min_{v \in \bigotimes_{\beta \in S(D)} U_\beta} \|u - v\|$$

define sequences of optimal and non necessarily nested subspaces $\{U_\alpha^{r_\alpha}; r_\alpha \geq 1\}$.

- Algorithms for the construction of suboptimal sequences of nested subspaces ... strategies of enrichment for non isotropic constructions ?

Outline

Low-rank approximation based on residual minimization

Problem to solve

$$A(u) = b \quad \text{with} \quad u \in V = \bigotimes_{\nu=1}^{d} V^{\nu}$$

Given a low-rank tensor subset $\mathcal{M}$, replace

$$\inf_{v \in \mathcal{M}} \|u - v\|$$

by the optimization of a criterium

$$\inf_{v \in \mathcal{M}} \mathcal{E}(v, u)$$

yielding an computable approximation of u in $\mathcal{M}$.

$$\mathcal{E}(v, u) = \|b - Av\|_*^2$$

Good residual norms ?

Ideal minimal residual formulation

Definition of best approximations based on minimal residual formulation

$$\boxed{\min_{v \in \mathcal{M}} \|Av - b\|_*}$$

with a residual norm such that

$$\boxed{\|Av - b\|_* \approx \|u - v\|_V}$$

with $\|\cdot\|_V$ a chosen norm on V.

Ideal minimal residual formulation

Definition of best approximations based on minimal residual formulation

$$\boxed{\min_{v \in \mathcal{M}} \|Av - b\|_*}$$

with a residual norm such that

$$\boxed{\|Av - b\|_* \approx \|u - v\|_V}$$

with $\|\cdot\|_V$ a chosen norm on V.

Example : weighted Sobolev norms

$$u \in H_0^1(\Omega) \otimes L_\mu^2(\Xi), \quad \|u\|_V^2 = \int_{\Omega \times \Xi} \alpha(x,\xi) \nabla u^2 dx\, d\mu(\xi)$$

$$\alpha(x,\xi) = \alpha_1(\xi) I_{\Omega_1}(x) + \alpha_2(\xi) I_{\Omega \setminus \Omega_1}(x) \textit{ with } \Omega_1 \subset \Omega$$

- Norm involving quantities of interest

$$\|u-v\|_{V,\alpha}^2 = \|u-v\|_{V,0}^2 + \alpha\|Lu-Lv\|_Z^2$$

with

$$L : V \to Z$$

- Norm involving quantities of interest

$$\|u - v\|_{V,\alpha}^2 = \|u - v\|_{V,0}^2 + \alpha \|Lu - Lv\|_Z^2$$

with

$$L : V \to Z$$

Example: quantities of interest in uncertainty quantification

$$u \in L^2_\mu(\Xi) \otimes \mathcal{V}$$

- Norm involving quantities of interest

$$\|u - v\|_{V,\alpha}^2 = \|u - v\|_{V,0}^2 + \alpha\|Lu - Lv\|_Z^2$$

with

$$L : V \to Z$$

Example: quantities of interest in uncertainty quantification

$$u \in L^2_\mu(\Xi) \otimes \mathcal{V}$$

- Variable of interest

$$Lu(\xi) = \ell(u(\xi); \xi), \quad Z = L^2_\mu(\Xi)$$

- Norm involving quantities of interest

$$\|u - v\|_{V,\alpha}^2 = \|u - v\|_{V,0}^2 + \alpha \|Lu - Lv\|_Z^2$$

with

$$L : V \to Z$$

Example: quantities of interest in uncertainty quantification

$$u \in L^2_\mu(\Xi) \otimes \mathcal{V}$$

- Variable of interest

$$Lu(\xi) = \ell(u(\xi); \xi), \quad Z = L^2_\mu(\Xi)$$

- Expectation

$$L(u) = \mathbb{E}_\mu(\ell(u(\xi); \xi)), \quad Z = \mathbb{R}$$

- Conditional expectation

$$L(u) = \mathbb{E}_\mu(\ell(u(\xi); \xi) | \xi_\nu), \quad Z = L^2_{\mu_\nu}(\Xi_\nu)$$

A strategy for weakly coercive problems

$$\boxed{Au = b}, \quad u \in V, \quad b \in W', \quad A : V \to W'$$

with A defining an isomorphism such that $\alpha \|v\|_V \leq \|Av\|_{W'} \leq \beta \|v\|_V$.

