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

Marie Billaud-Friess, Anthony Nouy, Olivier Zahm

▶ **To cite this version:**

Marie Billaud-Friess, Anthony Nouy, Olivier Zahm. A tensor approximation method based on ideal minimal residual formulations for the solution of high-dimensional problems. ESAIM: Mathematical Modelling and Numerical Analysis, 2014, 48, pp.1777-1806. ⟨10.1051/m2an/2014019⟩. ⟨hal-00861914⟩

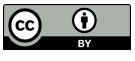
HAL Id: hal-00861914

https://hal.science/hal-00861914v1

Submitted on 6 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

arXiv:1304.6126v2 [math.NA] 26 Mar 2014

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

Marie Billaud-Friess† Anthony Nouy†‡ Olivier Zahm†

Abstract

In this paper, we propose a method for the approximation of the solution of high-dimensional weakly coercive problems formulated in tensor spaces using low-rank approximation formats. The method can be seen as a perturbation of a minimal residual method with a measure of the residual corresponding to the error in a specified solution norm. The residual norm can be designed such that the resulting low-rank approximations are optimal with respect to particular norms of interest, thus allowing to take into account a particular objective in the definition of reduced order approximations of high-dimensional problems. We introduce and analyze an iterative algorithm that is able to provide an approximation of the optimal approximation of the solution in a given low-rank subset, without any a priori information on this solution. We also introduce a weak greedy algorithm which uses this perturbed minimal residual method for the computation of successive greedy corrections in small tensor subsets. We prove its convergence under some conditions on the parameters of the algorithm. The proposed numerical method is applied to the solution of a stochastic partial differential equation which is discretized using standard Galerkin methods in tensor product spaces.

Introduction

Low-rank tensor approximation methods are receiving growing attention in computational science for the numerical solution of high-dimensional problems formulated in tensor spaces (see the recent surveys [30, 9, 29, 23] and monograph [24]). Typical problems include the solution of high-dimensional partial differential equations arising in stochastic calculus, or the solution of stochastic or parametric partial differential equations using functional approaches, where functions of multiple (random) parameters have to be approximated. These problems take the general form

$$A(u) = b, \quad u \in X = X_1 \otimes \ldots \otimes X_d, \tag{1}$$

where A is an operator defined on the tensor space X. Low-rank tensor methods then consist in searching an approximation of the solution u in a low-dimensional subset $\mathcal{S}_X$ of elements of X of the form

$$\sum_{i_1} \cdots \sum_{i_d} \alpha_{i_1 \ldots i_d} w^1_{i_1} \otimes \ldots \otimes w^d_{i_d}, \; w^\mu_{i_\mu} \in X_\mu, \tag{2}$$

where the set of coefficients $(\alpha_{i_1 \ldots i_d})$ possesses some specific structure. Classical low-rank tensor subsets include canonical tensors, Tucker tensors, Tensor Train tensors [40, 27], Hierarchical Tucker tensors [25] or more general tree-based Hierarchical Tucker tensors [18]. In practice, many tensors

*This work is supported by the ANR (French National Research Agency, grants ANR-2010- COSI-006 and ANR-2010-BLAN-0904)

†Ecole Centrale de Nantes, GeM UMR CNRS 6183, LUNAM Université, France.

‡Corresponding author (anthony.nouy@ec-nantes.fr).

arising in applications are observed to be efficiently approximable by elements of the mentioned subsets. Low-rank approximation methods are closely related to a priori model reduction methods in that they provide approximate representations of the solution on low-dimensional reduced bases $\{w_{i_1}^1 \otimes \ldots \otimes w_{i_d}^d\}$ that are not selected a priori.

The best approximation of $u \in X$ in a given low-rank tensor subset $\mathcal{S}_X$ with respect to a particular norm $\|\cdot\|_X$ in X is the solution of

$$\min_{v \in \mathcal{S}_X} \|u - v\|_X. \tag{3}$$

Low-rank tensor subsets are neither linear subspaces nor convex sets. However, they usually satisfy topological properties that make the above best approximation problem meaningful and allows the application of standard optimization algorithms [41, 14, 44]. Of course, in the context of the solution of high-dimensional problems, the solution u of problem (1) is not available, and the best approximation problem (3) cannot be solved directly. Tensor approximation methods then typically rely on the definition of approximations based on the residual of equation (1), which is a computable quantity. Different strategies have been proposed for the construction of low-rank approximations of the solution of equations in tensor format. The first family of methods consists in using classical iterative algorithms for linear or nonlinear systems of equations with low-rank tensor algebra (using low-rank tensor compression) for standard algebraic operations [31, 28, 34, 4]. The second family of methods consists in directly computing an approximation of u in $\mathcal{S}_X$ by minimizing some residual norm [5, 35, 12]:

$$\min_{v \in \mathcal{S}_X} \|Av - b\|_\star. \tag{4}$$

In the context of approximation, where one is interested in obtaining an approximation with a given precision rather than obtaining the best low-rank approximation, constructive greedy algorithms have been proposed that consist in computing successive corrections in a small low-rank tensor subset, typically the set of rank-one tensors [32, 1, 35]. These greedy algorithms have been analyzed in several papers [2, 6, 15, 17, 7, 19] and a series of improved algorithms have been introduced in order to increase the quality of suboptimal greedy constructions [37, 38, 33, 17, 21].

Although minimal residual based approaches are well founded, they generally provide low-rank approximations that can be very far from optimal approximations with respect to the natural norm $\|\cdot\|_X$, at least when using usual measures of the residual. If we are interested in obtaining an optimal approximation with respect to the norm $\|\cdot\|_X$, e.g. taking into account some particular quantity of interest, an ideal approach would be to define the residual norm such that

$$\|Av - b\|_\star = \|u - v\|_X,$$

where $\|\cdot\|_X$ is the desired solution norm, that corresponds to solve an ideally conditioned problem. Minimizing the residual norm would therefore be equivalent to solving the best approximation problem (3). However, the computation of such a residual norm is in general equivalent to the solution of the initial problem (1).

In this paper, we propose a method for the approximation of the ideal approach. This method applies to a general class of weakly coercive problems. It relies on the use of approximations $r_\delta(v)$ of the residual $r(v) = Av - b$ such that $\|r_\delta(v)\|_\star$ approximates the ideal residual norm $\|r(v)\|_\star = \|u - v\|_X$. The resulting method allows for the construction of low-rank tensor approximations which are quasi-optimal with respect to a norm $\|\cdot\|_X$ that can be designed according to some quantity of interest. We first introduce and analyze an algorithm for minimizing the approximate residual norm $\|r_\delta(v)\|_\star$ in a given subset $\mathcal{S}_X$. This algorithm can be seen as an extension of the algorithms introduced in [8, 10] to the context of nonlinear approximation in subsets $\mathcal{S}_X$. It consists in a perturbation of a gradient algorithm for minimizing in $\mathcal{S}_X$ the ideal residual norm

$\|r(v)\|_\star$, using approximations $r_\delta(v)$ of the residual $r(v)$. An ideal algorithm would consist in computing an approximation $r_\delta(v)$ such that

$$(1-\delta)\|u-v\|_X \leq \|r_\delta(v)\|_\star \leq (1+\delta)\|u-v\|_X, \tag{5}$$

for some precision δ, that requires the use of guaranteed error estimators. In the present paper, (5) is not exactly satisfied since we only use heuristic error estimates. However, these estimates seem to provide an acceptable measure of the error for the considered applications. The resulting algorithm can be interpreted as a preconditioned gradient algorithm with an implicit preconditioner that approximates the ideal preconditioner. Also, we propose a weak greedy algorithm for the adaptive construction of an approximation of the solution of problem (1), using the perturbed ideal minimal residual approach for the computation of greedy corrections. A convergence proof is provided under some conditions on the parameters of the algorithm.

The outline of the paper is as follows. In section 1, we introduce a functional framework for weakly coercive problems. In section 2, we briefly recall some definitions and basic properties of tensor spaces and low-rank tensor subsets. In section 3, we present a natural minimal residual based method for the approximation in a nonlinear subset $\mathcal{S}_X$, and we analyze a simple gradient algorithm in $\mathcal{S}_X$. We discuss the conditioning issues that restrict the applicability of such algorithms when usual residual norms are used, and the interest of using an ideal measure of the residual. In section 4, we introduce the perturbed ideal minimal residual approach. A gradient-type algorithm is introduced and analyzed and we prove the convergence of this algorithm towards a neighborhood of the best approximation in $\mathcal{S}_X$. Practical computational aspects are detailed in section 5. In section 6, we analyze a weak greedy algorithm using the perturbed ideal minimal residual method for the computation of greedy corrections. In section 7, a detailed numerical example will illustrate the proposed method. The example is a stochastic reaction-advection-diffusion problem which is discretized using Galerkin stochastic methods. In particular, this example will illustrate the possibility to introduce norms that are adapted to some quantities of interest and the ability of the method to provide (quasi-)best low-rank approximations in that context.

1 Functional framework for weakly coercive problems

1.1 Notations

For a given Hilbert space H, we denote by $\langle\cdot,\cdot\rangle_H$ the inner product in H and by $\|\cdot\|_H$ the associated norm. We denote by H' the topological dual of H and by $\langle\cdot,\cdot\rangle_{H',H}$ the duality pairing between H and H'. For $v\in H$ and $\varphi\in H'$, we denote $\varphi(v)=\langle\varphi,v\rangle_{H',H}$. We denote by $R_H : H\to H'$ the Riesz isomorphism defined by

$$\langle v,w\rangle_H = \langle v,R_H w\rangle_{H,H'} = \langle R_H v,w\rangle_{H',H} = \langle R_H v,R_H w\rangle_{H'} \quad \forall v,w\in H.$$

1.2 Weakly coercive problems

We denote by X (resp. Y) a Hilbert space equipped with inner product $\langle\cdot,\cdot\rangle_X$ (resp. $\langle\cdot,\cdot\rangle_Y$) and associated norm $\|\cdot\|_X$ (resp. $\|\cdot\|_Y$). Let $a : X\times Y\to\mathbb{R}$ be a bilinear form and let $b\in Y'$ be a continuous linear form on Y. We consider the variational problem: find $u\in X$ such that

$$a(u,v)=b(v) \quad \forall v\in Y. \tag{6}$$

We assume that a is continuous and weakly coercive, that means that there exist constants α and β such that

$$\sup_{v\in X}\sup_{w\in Y}\frac{a(v,w)}{\|v\|_X\|w\|_Y}=\beta<+\infty, \tag{7}$$

$$\inf_{v\in X}\sup_{w\in Y}\frac{a(v,w)}{\|v\|_X\|w\|_Y}=\alpha>0, \tag{8}$$

and

$$\sup_{v\in X}\frac{a(v,w)}{\|v\|_X}>0 \quad \forall w\neq 0 \text{ in } Y. \tag{9}$$

We introduce the linear continuous operator $A : X \to Y'$ such that for all $(v, w) \in X \times Y$,

$$a(v,w)=\langle Av,w\rangle_{Y',Y}.$$

We denote by $A^* : Y \to X'$ the adjoint of A, defined by

$$\langle Av,w\rangle_{Y',Y}=\langle v,A^*w\rangle_{X,X'} \quad \forall (v,w)\in X\times Y.$$

Problem (6) is therefore equivalent to find $u \in X$ such that

$$Au=b. \tag{10}$$

Properties (7),(8) and (9) imply that A is a norm-isomorphism from X to Y' such that for all $v \in X$,

$$\alpha\|v\|_X\leq\|Av\|_{Y'}\leq\beta\|v\|_X \tag{11}$$

ensuring the well-posedness of problem (10)[13]. The norms of A and its inverse A^{-1} are such that $\|A\|_{X\to Y'}=\beta$ and $\|A^{-1}\|_{Y'\to X}=\alpha^{-1}$. Then, the condition number of the operator A is

$$\kappa(A)=\|A\|_{X\to Y'}\|A^{-1}\|_{Y'\to X}=\frac{\beta}{\alpha}\geq 1.$$

2 Approximation in low-rank tensor subsets

2.1 Hilbert tensor spaces

We here briefly recall basic definitions on Hilbert tensor spaces (see [24]). We consider Hilbert spaces X_μ, $1 \leq \mu \leq d$, equipped with norms $\|\cdot\|_{X_\mu}$ and associated inner products $\langle\cdot,\cdot\rangle_\mu$[1]. We denote by $\otimes_{\mu=1}^d v^\mu = v^1 \otimes \ldots \otimes v^d$, $v^\mu \in X_\mu$, an elementary tensor. We then define the algebraic tensor product space as the linear span of elementary tensors:

$$\underset{a}{\bigotimes_{\mu=1}^{d}} X_\mu=\operatorname{span}\{\otimes_{\mu=1}^d v^\mu \ : \ v^\mu\in X_\mu, 1\leq\mu\leq d\}.$$

A Hilbert tensor space X equipped with the norm $\|\cdot\|_X$ is then obtained by the completion with respect to $\|\cdot\|_X$ of the algebraic tensor space, i.e.

$$X=\overline{\underset{a}{\bigotimes_{\mu=1}^{d}} X_\mu}^{\|\cdot\|_X}=\underset{\|\cdot\|_X}{\bigotimes_{\mu=1}^{d}} X_\mu.$$

Note that for finite dimensional tensor spaces, the resulting space X is independent of the choice of norm and coincides with the normed algebraic tensor space.

A natural inner product on X is induced by inner products $\langle\cdot,\cdot\rangle_\mu$ in X_μ, $1 \leq \mu \leq d$. It is defined for $v=\otimes_{\mu=1}^d v^\mu$ and $w=\otimes_{\mu=1}^d w^\mu$ by

$$\langle v,w\rangle_X=\prod_{\mu=1}^d\langle v^\mu,w^\mu\rangle_\mu$$

and extended by linearity on the whole algebraic tensor space. This inner product is called the *induced (or canonical) inner product* and the associated norm the *induced (or canonical) norm.*

[1] e.g. $X_\mu=\mathbb{R}^{n_\mu}$ equipped with the Euclidian norm, or $X_\mu=H_0^k(\Omega_\mu)$, $k\geq 0$, a Sobolev space of functions defined on a domain Ω_μ.

2.2 Classical low-rank tensor subsets

Low-rank tensor subsets $\mathcal{S}_X$ of a tensor space $X = {}_{\|\cdot\|}\bigotimes_{\mu=1}^{d} X_\mu$ are subsets of the algebraic tensor space ${}_a\bigotimes_{\mu=1}^{d} X_\mu$, which means that elements $v \in \mathcal{S}_X$ can be written under the form

$$v = \sum_{i_1 \in I_1} \cdots \sum_{i_d \in I_d} \alpha_{i_1,\ldots,i_d} \bigotimes_{\mu=1}^{d} v_{i_\mu}^{\mu}, \tag{12}$$

where $\alpha = (\alpha_i)_{i \in I} \in \mathbb{R}^I$, with $I := I_1 \times \ldots \times I_d$, is a set of real coefficients that possibly satisfy some constraints, and $(v_{i_\mu}^{\mu})_{i_\mu \in I_\mu} \in (X_\mu)^{I_\mu}$, for $1 \leq \mu \leq d$, is a set of vectors that also possibly satisfy some constraints (e.g. orthogonality).

