Proceedings of the ASME 2010 3rd Joint US-European Fluids Engineering Summer Meeting and 8th International Conference on Nanochannels, Microchannels, and Minichannels
FEDSM-ICNMM2010
August 1-5, 2010, Montreal, Canada

FEDSM-ICNMM2010-30373

CONVERGENCE OF A GREEDY ALGORITHM ON NONLINEAR CONVEX PROBLEMS AND APPLICATION TO UNCERTAINTY QUANTIFICATION ON OBSTACLE PROBLEMS

Virginie Ehrlacher
Université Paris-Est, Ecole nationale des ponts et chaussées, CERMICS
6 et 8 avenue Blaise Pascal
77 455 Marne-la-Vallée Cedex 2
FRANCE
Email: ehrlachv@cermics.enpc.fr

ABSTRACT

In this article, a method is proposed to study uncertainty propagation on high-dimensional obstacle problems. A greedy algorithm, based on variable decomposition, is used to approximate the solution of regularized problems obtained by penalization of the initial problem. The convergence of this algorithm is a consequence of a more general theorem. Indeed, the algorithm converges for the minimization of a strongly convex functional whose derivative is Lipschitz on bounded sets. We describe how this algorithm was numerically implemented and present the results which were obtained with a one-dimensional membrane problem.

INTRODUCTION

In this article, we are interested in uncertainty quantification methods applied to high-dimensional obstacle problems. More precisely, the focus of this paper is rather on the development of methods to compute efficiently a *reduced model* which rapidly gives the output of interest as a function of the random variable which enters the input parameters, in the context of contact problems in continuum mechanics. Such a model can then be used to evaluate the distribution of the outputs (for a given distribution of the input parameters), or to reduce the variance in a Monte Carlo computation for example.

More precisely, let us assume that the noise on the parameters of the model can be modeled by a possibly large number of random variables $T = (T_1, \dots, T_p) \in \mathbb{R}^p$, so that the quantity of interest (say the deformation tensor) $u(t,x)$ is a function of $(p+d)$ variables, where d is the dimension of the physical space. The question is then how to approximate a function on such a high-dimensional space. The idea at the basis of many classical approaches [1, 2] such as stochastic collocation methods, Galerkin methods, perturbation methods, etc. is to look for the solution to this problem as a linear combination of tensor products:

$$u(t,x) = \sum_{k=1}^{K} \sum_{l=1}^{L} \lambda_{k,l}\, r_k(t) s_l(x), \tag{1}$$

where $(s_l)_{1\le l\le L}$, $(r_k)_{1\le k_1\le K}$ are bases of functions which are known *a priori* and where $(\lambda_{k,l})$ are scalars which are to be computed. In this case, the dimension of the problem $N = KL$ will be too large for a classical discretization method.

The method we are studying is a way to circumvent this difficulty. The principle is the following: (i) to rewrite the original problem as a minimization problem:

$$u \in \operatorname*{argmin}_{v\in V} \mathscr{E}(v)$$

where $\mathscr{E}$ is a functional defined on a Hilbert space V and (ii) to expand the solution in tensor products of lower-dimensional

functions

$$u_n(t,x) = \sum_{k=1}^{n} r_k(t)s_k(x)$$

which are *sequentially* determined:

$$(r_n,s_n) \in \underset{(r,s)\in V_t\times V_x}{\operatorname{argmin}} \mathscr{E}\left(\sum_{k=1}^{n-1} r_k(t)s_k(x) + r(t)s(x)\right). \qquad (2)$$

To rewrite the two problems mentioned above as minimization problems on Hilbert spaces, we penalize the constraints, namely the presence of the obstacle for the contact problem, which leads to the minimization of nonlinear convex functional. In this case, for each iteration of the algorithm, the computation of the pair of functions (r_n,s_n) is a problem of dimension $N' = K+L$. If the numbers of terms in the above expansion remains small enough, this method provides a way to circumvent the curse of dimensionality.

The method described above has been introduced by Chinesta [3] for solving high-dimensional Fokker-Planck equations, by Nouy [4] in the context of uncertainty quantification in mechanics, and is very much related to so-called greedy algorithms [5, 6] used in nonlinear approximation theory. A greedy algorithm is any algorithm that follows the problem solving approach of making the locally optimal choice at each stage with the hope of finding the global minimum [7]. We are able to prove [8] that this algorithm converges in a more general theoretical setting, which contains the penalization problem we consider.

We will first describe the obstacle problem we consider, and then state our convergence result in the general theoretical setting we mentioned above. Lastly, details will be given on how our algorithm was numerically implemented and the results we obtain on a one-dimensional membrane problem will be presented.

PRESENTATION OF THE PROBLEM

In this paper, we are interested in the convergence of a greedy algorithm for the study of uncertainty propagation on a high-dimensional obstacle problem. The prototypical example we consider is described in this section.

Throughout this article, p and d will denote some positive integers, and $\mathscr{T}$ and $\mathscr{X}$ some open sets of $\mathbb{R}^p$ and $\mathbb{R}^d$ respectively.