Ideal approach

Work with two different norms on V and W such that

$$\| \cdot \|_V = \|A(\cdot)\|_{W'}$$

corresponding to $\alpha = \beta = 1$, and therefore

$$\|Av - b\|_{W'} = \|v - u\|_V$$

Denoting by $R_V : V \to V'$ and $R_W : W \to W'$ the Riesz maps, it is equivalent to imposing

$$R_W = AR_V^{-1}A^*$$

Approximation of the ideal approach [Cohen-Dahmen-Welper 2012]

Let $\Lambda^\delta : W \to W$ be such that for all y (in a particular subset),

$$\|\Lambda^\delta(y) - y\|_W \leq \delta \|y\|_W.$$

Then, letting

$$\boxed{\|Av - b\|_* = \|\Lambda^\delta R_W^{-1}(Av - b)\|_W},$$

we have

$$\boxed{(1-\delta)\|u - v\|_V \leq \|Av - b\|_* \leq (1+\delta)\|u - v\|_V}$$

Approximation in a tensor subset

Let $\mathcal{M} \subset V$ be a given approximation subset. We would like to solve

$$u^\delta \in \min_{v \in \mathcal{M}} \|\Lambda^\delta R_W^{-1}(Av - b)\|_W$$

Quasi-best approximation with respect to $\| \cdot \|_V$

$$\|u - u^\delta\|_V \leq \frac{1+\delta}{1-\delta} \inf_{v \in \mathcal{M}} \|u - v\|_V$$

Algorithm

Gradient-type algorithm

Initialize $u^0 = 0$ and construct a sequence $\{u^k\}_{k\geq 1}$ such that

1. $y^k = \Lambda^\delta(R_W^{-1}(Au^k - b))$
2. $u^k \in \Pi_{\mathcal{M}}^{\epsilon}(u^{k-1} - R_V^{-1}A^* y^k)$

with

- $y^k = \Lambda^\delta(z^k)$ with

$$\left\langle R_W z^k, \delta y \right\rangle_{W',W} = \left\langle Au^k - b, \delta y \right\rangle_{W',W} \quad \forall \delta y \in W$$

  solved with a tensor approximation method (with precision δ).

- Quasi-best approximations in $\mathcal{M}$:

$$\Pi_{\mathcal{M}}^{\epsilon}(u) = \left\{ w \in \mathcal{M} \ : \ \|u - w\|_V \leq (1+\epsilon) \min_{v \in \mathcal{M}} \|u - v\|_V \right\}$$

Properties of the gradient-type algorithm

- If $\delta(2+\epsilon) < 1$, then

$$\limsup_{k\to\infty} \|u^k - u\|_V \leq \frac{1+\epsilon}{1-\delta(2+\epsilon)} \inf_{v\in\mathcal{M}} \|u - v\|_V$$

- For $\delta = 0$, convergence in one iteration ($u^1 \in \Pi^\epsilon_{\mathcal{M}}(u)$).
- For $\delta \neq 0$, quite fast convergence.

Properties of the gradient-type algorithm

- If $\delta(2+\epsilon) < 1$, then

$$\limsup_{k\to\infty} \|u^k - u\|_V \leq \frac{1+\epsilon}{1-\delta(2+\epsilon)} \inf_{v\in\mathcal{M}} \|u-v\|_V$$

- For $\delta = 0$, convergence in one iteration ($u^1 \in \Pi^{\epsilon}_{\mathcal{M}}(u)$).
- For $\delta \neq 0$, quite fast convergence.