Basic low-rank tensor subsets are the set of tensors with canonical rank bounded by r:

$$\mathcal{R}_r(X) = \left\{ v = \sum_{i=1}^{r} \otimes_{\mu=1}^{d} v_i^{\mu} \ : \ v_i^{\mu} \in X_\mu \right\},$$

and the set of Tucker tensors with multilinear rank bounded by $r = (r_1, \ldots, r_d)$:

$$\mathcal{T}_r(X) = \left\{ v = \sum_{i_1=1}^{r_1} \cdots \sum_{i_d=1}^{r_d} \alpha_{i_1,\ldots,i_d} \otimes_{\mu=1}^{d} v_{i_\mu}^{\mu} \ : \ v_{i_\mu}^{\mu} \in X_\mu, \ \alpha_{i_1,\ldots,i_d} \in \mathbb{R} \right\}$$

Other low-rank tensor subsets have been recently introduced, such as Tensor Train tensors [40, 27] or more general tree-based Hierarchical Tucker tensors [25, 18], these tensor subsets corresponding to a form (12) with a particular structure of tensor α. Note that for the case $d = 2$, all the above tensor subsets coincide.

Remark 2.1. *From a numerical point of view, the approximate solution of the variational problem* (6) *requires an additional discretization which consists in introducing an approximation space* $\widetilde{X} = \otimes_{\mu=1}^{d} \widetilde{X}_\mu$*, where the* $\widetilde{X}_\mu \subset X_\mu$ *are finite dimensional approximation spaces (e.g. finite element spaces). Then, approximations are searched in low-rank tensor subsets* $\mathcal{S}_X$ *of* $\widetilde{X}$ *(e.g.* $\mathcal{R}_r(\widetilde{X})$ *or* $\mathcal{T}_r(\widetilde{X})$*), thus introducing two levels of discretizations. In the following, we adopt a general point of view where* X *can either denote an infinite dimensional space, an approximation space obtained after the discretization of the variational problem, or even finite dimensional Euclidian spaces for problems written in an algebraic form.*

2.3 Best approximation in tensor subsets

Low-rank tensor approximation methods consist in computing an approximation of a tensor $u \in X$ in a suitable low-rank subset $\mathcal{S}_X$ of X. The best approximation of u in $\mathcal{S}_X$ is defined by

$$\min_{v \in \mathcal{S}_X} \|u - v\|_X. \tag{13}$$

The previously mentioned classical tensor subsets are neither linear subspaces nor convex sets. However, they usually satisfy properties that give sense to the above best approximation problem. We consider the case that $\mathcal{S}_X$ satisfies the following properties:

$$\mathcal{S}_X \text{ is weakly closed (or simply closed in finite dimension),} \tag{14}$$

$$\mathcal{S}_X \subset \gamma \mathcal{S}_X \text{ for all } \gamma \in \mathbb{R}. \tag{15}$$

Property (15) is satisfied by all the classical tensor subsets mentioned above (canonical tensors, Tucker and tree-based Hierarchical Tucker tensors). Property (14) ensures the existence of solutions to the best approximation problem (13). This property, under some suitable conditions on the norm $\|\cdot\|_X$ (which is naturally satisfied in finite dimension), is verified by most tensor subsets

used for approximation (e.g. the set of tensors with bounded canonical rank for $d = 2$, the set of elementary tensors $\mathcal{R}_1$ for $d \geq 2$ [15], the sets of Tucker or tree-based Hierarchical Tucker tensors [16]).

We then introduce the set-valued map $\Pi_{\mathcal{S}_X} : X \to 2^{\mathcal{S}_X}$ that associates to an element $u \in X$ the set of best approximations of u in $\mathcal{S}_X$:

$$\Pi_{\mathcal{S}_X}(u) = \arg\min_{v \in \mathcal{S}_X} \|u - v\|_X. \tag{16}$$

Note that if $\mathcal{S}_X$ were a closed linear subspace or a closed convex set of X, then $\Pi_{\mathcal{S}_X}(u)$ would be a singleton and $\Pi_{\mathcal{S}_X}$ would coincide with the classical definition of the metric projection on $\mathcal{S}_X$. Property (15) still implies the following property of projections: for all $v \in X$ and for all $w \in \Pi_{\mathcal{S}_X}(v)$,

$$\|v - w\|_X^2 = \|v\|_X^2 - \|w\|_X^2 \quad \text{with} \quad \|w\|_X = \sigma(v; \mathcal{S}_X) = \max_{z \in \mathcal{S}_X} \frac{\langle v, z \rangle_X}{\|z\|_X}. \tag{17}$$

$\Pi_{\mathcal{S}_X}(v)$ is therefore a subset of the sphere of radius $\sigma(v; \mathcal{S}_X)$ in X. In the following, we will use the following abuse of notation: for a subset $S \subset X$ and for $w \in X$, we define

$$\|S - w\|_X := \sup_{v \in S} \|v - w\|_X$$

With this convention, we have $\|\Pi_{\mathcal{S}_X}(v)\|_X = \sigma(v; \mathcal{S}_X)$ and

$$\|\Pi_{\mathcal{S}_X}(v) - v\|_X^2 = \|v\|_X^2 - \|\Pi_{\mathcal{S}_X}(v)\|_X^2. \tag{18}$$

3 Minimal residual based approximation

We now consider that problem (10) is formulated in tensor Hilbert spaces $X = {}_{\|\cdot\|_X}\bigotimes_{\mu=1}^d X_\mu$ and $Y = {}_{\|\cdot\|_Y}\bigotimes_{\mu=1}^d Y_\mu$. The aim is here to find an approximation of the solution u of problem (10) in a given tensor subset $\mathcal{S}_X \subset X$.

3.1 Best approximation with respect to residual norms

Since the solution u of problem (10) is not available, the best approximation problem (13) cannot be solved directly. However, tensor approximations can be defined using the residual of equation (10), which is a computable information. An approximation of u in $\mathcal{S}_X$ is then defined by the minimization of a residual norm:

$$\min_{v \in \mathcal{S}_X} \|Av - b\|_{Y'} = \min_{v \in \mathcal{S}_X} \|A(v - u)\|_{Y'}. \tag{19}$$

Assuming that we can define a tangent space $T_v(\mathcal{S}_X)$ to $\mathcal{S}_X$ at $v \in \mathcal{S}_X$, the stationarity condition of functional $J : v \mapsto \|A(v - u)\|_{Y'}^2$ at $v \in \mathcal{S}_X$ is

$$\langle J'(v), \delta v \rangle_{X',X} = 0 \quad \forall \delta v \in T_v(\mathcal{S}_X),$$

or equivalently, noting that the gradient of J at v is $J'(v) = A^* R_Y^{-1}(Av - b) \in X'$,

$$\langle Av - b, A\delta v \rangle_{Y'} = 0 \quad \forall \delta v \in T_v(\mathcal{S}_X).$$

3.2 Ideal choice of the residual norm

When approximating u in $\mathcal{S}_X$ using (19), the obtained approximation depends on the choice of the residual norm. If we want to find a best approximation of u with respect to the norm $\|\cdot\|_X$, then the residual norm should be chosen [8, 10] such that

$$\|A(v-u)\|_{Y'} = \|v-u\|_X \quad \forall v \in X,$$

or equivalently such that the following relation between inner products holds:

$$\langle v, w\rangle_X = \langle Av, Aw\rangle_{Y'} \quad \forall v, w \in X. \tag{20}$$

This implies

$$\langle v, w\rangle_X = \langle Av, R_Y^{-1}Aw\rangle_{Y',Y} = \langle v, R_X^{-1}A^*R_Y^{-1}Aw\rangle_X,$$

for all $v, w \in X$, and therefore, by identification,

$$I_X = R_X^{-1}A^*R_Y^{-1}A \Leftrightarrow R_Y = AR_X^{-1}A^* \Leftrightarrow R_X = A^*R_Y^{-1}A. \tag{21}$$

Also, since

$$\begin{aligned}\langle v, w\rangle_Y &= \langle R_Y v, w\rangle_{Y',Y} = \langle AR_X^{-1}A^*v, w\rangle_{Y',Y} \\ &= \langle R_X^{-1}A^*v, A^*w\rangle_{X,X'} = \langle A^*v, A^*w\rangle_{X'}\end{aligned}$$

for all $v, w \in Y$, we also have that (20) is equivalent to the following relation:

$$\langle v, w\rangle_Y = \langle A^*v, A^*w\rangle_{X'} \quad \forall v, w \in Y. \tag{22}$$

Note that (20) and (22) respectively impose

$$\|v\|_X = \|Av\|_{Y'} \text{ and } \|w\|_Y = \|A^*w\|_{X'}. \tag{23}$$

This choice implies that the weak coercivity and continuity constants are such that $\alpha = \beta = 1$, and therefore

$$\kappa(A) = 1,$$

meaning that problem (10) is ideally conditioned.

In practice, we will first define the inner product $\langle\cdot,\cdot\rangle_X$ and the other inner product $\langle\cdot,\cdot\rangle_Y$ will be deduced from (22).

Example 3.1. *Consider that $X = Y$ and let $A = B + C$ with B a symmetric coercive and continuous operator and C a skew-symmetric operator. We equip X with inner product $\langle v, w\rangle_X = \langle Bv, w\rangle_{X',X}$, which corresponds to $R_X = B$. Therefore,*

$$\|v\|_Y^2 = \|A^*v\|_{X'}^2 = \|Bv\|_{X'}^2 + \|Cv\|_{X'}^2 = \|v\|_X^2 + \|Cv\|_{X'}^2.$$

$\|v\|_Y$ corresponds to the graph norm of the skew-symmetric part C of the operator A. When $C = 0$, we simply have $\|v\|_Y^2 = \|v\|_X^2$.

Example 3.2 (Finite dimensional problem)**.** *Consider the case of finite dimensional tensor spaces $X = Y = \mathbb{R}^{n_1\times\cdots\times n_d}$, e.g. after a discretization step for the solution of a high-dimensional partial differential equation. The duality pairings are induced by the standard canonical inner product. We can choose for $\langle\cdot,\cdot\rangle_X$ the canonical inner product on $\mathbb{R}^{n_1\times\cdots\times n_d}$, which corresponds to $R_X = I_X$, the identity on X. Then, inner product on Y is defined by relation* (22)*, which implies*

$$\langle v, w\rangle_Y = \langle A^*v, A^*w\rangle_X \quad \textit{and} \quad R_Y = AA^*.$$

3.3 Gradient-type algorithm

For solving (19), we consider the following basic gradient-type algorithm: letting $u^0 = 0$, we construct a sequence $\{u^k\}_{k\geq 0}$ in $\mathcal{S}_X$ and a sequence $\{y^k\}_{k\geq 0}$ in Y defined for $k \geq 0$ by

$$\begin{cases} y^k = R_Y^{-1}(Au^k - b) \\ u^{k+1} \in \Pi_{\mathcal{S}_X}(u^k - \rho R_X^{-1} A^* y^k) \end{cases} \tag{24}$$

with $\rho > 0$. Equations (24) yield

$$u^{k+1} \in \Pi_{\mathcal{S}_X}(u + B_\rho(u^k - u)),$$

with $B_\rho = I_X - \rho R_X^{-1} A^* R_Y^{-1} A$ a symmetric operator from X to X. For all $v \in X$,

$$\frac{\langle B_\rho v, v\rangle_X}{\|v\|_X^2} = 1 - \rho \frac{\|Av\|_{Y'}^2}{\|v\|_X^2}.$$

Here, we assume that $\|\cdot\|_X$ and $\|\cdot\|_Y$ do not necessarily satisfy the relation (23) (i.e. $\frac{\alpha}{\beta} \neq 1$). From (11), we deduce that the eigenvalues of B_ρ are in the interval $[1-\rho\beta^2, 1-\rho\alpha^2]$. The spectral radius of B_ρ is therefore bounded by

$$\gamma(\rho) = \max\{|1-\rho\beta^2|, |1-\rho\alpha^2|\}.$$

Proposition 3.3. *Assuming $\gamma(\rho) < 1/2$, the sequence $\{u^k\}_{k\geq 1}$ defined by (24) is such that*

$$\|u^k - u\|_X \leq (2\gamma)^k \|u^0 - u\|_X + \frac{1}{1-2\gamma}\|u - \Pi_{\mathcal{S}_X}(u)\|_X \tag{25}$$

and

$$\limsup_{k\to\infty} \|u^k - u\|_X \leq \frac{1}{1-2\gamma}\|u - \Pi_{\mathcal{S}_X}(u)\|_X \tag{26}$$

Proof. Denoting $v^k = u^k - u$, we have

$$\begin{aligned} \|u^{k+1} - u\|_X &\leq \|\Pi_{\mathcal{S}_X}(u + B_\rho v^k) - u\|_X \\ &\leq \|\Pi_{\mathcal{S}_X}(u + B_\rho v^k) - (u + B_\rho v^k)\|_X + \|B_\rho v^k\|_X \\ &\leq \|w - (u + B_\rho v^k)\|_X + \|B_\rho v^k\|_X \end{aligned}$$

for all $w \in \mathcal{S}_X$. In particular, this inequality is true for all $w \in \Pi_{\mathcal{S}_X}(u)$, and therefore, taking the supremum over all $w \in \Pi_{\mathcal{S}_X}(u)$, we obtain

$$\begin{aligned} \|u^{k+1} - u\|_X &\leq \|\Pi_{\mathcal{S}_X}(u) - (u + B_\rho v^k)\|_X + \|B_\rho v^k\|_X \\ &\leq \|\Pi_{\mathcal{S}_X}(u) - u\|_X + 2\|B_\rho v^k\|_X \end{aligned}$$

Since $\|B_\rho v\|_X \leq \gamma\|v\|_X$ for all $v \in X$ and since $2\gamma < 1$, we have

$$\begin{aligned} \|u^{k+1} - u\|_X &\leq \|\Pi_{\mathcal{S}_X}(u) - u\|_X + 2\gamma\|u - u^k\|_X \\ &\leq (2\gamma)^{k+1}\|u^0 - u\|_X + \frac{1-(2\gamma)^{k+1}}{1-2\gamma}\|u - \Pi_{\mathcal{S}_X}(u)\|_X \end{aligned}$$

from which we deduce (25) and (26). □

The condition $\gamma(\rho) < 1/2$ imposes $\frac{\beta}{\alpha} < \sqrt{3}$ and $\rho \in (\frac{1}{2\alpha^2}, \frac{3}{2\beta^2})$. The condition $\frac{\beta}{\alpha} < \sqrt{3}$ is a very restrictive condition which is in general not satisfied without an excellent preconditioning of the operator A.