We assume that uncertainty can be modeled by a set of p random variables T_1, T_2, ..., T_p, and that the random vector $T = (T_1,...,T_p)$ takes its values in $\mathscr{T}$.

We consider also that the physical problem is defined over the domain $\mathscr{X}$, which is supposed to be a bounded subset of $\mathbb{R}^d$.

If H is a Hilbert space of functions defined over $\mathscr{X}$, we denote by

$$L^2_T(\mathscr{T},H) = \left\{v : \mathscr{T} \to H \mid \mathbb{E}\left[\|v(T)\|^2_H\right] < +\infty\right\},$$

where $\mathbb{E}$ denotes the expectation with respect to the probability law of T, and $\|.\|_H$ denotes the norm of H. We denote $\|v\|_{L^2_T(\mathscr{T},H)} = \left(\mathbb{E}\left[\|v(T)\|^2_H\right]\right)^{1/2}$ for $v \in L^2_T(\mathscr{T},H)$. The following assumption is done $\mathscr{X}$ is a regular bounded domain and let us denote by $\partial\mathscr{X}$ its boundary.

We consider the following physical problem: a membrane is stretched over $\mathscr{X}$ and is deflecting by some force having random pointwise density $f(x,T)$ for $x \in \mathscr{X}$. The membrane is fixed at a zero altitude at the boundary $\partial\mathscr{X}$. Moreover, the membrane is assumed to be hanging an obstacle whose altitude is also given by a random function $g(x,T)$. Let $g \in L^2_T(\mathscr{T},H^1_0(\mathscr{X}))$ and $f \in L^2_T(\mathscr{T},H^{-1}(\mathscr{X}))$. If we denote by $u(x,T)$ the altitude of the membrane at point $x \in \mathscr{X}$, we can show that u is the unique solution to the following obstacle problem with uncertainty [9](see Fig. 1):

$$\left\{\begin{array}{c} -\Delta_x u(T,x) \geq f(T,x) \\ u(T,x) \geq g(T,x) \\ (\Delta_x u(T,x) + f(T,x))\,(u(T,x) - g(T,x)) = 0 \end{array}\right\} \quad \forall x \in \mathscr{X},$$
$$u(T,x) = 0 \qquad \forall x \in \partial\mathscr{X}. \qquad (3)$$

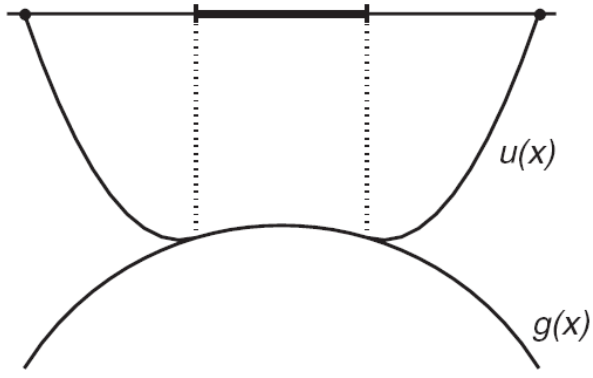

FIGURE 1. OBSTACLE PROBLEM.

The following formulation is equivalent. Let us denote

$$\mathscr{K}_g = \left\{v \in L^2_T(\mathscr{T},H^1_0(\mathscr{X})) \mid \forall (t,x) \in \mathscr{T}\times\mathscr{X},\ v(t,x) \geq g(t,x)\right\}.$$

Then, solving the obstacle problem consists in solving the minimization problem

$$\inf_{v\in\mathscr{K}_g} \mathscr{J}(v), \tag{4}$$

where

$$\mathscr{J}(v) = \begin{array}{c} \mathbb{E}\left[\frac{1}{2}\int_{\mathscr{X}} |\nabla_x v(T,x)|^2 dx\right] \\ -\mathbb{E}\left[\langle f(T,.), v(T,.)\rangle_{H^{-1}(\mathscr{X}),H_0^1(\mathscr{X})}\right]. \end{array}$$

One of the main difficulty of this kind of problems is their very high nonlinearity. Many methods have been proposed to approach the solution of these problems. See for example [10], [11], [12] and [9]. Among them, penalization methods (see [9] and [10]) are among the most widely used. They consist in approximating the solution of an initial obstacle problem by a sequence of solutions of penalized problems defined over the entire Hilbert space.

Let ρ be a parameter in $\mathbb{R}_+$. Such a penalized problem associated with problem (4) may be defined as

$$\inf_{v\in L_T^2(\mathscr{T},H_0^1(\mathscr{X}))} \mathscr{J}_\rho(v), \tag{5}$$

where $\mathscr{J}_\rho(v) = \mathscr{J}(v) + \mathbb{E}\left[\frac{\rho}{2}\int_{\mathscr{X}} [g(T,x)-v(T,x)]_+^2 dx\right]$.