Error estimation

$$(1-\delta)\|u^k - u\|_V \leq \|y^k\|_W \leq (1+\delta)\|u^k - u\|_V$$

Furthermore, if Λ^{δ} is an W-orthogonal projection onto some subspace of W, then

$$\sqrt{1-\delta^2}\|u^k - u\|_V \leq \|y^k\|_W \leq \|u^k - u\|_V$$

Progressive (greedy-type) algorithm for solving $Au = b$

Greedy-type algorithm

Let $u_0 = 0$ and $U_0 = 0$. For $m \geq 1$,

1. Set $b_m = b - Au_{m-1}$.
2. Compute a correction $w_m \in \mathcal{M}$ of $u - u_{m-1}$ such that

$$\|Aw_m - b_m\|_W \leq (1 + \lambda_m) \min_{w \in \mathcal{M}} \|Aw - b_m\|_W$$

   with $\|Aw - b_m\|_W = \|u - u_{m-1} - w_m\|_V$. Use a gradient type algorithm with precision δ_m ($\lambda_m \geq \delta_m$).
3. Define a subspace U_m such that $U_m \supset U_{m-1}$ and $w_m \in U_m$, and compute $u_m \in U_m$ such that

$$\|u - u_m\|_V \leq (1 + \delta'_m) \min_{v \in U_m} \|u - v\|_V$$

Progressive (greedy-type) algorithm for solving $Au = b$

Define

$$\kappa_m = 1 - \beta_m \alpha_m^2$$

with

$$\beta_m = (1 - \delta'_m)^{-2}(1 + \lambda_m)^2 - 1 \quad \text{and} \quad \alpha_m = \frac{\|u - u_{m-1} - w_m\|_V}{\|w_m\|_V}$$

Convergence result

Assume

$$\delta'_m \to 0$$

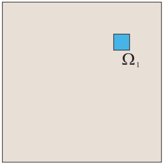
and

$$\kappa_m \geq 0, \quad \sum_m \kappa_m = +\infty$$

Then, the sequence $\{u_m\}_{m \geq 1}$ converges towards the solution u.

Illustration: advection-diffusion-reaction equation

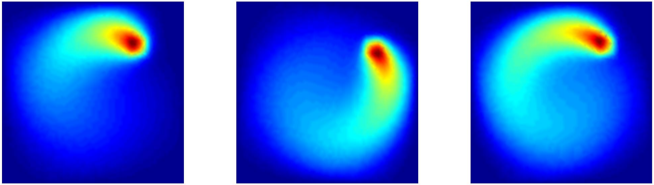
$$A(u) = -a_1 \Delta u + a_2 u + a_3 c \cdot \nabla u = I_{\Omega_1} \quad on \quad \Omega$$
$$u = 0 \quad on \quad \partial\Omega$$



Uncertain parameters

$$a_1 = \mu_1(1 + 0.2\xi_1), \quad a_2 = \mu_2(1 + 0.2\xi_2), \quad a_3 = \xi_3$$
$$\xi_i \in U(-1, 1), \ \Xi = (-1, 1)^3$$

Samples of the solution $u(x, \boldsymbol{\xi})$



Description of the solution method

$$A(u) = b, \quad A : V \to W', \quad V = W = \mathcal{V}_N \otimes \mathcal{S}_P$$

with $\mathcal{V}_N$ a finite element space and $\mathcal{S}_P = \mathbb{P}_4(-1,1) \otimes \mathbb{P}_4((-1,0),(0,1)) \otimes \mathbb{P}_4(-1,1)$.

Progressive construction of tensor approximation

- We choose the canonical norm

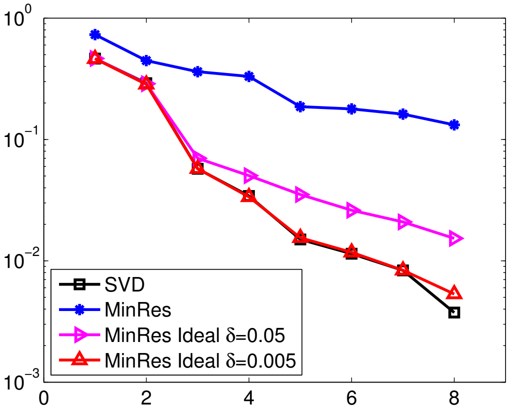
$$\|u\|_V^2 = \int_\Xi \|u(\xi)\|_{\mathcal{V}_N}^2 d\mu(\xi)$$

- We choose $\mathcal{M} = \mathcal{R}_1 = \{v \otimes \phi; v \in \mathcal{V}_N, \phi \in \mathcal{S}_P\}$, the set of rank-one elements
- Progressive construction of $u_m = u_{m-1} + w_m$ with $w_m \in \mathcal{M}$ an approximate solution of

$$\min_{w \in \mathcal{M}} \|A(u_{m-1} + w) - b\|_*$$

For $\|A(\cdot)\|_* = \|\cdot\|_V$, u_m is the ideal rank-m approximation (SVD).