However, with the ideal choice of norms introduced in the previous section (equation (23)), we have $\alpha = \beta = 1$ and $B_\rho = (1-\rho)I_X$. That means that the problem is ideally conditioned and we have convergence for all $\rho \in [\frac{1}{2}, \frac{3}{2}]$ towards a neighborhood of $\Pi_{\mathcal{S}_X}(u)$ of size $\frac{2\gamma}{1-2\gamma}\|u - \Pi_{\mathcal{S}_X}(u)\|_X$ with $\gamma = |1-\rho|$.

Corollary 3.4. *Assume that* (23) *is satisfied. Then, if* $\rho \in [\frac{1}{2}, \frac{3}{2}]$, *the sequence* $\{u^k\}_{k\geq 1}$ *defined by* (24) *verifies* (25) *and* (26) *with* $\gamma(\rho) = |1-\rho|$. *Moreover, if* $\rho = 1$, *then* $u^1 \in \Pi_{\mathcal{S}_X}(u)$, *which means that the algorithm converges in one iteration for any initialization* u^0.

4 Perturbation of the ideal approximation

We now consider that function spaces X and Y are equipped with norms satisfying the ideal condition

$$\|Av\|_{Y'} = \|v\|_X \quad \forall v \in X. \tag{27}$$

The solution of problem (19) using this ideal choice of norms is therefore equivalent to the best approximation problem (13), i.e.

$$\min_{v\in\mathcal{S}_X} \|Av - b\|_{Y'} = \min_{v\in\mathcal{S}_X} \|v - u\|_X. \tag{28}$$

Unfortunately, the computation of the solution of (28) would require the solution of the initial problem. We here propose to introduce a computable perturbation of this ideal approach.

4.1 Approximation of the ideal approach

Following the idea of [8], the problem (28) is replaced by the following problem:

$$\min_{v\in\mathcal{S}_X} \|\Lambda^\delta(R_Y^{-1}(Av - b))\|_Y, \tag{29}$$

where $\Lambda^\delta : Y \to Y$ is a mapping that provides an approximation $\Lambda^\delta(r)$ of the residual $r = R_Y^{-1}(Av - b) \in Y$ with a controlled relative precision $\delta > 0$, i.e. $\|\Lambda^\delta(r) - r\|_Y \leq \delta\|r\|_Y$. We will then assume that the mapping Λ^δ is such that:

$$\|\Lambda^\delta(y) - y\|_Y \leq \delta\|y\|_Y, \quad \forall y \in \mathcal{D}_Y = \left\{R_Y^{-1}(Av - b); v \in \mathcal{S}_X\right\}. \tag{30}$$

As we will see in the following algorithm, it is sufficient for Λ^δ to well approximate residuals that are in the subset $\mathcal{D}_Y$ whose content depends on the chosen subset $\mathcal{S}_X$ and on the operator and right-hand side of the problem.

4.2 Quasi-optimal approximations in $\mathcal{S}_X$

Here we consider the case where we are not able to solve the best approximation problem in $\mathcal{S}_X$ exactly, because there is no available algorithm for computing a global optimum, or because the algorithm has been stopped at a finite precision (see section 5.1 for practical comments). We introduce a set of quasi-optimal approximations $\Pi^\eta_{\mathcal{S}_X}(u) \subset \mathcal{S}_X$ such that

$$\|u - \Pi^\eta_{\mathcal{S}_X}(u)\|_X \leq \eta\|u - \Pi_{\mathcal{S}_X}(u)\|_X \quad (\eta \geq 1). \tag{31}$$

Remark 4.1. *Note that by introducing this new perturbation, we are able to remove the assumption that* $\mathcal{S}_X$ *is closed and to handle the case where the problem* (28) *does not have a solution, i.e.* $\Pi_{\mathcal{S}_X}(u) = \emptyset$. *In this case, we have to replace* $\|u - \Pi_{\mathcal{S}_X}(u)\|_X$ *by* $\inf_{w\in\mathcal{S}_X} \|u - w\|_X$ *in equation* (31).

Remark 4.2. *Note that if* $\mathcal{S}_X$ *denotes a low-rank subset of an infinite dimensional space* X, *additional approximations have to be introduced from a numerical point of view (see remark 2.1). These additional approximations could be also considered as a perturbation leading to quasi-optimal approximations, where* η *takes into account the approximation errors. In numerical examples, we will not adopt this point of view and we will consider* X *as the approximation space and the approximate solution in* X *of the variational problem will serve as a reference solution.*

4.3 Perturbed gradient-type algorithm

For solving (29), we now introduce an algorithm which can be seen as a perturbation of the ideal gradient-type algorithm (24) introduced in section 3.3. Letting $u^0 = 0$, we construct a sequence $\{u^k\}_{k\geq 0} \subset \mathcal{S}_X$ and a sequence $\{y^k\}_{k\geq 0} \subset Y$ defined for $k \geq 0$ by

$$\begin{cases} y^k = \Lambda^\delta(R_Y^{-1}(Au^k - b)) \\ u^{k+1} \in \Pi^\eta_{\mathcal{S}_X}(u^k - R_X^{-1}A^*y^k) \end{cases} \tag{32}$$

Proposition 4.3. *Assume* (27), (30), *and* (31), *with* $\delta(1+\eta) < 1$. *Then, the sequence* $\{u^k\}_{k\geq 1}$ *defined by* (32) *is such that*

$$\|u^k - u\|_X \leq ((1+\eta)\delta)^k \|u^0 - u\|_X + \frac{\eta}{1-\delta(1+\eta)} \|u - \Pi_{\mathcal{S}_X}(u)\|_X. \tag{33}$$

Proof. Equation (32) can also be written

$$u^{k+1} \in \Pi^\eta_{\mathcal{S}_X}(u + B^\delta(u^k - u))$$

with $B^\delta(v) = v - R_X^{-1}A^*\Lambda^\delta(R_Y^{-1}A(v))$. Denoting $v^k = u^k - u$, and following the proof of Proposition 3.3, we obtain

$$\begin{aligned} \|u^{k+1} - u\|_X &\leq \|\Pi^\eta_{\mathcal{S}_X}(u + B^\delta v^k) - (u + B^\delta v^k)\|_X + \|B^\delta v^k\|_X \\ &\leq \eta\|\Pi_{\mathcal{S}_X}(u) - (u + B^\delta v^k)\|_X + \|B^\delta v^k\|_X \\ &\leq \eta\|\Pi_{\mathcal{S}_X}(u) - u\|_X + (1+\eta)\|B^\delta v^k\|_X \end{aligned}$$

Moreover, using (27) and (21), we have

$$\begin{aligned} \|B^\delta v^k\|_X &= \|v^k - R_X^{-1}A^*\Lambda^\delta(R_Y^{-1}Av^k)\|_X \\ &= \|Av^k - AR_X^{-1}A^*\Lambda^\delta(R_Y^{-1}Av^k)\|_{Y'} \\ &= \|R_Y^{-1}Av^k - \Lambda^\delta(R_Y^{-1}Av^k)\|_Y. \end{aligned}$$

Noting that $R_Y^{-1}Av^k = R_Y^{-1}(Au^k - b)$ belongs to the subset $\mathcal{D}_Y$, we deduce from assumption (30) and equation (27) that

$$\|B^\delta v^k\|_X \leq \delta\|R_Y^{-1}Av^k\|_Y = \delta\|v^k\|_X.$$

Denoting $\delta_\eta = \delta(1+\eta) < 1$, we finally have

$$\begin{aligned} \|u^{k+1} - u\|_X &\leq \eta\|\Pi_{\mathcal{S}_X}(u) - u\|_X + \delta_\eta\|u^k - u\|_X \\ &\leq \delta_\eta^{k+1}\|u^0 - u\|_X + \eta\frac{1-\delta_\eta^{k+1}}{1-\delta_\eta}\|u - \Pi_{\mathcal{S}_X}(u)\|_X, \end{aligned}$$

from which we deduce (33). □

Comments We note the sequence converges towards a neighborhood of $\Pi_{\mathcal{S}_X}(u)$ whose size is $\frac{\eta-1+(1+\eta)\delta}{1-(1+\eta)\delta}\|u - \Pi_{\mathcal{S}_X}(u)\|_X$. Indeed, (33) implies that

$$\|u - \Pi_{\mathcal{S}_X}(u)\|_X \leq \|u - u^k\|_X \leq (1+\gamma_k)\|u - \Pi_{\mathcal{S}_X}(u)\|_X, \tag{34}$$

with $\limsup_{k\to\infty}\gamma_k \leq \frac{\eta-1+(1+\eta)\delta}{1-(1+\eta)\delta}$. Therefore, the sequence tends to provide a good approximation of the best approximation of u in $\mathcal{S}_X$, and the parameters δ and η control the quality of this approximation. Moreover, equation (33) indicates that the sequence converges quite rapidly to this neighborhood. Indeed, in the first iterations, when the error $\|u - u^k\|_X$ is dominated by the

first term $((1+\eta)\delta)^k\|u-u^0\|_X$, the algorithm has at least a linear convergence with convergence rate $(1+\eta)\delta$ (note that for $\eta \approx 1$, the convergence rate is very high for small δ). Once both error terms are balanced, the error stagnates at the value $\frac{\eta}{1-(1+\eta)\delta}\|u-\Pi_{\mathcal{S}_X}(u)\|_X$. Note that when $\delta \to 0$, we recover an ideal algorithm with a convergence in only one iteration to an element of the set $\Pi^{\eta}_{\mathcal{S}_X}(u)$ of quasi-best approximations of u in $\mathcal{S}_X$.

Remark 4.4. *Even if $\mathcal{S}_X$ is chosen as a subset of low-rank tensors, the subset $\mathcal{D}_Y$ defined in* (30) *possibly contains tensors with high ranks (or even tensors with full rank) that are not easy to approximate with a small precision δ using low-rank tensor representations. However, the algorithm only requires to well approximate the sequence of residuals $\{R_Y^{-1}(Au^k-b)\}_{k\geq 0} \subset \mathcal{D}_Y$, which may be achievable in practical applications.*

4.4 Error indicator

Along the iterations of algorithm (32), an estimation of the true error $\|u-u^k\|_X$ can be simply obtained by evaluating the norm $\|y^k\|_Y$ of the iterate $y^k = \Lambda^\delta(r^k)$ with $r^k = R_Y^{-1}(Au^k-b)$. Indeed, from property (30), we have

$$(1-\delta)\|y\|_Y \leq \|\Lambda^\delta(y)\|_Y \leq (1+\delta)\|y\|_Y, \tag{35}$$

for all $y \in \mathcal{D}_Y$. Therefore, noting that $r^k \in \mathcal{D}_Y$ and $\|r^k\|_Y = \|A(u-u^k)\|_{Y'} = \|u-u^k\|_X$, we obtain

$$(1-\delta)\|u-u^k\|_X \leq \|y^k\|_Y \leq (1+\delta)\|u-u^k\|_X. \tag{36}$$

In other words,

$$\epsilon^k = \frac{1}{1-\delta}\|y^k\|_Y \tag{37}$$

provides an error indicator of the true error $\|u-u^k\|_X$ with an effectivity index $\tau^k = \frac{\epsilon^k}{\|u-u^k\|_X} \in (1, \frac{1+\delta}{1-\delta})$, which is very good for small δ.

Moreover, if Λ^δ is an orthogonal projection onto some subspace $Y^\delta \subset Y$, we easily obtain the following improved lower and upper bounds:

$$\sqrt{1-\delta^2}\|u-u^k\|_X \leq \|y^k\|_Y \leq \|u-u^k\|_X, \tag{38}$$

that means that the following improved error estimator can be chosen:

$$\hat{\epsilon}^k = \frac{1}{\sqrt{1-\delta^2}}\|y^k\|_Y, \tag{39}$$

with effectivity index $\hat{\tau}^k = \frac{\hat{\epsilon}^k}{\|u-u^k\|_X} \in (1, \frac{1}{\sqrt{1-\delta^2}})$.

5 Computational aspects

5.1 Best approximation in tensor subsets

We here discuss the available algorithms for computing an element in $\Pi_{\mathcal{S}_X}(v)$, that means for solving

$$\min_{w\in\mathcal{S}_X} \|v-w\|_X, \tag{40}$$

where v is a given tensor in the tensor space $X = {}_{\|\cdot\|_X}\bigotimes_{\mu=1}^d X_\mu$ equipped with norm $\|\cdot\|_X$, and where $\mathcal{S}_X$ is a given tensor subset. Note that except for the case where $d=2$ and $\|\cdot\|_X$ is the induced (canonical) norm, the computation of a global optimum is still an open problem.

Canonical norm, $d = 2$**.** For the case $d = 2$, we first note that all classical low-rank tensor formats coincide with the canonical format, that means $\mathcal{S}_X = \mathcal{R}_m(X)$ for some rank m. When the norm $\|\cdot\|_X$ is the canonical norm, then $u_m \in \Pi_{\mathcal{S}_X}(u)$ coincides with a rank-m singular value decomposition (SVD) of u (which is possibly not unique in the case of multiple singular values). Moreover, $\sigma(u;\mathcal{S}_X)^2 = \|\Pi_{\mathcal{S}_X}(u)\|_X^2$ is the sum of the squares of the m dominant singular values of u (see e.g. [15]). Efficient algorithms for computing the SVD can therefore be applied to compute an element in $\Pi_{\mathcal{S}_X}(v)$ (a best approximation). That means that the algorithm (32) can be applied with $\eta = 1$.

Canonical norm, $d > 2$**.** For $d > 2$ and when the norm $\|\cdot\|_X$ is the canonical norm, different algorithms based on optimization methods have been proposed for the different tensor formats (see e.g. [14, 26] or [24] for a recent review). Very efficient algorithms based on higher order SVD have also been proposed in [11], [22] and [39], respectively for Tucker, Hierarchical Tucker and Tensor Train tensors. Note that these algorithms provide quasi-best approximations (but not necessarily best approximations) satisfying (31) with a η bounded by a function of the dimension d: $\eta \leq \sqrt{d}$, $\eta \leq \sqrt{2d-3}$ respectively for Tucker and Hierarchical Tucker formats (see [24]). For a high dimension d, such bounds for η would suggest taking very small values for parameter δ in order to satisfy the assumption of Proposition 4.3. However, in practice, these a priori bounds are rather pessimistic. Moreover, quasi-best approximations obtained by higher order SVD can be used as initializations of optimization algorithms yielding better approximations, i.e. with small values of η.

General norms, $d \geq 2$**.** For a general norm $\|\cdot\|_X$, the computation of a global optimum to the best approximation problem is still an open problem for all tensor subsets, and methods based on SVD cannot be applied anymore. However, classical optimization methods can still be applied (such as Alternating Minimization Algorithm (AMA)) in order to provide an approximation of the best approximation [41, 44, 14]. We do not detail further these computational aspects and we suppose that algorithms are available for providing an approximation of the best approximation in $\mathcal{S}_X$ such that (31) holds with a controlled precision η, arbitrarily close to 1.