Here and below, we denote by $[a]_+$ the positive part of the real number a, i.e. $[a]_+ = 0$ if $a \leq 0$ and $[a]_+ = a$ if $a \geq 0$.

When ρ goes to infinity, the solution u_ρ of problem (5) strongly converges to the solution u of problem (4). The goal of the algorithm we described in the previous section is to calculate the solution u_ρ of this regularized problem for a given value of the parameter ρ.

There exist several variants of the obstacle problem which could be tackled with our algorithm. We refer to [9] or [12] for such examples.

DESCRIPTION OF THE ALGORITHM

We denote for all $(r,s) \in L_T^2(\mathscr{T},\mathbb{R}) \times H_0^1(\mathscr{X})$,

$$r\otimes s : \left\{ \begin{array}{ccc} \mathscr{T}\times\mathscr{X} & \to & \mathbb{R} \\ (t,x) & \mapsto & r(t)s(x) \end{array} \right. .$$

In this case, our algorithm can be rewritten in the following form. Set $f_0 = f$ and $g_0 = g$ and define recursively $(r_n,s_n) \in L_T^2(\mathscr{T},\mathbb{R}) \times H_0^1(\mathscr{X})$ as

$$(r_n,s_n) \in \underset{(r,s)\in L_T^2(\mathscr{T},\mathbb{R})\times H_0^1(\mathscr{X})}{\operatorname{argmin}} \mathscr{E}_n(r\otimes s), \tag{6}$$

with

$$\begin{aligned} \mathscr{E}_n(r\otimes s) = \mathbb{E}&\left[\frac{1}{2}\int_{\mathscr{X}} |\nabla_x r\otimes s(T,x)|^2 dx\right] \\ &-\mathbb{E}\left[\int_{\mathscr{X}} f_{n-1}(T,x) r\otimes s(T,x) dx\right] \\ &+\frac{\rho}{2}\mathbb{E}\left[\int_{\mathscr{X}} [g_{n-1}(T,x) - r\otimes s(T,x)]_+^2 dx\right], \end{aligned}$$

where

$$\begin{aligned} f_n &= f_{n-1} + \Delta_x(r_n\otimes s_n), \\ g_n &= g_{n-1} - r_n\otimes s_n. \end{aligned}$$

The fact that the iterations defined in (6) are well-defined and that the algorithm converges towards the desired result is a consequence of a more general result which we present in the next section.

GENERAL THEORETICAL SETTING

Let us introduce a general setting which will contain as particular example the prototypical problem introduced in the previous section.

Let V_t and V_x be Hilbert spaces of real-valued functions respectively defined over $\mathscr{T}$ and $\mathscr{X}$. Let $\|.\|_t$ and $\|.\|_x$ be the norms of V_t and V_x.

We define the following tensor product for all $(r,s) \in V_t \times V_x$,

$$r\otimes s : \left\{ \begin{array}{ccc} \mathscr{T}\times\mathscr{X} & \to & \mathbb{R} \\ (t,x) & \mapsto & r(t)s(x) \end{array} \right. , \tag{7}$$

which defines a real-valued function defined over $\mathscr{T}\times\mathscr{X}$.

We also denote $\Sigma = \{r\otimes s \mid (r,s) \in V_t \times V_x\}$.

Let V be a Hilbert space of real-valued functions defined over $\mathscr{T}\times\mathscr{X}$. The scalar product of V is denoted $\langle .,. \rangle$ and the associated norm is denoted $\|.\|_V$.

Let $\mathscr{E}$ be a differentiable real-valued functional defined on V. For all $v \in V$, we denote by $\mathscr{E}'(v)$ the gradient of $\mathscr{E}$ at v.

We make the following assumptions:

$(A1)$ Span(Σ) is a dense subset of V for $\|.\|_V$.

$(A2)$ For all sequences of Σ bounded in V, there exists a subsequence which weakly converges in V towards an element of Σ.

$(A3)$ The functional $\mathscr{E}$ is strongly convex for $\|.\|_V$, i.e. there exists a constant $\alpha \in \mathbb{R}_+^*$ for which

$$\forall v, w \in V,\ \mathscr{E}(v) \geq \mathscr{E}(w) + \langle \mathscr{E}'(w), v-w \rangle + \frac{\alpha}{4}\|v-w\|_V^2. \quad (8)$$

The functional $\mathscr{E}$ is also said to be α-convex.

$(A4)$ The gradient of $\mathscr{E}$ is Lipschitz on bounded sets: for each bounded subset K of V, there exists a non-negative constant $L_K \in \mathbb{R}_+$ such that

$$\forall v, w \in V,\ \ \|\mathscr{E}'(v) - \mathscr{E}'(w)\|_V \leq L_K \|v-w\|_V. \quad (9)$$

The unique global minimizer of $\mathscr{E}$ on V is denoted by $u \in V$. Its existence and uniqueness are ensured by the α-convexity of the functional $\mathscr{E}$.