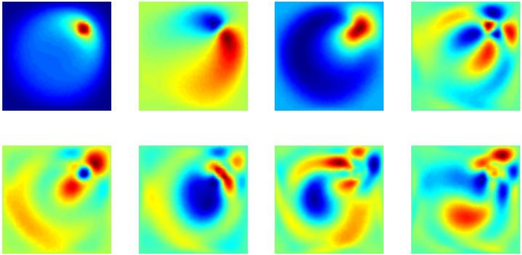
Convergence of progressive rank-m decompositions



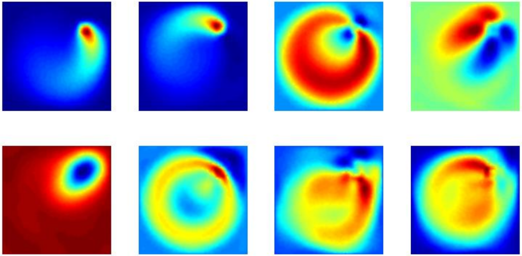
- For $\|A(\cdot)\|_* = \|A(\cdot)\|_V$, classical Minimal Residual formulation
- As $\delta \to 0$, convergence to the SVD
- For a fixed δ, it coincides with SVD up to precision δ.

8 first spatial modes v_i of the decomposition $u_m = \sum_{i=1}^{m} v_i \otimes \phi_i$

- Singular Value Decomposition of u

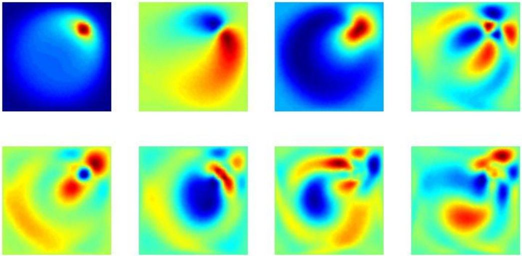


- Progressive approximation based on Minimal residual (canonical norm)

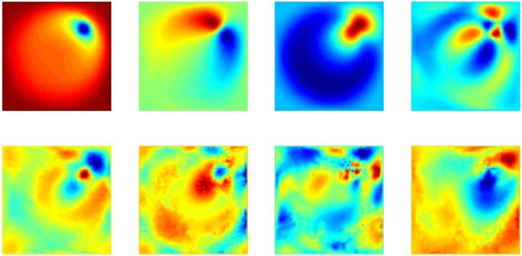


8 first spatial modes v_i of the decomposition $u_m = \sum_{i=1}^{m} v_i \otimes \phi_i$

- Singular Value Decomposition of u

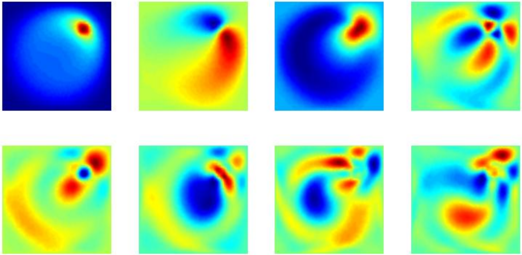


- Progressive approximation based on Ideal Minimal residual ($\delta = 0.05$)

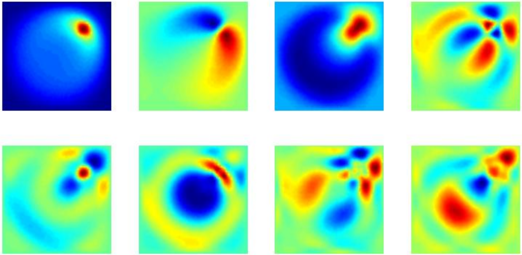


8 first spatial modes v_i of the decomposition $u_m = \sum_{i=1}^{m} v_i \otimes \phi_i$