5.2 Construction of an approximation of $\Lambda^\delta(r)$

At each iteration of the algorithm (32), we have to compute $y^k = \Lambda^\delta(r^k)$, with $r^k = R_Y^{-1}(Au^k - b) \in Y$, such that it satisfies

$$\|y^k - r^k\|_Y \leq \delta \|r^k\|_Y. \tag{41}$$

First note that r^k is the unique solution of

$$\min_{r \in Y} \|r - R_Y^{-1}(Au^k - b)\|_Y^2. \tag{42}$$

Therefore, computing y^k is equivalent to solving the best approximation problem (42) with a relative precision δ. One can equivalently characterize $r^k \in Y$ by the variational equation

$$\langle r^k, \delta r\rangle_Y = \langle Au^k - b, \delta r\rangle_{Y',Y} \quad \forall \delta r \in Y,$$

or in an operator form:

$$R_Y r^k = Au^k - b, \tag{43}$$

where the Riesz map $R_Y = AR_X^{-1}A^*$ is a positive symmetric definite operator.

Remark 5.1. *For A symmetric and positive definite, it is possible to choose $R_X = R_Y = A$ (see example 3.2) that corresponds to the energy norm on X. For this choice, the auxiliary problem* (42) *has the same structure as the initial problem, with an operator A and a right-hand side $Au^k - b$.*

5.2.1 Low-rank tensor methods

For solving (42), we can also use low-rank tensor approximation methods. Note that in general, $\|\cdot\|_Y$ is not an induced (canonical) norm in Y, so that classical tensor algorithms (e.g. based on SVD) cannot be applied for solving (42) (even approximatively). Different strategies have been proposed in the literature for constructing tensor approximations of the solution of optimization problems. We can either use iterative solvers using classical tensor approximations applied to equation (43) [31, 28, 34, 4], or directly compute an approximation y^k of r^k in low-rank tensor subsets using optimization algorithms applied to problem (42). Here, we adopt the latter strategy and rely on a greedy algorithm which consists in computing successive corrections of the approximation in a fixed low-rank subset.

5.2.2 A possible (heuristic) algorithm

We use the following algorithm for the construction of a sequence of approximations $\{y_m^k\}_{m\geq 0}$.

Let $y_0^k = 0$. Then, for each $m \geq 1$, we proceed as follows:

1. compute an optimal correction w_m^k of y_{m-1}^k in $\mathcal{S}_Y$:
$$w_m^k \in \arg\min_{w\in\mathcal{S}_Y} \|y_{m-1}^k + w - r^k\|_Y,$$
2. define a linear subspace Z_m^k such that $y_{m-1}^k + w_m^k \in Z_m^k$,
3. compute y_m^k as the best approximation of r^k in Z_m^k,
$$y_m^k = \arg\min_{y\in Z_m^k} \|y - r^k\|_Y,$$
4. return to step (2) or (1).

Remark 5.2. *The convergence proof for this algorithm can be found in [17]. The convergence ensures that the precision δ can be achieved after a certain number of iterations.*[2] *However, in practice, best approximation problems at step (1) can not be solved exactly except for particular situations (see section 5.1), so that the results of [17] do not guaranty anymore the convergence of the algorithm. If quasi-optimal solutions can be obtained, this algorithm is a modified version of weak greedy algorithms (see [43]) for which convergence proofs can also be obtained. Available algorithms for obtaining quasi-optimal solutions of best low-rank approximation problem appearing at step (1) are still heuristic but seem to be effective.*

In this paper, we will only rely on the use of low-rank canonical formats for numerical illustrations. At step (1), we introduce rank-one corrections $w_m^k \in \mathcal{S}_Y = \mathcal{R}_1(Y)$, where $Y = {}_{\|\cdot\|_Y}\bigotimes_{\mu=1}^d Y^\mu$. The auxiliary variable $y_m^k \in \mathcal{R}_m(Y)$ can be written in the form $y_m^k = \sum_{i=1}^m \otimes_{\mu=1}^d w_i^{k,\mu}$. At step (2), we select a particular dimension $\mu \in \{1,\dots,d\}$ and define

$$Z_m^k = \left\{ \sum_{i=1}^m w_i^{k,1} \otimes \cdots \otimes v_i^\mu \otimes \cdots \otimes w_i^{k,d} \ , \ v_i^\mu \in Y^\mu \right\},$$

where $\dim(Z_m^k) = m \dim(Y^\mu)$. Step (3) therefore consists in updating functions $w_i^{k,\mu}$, $i = 1 \dots d$, in the representation of y_m^k. Before returning to step (1), the updating steps (2)-(3) can be performed several times for a set of dimension $\mu \in I \subset \{1,\dots,d\}$.

[2] Note however that a slow convergence of these algorithms may yield to high rank representations of iterates y_m^k, even for a low-rank subset $\mathcal{S}_Y$.

Remark 5.3. *Note that the solution of minimization problems at steps (1) and (3) do not require to know r^k explicitly. Indeed, the stationary conditions associated with these optimization problems only require the evaluation of $\langle r^k, \delta y\rangle_Y = \langle Au^k - b, \delta y\rangle_{Y',Y}$, for $\delta Y \in Y$. For step (1), the stationary equation reads $\langle R_Y w_m^k, \delta y\rangle_{Y',Y} = \langle R_Y y_{m-1}^k + Au^k - b, \delta y\rangle_{Y',Y}$ for all δy in the tangent space to $\mathcal{S}_Y$, while the variational form of step (3) reads $\langle R_Y y_m^k, \delta y\rangle_{Y',Y} = \langle Au^k - b, \delta y\rangle_{Y',Y}$ for all δy in Z_m^k.*

Finally, as a stopping criterion, we use a heuristic error estimator based on stagnation. The algorithm is stopped at iteration m if

$$e_m^p = \frac{\|y_m^k - y_{m+p}^k\|_Y}{\|y_{m+p}^k\|_Y} \leq \delta, \tag{44}$$

for some chosen $p \geq 1$ (typically $p = 1$). Note that for p sufficiently large, y_{m+p}^k can be considered as a good estimation of the residual r^k and the criterion reads $\|r^k - y_m^k\|_Y \leq \delta \|r^k\|_Y$, which is the desired property. This stopping criterion is quite rudimentary and should be improved for a real control of the algorithm. Although numerical experiments illustrate that this heuristic error estimator provides a rather good approximation of the true error, an upper bound of the true error should be used in order to guarantee that the precision δ is really achieved. However, a tight error bound should be used in order to avoid a pessimistic overestimation of the true error which may yield an (unnecessary) increase of the computational costs for the auxiliary problem. This key issue will be addressed in a future work.

Remark 5.4. *Other updating strategies could be introduced at steps (2)-(3). For example, we could choose $Z_m^k = span\{w_1^k, \ldots, w_m^k\}$, thus making the algorithm an orthogonal greedy algorithm with a dictionary $\mathcal{S}_Y$ [42]. Nevertheless, numerical simulations demonstrate that when using rank-one corrections (i.e. $\mathcal{S}_Y = \mathcal{R}_1(Y)$), this updating strategy do not significantly improve the convergence of pure greedy constructions. When it is used for obtaining an approximation y_m^k of r^k with a small relative error δ, it usually requires a very high rank m. A more efficient updating strategy consists in defining Z_m^k as the tensor space $\bigotimes_{\mu=1}^d Z_m^{k,\mu}$ with $Z_m^{k,\mu}$ $span\{w_1^{k,\mu}, \ldots, w_m^{k,\mu}\}$. Since $\dim(Z_m^k) = m^d$, the projection of r^k in Z_m^k can not be computed exactly for high dimensions d. However, approximations of this projection can be obtained using again low-rank formats (see [20]).*

5.2.3 Remark on the tensor structure of Riesz maps

We consider that operator A and right-hand side b admit low-rank representations

$$A = \sum_{i=1}^{r_A} \otimes_{\mu=1}^d A_i^\mu \quad \text{and} \quad b = \sum_{i=1}^{r_b} \otimes_{\mu=1}^d b_i^\mu.$$

We suppose that a norm $\|\cdot\|_X$ has been selected and corresponds to a Riesz map R_X with a low-rank representation:

$$R_X = \sum_{i=1}^{r_X} \otimes_{\mu=1}^d R_i^\mu.$$

The ideal choice of norm $\|\cdot\|_Y$ then corresponds to the following expression of the Riesz map R_Y:

$$R_Y = AR_X^{-1}A^* = (\sum_{i=1}^{r_A} \otimes_{\mu=1}^d A_i^\mu)(\sum_{i=1}^{r_X} \otimes_{\mu=1}^d R_i^\mu)^{-1}(\sum_{i=1}^{r_A} \otimes_{\mu=1}^d A_i^{\mu *}).$$

Note that the expression of R_Y cannot be computed explicitly (R_Y is generally a full rank tensor). Therefore, in the general case, algorithms for solving problem (43) have to be able to handle an implicit formula for R_Y. However, in the particular case where the norm $\|\cdot\|_X$ is a canonical norm induced by norms $\|\cdot\|_\mu$ on X_μ, the mapping R_X is a rank one tensor $R_X = \otimes_{\mu=1}^d R_{X_\mu}$, where R_{X_μ}

is the Riesz map associated with the norm $\|\cdot\|_\mu$ on X_μ. R_Y then admits the following explicit expression:

$$R_Y = AR_X^{-1}A^* = \sum_{i=1}^{r_A}\sum_{j=1}^{r_A} \otimes_{\mu=1}^{d} (A_i^\mu R_{X_\mu}^{-1} A_j^{\mu *}).$$

In the numerical examples, we only consider this simple particular case. Efficient numerical methods for the general case will be proposed in a subsequent paper.

5.3 Summary of the algorithm

Algorithm 1 provides a step-by-step outline of the overall iterative method for the approximation of the solution of (28) in a fixed subset $\mathcal{S}_X$ and with a chosen metric $\|\cdot\|_X$. Given a precision δ, an approximation of the residual is obtained with a greedy algorithm using a fixed subset $\mathcal{S}_Y$ for computing successive corrections. We denote by $e(y_m^k, r^k)$ an estimation of the relative error $\|y_m^k - r^k\|_Y / \|r^k\|_Y$, where $r^k = R_Y^{-1}(Au^k - b)$.

Algorithm 1 Gradient-type algorithm

Set $u^0 = 0$;
for $k = 0$ to K **do**
  Set $m = 0$;
  while $e(y_m^k, r^k) \leq \delta$ **do**
    $m = m + 1$;
    Compute a correction $w_m^k \in \arg\min_{w \in \mathcal{S}_Y} \|y_{m-1}^k + w - r^k\|_Y$;
    Set $y_m^k = y_{m-1}^k + w_m^k$;
    Define Z_m^k containing y_m^k ;
    Compute the projection $y_m^k = \arg\min_{y \in Z_m^k} \|y - r^k\|_Y$;
    Return to step 7 or continue ;
  end while
  Compute $u^{k+1} \in \Pi_{\mathcal{S}_X}^{\eta}(u^k - R_X^{-1}A^* y_m^k)$;
end for

6 Greedy algorithm

In this section, we introduce and analyze a greedy algorithm for the progressive construction of a sequence $\{u_m\}_{m\geq 0}$, where u_m is obtained by computing a correction of u_{m-1} in a given low-rank tensor subset $\mathcal{S}_X$ (typically a small subset such as the set of rank-one tensors $\mathcal{R}_1(X)$). Here, we consider that approximations of optimal corrections are available with a certain precision. It results in an algorithm which can be considered as a modified version of weak greedy algorithms [43]. This weak greedy algorithm can be applied to solve the best approximation problem (19) where approximations of optimal corrections are obtained using Algorithm 1 with an updated right-hand side at each greedy step. The interest of such a global greedy strategy is twofold. First, an adaptive approximation strategy which would consist in solving approximation problems in an increasing sequence of low-rank subsets $\mathcal{S}_X$ is often unpractical since for high dimensional problems and subspace based tensor formats, computational complexity drastically increases with the rank. Second, it simplifies the solution of auxiliary problems (i.e. the computation of the sequence of y^k) when solving best low-rank approximation problems using Algorithm 1. Indeed, if the sequence u^k in Algorithm 1 belongs to a low rank tensor subset (typically a rank-one tensor subset), the residual r^k in Algorithm 1 admits a moderate rank or can be obtained by a low-rank correction of the residual of the previous greedy iteration.

Here, we assume that the subset $\mathcal{S}_X$ verifies properties (14) and (15), and that $\text{span}(\mathcal{S}_X)$ is dense in X (which is verified by all classical tensor subsets presented in section 2.2).

6.1 A weak greedy algorithm

We consider the following greedy algorithm. Given $u_0 = 0$, we construct a sequence $\{u_m\}_{m\geq 1}$ defined for $m \geq 1$ by

$$u_m = u_{m-1} + \widetilde{w}_m, \tag{45}$$

where $\widetilde{w}_m \in \mathcal{S}_X$ is a correction of u_{m-1} satisfying

$$\|u - u_{m-1} - \widetilde{w}_m\|_X \leq (1+\gamma_m) \min_{w\in\mathcal{S}_X} \|u - u_{m-1} - w\|_X, \tag{46}$$

with γ_m a sequence of small parameters.

Remark 6.1. *A $\widetilde{w}_m$ satisfying* (46) *can be obtained using the gradient type algorithm of section 4 that provides a sequence that satisfies* (34)*. Given the parameter $\delta = \delta_m$ in* (32)*, property* (46) *can be achieved with any $\gamma_m > \frac{2\delta_m}{1-2\delta_m}$.*

6.2 Convergence analysis

Here, we provide a convergence result for the above greedy algorithm whose proof follows the lines of [43] for the convergence proof of weak greedy algorithms[3].

In the following, we denote by $f_m = u - u_m$. For the sake of simplicity, we denote by $\|\cdot\| = \|\cdot\|_X$ and $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_X$ and we let $w_m \in \Pi_{\mathcal{S}_X}(f_{m-1})$, for which we have the following useful relations coming from properties of best approximation problems in tensor subsets (see section 2.2):

$$\|f_{m-1} - w_m\|^2 = \|f_{m-1}\|^2 - \|w_m\|^2 \quad \text{and} \quad \|w_m\|^2 = \langle f_{m-1}, w_m\rangle. \tag{47}$$

We introduce the sequence $\{\alpha_m\}_{m\geq 1}$ defined by

$$\alpha_m = \frac{\|f_{m-1} - w_m\|}{\|f_{m-1}\|} \in [0,1[. \tag{48}$$

It can be also useful to introduce the computable sequence $\{\widetilde{\alpha}_m\}_{m\geq 1}$ such that

$$\widetilde{\alpha}_m = \frac{\|f_{m-1} - \widetilde{w}_m\|}{\|f_{m-1}\|}. \tag{49}$$

that satisfies for all $m \leq 0$

$$\alpha_m \leq \widetilde{\alpha}_m \leq (1+\gamma_m)\alpha_m. \tag{50}$$

Lemma 6.2. *Assuming that for all $m \geq 1$ we have*

$$(1+\gamma_m)\alpha_m < 1, \tag{51}$$

the sequence $\{\|f_m\|\}_{m\geq 1}$ converges. Furthermore, it is possible to define a positive sequence $\{\kappa_m\}_{m\geq 1}$ as

$$\kappa_m^2 = 2\frac{\langle f_{m-1}, \widetilde{w}_m\rangle}{\|\widetilde{w}_m\|^2} - 1, \tag{52}$$

and we have $\{\kappa_m\|\widetilde{w}_m\|\}_{m\geq 1} \in \ell^2$.