The algorithm introduced in (6) can then be rewritten in this context: the sequence $((r_n, s_n))_{n\in\mathbb{N}^*} \in (V_t \times V_x)^{\mathbb{N}^*}$ is defined recursively by

$$(r_n, s_n) \in \underset{(r,s)\in V_t\times V_x}{\operatorname{argmin}}\ \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k + r \otimes s\right). \quad (10)$$

Throughout this article, we will denote for all $n \in \mathbb{N}^*$,

$$u_n = \sum_{k=1}^{n} r_k \otimes s_k. \quad (11)$$

We have then the following theorem [8].

Theorem 1. *Under the assumptions $(A1)$, $(A2)$, $(A3)$ and $(A4)$, the iterations of the algorithm are well-defined, in the sense that (10) has at least one minimizer (r_n, s_n). Moreover, the sequence $(u_n)_{n\in\mathbb{N}}$ strongly converges in V towards u.*

We have an even more general result in the case when the norm of V satisfies the following property [8]:

$(A5)$ There exist $\beta, \gamma \in \mathbb{R}_+$ such that

$$\forall (r,s) \in V_t \times V_x,\ \ \beta\|r\|_{V_t}\|s\|_{V_x} \leq \|r\otimes s\|_V \leq \gamma\|r\|_{V_t}\|s\|_{V_x}. \quad (12)$$

Theorem 2. *Let us suppose that the assumptions $(A1)$, $(A2)$, $(A3)$, $(A4)$ and $(A5)$ hold true. We define recursively $(r_n, s_n) \in V_t \times V_x$ as a **local** minimum of $V_t \times V_x \ni (r,s) \mapsto \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k + r \otimes s\right)$ such that $\mathscr{E}\left(\sum_{k=1}^{n} r_k \otimes s_k\right) < \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right)$.*

In this case, the sequence $\left(\sum_{k=1}^{n} r_k \otimes s_k\right)_{n\in\mathbb{N}^*}$ still converges strongly in V towards u.

Besides, in the finite dimensional case, we are able to prove that this algorithm converges exponentially fast [8].

Theorem 3. *We consider the algorithm described in Theorem 2. We assume that V_t and V_x are finite-dimensional and that assumptions $(A1)$, $(A2)$, $(A3)$ and $(A4)$ are fulfilled. Then there exist two constants $\tau > 0$ and $\sigma \in (0,1)$ such that,*

$$0 \leq \mathscr{E}(u_n) - \mathscr{E}(u) \leq \tau\sigma^n, \quad (13)$$

and

$$\|u - u_n\|_V \leq 2\alpha^{-1/2}\tau^{1/2}\sigma^{n/2}. \quad (14)$$

NUMERICAL RESULTS

In this section, we will describe how we implemented the algorithm we described in section 3 for the resolution of problem (5). We will then present the numerical results we obtained on a one-dimensional membrane problem with uncertainty.

Implementation of the algorithm

Let $D \in \mathbb{R}^{l\times l}$ and $F, G \in \mathbb{R}^{k\times l}$ represent respectively the discretization by finite elements of the one-dimensional operator $-\partial_{xx}$ and the discretization by finite elements of functions f and g. Problem (5) can be rewritten:

$$\text{Find } U \in \mathbb{R}^{k\times l} \text{ such that}$$
$$U \in \underset{V\in\mathbb{R}^{k\times l}}{\operatorname{argmin}} \frac{1}{2}VD : V - F : V + \frac{\rho}{2}[G-V]_+ : [G-V]_+,$$

where for $A, B \in \mathbb{R}^{k\times l}$, $A : B = \mathrm{Tr}(AB^T) = \sum_{1\leq i\leq k}\sum_{1\leq j\leq l} A_{ij}B_{ij}$.

This problem is equivalent to:

$$\text{Find } U \in \mathbb{R}^{k\times l} \text{ such that } UD = F + \rho[G-U]_+.$$

Our algorithm can then be rewritten as:

Choose a threshold $\varepsilon > 0$ and set $F_0 = F$, $G_0 = G$. At iteration $n \geq 1$:

1) Find R_n and S_n two vectors respectively in $\mathbb{R}^k$ and $\mathbb{R}^l$ such that:

$$(R_n,S_n) \in \underset{(R,S)\in\mathbb{R}^k\times\mathbb{R}^l}{\operatorname{argmin}} \mathscr{E}_n(R,S),$$

with

$$\mathscr{E}_n(R,S) = \frac{1}{2}(RS^T)D : (RS^T) - F_{n-1} : (RS^T) + \frac{\rho}{2}[G_{n-1} - RS^T]_+ : [G_{n-1} - RS^T]_+.$$

2) Set $F_n = F_{n-1} - (R_n S_n^T)D$, and $G_n = G_{n-1} - R_n S_n^T$.
3) If $\|F_n - \rho[G_n]_+\| \geq \varepsilon$, proceed to iteration $n+1$. Otherwise, stop.

The remaining question is how can we compute the solution (R_n,S_n) at step 1? The following section will describe this critical step.