- Singular Value Decomposition of u



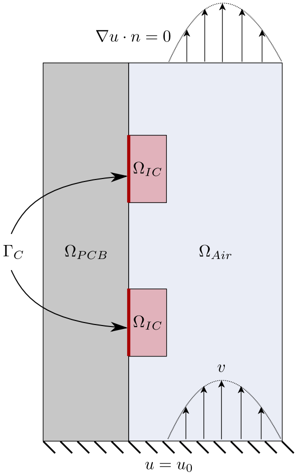
- Progressive approximation based on Ideal Minimal residual ($\delta = 0.005$)



Cooling of electronic components (benchmark OPUS : http://www.opus-project.fr)

$$-\nabla \cdot \kappa \nabla u + D v \cdot \nabla u = f$$

- $\kappa_{IC} \sim \log \mathcal{U}\left(0.2, 2\right) \rightsquigarrow$ diffusion coefficient
- $r \sim \log \mathcal{U}\left(0.1, 100\right) \rightsquigarrow$ thermal contact conductance
- $D \sim \log \mathcal{U}\left(5.10^{-4}, 10^{-2}\right) \rightsquigarrow$ advection intensity



Variable of interest :

$$\ell(u(\xi);\xi) = \int_{\Omega_{IC}} u(x,\xi)dx$$

Quantity of interest :

$$Lu = \mathbb{E}_\mu \left(\ell(u(\xi);\xi)|\kappa_{IC}\right)$$

Numerical results : $\delta = 0.5$, $\varepsilon = 10^{-3}$

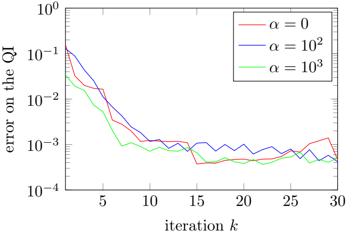


Figure: Evolution of the error on the quantity of interest during iteration process

Numerical results : $\delta = 0.5$, $\varepsilon = 10^{-3}$

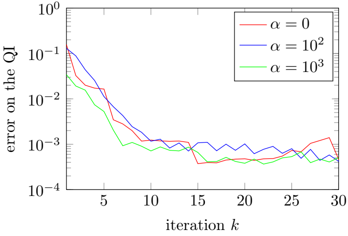


Figure: Evolution of the error on the quantity of interest during iteration process

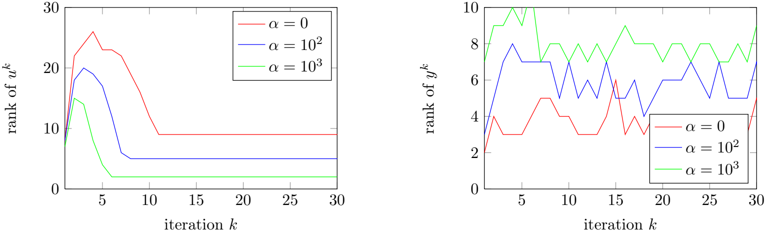


Figure: Evolution of the rank of u^k and y^k during iteration process

Numerical results : rank of the approximation

α \ ε	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
0	3	5	9	15	21
10^2	2	4	5	10	17
10^3	1	1	2	8	14
10^4	1	1	1	3	7
10^5	1	1	1	1	4
CMR	5	9	14	20	36

Figure: Final rank of the approximation u^k

Comparison with the canonical minimal residual method (CMR) :

$$\min_{v \in \mathcal{M}_r} \|Av - b\|_X \quad \rightsquigarrow \quad \text{error on the QI evaluated } \textit{a posteriori}$$

On-going works and open questions

- A priori results for greedy constructions of subspaces: optimal vs nested subspaces
- Strategy for general tree-based formats ?
- Optimal approximation with respect to quantities of interest (not norms)

References

A. Nouy.

Low-rank tensor approximation methods and Proper Generalized Decompositions.

Book Chapter. Preprint.

L. Giraldi, A. Nouy and G. Legrain.

Low-rank approximate inverse for preconditioning tensor-structured linear systems.

SIAM Journal on Scientific Computing, 36(4):A1850–A1870, 2014.

M. Billaud-Friess, A. Nouy, and O. Zahm.

A tensor approximation method based on ideal minimal residual formulations for the solution of high-dimensional problems.

ESAIM: Mathematical Modelling and Numerical Analysis, eFirst, 4 2014.