[3]Note that the condition (46) on the successive corrections does not allow to directly apply the results on classical weak greedy algorithms.

Proof. From (45) and (46), we have

$$\|f_m\| = \|f_{m-1} - \widetilde{w}_m\| \leq (1+\gamma_m)\|f_{m-1} - w_m\| = (1+\gamma_m)\alpha_m\|f_{m-1}\|.$$

Under assumption (51), $\{\|f_m\|\}_{m\geq 1}$ is a strictly decreasing and positive sequence and therefore converges. Moreover, this implies that $\widetilde{w}_m \neq 0$ and since

$$\|f_{m-1} - \widetilde{w}_m\|^2 = \|f_{m-1}\|^2 - \left(2\langle f_{m-1}, \widetilde{w}_m\rangle - \|\widetilde{w}_m\|^2\right) \leq \|f_{m-1}\|^2,$$

it follows that $2\langle f_{m-1}, \widetilde{w}_m\rangle > \|\widetilde{w}_m\|^2$. Therefore, κ_m is positive and can be defined by (52) and we have

$$\|f_{m-1} - \widetilde{w}_m\|^2 = \|f_{m-1}\|^2 - \kappa_m^2\|\widetilde{w}_m\|^2 = \|f_0\|^2 - \sum_{i=1}^{m} \kappa_i^2\|\widetilde{w}_i\|^2,$$

that completes the proof. □

We now provide a result giving a relation between $\|w_m\|$ and $\|\widetilde{w}_m\|$.

Lemma 6.3. *Assume* (51) *holds and let* $\mu_m^2 = \dfrac{1-(1+\gamma_m)^2\alpha_m^2}{1-\alpha_m^2} \in [0,1]$. *Then, we have*

$$\mu_m\|w_m\| \leq \kappa_m\|\widetilde{w}_m\| \leq \|w_m\|, \tag{53}$$

and

$$\frac{\mu_m}{2} \leq \kappa_m. \tag{54}$$

Proof. From inequality (46) and from the optimality of w_m, it follows that

$$\begin{aligned}
&\|f_{m-1} - w_m\|^2 \leq \|f_{m-1} - \widetilde{w}_m\|^2 \leq (1+\gamma_m)^2\|f_{m-1} - w_m\|^2 \\
\Rightarrow\ &\|f_{m-1}\|^2 - \|w_m\|^2 \leq \|f_{m-1}\|^2 - \kappa_m^2\|\widetilde{w}_m\|^2 \leq (1+\gamma_m)^2\alpha_m^2\|f_{m-1}\|^2 \\
\Rightarrow\ &(1-(1+\gamma_m)^2\alpha_m^2)\|f_{m-1}\|^2 \leq \kappa_m^2\|\widetilde{w}_m\|^2 \leq \|w_m\|^2
\end{aligned}$$

Using $\|f_{m-1}\|^2 = \|f_{m-1} - w_m\|^2 + \|w_m\|^2 = \alpha_m^2\|f_{m-1}\|^2 + \|w_m\|^2$, and using the definition of μ_m, we obtain (53). In addition, from the optimality of w_m, we have $\langle \frac{\widetilde{w}_m}{\|\widetilde{w}_m\|}, f_{m-1}\rangle \leq \langle \frac{w_m}{\|w_m\|}, f_{m-1}\rangle = \|w_m\|$, or equivalently $\frac{\kappa_m^2+1}{2}\|\widetilde{w}_m\| \leq \|w_m\|$. Combined with (53), it gives $\frac{\kappa_m^2+1}{2} \leq \frac{\|w_m\|}{\|\widetilde{w}_m\|} \leq \frac{\kappa_m}{\mu_m}$, which implies (54). □

Proposition 6.4. *Assume* (51) *and that* $\left\{\mu_m^2\right\}_{m\geq 1}$ *is such that* $\sum_{m=1}^{\infty}\mu_m^2 = \infty$. *Then, if* $\{f_m\}_{m\geq 1}$ *converges, it converges to zero.*

Proof. Let us use a proof by contradiction. Assume that $f_m \to f \neq 0$ as $m \to \infty$, with $f \in X$. As $\mathrm{span}(\mathcal{S}_X)$ is dense in X, there exists $\epsilon > 0$ such that $\sup_{v\in\mathcal{S}_X} |\langle f, \frac{v}{\|v\|}\rangle| \geq 2\varepsilon$. Using the definition of w_m and of f as a limit of f_m, we have that there exists $N > 0$ such that

$$\|w_m\| = \sup_{v\in\mathcal{S}_X} |\langle f_{m-1}, \frac{v}{\|v\|}\rangle| \geq \varepsilon, \quad \forall m \geq N. \tag{55}$$

Thanks to (53), we have

$$\begin{aligned}
\|f_m\|^2 &= \|f_{m-1}\|^2 - \|\widetilde{w}_m\|^2\kappa_m^2 \leq \|f_{m-1}\|^2 - \|w_m\|^2\mu_m^2, \\
&\leq \|f_N\|^2 - \sum_{i=N+1}^{m}\mu_i^2\|w_i\|^2 \leq \|f_N\|^2 - \varepsilon^2\sum_{i=N+1}^{m}\mu_i^2,
\end{aligned}$$

which implies that $\{\mu_m\}_{m\geq 0} \in \ell^2$, a contradiction to the assumption.

□

Proposition 6.5. *Assume* (51). *Further assume that the sequence* μ_m *is non increasing and verifies*

$$\sum_{m=1}^{\infty} \frac{\mu_m^2}{m} = \infty. \tag{56}$$

Then the sequence $\{u_m\}_{m\geq 1}$ *converges to* u.

Proof. Let two integers $n < m$ and consider

$$\|f_n - f_m\|^2 = \|f_n\|^2 - \|f_m\|^2 - 2\langle f_n - f_m, f_m\rangle.$$

Defining $\theta_{n,m} = |\langle f_n - f_m, f_m\rangle|$ and using Lemma 6.3, we obtain

$$\theta_{n,m} \leq \sum_{i=n+1}^{m} |\langle \widetilde{w}_i, f_m\rangle| \leq \|w_{m+1}\| \sum_{i=1}^{m} \|\widetilde{w}_i\| \leq 2\frac{\kappa_{m+1}\|\widetilde{w}_{m+1}\|}{\mu_{m+1}^2} \sum_{i=1}^{m} \kappa_i \|\widetilde{w}_i\|.$$

Lemma 6.2 implies that $\kappa_m\|\widetilde{w}_m\| \in \ell^2$. Together with assumption (56), and using Lemma 2.7 in [42], we obtain that $\liminf_{m\to\infty} \max_{n<m} \theta_{n,m} = 0$. Lemma 2.8 in [42] then proves that the sequence $\{f_m\}_{m\geq 1}$ converges. Noting that (56) implies that $\{\mu_m\}_{m=1}^{\infty} \notin \ell^2$, Lemma 6.4 allows to conclude the proof. □

In practice, condition (56) can be satisfied by the following sufficient condition on the sequence $\widetilde{\alpha}_m$, which is a computable sequence.

Corollary 6.6. *If there exists a constant* $0 < \epsilon < 1$, *independent of* m, *such that*

$$\widetilde{\alpha}_m^2 \leq \frac{1-\epsilon}{(1+\gamma_m)^2 - \epsilon}, \tag{57}$$

then the sequence $\{u_m\}_{m\geq 1}$ *converges to* u.

Proof. Under assumption (57) and using relation (50), it holds that for all $m \geq 0$

$$\alpha_m^2 \leq \frac{1-\epsilon}{(1+\gamma_m)^2 - \epsilon} \quad \Rightarrow \quad (1+\gamma_m)^2\alpha_m^2 \leq 1 - \epsilon(1-\alpha_m^2) < 1.$$

which implies condition (51). Moreover, we have

$$\epsilon(1-\alpha_m^2) \leq 1 - (1+\gamma_m)^2\alpha_m^2 \quad \Rightarrow \quad \epsilon \leq \frac{1-(1+\gamma_m)^2\alpha_m^2}{(1-\alpha_m^2)} = \mu_m^2,$$

which implies condition (56). Proposition 6.5 ends the proof. □

Remark 6.7. *From a practical point of view, condition* (57) *provides a sufficient criterion on* γ_m *(or equivalently on* δ_m*). Note that* $\widetilde{\alpha}_m$ *depends on* $\widetilde{w}_m$ *which depends on the choice of the precision* γ_m. *Therefore,* (57) *is an implicit condition on* γ_m *which suggests an iterative strategy for the control of the condition. A possible strategy would be to adapt the parameter* γ_m *during the iterations of the gradient type algorithm used to compute the* $\widetilde{w}_m$.

7 Numerical example

In this section, we apply the proposed method to the numerical solution of a stochastic steady reaction-advection-diffusion problem.

7.1 Stochastic reaction-advection-diffusion problem

We consider the following steady reaction-advection-diffusion problem on a two-dimensional unit square domain $\Omega = [0,1]^2$ (see Figure 1):

$$\begin{aligned} -\nabla \cdot (\kappa \nabla u) + c \cdot \nabla u + au &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{58}$$

First, we consider a constant diffusion $\kappa = 1$. The advection coefficient c and the reaction coefficient a are considered as random and are given by $c = \xi_1 c_0$ and $a = \exp(\xi_2)$, where $\xi_1 \sim U(-350, 350)$ and $\xi_2 \sim U(\log(0.1), \log(10))$ are independent uniform random variables, and $c_0(x) = (x_2 - 1/2, 1/2 - x_1)$, $x = (x_1, x_2) \in \Omega$. We denote by $\Xi_1 =]\text{-}350, 350[$ and $\Xi_2 =]\log(0.1), \log(10)[$, and we denote by $(\Xi, \mathcal{B}(\Xi), P_\xi)$ the probability space induced by $\xi = (\xi_1, \xi_2)$, with $\Xi = \Xi_1 \times \Xi_2$ and P_ξ the probability law of ξ. The external source term f is given by $f(x) = I_{\Omega_1}(x) - I_{\Omega_2}(x)$, where $\Omega_1 =]0.45, 0.55[\times]0.15, 0.25[$ and $\Omega_2 =]0.45, 0.55[\times]0.75, 0.85[$, and where I_{Ω_k} denotes the indicator function of Ω_k.

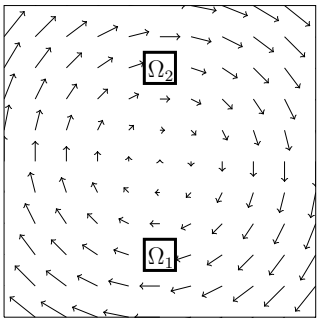


Figure 1: Example : reaction-advection-diffusion problem.

Let $V = \mathrm{H}^1{}_0(\Omega)$ and $S = \mathrm{L}^2(\Xi, dP_\xi)$. We introduce approximation spaces $V_N \subset V$ and $S_P \subset S$, with $N = \dim(V_N)$ and $P = \dim(S_P)$. V_N is a $\mathbb{Q}_1$ finite element space associated with a uniform mesh of 1600 elements such that $N = 1521$. We choose $S_P = S_{p_1}^{\xi_1} \otimes S_{p_2}^{\xi_2}$, where $S_{p_1}^{\xi_1}$ is the space of piecewise polynomials of degree 5 on Ξ_1 associated with the partition $\{]\text{-}350, 0[,]0, 350[\}$ of Ξ_1, and $S_{p_2}^{\xi_2}$ is the space of polynomials of degree 5 on Ξ_2. This choice results in $P = 72$. The Galerkin approximation $u \in V_N \otimes S_P \subset V \otimes S$ of the solution of (58) is defined by the following equation[4]:

$$\int_\Xi \int_\Omega (\nabla u \cdot \nabla v + c \cdot \nabla u v + auv)\, dx\, dP_\xi = \int_\Xi \int_\Omega f\, v\, dx\, dP_\xi, \tag{59}$$

for all $v \in V_N \otimes S_P$. Letting $V_N \otimes S_P = span\{\varphi_i \otimes \psi_j; 1 \le i \le N, 1 \le j \le P\}$, the Galerkin approximation $u = \sum_{i=1}^N \sum_{j=1}^P u_{ij} \varphi_i \otimes \psi_j$ can be identified with its set of coefficients on the chosen basis, still denoted u, which is a tensor

$$u \in X = \mathbb{R}^N \otimes \mathbb{R}^P \quad \text{such that} \quad Au = b, \tag{60}$$

where $b = b^x \otimes b^\xi$, with $b_i^x = \int_\Omega f \varphi_i$ and $b_j^\xi = \int_\Xi \psi_j dP_\xi$, and where A is a rank-3 operator such that $A = D^x \otimes M^\xi + C^x \otimes H^{\xi_1} + R^x \otimes H^{\xi_2}$, with $D_{ik}^x = \int_\Omega \nabla \varphi_i \cdot \nabla \varphi_k dx$, $C_{ik}^x = \int_\Omega \varphi_i c_0 \cdot \nabla \varphi_k dx$, $R_{ik}^x = \int_\Omega \varphi_i \varphi_k dx$, $M_{jl}^\xi = \int_\Xi \psi_j(y)\psi_l(y) dP_\xi(y)$, $H_{jl}^{\xi_n} = \int_\Xi y_n \psi_j(y)\psi_l(y) dP_\xi(y)$, $n = 1, 2$. Here, we use orthonormal basis functions $\{\psi_j\}$ in S_P, so that $M^\xi = I_P$, the identity matrix in $\mathbb{R}^P$.

[4]The mesh Péclet number is sufficiently small so that an accurate Galerkin approximation can be obtained without introducing a stabilized formulation.

7.2 Comparison of minimal residual methods

In this section, we present numerical results concerning the approximate ideal minimal residual method (A-IMR) applied to the algebraic system of equations (60) in tensor format. This method provides an approximation of the best approximation of u with respect to a norm $\|\cdot\|_X$ that can be freely chosen a priori. Here, we consider the application of the method for two different norms. We first consider the natural canonical norm on X, denoted $\|\cdot\|_2$ and defined by

$$\|v\|_2^2 = \sum_{i=1}^{N}\sum_{j=1}^{P}(v_{ij})^2. \tag{61}$$

This choice corresponds to an operator $R_X = I_X = I_N \otimes I_P$, where I_N (resp. I_P) is the identity in $\mathbb{R}^N$ (resp. $\mathbb{R}^P$). We also consider a weighted canonical norm, denoted $\|\cdot\|_w$ and defined by

$$\|v\|_w^2 = \sum_{i=1}^{N}\sum_{j=1}^{P}\left(w(x_i)v_{ij}\right)^2, \tag{62}$$

where $w : \Omega \to \mathbb{R}$ is a weight function and the x_i are the nodes associated with finite element shape functions φ_i. This norm allows to give a more important weight to a particular region $D \subset \Omega$, that may be relevant if one is interested in the prediction of a quantity of interest that requires a good precision of the numerical solution in this particular region (see section 7.2.3). This choice corresponds to an operator $R_X = D_w \otimes I_P$, with $D_w = \mathrm{diag}(w(x_1)^2, \ldots, w(x_N)^2)$.