Computing (R_n,S_n)

Fixed-point procedure Let us first describe the method which is generally [6] used in the singular value decomposition case. We fix a given matrix $M \in \mathbb{R}^{k\times l}$. The goal of the singular value decomposition is to apply our greedy algorithm to the energy $\mathscr{E}(V) = \|M-V\|_V^2$ for $V \in \mathbb{R}^{k\times l}$.

The algorithm for singular value decomposition can be rewritten in the following form.

Choose a threshold $\varepsilon > 0$ and set $M_0 = M$. At iteration $n \geq 1$,

1) Find two vectors R_n and S_n respectively in $\mathbb{R}^k$ and $\mathbb{R}^l$ such that

$$(R_n,S_n) \in \underset{(R,S)\in\mathbb{R}^k\times\mathbb{R}^l}{\operatorname{argmin}} \left\|M_{n-1} - RS^T\right\|_V^2. \qquad (15)$$

2) Set $M_n = M_{n-1} - R_n S_n^T$.
3) If $\|M_n\|_V \geq \varepsilon$, proceed to iteration $n+1$. Otherwise, stop.

In the case when we have for all $(R,S) \in \mathbb{R}^k \times \mathbb{R}^l$, $\|RS^T\|_V = \|R\|_{V_t}\|S\|_{V_x}$, the Euler-Lagrange equations associated to this problem can be rewritten as

$$\begin{cases} \|S_n\|_{V_x}^2 R_n = M_{n-1}S_n, \\ \|R_n\|_{V_t}^2 S_n = (M_{n-1})^T R_n. \end{cases}$$

The method which is generally used (see for example [6]) to solve these Euler-Lagrange equations is a fixed-point algorithm. The fixed-point procedure then simply reads (for a fixed n): at iteration $m \geq 0$, compute two vectors $(R_n^m, S_n^m) \in \mathbb{R}^k \times \mathbb{R}^l$ such that

$$\begin{cases} \|S_n^m\|_{V_x}^2 R_n^{m+1} = M_{n-1}S_n^m, \\ \|R_n^{m+1}\|_{V_t}^2 S_n^{m+1} = (M_{n-1})^T R_n^{m+1}. \end{cases} \qquad (16)$$

One can check that this procedure is similar to the power method to compute the largest eigenvalues (and associated eigenvectors) of the matrix $(M_{n-1})^T M_{n-1}$.

One could think of transposing this fixed-point procedure to the case of the obstacle problem we consider in this article. In our case, the Euler-Lagrange equations

$$\begin{cases} (R_n : R_n)DS_n = F_{n-1}^T R_n + \rho[G_{n-1} - R_n S_n^T]_+^T R_n, \\ (DS_n : S_n)R_n = F_{n-1}S_n + \rho[G_{n-1} - R_n S_n^T]_+ S_n. \end{cases}$$

could be solved with a fixed point algorithm, which, at iteration m, might be written as

$$\begin{cases} (R_n^m : R_n^m)DS_n^{m+1} = F_{n-1}^T R_n^m + \rho[G_{n-1} - R_n^m S_n^{mT}]_+^T R_n^m, \\ (DS_n^{m+1} : S_n^{m+1})R_n^{m+1} = F_{n-1}S_n^{m+1} + \rho[G_{n-1} - R_n^m S_n^{m+1,T}]_+ S_n^{m+1}. \end{cases}$$

However, we have checked numerically that such a fully-explicit algorithm did not converge for large values of the parameter ρ. Fig 2 represents an example of the evolution of the convergence criterion $\|R_n^m S_n^{mT} - R_n^{m+1} S_n^{m+1,T}\|$ in the numerical case which is detailed in the subsection *One-dimensional membrane problem*. This issue might be solved by considering an implicit formulation of the algorithm but we did not test such a formulation. We prefered a more general approach (in the sense that it can be easily computed whatever the functional $\mathscr{E}$) and that it respects the general setting of Theorem 2.

Minimization procedure The approach we adopted then was the following. We choose to set an initial pair $(R_n^0, S_n^0) \in \mathbb{R}^k \times \mathbb{R}^d$ and then perform a gradient algorithm to find a local minimum of the function

$$\frac{1}{2}(RS^T)D : (RS^T) - F_{n-1} : (RS^T) + \frac{\rho}{2}[G_{n-1} - RS^T]_+ : [G_{n-1} - RS^T]_+.$$

The main difficulty is to find a proper initial pair (R_n^0, S_n^0) such that

$$\begin{aligned} &\tfrac{1}{2}(R_n^0 S_n^{0T})D : (R_n^0 S_n^{0T}) - F_{n-1} : (R_n^0 S_n^{0T}) \\ &+ \tfrac{\rho}{2}[G_{n-1} - R_n^0 S_n^{0T}]_+ : [G_{n-1} - R_n^0 S_n^{0T}]_+ \\ &< \tfrac{\rho}{2}[G_{n-1}]_+ : [G_{n-1}]_+, \end{aligned}$$

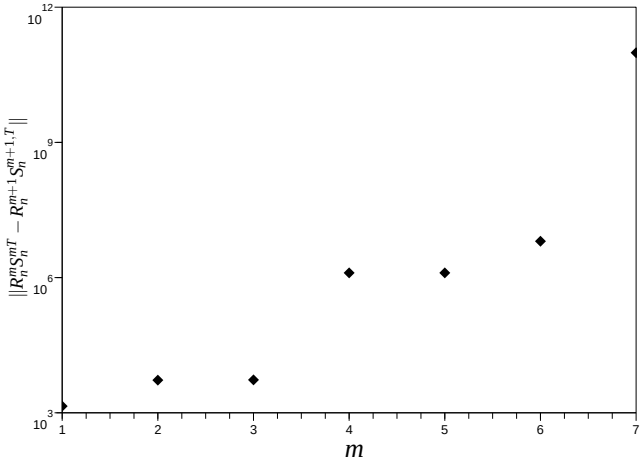

FIGURE 2. CONVERGENCE CRITERION OF THE FULLY-EXPLICIT FIXED-POINT PROCEDURE.

to ensure that the energy decreases.

Let us describe our approach in the continuous setting with the notations used in the previous sections. It then consists in finding a pair $(r_n^0, s_n^0) \in V_t \times V_x$ such that

$$\mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k + r_n^0 \otimes s_n^0\right) < \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right).$$

We notice that for $(r,s) \in V_t \times V_x$, and $\varepsilon > 0$, we have

$$\begin{array}{c} \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k + \varepsilon r \otimes s\right) - \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right) \\ = \varepsilon \left\langle \mathscr{E}'\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right), r \otimes s \right\rangle + o(\varepsilon), \end{array}$$

for ε small enough.

The idea is then to find a pair $(r,s) \in V_t \times V_x$ such that $\mathscr{E}'(\sum_{k=1}^{n-1} r_k \otimes s_k, r \otimes s\rangle < 0$ so that there exists $\varepsilon > 0$ small enough so that $\mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k + \varepsilon r \otimes s\right) - \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right) < 0$. Then, $r_n^0 \otimes s_n^0 = \varepsilon r \otimes s$ is a good initial guess.

Let us first consider the pair $(\overline{r_n^0}, \overline{s_n^0}) \in V_t \times V_x$ such that

$$(\overline{r_n^0}, \overline{s_n^0}) \in \underset{(r,s)\in V_t \times V_x}{\operatorname{argmin}} \frac{1}{2} \left\| \mathscr{E}'\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right) - r \otimes s \right\|_V^2.$$

In other words, we consider $(\overline{r_n^0}, \overline{s_n^0})$ the first term of the singular value decomposition of $\mathscr{E}'\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right)$ in V. The Euler-Lagrange equations then imply

$$-\langle \mathscr{E}'(u_{n-1}) - \overline{r_n^0} \otimes \overline{s_n^0}, \overline{r_n^0} \otimes \overline{s_n^0} \rangle = 0,$$

$$\langle \mathscr{E}'(u_{n-1}), \overline{r_n^0} \otimes \overline{s_n^0} \rangle = \|\overline{r_n^0} \otimes \overline{s_n^0}\|_V^2 > 0.$$

By taking $r_n^0 \otimes s_n^0 = -\varepsilon \overline{r_n^0} \otimes \overline{s_n^0}$, there exists then $\varepsilon > 0$ small enough such that

$$\mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k - \varepsilon \overline{r_n^0} \otimes \overline{s_n^0}\right) - \mathscr{E}\left(\sum_{k=1}^{n-1} r_k \otimes s_k\right) < 0.$$

In the discrete case associated to problem (5), $(\overline{R_n^0}, \overline{S_n^0})$ is obtained by taking the first term of the singular value decomposition of the matrix $F_{n-1} + \rho[G_{n-1}]_+$. This can be done with a fixed point procedure similar to (16).

One-dimensional membrane problem

In this section, we present the results we obtained with this algorithm on the following membrane problem.

We suppose $\mathscr{X} = \mathscr{T} = (0,1)$. We consider a random variable T following a uniform law of probability on the interval $(0,1)$. We wish to study problem (4) with the following values for f and g,

$$\begin{array}{c} \forall (t,x) \in (0,1)^2\ ,\ f(t,x) = -1 \\ \text{and } g(t,x) = t[\sin(3\pi x)]_+ + (t-1)[\sin(3\pi x)]_-. \end{array}$$

We denoted by $[a]_-$ the negative part of $a \in \mathbb{R}$, i.e. $[a]_- = 0$ if $a \geq 0$, and $[a]_- = -a$ if $a \leq 0$.

This represents the problem of a rope attached at $x = 0$ and $x = 1$ and subjected to gravity. The term $f(t,x)$ represents the external forces that are applied to the rope. Here they are chosen to be constant in order to model gravity. This rope rests upon obstacles whose altitudes are given by $g(t,x)$. The quantity $u(t,x)$ then represents the altitude of the rope at absciss x when the random parameter $T = t$.