The A-IMR provides an approximation $\widetilde{u} \in \mathcal{S}_X$ of the $\|\cdot\|_X$-best approximation of u in $\mathcal{S}_X$ (that means an approximation of an element in $\Pi_{\mathcal{S}_X}(u)$), where $\|\cdot\|_X$ is either $\|\cdot\|_2$ or $\|\cdot\|_w$. The set $\mathcal{S}_X$ is taken as the set $\mathcal{R}_r(X)$ of rank-r tensors in $X = \mathbb{R}^N \otimes \mathbb{R}^P$. The dimension of X is about 75,000 so that the exact solution u of (60) can be computed and used as a reference solution. We note that both norms are induced norms in $\mathbb{R}^N \otimes \mathbb{R}^P$ (associated with rank one operators R_X) so that the $\|\cdot\|_X$-best approximation of u in $\mathcal{S}_X$ is a rank-r SVD that can be computed exactly using classical algorithms (see section 5.1).[5] For the construction of an approximation in $\mathcal{R}_r(X)$ using A-IMR, we consider two strategies: the direct approximation in $\mathcal{R}_r(X)$ using Algorithm 1 with $\mathcal{S}_X = \mathcal{R}_r(X)$, and a greedy algorithm that consists in a series of r corrections in $\mathcal{R}_1(X)$ computed using Algorithm 1 with $\mathcal{S}_X = \mathcal{R}_1(X)$ and with an updated residual b at each correction.

The A-IMR will be compared to a standard approach, denoted CMR, which consists in minimizing the canonical norm of the residual of equation (60), that means in solving

$$\min_{v \in \mathcal{S}_X} \|Av - b\|_2. \tag{63}$$

This latter approach has been introduced and analyzed in different papers, using either direct minimization or greedy rank-one algorithms [5, 12, 2], and is known to suffer from ill-conditioning of the operator A. We note that this approach corresponds to choosing $R_X = A^*A$ and $R_Y = I_X = I_N \otimes I_P$.

7.2.1 Natural canonical norm $\|\cdot\|_2$

First, we compare both greedy and direct algorithms for $\|\cdot\|_X = \|\cdot\|_2$, using either CMR or A-IMR with different precisions δ. The convergence curves with respect to the rank are shown in Figure 2, where the error is measured in the $\|\cdot\|_2$ norm. Concerning the direct approach, we observe that the different algorithms have roughly the same rate of convergence. The A-IMR convergence curves are close to the optimal SVD (corresponding to $\widetilde{u}_2$) for a wide range of values of δ. One should note that A-IMR seems to provide good approximations also for the value $\delta = 0.9$ which is greater than the theoretical bound 0.5 ensuring the convergence of the gradient-type algorithm. Concerning the greedy approach, we observe a significant difference between A-IMR and CMR. We note that A-IMR is close to the optimal SVD up to a certain rank (depending on δ) after which

[5] Note that different truncated SVD are obtained when $\mathbb{R}^N$ is equipped with different norms.

the convergence rate decreases but remains better than the one of CMR. Finally, one should note that using a precision $\delta = 0.9$ for A-IMR yields less accurate approximations than CMR. However, A-IMR provides better results than CMR once the precision δ is lower than 0.5.

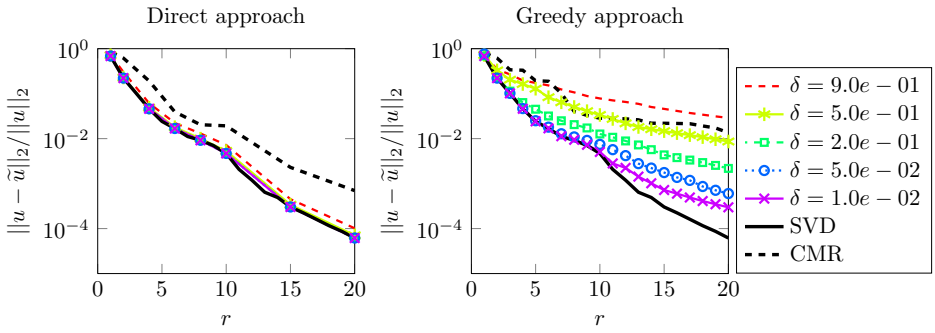


Figure 2: Comparison of minimal residual methods for $\mathcal{S}_X = \mathcal{R}_r(X)$ and $\|\cdot\|_X = \|\cdot\|_2$. Convergence with the rank r of the approximations obtained with CMR or A-IMR with different precisions δ, and with direct (left) or greedy rank-one (right) approaches.

7.2.2 Weighted norm $\|\cdot\|_w$

Here, we perform the same numerical experiments as previously using the weighted norm $\|\cdot\|_X = \|\cdot\|_w$, with w equal to 10^3 on $D = [0.15, 0.25] \times [0.45, 0.55]$ and $w = 1$ on $\Omega \setminus D$. The convergence curves with respect to the rank are plotted on Figure 3. The conclusions are similar to the case $\|\cdot\|_X = \|\cdot\|_2$, although the use of the weighted norm seems to slightly deteriorate the convergence properties of A-IMR. However, the direct A-IMR still provides better approximations than the direct CMR, closer to the reference SVD (denoted by $\tilde{u}_w$) for different values of precision δ.

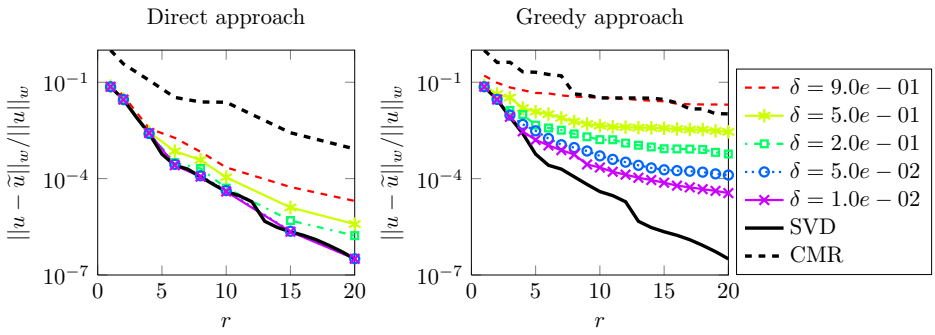


Figure 3: Comparison of minimal residual methods for $\mathcal{S}_X = \mathcal{R}_r(X)$ and $\|\cdot\|_X = \|\cdot\|_w$. Convergence with the rank of the approximations obtained with CMR or A-IMR with different precisions δ, and with direct (left) or greedy rank-one (right) approaches.

7.2.3 Interest of using a weighted norm

Here, we illustrate the interest of using the weighted norm rather than the natural canonical norm when one is interested in computing a quantity of interest. For the sake of readability, we let $\widetilde{u}_w$ (resp. $\widetilde{u}_2$) denote the best approximation of u in $\mathcal{R}_r(X)$ with respect to the norm $\|\cdot\|_w$ (resp. $\|\cdot\|_2$). Figure 4 illustrates the convergence with r of these approximations. We observe that approximations $\widetilde{u}_w$ and $\widetilde{u}_2$ are of the same quality when the error is measured with the norm $\|\cdot\|_2$, while $\widetilde{u}_w$ is a far better approximation than $\widetilde{u}_2$ (almost two orders of magnitude) when the error is measured with the norm $\|\cdot\|_w$. We observe that $\widetilde{u}_w$ converges faster to u with $\|\cdot\|_w$ than $\widetilde{u}_2$

with $\|\cdot\|_2$. For example, with a rank $r = 9$, $\widetilde{u}_w$ has a $\|\cdot\|_w$-error of 10^4 while $\widetilde{u}_2$ has a $\|\cdot\|_2$-error of 10^2. On Figure 5, plotted are the spatial modes of the rank-r approximations $\widetilde{u}_2$ and $\widetilde{u}_w$. These spatial modes are significantly different and obviously capture different features of the solution.

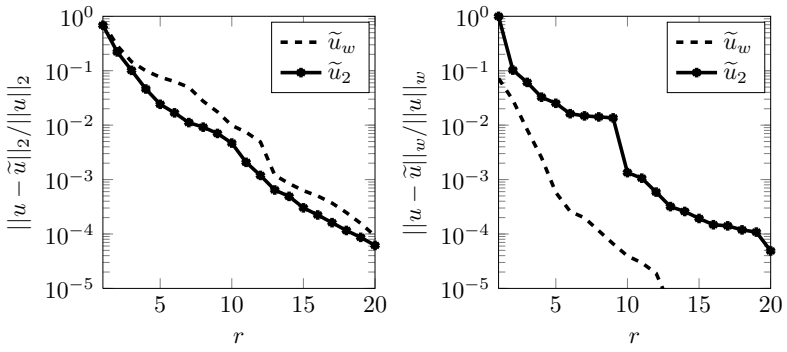


Figure 4: Convergence of best rank-r approximations $\widetilde{u}_2$ and $\widetilde{u}_w$ of the solution u measured with the natural canonical norm $\|\cdot\|_2$ or the weighted norm $\|\cdot\|_w$.

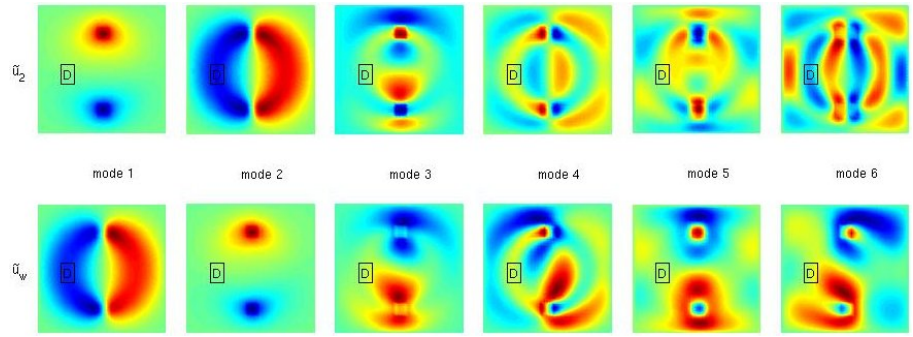


Figure 5: Comparison of the first spatial modes of the rank-r approximations $\widetilde{u}$ and $\widetilde{u}_w$.

Now, we introduce a quantity of interest Q which is the spatial average of u on subdomain D:

$$Q(u) = \frac{1}{|D|}\int_D u \, \mathrm{d}x. \tag{64}$$

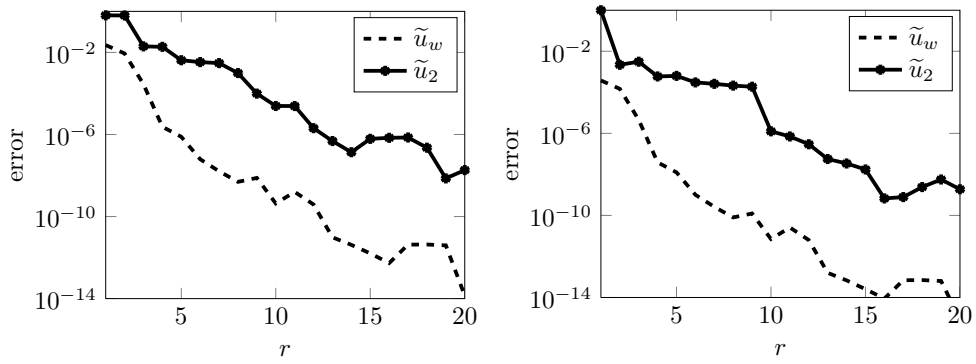
Due to the choice of norm, $\widetilde{u}_w$ is supposed to be more accurate than $\widetilde{u}_2$ in the subdomain D, and therefore, $Q(\widetilde{u}_w)$ is supposed to provide a better estimation of $Q(u)$ than $Q(\widetilde{u}_2)$. This is confirmed by Figure 6, where we have plotted the convergence with the rank of the statistical mean and variance of $Q(\widetilde{u}_w)$ and $Q(\widetilde{u}_2)$. With only a rank $r = 5$, $\widetilde{u}_w$ gives a precision of 10^{-7} on the mean, whereas $\widetilde{u}_2$ gives only a precision of 10^{-2}. In conclusion, we observe that a very low-rank approximation $\widetilde{u}_w$ is able to provide a very good approximation of the quantity of interest.

7.3 Properties of the algorithms

Now, we detail some numerical aspects of the proposed methodology. We first focus on the gradient-type algorithm, and then on evaluations of the map Λ^δ for the approximation of residuals.

7.3.1 Analysis of the gradient-type algorithm

The behavior of the gradient-type algorithm for different choices of norms $\|\cdot\|_X$ is very similar, so we only illustrate the case where $\|\cdot\|_X = \|\cdot\|_2$. The convergence of this algorithm is plotted in Figure 7 for the case $\mathcal{S}_X = \mathcal{R}_{10}(X)$. It is in very good agreement with theoretical expectations



(a) Error on the mean value of $Q(\widetilde{u})$ (b) Error on the variance of $Q(\widetilde{u})$

Figure 6: Convergence with the rank of the mean (left) and variance (right) of $Q(\widetilde{u}_2)$ and $Q(\widetilde{u}_w)$. Relative error with respect to the mean and variance of the reference solution $Q(u)$.

(Proposition 4.3): we first observe a linear convergence with a convergence rate that depends on δ, and then a stagnation within a neighborhood of the solution with an error depending on δ. The gradient-type algorithm is then applied for subsets $\mathcal{S}_X = \mathcal{R}_r(X)$ with different ranks

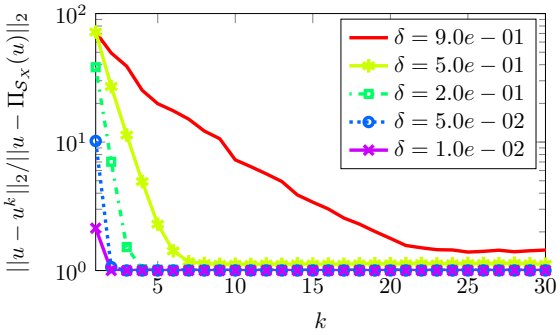


Figure 7: Convergence of the gradient-type algorithm for different values of the relative precision δ, for $\mathcal{S}_X = \mathcal{R}_{10}(X)$ and $\|\cdot\|_X = \|\cdot\|_2$.

r. The estimate of the linear convergence rate ρ is given in Table 1. We observe that for all values of r, ρ takes values closer to δ than to the theoretical bound 2δ of Proposition 4.3. This means that the theoretical bound of the convergence rate overestimates the effective one, and the algorithm converges faster than expected. Now, in order to evaluate the quality of the resulting

δ	0.90	0.50	0.20	0.05	0.01
$r = 4$	0.78	0.36	≈ 0	≈ 0	≈ 0
$r = 6$	0.83	0.45	0.165	≈ 0	≈ 0
$r = 10$	0.82	0.42	0.183	≈ 0	≈ 0
$r = 15$	0.84	0.47	0.189	0.047	≈ 0
$r = 20$	0.86	0.48	0.197	0.051	0.011

Table 1: Estimation of the convergence rate ρ of the gradient-type algorithm (during the linear convergence phase) for different subsets $\mathcal{S}_X = \mathcal{R}_r(X)$, and for $\|\cdot\|_X = \|\cdot\|_2$.

approximation, we compute the error after the stagnation phase has been reached. More precisely, we compute the value

$$\widetilde{\gamma}_k = \frac{\|u^k - u\|_X}{\|u - \Pi_{\mathcal{S}_X}(u)\|_X} - 1,$$

for $k = 100$. Values of $\widetilde{\gamma}_{100}$ are summarized in Table 2 and are compared to the theoretical upper bound $\gamma = 2\delta/(1-2\delta)$ given by Proposition 4.3. Once again, one can observe that the effective error of the resulting approximation is lower than the predicted value regardless of the choice of $\mathcal{R}_r(X)$.

δ	0.90	0.50	0.20	0.05	0.01
$2\delta/(1-2\delta)$	-	-	6.6e-1	1.1e-1	2.1e-2
$r = 4$	3.3e-1	5.6e-2	4.9e-3	3.5e-4	3.0e-5
$r = 6$	3.0e-1	6.8e-2	1.1e-2	8.6e-4	8.0e-5
$r = 10$	5.2e-1	1.3e-1	1.7e-2	1.8e-3	3.3e-5
$r = 15$	4.9e-1	1.1e-1	1.5e-2	1.0e-3	7.5e-5
$r = 20$	6.4e-1	1.5e-1	1.9e-2	1.2e-3	7.3e-5

Table 2: Final approximation errors (estimated by $\widetilde{\gamma}_{100}$) for different subsets $\mathcal{S}_X = \mathcal{R}_r(X)$ and different precisions δ. Comparison with the theoretical upper bound $2\delta/(1-2\delta)$.

Now, we focus on numerical estimations of the error $\|u - u^k\|_X$. It has been pointed out in Section 4.4 that $\hat{\epsilon}^k$, defined in Eq. (39), should provide a good error estimator with effectivity index $\hat{\tau}^k \in (1, (1-\delta^2)^{-1/2})$. For $\delta = 0.2$ and $\mathcal{S}_X = \mathcal{R}_{10}(X)$, numerical values taken by $\hat{\tau}^k$ during the gradient-type algorithm are plotted on Figure 8 and are compared to the expected theoretical values of its lower and upper bounds 1 and $(1-\delta^2)^{-1/2}$ respectively. We observe that the theoretical upper bound is strictly satisfied, while the lower bound is almost but not exactly satisfied. This violation of the theoretical lower bound is explained by the fact that the precision δ is not satisfied at each iteration of the gradient-type algorithm due to the use of a heuristic convergence criterion in the computation of residuals (see next section 7.3.2). However, although it does not provide a controlled error estimation, the error indicator based on the computed residuals is of very good quality.

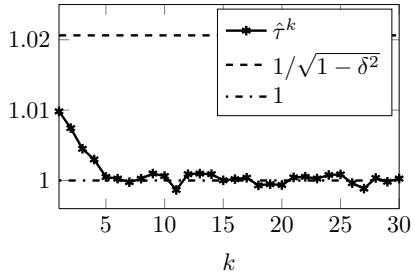


Figure 8: Effectivity index $\hat{\tau}^k$ of the error estimator $\hat{\epsilon}^k$ at different iterations k of the gradient-type algorithm, with $\mathcal{S}_X = \mathcal{R}_{10}(X)$ and $\delta = 0.2$.

7.3.2 Application of Λ^δ for the approximation of residuals

We study the behavior of the updated greedy algorithm described in Section 5.2.2 for the computation of an approximation $y_m^k = \Lambda^\delta(r^k)$ of the residual r^k during the gradient-type algorithm. Here, we use the particular strategy which consists in updating functions associated to each dimension $\mu \in I = \{1, 2\}$ (steps (2)-(3) are performed two times per iteration). We first validate the ability of the heuristic stopping criterion (44) to ensure a prescribed relative precision δ. Let $M = M(\delta)$ denote the iteration for which the condition $e_M^p \leq \delta$ is satisfied. The exact relative error $e_M = \|y_M^k - r^k\|_Y / \|r^k\|_Y$ is computed using a reference computation of r^k, and we define the effectivity index $\lambda_M^p = e_M^p / e_M$. Figure 9 shows the convergence of this effectivity index with respect to p, when using the natural canonical norm $\|\cdot\|_2$ or the weighted norm $\|\cdot\|_w$. We observe that λ_M^p tends to 1 as $p \to \infty$, as it was expected since the sequence $\{y_m^k\}_{m \geq 1}$ converges to r^k. However, we clearly observe that the quality of the error indicator differs for the two different

norms. When using the weighted norm, it appears that a large value of p (say $p \geq 20$) is necessary to ensure $\lambda_M^p \in [0.9, 1]$, while $p \leq 10$ seems sufficiently large when using the natural canonical norm. That simply reflects a slower convergence of the greedy algorithm when using the weighted norm.

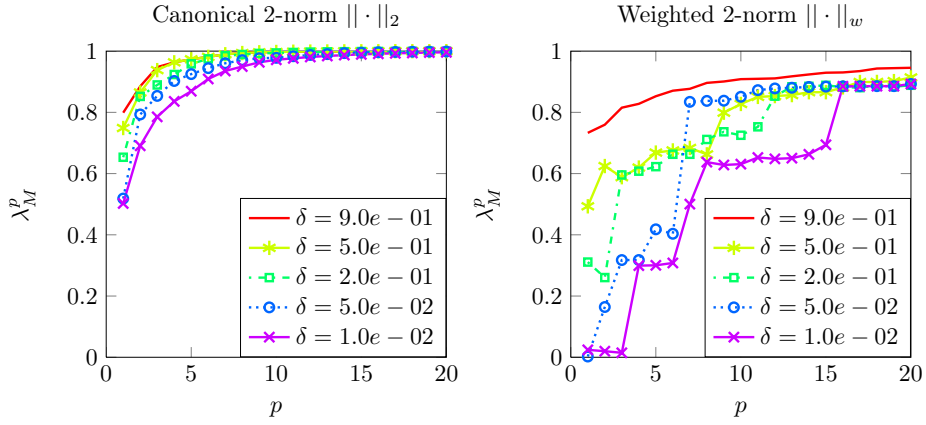


Figure 9: Evolution with p of the effectivity index λ_M^p for different δ at step $k = 1$ of the gradient-type algorithm with $\mathcal{S}_X = \mathcal{R}_{10}(X)$ and for the natural canonical norm (left) or the weighted norm (right).

Remark 7.1. *One can prove that at step k of the gradient-type algorithm, when computing an approximation y_M^k of r^k with a greedy algorithm stopped using the heuristic stopping criterion* (44)*, the effectivity index $\hat{\tau}^k$ of the computed error estimator $\hat{\epsilon}^k$ is such that*

$$\hat{\tau}^k \in \left(\sqrt{\frac{1 - (\delta/\lambda_M^p)^2}{1 - \delta^2}}, \sqrt{\frac{1}{1 - \delta^2}} \right).$$

where λ_M^p is the effectivity index of error indicator e_M^p (supposed such that $\delta/\lambda_M^p < 1$). That provides an explanation for the observations made on Figure 8.

Now, we observe in Table 3 the number of iterations of the greedy algorithm for the approximation of the residual r^k with a relative precision δ, with a fixed value $p = 20$ for the evaluation of the stopping criterion. The number of iterations corresponds to the rank of the resulting approximation. We note that the required rank is higher when using the weighted norm. It reflects the fact that it is more difficult to reach precision δ when using the weighted norm rather than the natural canonical norm.

7.4 Higher dimensional case

Now, we consider a diffusion coefficient of the form $\kappa(x, \xi) = \kappa_0 + \sum_{i=1}^{8} \xi_i \kappa_i(x)$ where $\kappa_0 = 10$, $\xi_i \sim U(-1, 1)$ are independent uniform random variables, and the functions $\kappa_i(x)$ are given by:

$$\begin{array}{llll} \kappa_1(x) = \cos(\pi x_1), & \kappa_3(x) = \sin(\pi x_1), & \kappa_5(x) = \cos(\pi x_1)\cos(\pi x_2), & \kappa_7(x) = \cos(\pi x_1)\sin(\pi x_2), \\ \kappa_2(x) = \cos(\pi x_2), & \kappa_4(x) = \sin(\pi x_2), & \kappa_6(x) = \sin(\pi x_1)\sin(\pi x_2), & \kappa_8(x) = \sin(\pi x_1)\cos(\pi x_2). \end{array}$$

In addition, the advection coefficient is given by $c = \xi_0 c_0$, where $\xi_0 \sim U(0, 4000)$ is a uniform random variable. We denote $V = \mathrm{H}^1{}_0(\Omega)$ and $S = \mathrm{L}^2(\Xi, dP_\xi)$ where $(\Xi, \mathcal{B}(\Xi), P_\xi)$ is a probability space with $\Xi =]-1, 1[^8 \times]0, 4000[$ and P_ξ the uniform measure. Here $V_N \subset V$ is a $\mathbb{Q}_1$ finite element space associated with a uniform mesh of 3600 elements, with a dimension $N = 3481$. We take $S_P = \otimes_{i=0}^{8} S_P^{\xi_i} \subset S$, where $S_P^{\xi_i}$ are polynomial function spaces of degree 7 on Ξ_i with $P = \dim(S_P^{\xi_i}) = 8$. Then, the Galerkin approximation in $V_N \otimes S_P$ (solution of (59)) is searched under

	Canonical 2-norm $\|\cdot\|_2$					Weighted 2-norm $\|\cdot\|_w$				
k\δ	0.9	0.5	0.2	0.05	0.01	0.9	0.5	0.2	0.05	0.01
1	1	1	3	7	11	8	21	31	35	51
2	1	3	7	16	27	5	22	14	24	42
3	1	5	11	19	24	4	15	24	23	43
4	1	3	11	14	24	8	11	19	37	42
5	1	6	7	15	24	6	19	23	14	38
6	1	8	8	16	24	3	12	47	25	63
7	1	5	7	17	24	7	14	16	29	47
8	1	4	8	16	24	5	12	22	21	40
9	1	4	8	16	24	7	13	18	36	45

Table 3: Computation of $\Lambda^\delta(r^k)$ for different precisions δ and at different steps k of the gradient-type algorithm, with $\mathcal{S}_X = \mathcal{R}_{10}$ (direct approach). The table indicates the number of greedy corrections computed for reaching the precision δ using the heuristic stopping criterion (44) with $p = 20$.

the form $u = \sum_{i=1}^N \sum_{j_0=1}^P \cdots \sum_{j_8=1}^P (u_{i,j_0,\cdots,j_9})\phi_j \otimes (\otimes_{\mu=0}^8 \psi_{j_\mu}^\mu)$. This Galerkin approximation can be identified with its set of coefficients, still denoted by u which is a tensor

$$u \in X = \mathbb{R}^N \otimes (\otimes_{\mu=0}^8 \mathbb{R}^P) \quad \text{such that} \quad Au = b, \tag{65}$$

where A and b are the algebraic representations on the chosen basis of $V_N \otimes S_P$ of the bilinear and linear forms in (59). The obtained algebraic system of equations has a dimension larger than 10^{11} and its solution clearly requires the use of model reduction methods.

Here, we compute low rank approximations of the solution of (65) in the canonical tensor subset $\mathcal{R}_r(X)$ with $r \geq 1$. Since best approximation problems in $\mathcal{R}_r(X)$ are well posed for $r = 1$ but ill posed for $d > 2$ and $r > 1$, we rely on the greedy algorithm presented in section 6 with successive corrections in $\mathcal{S}_X = \mathcal{R}_1(X)$ computed with Algorithm 1.

Remark 7.2. *Low-rank approximations could have been computed directly with Algorithm 1 by choosing for $\mathcal{S}_X$ other stable low-rank formats adapted to high-dimensional problems, such as Hierarchical Tucker (or Tensor Train) low-rank formats.*

7.4.1 Convergence study

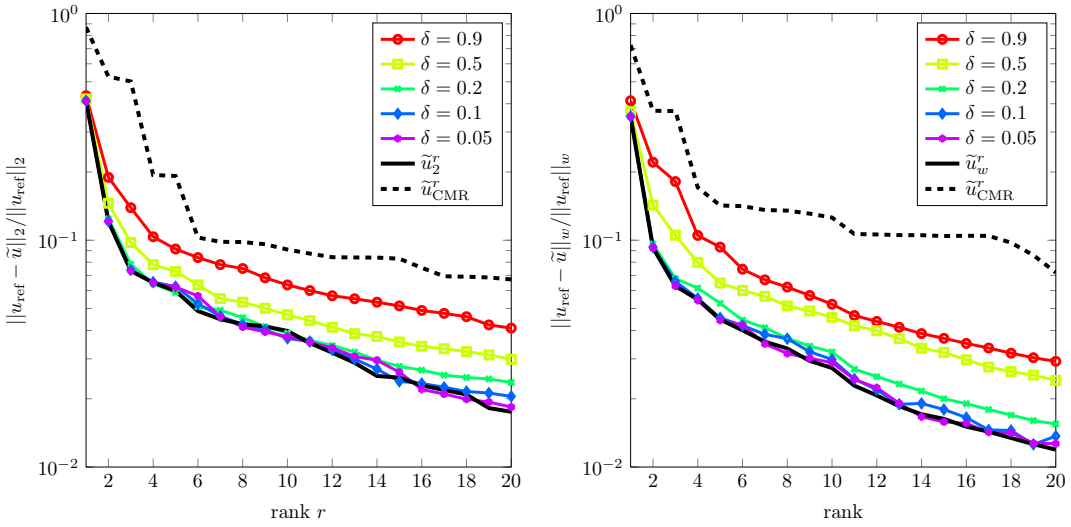
In this section, low rank approximations of the solution u of (65) are computed for the two different norms $\|\cdot\|_2$ and $\|\cdot\|_w$ defined as in section 7.2. Here, we assume that the weighting function w is equal to 100 in the subdomain $D \subset \Omega$, and 1 elsewhere.
Since $\dim(X) \geq 10^{11}$, the exact Galerkin approximation u in X is no more computable. As a reference solution, we consider a low-rank approximation u_{ref} of u computed using a greedy rank-one algorithm based on a canonical minimal residual formulation. We introduce an estimation $\hat{E}_K$ of $\frac{\|u-u_{\text{ref}}\|_2}{\|u\|_2}$ based on Monte-Carlo integrations using K realizations $\{\xi_k\}_{k=1}^K$ of the random variable ξ, defined by

$$\hat{E}_K^2 = \frac{\frac{1}{K}\sum_{k=1}^K \|u(\xi_k) - u_{\text{ref}}(\xi_k)\|_V^2}{\frac{1}{K}\sum_{k=1}^K \|u(\xi_k)\|_V^2},$$

with a number of samples K such that the Monte-Carlo estimates has a relative standard deviation (estimated using the statistical variance of the sample) lower than 10^{-1}. The rank of u_{ref} is here selected such that $\hat{E}_K < 10^{-4}$, which gives a reference solution with a rank of 212.