We choose to approximate this problem by problem (5) with parameter $\rho = 2500$. Unfortunately, we were not able to perform simulations for larger values of ρ. Indeed, if we try to increase more the parameter ρ, we will encounter classical problems of penalization methods, that is the ill-conditioning of the computed matrices.

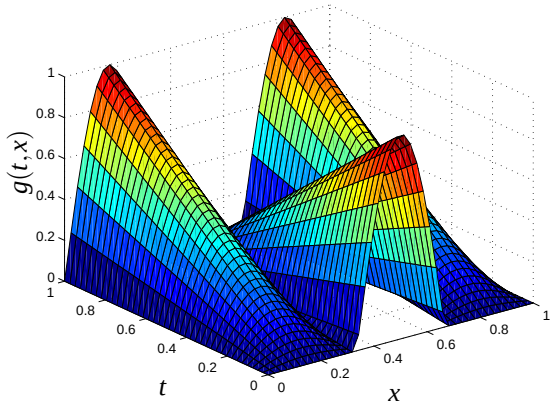

FIGURE 3. ALTITUDE OF THE OBSTACLES.

The solution we obtain for u is then not exactly equal to the solution of problem (4). Indeed, this is due to the fact that we approximate a solution u_ρ of the penalized problem (5) for $\rho = 2500$. This is the main drawback of our method: we do not approximate directly the solution of the initial obstacle problem but the solution of a close regularized problem. In fact, error estimations are well-known in the case of such penalized problems [10] [9]. There exists a constant $C > 0$ independent of ρ such that $\|u - u_\rho\|_V \leq C\frac{1}{\rho}$.

The problems which are related to penalization methods (that is the ill-conditioning of matrices and the computation of an approximate solution) are well-known. To tackle this problem, other iterative methods have been developped [11] [12] such as augmented-lagrangian algorithms, which avoid the problem of ill-conditioned matrices and converge towards the real solution of the initial obstacle problem. The extension of our work which consists in combining our greedy algorithm with such methods is work in progress.

Here, we provide the numerical results we obtained by using the penalization setting we described earlier. The problem was discretized with a regular mesh and $\mathbb{P}_1$ finite elements in each direction. We choose $k = l = 40$ as discretization parameters.

Fig. 3 represents the altitude of the obstacles given by $g(t,x)$ for $(t,x) \in [0,1]^2$.

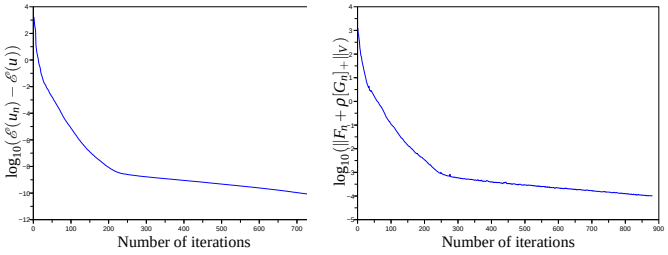

FIGURE 4. EVOLUTION OF THE ENERGY AND OF THE V-NORM OF THE RESIDUAL.

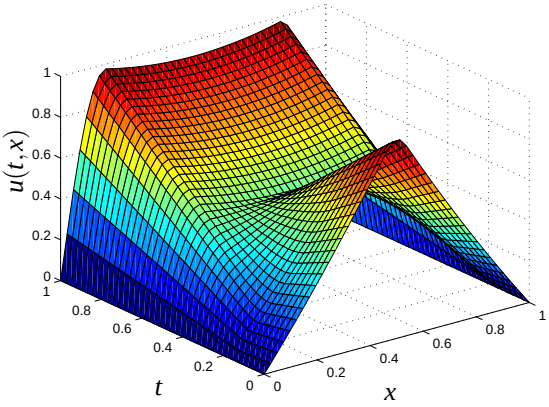

FIGURE 5. ALTITUDE OF THE ROPE.

We then applied the algorithm described in the previous sections with the following stopping criterion: $\|F_n + \rho[G_n]_+\|_V < 10^{-4}$ with $\|A\|_V = \sqrt{\mathrm{Tr}(AA^T)} = \sqrt{\sum_{i=1}^k \sum_{j=1}^l A_{ij}^2}$ for $A \in \mathbb{R}^{k\times l}$.

Fig. 4 represents the evolution of $\log_{10}(\mathscr{E}(u_n) - \mathscr{E}(u))$ and of $\log_{10}(\|F_n + \rho[G_n]_+\|_V)$.

We can see that our algorithm captures very quickly the main modes of the solution and that both the energy and the V-norm of the residue $\|F_n + \rho[G_n]_+\|_V$ converges exponentially fast, as predicted by Theorem 3.

Fig. 5 represents the results obtained for the solution $u(t,x)$.

Fig. 6 represents the number of terms that are computed in the expansion (11) for different values of $k = l$ and of ρ. We can see that, the larger ρ is the more terms need to be computed in order to obtain a fixed level of precision $\varepsilon = 10^{-4}$ in the approximation. This is another issue which would be solved with the use of augmented-lagrangian methods instead of penalized formulations.