On Figure 10, we plot the convergence with the rank r of the approximations computed by both A-IMR and CMR algorithms and of the greedy approximations $\widetilde{u}_2^r$ and $\widetilde{u}_w^r$ of the reference solution u_{ref} for both norms. We observe (as for the lower-dimensional example) that for both norms, with different values of the parameter δ (up to 0.9), the A-IMR method provides a better

approximation of the solution in comparison to the CMR method. When decreasing δ, the proposed algorithm seems to provide approximations that tend to present the same convergence as the greedy approximations $\widetilde{u}_2^r$ and $\widetilde{u}_w^r$.



(a) Canonical norm (b) Weighted norm

Figure 10: Convergence with the rank of approximations obtained with the greedy CMR or A-IMR algorithms for different precisions δ. On the left (resp. right) plot, convergence is plotted with respect to the norm $\|\cdot\|_2$ (resp. $\|\cdot\|_w$) and A-IMR is used with the objective norm $\|\cdot\|_2$ (resp. $\|\cdot\|_w$).

7.4.2 Study of the greedy algorithm for Λ^δ

Now, we study the behavior of the updated greedy algorithm described in Section 5.2.2 for the computation of an approximation $y_m^k = \Lambda^\delta(r^k)$ of the residual r^k during the gradient-type algorithm. Here, we use the particular strategy which consists in updating functions associated to each dimension $\mu \in I = \{2, \ldots, 10\}$ (steps (2)-(3) are performed 9 times per iteration). The update of functions associated with the first dimension is not performed since it would involve the expensive computation of approximations in a space Z_m^k with a large dimension mN.

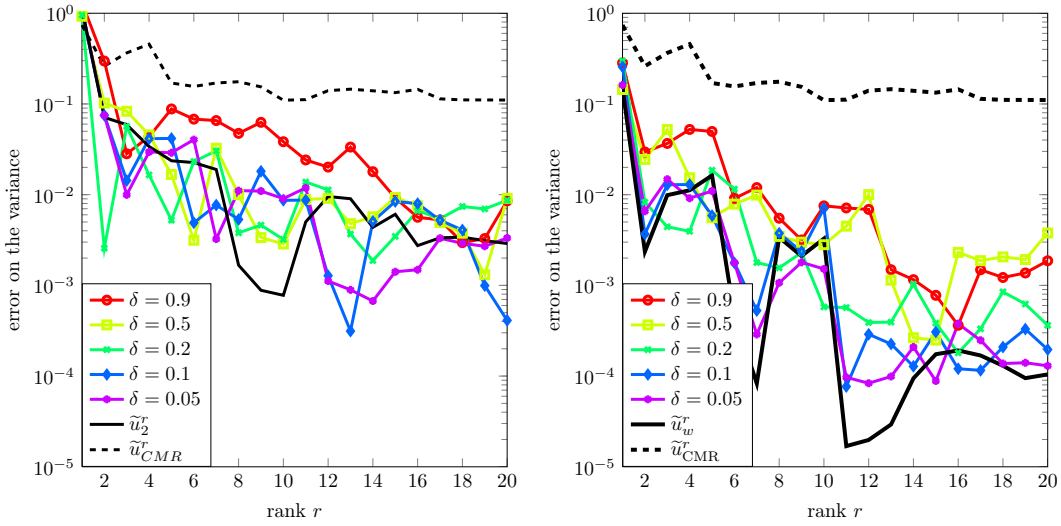
In table 4, we summarize the required number of greedy corrections needed at each iteration of the gradient type algorithm for reaching the precision δ with the heuristic stagnation criterion (44) with $p = 20$. As for the previous lower-dimensional test case, the number of corrections increases as δ decreases and is higher for the weighted norm than for the canonical norm. However, we observe that this number of corrections remains reasonable even for small δ.

7.4.3 Estimation of a quantity of interest

Finally, we study the quality of the low rank approximations $\widetilde{u}$ obtained with both CMR and A-IMR algorithms for the canonical and weighted norms. To this end, we compute the quantity of interest $Q(\widetilde{u})$ defined by (64). Figure 11 illustrates the convergence with the rank of the variance of the approximate quantities of interest. Note that the algorithm do not guarantee the monotone convergence of the quantity of interest with respect to the rank, that is confirmed by the numerical results. However, we observe that the approximations provided by the A-IMR algorithm are better than the ones given by the CMR, even for large δ. Also, when using the weighted norm in the A-IMR algorithm, the quantity of interest is estimated with an better precision. Similar behaviors are observed for the convergence of the mean.

	Canonical 2-norm $\|\cdot\|_2$					Weighted 2-norm $\|\cdot\|_w$				
k\δ	0.9	0.5	0.2	0.05	0.01	0.9	0.5	0.2	0.05	0.01
1	1	1	3	6	14	3	12	53	65	91
2	1	3	5	13	24	2	11	49	62	91
3	1	3	5	12	17	3	12	49	62	91
4	1	3	5	13	26	3	12	53	62	91
5	1	3	6	12	24	2	11	47	65	89
6	1	3	5	13	27	3	11	42	63	88
7	1	3	5	12	27	3	10	50	65	88
8	1	3	5	12	26	3	10	49	60	87
9	1	3	6	12	26	3	13	49	65	80

Table 4: Computation of $\Lambda^\delta(r^k)$ for different precisions δ and at different steps k of the gradient-type algorithm (first iteration $r = 1$ of the greedy approach with $\mathcal{S}_X = \mathcal{R}_1$). The table indicates the number of greedy corrections computed for reaching the precision δ using the heuristic stopping criterion (44) with $p = 20$.



(a) Using the canonical norm.

(b) Using the weighted norm.

Figure 11: Relative error with respect to the variance of the reference solution $Q(u_{\mathrm{ref}})$ with the canonical (left) and weighted (right) norms.

8 Conclusion

In this paper, we have proposed a new algorithm for the construction of low-rank approximations of the solutions of high-dimensional weakly coercive problems formulated in a tensor space X. This algorithm is based on the approximate minimization (with a certain precision δ) of a particular residual norm on given low-rank tensor subsets $\mathcal{S}_X$, the residual norm coinciding with some measure of the error in solution. Therefore, the algorithm is able to provide a quasi-best low-rank approximation with respect to a norm $\|\cdot\|_X$ that can be designed for a certain objective. A weak greedy algorithm using this minimal residual approach has been introduced and its convergence has been proved under some conditions. A numerical example dealing with the solution of a stochastic partial differential equation has illustrated the effectivity of the method and the properties of the proposed algorithms. Some technical points have to be addressed in order to apply the method to a more general setting and to improve its efficiency and robustness: the development of efficient solution methods for the computation of residuals when using general norms $\|\cdot\|_X$ (that are not induced norms in the tensor space X), the introduction of robust error estimators during the computation of residuals (for the robust control of the precision δ, which

is the key point for controlling the quality of the obtained approximations), the application of the method for using tensor formats adapted to high-dimensional problems (such as Hierarchical formats). Also, a challenging perspective consists in coupling low-rank approximation techniques with adaptive approximations in infinite-dimensional tensor spaces (as in [3]) in order to provide approximations of high-dimensional equations (PDEs or stochastic PDEs) with a complete control on the precision of quantities of interest.

References

[1] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modelling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139(3):153–176, 2006.

[2] A. Ammar, F. Chinesta, and A. Falco. On the convergence of a greedy rank-one update algorithm for a class of linear systems. *Archives of Computational Methods in Engineering*, 17(4):473–486, 2010.

[3] M. Bachmayr, and W. Dahmen. Adaptive Near-Optimal Rank Tensor Approximation for High-Dimensional Operator Equations *ArXiv e-print*, November 2013.

[4] J. Ballani and L. Grasedyck. A projection method to solve linear systems in tensor format. *Numerical Linear Algebra with Applications*, 20(1):27-43, 2013.

[5] G. Beylkin and M. J. Mohlenkamp. Algorithms for numerical analysis in high dimensions. *SIAM J. Sci. Comput.*, 26(6):2133–2159, 2005.

[6] E. Cances, V. Ehrlacher, and T. Lelievre. Convergence of a greedy algorithm for high-dimensional convex nonlinear problems. *Mathematical Models & Methods In Applied Sciences*, 21(12):2433–2467, 2011.

[7] E. Cances, V. Ehrlacher, and T. Lelievre. Greedy algorithms for high-dimensional non-symmetric linear problems. *ArXiv e-prints*, October 2012.

[8] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and variational stabilization for convection-diffusion equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:1247–1273, 8 2012.

[9] F. Chinesta, P. Ladeveze, and E. Cueto. A short review on model order reduction based on proper generalized decomposition. *Archives of Computational Methods in Engineering*, 18(4):395–404, 2011.

[10] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive petrov–galerkin methods for first order transport equations. *SIAM Journal on Numerical Analysis*, 50(5):2420–2445, 2012.

[11] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.

[12] A. Doostan and G. Iaccarino. A least-squares approximation of partial differential equations with high-dimensional random inputs. *Journal of Computational Physics*, 228(12):4332–4345, 2009.

[13] A. Ern and J.-L. Guermond. Theory and practice of finite elements, volume 159 of applied mathematical sciences, 2004.

[14] M. Espig and W. Hackbusch. A regularized newton method for the efficient approximation of tensors represented in the canonical tensor format. *Numerische Mathematik*, 122:489–525, 2012.

[15] A. Falcó and A. Nouy. A Proper Generalized Decomposition for the solution of elliptic problems in abstract form by using a functional Eckart-Young approach. *Journal of Mathematical Analysis and Applications*, 376(2):469–480, 2011.

[16] A. Falcó and W. Hackbusch. On minimal subspaces in tensor representations. *Foundations of Computational Mathematics*, 12:765–803, 2012.

[17] A. Falcó and A. Nouy. Proper generalized decomposition for nonlinear convex problems in tensor banach spaces. *Numerische Mathematik*, 121:503–530, 2012.

[18] A. Falcó, W. Hackbusch, and A. Nouy. Geometric structures in tensor representations. Preprint 9/2013, MPI MIS.

[19] L. Figueroa and E. Suli. Greedy approximation of high-dimensional Ornstein-Uhlenbeck operators. *Foundations of Computational Mathematics*, 12:573–623, 2012.

[20] L. Giraldi. *Contributions aux Méthodes de Calcul Basées sur l'Approximation de Tenseurs et Applications en Mécanique Numérique.* PhD thesis, Ecole Centrale Nantes, 2012.

[21] L. Giraldi, A. Nouy, G. Legrain, and P. Cartraud. Tensor-based methods for numerical homogenization from high-resolution images. *Computer Methods in Applied Mechanics and Engineering*, 254(0):154 – 169, 2013.

[22] L. Grasedyck. Hierarchical singular value decomposition of tensors. *SIAM J. Matrix Anal. Appl.*, 31:2029–2054, 2010.

[23] L. Grasedyck, D. Kressner, and C. Tobler. A literature survey of low-rank tensor approximation techniques. *GAMM-Mitteilungen*, 2013.

[24] W. Hackbusch. Tensor Spaces and Numerical Tensor Calculus, *Series in Computational Mathematics* volume 42. Springer, 2012.

[25] W. Hackbusch and S. Kuhn. A New Scheme for the Tensor Representation. *Journal of Fourier analysis and applications*, 15(5):706–722, 2009.

[26] S. Holtz, T. Rohwedder, and R. Schneider. The Alternating Linear Scheme for Tensor Optimisation in the TT format. *SIAM Journal on Scientific Computing*, 34(2):683–713, 2012.

[27] S. Holtz, T. Rohwedder, and R. Schneider. On manifolds of tensors with fixed TT rank. *Numer. Math.*, 120(4):701–731, 2012.

[28] B. N. Khoromskij and C. Schwab. Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs. *SIAM Journal on Scientific Computing*, 33(1):364–385, 2011.

[29] B. N. Khoromskij. Tensors-structured numerical methods in scientific computing: Survey on recent advances. *Chemometrics and Intelligent Laboratory Systems*, 110(1):1–19, 2012.

[30] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, September 2009.

[31] D. Kressner and C. Tobler. Low-rank tensor krylov subspace methods for parametrized linear systems. *SIAM Journal on Matrix Analysis and Applications*, 32(4):1288–1316, 2011.

[32] P. Ladevèze. *Nonlinear Computational Structural Mechanics - New Approaches and Non-Incremental Methods of Calculation.* Springer Verlag, 1999.

[33] P. Ladevèze, J.C. Passieux, and D. Néron. The LATIN multiscale computational method and the Proper Generalized Decomposition. *Computer Methods in Applied Mechanics and Engineering*, 199(21-22):1287–1296, 2010.

[34] H. G. Matthies and E. Zander. Solving stochastic systems with low-rank tensor compression. *Linear Algebra and its Applications*, 436(10), 2012.

[35] A. Nouy. A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, *Computer Methods in Applied Mechanics and Engineering*, 196(45-48) (2007) 4521-4537.

[36] A. Nouy, Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations, Archives of Computational Methods in Engineering, 16(3) (2009) 251-285.

[37] A. Nouy. Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Archives of Computational Methods in Engineering*, 17(4):403–434, 2010.

[38] A. Nouy. A priori model reduction through proper generalized decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1603–1626, 2010.

[39] I. V. Oseledets and E. E. Tyrtyshnikov. Breaking the curse of dimensionality, or how to use SVD in many dimensions. *SIAM J. Sci. Comput.*, 31(5):3744–3759, 2009.

[40] I. V. Oseledets. Tensor-train decomposition. *SIAM J. Sci. Comput.*, 33(5):2295–2317, 2011.

[41] T. Rohwedder and A. Uschmajew. On local convergence of alternating schemes for optimization of convex problems in the tensor train format. *SIAM J. Numer. Anal.*, 51(2):1134-1162, 2013.

[42] V. Temlyakov. *Greedy Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2011.

[43] V. Temlyakov. Greedy approximation. *Acta Numerica*, 17:235–409, 2008.

[44] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. Technical report, ANCHP-MATHICSE, Mathematics Section, EPFL, 2012.