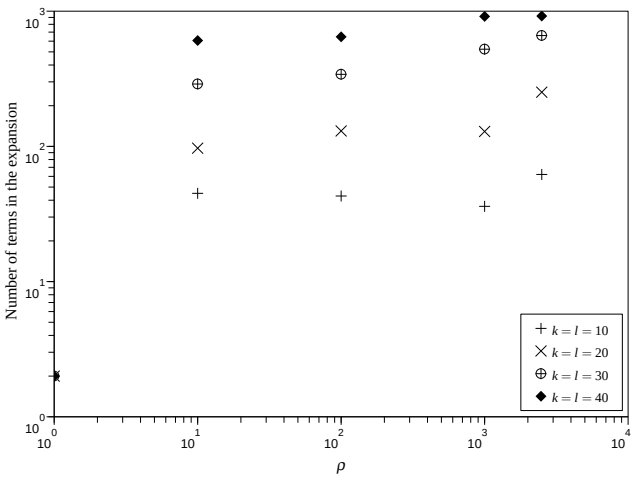

FIGURE 6. RATE OF CONVERGENCE AS A FUNCTION OF ρ.

CONCLUSION

In this article, we presented a greedy algorithm based on variable decomposition aiming at computing the global minimum of a strongly convex energy functional. We proved that, provided that the gradient of the energy is Lipschitz on bounded sets, and that the Hilbert spaces considered satisfy assumptions $(A1)$ and $(A2)$, then the approximation given by our algorithm strongly converges towards the desired result. One of the main advantage of the algorithm is that it can deal with highly nonlinear problems. We also proved that in finite dimension, this algorithm converges exponentially fast.

We applied this algorithm in the context of uncertainty quantification in obstacle problems. In this frame, we considered regularizations of this kind of problems by penalization methods. Indeed, the obstacle problem can be approximated by a global minimization problem defined over the entire Hilbert space of some strongly convex energy functional where the constraints of the initial problem are replaced by penalization terms in the expression of the functional. Our algorithm gives a good approximation of the solutions of the regularized problem. However, the problem of ill-conditioned matrices, which is inherent to penalization methods, limits the accuracy with which we can approach the solution of the initial obstacle problem.

A way to circumvent this problem could be to use augmented Lagrangian methods (see [11], [12] or [9]) instead of penalization methods. Indeed, the former algorithms converge towards the true solution of the initial obstacle problems. The adaptation of our algorithm to such methods is actually work in progress.

Another extension of our work would be to consider other problems than obstacle problems. In [4], a similar algorithm based on Proper Generalized Decomposition is used to study uncertainty quantification upon a Burger type equation. We think that it could be possible to extend our proof of convergence in the case of such hyperbolic systems.

ACKNOWLEDGMENT

I gratefully thank Tony Lelièvre and Eric Cancès with whom this work has been done. Financial support from Michelin is acknowledged.

REFERENCES

[1] Nouy, A., 2009. "Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations". *Archives of Computational Methods in Engineering,* ***16***, pp. 251–285.

[2] Ghanem, R., and Spanos, P., 1991. *Stochastic Finite Elements - A Spectral Approach*. Springer-Verlag.

[3] Ammar, A., Mokdad, B., Chinesta, F., and Keunings, R., 2002. "A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids". *Journal of Non-Newtonian Fluid Mechanics,* ***139***, pp. 153–176.

[4] Nouy, A., and Maitre, O. L., 2009. "Generalized spectral decomposition for stochastic non linear problems". *Journal of Computational Physics,* ***228***(1), pp. 202–235.

[5] Temlyakov, V., 2008. "Greedy approximation". *Acta Numerica,* ***17***, pp. 235–409.

[6] Bris, C. L., Lelièvre, T., and Maday, Y., 2009. "Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations". *Constructive Approximation,* ***30***(3), pp. 621–651.

[7] Cormen, T., Leiserson, C., and Rivest, R., 1990. *Introduction to algorithms*. MIT Press, Cambridge USA.

[8] E.Cancès, Ehrlacher, V., and Lelièvre, T., 2010. "Convergence of a greedy algorithm on convex non linear problems". *in preparation*.

[9] Grossmann, C., Roos, H.-G., and Stynes, M., 2007. *Numerical Treatment of Partial Differential Equations*. Springer, Berlin Heidelberg.

[10] Glowinski, R., Lions, J., and Trémolières, R., 1976. *Analyse numérique des inéquations variationnelles - Théorie générale et premières applications*. Editions Dunod, Paris.

[11] Fortin, M., and Glowinski, R., 1982. *Méthodes de Lagrangien augmenté - Application à la résolution numérique de problèmes aux limites*. Editions Dunod, Paris.

[12] Glowinski, R., and Tallec, P. L., 1989. *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. Society for Industrial and Applied Mathematics, Philadelphia.