



Sensitivity analysis of energy contracts management problem by stochastic programming techniques

Zhihao Cen, J. Frederic Bonnans, Thibault Christel

► To cite this version:

Zhihao Cen, J. Frederic Bonnans, Thibault Christel. Sensitivity analysis of energy contracts management problem by stochastic programming techniques. [Research Report] RR-7574, 2011. inria-00579668v1

HAL Id: inria-00579668

<https://inria.hal.science/inria-00579668v1>

Submitted on 24 Mar 2011 (v1), last revised 1 Aug 2011 (v2)

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.



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

***Sensitivity analysis of energy contracts management
problem by stochastic programming techniques***

Zhihao Cen — J. Frédéric Bonnans — Thibault Christel

N° 7574

Mar 2011

Thème NUM

 ***rapport
de recherche***

Sensitivity analysis of energy contracts management problem by stochastic programming techniques

Zhihao Cen^{*}, J. Frédéric Bonnans[†], Thibault Christel[‡]

Thème NUM — Systèmes numériques
Équipes-Projets Commands

Rapport de recherche n° 7574 — Mar 2011 — 21 pages

Abstract: We consider a model of medium-term commodity contracts management. Randomness takes place only in the prices on which the commodities are exchanged whilst state variable is multi-dimensional. In [8], we proposed an algorithm to deal with such problem, based on quantization of random process and a dual dynamic programming type approach. We obtained accurate estimates of the optimal value and a suboptimal strategy from this algorithm. In this paper, we analyse the sensitivity with respect to parameters driving the price model. We discuss the estimate of marginal price based on the Danskin's theorem. Finally, some numerical results applied to realistic energy market problems have been performed. Comparisons between results obtained in [8] and other classical methods are provided and evidence the accuracy of the estimate of marginal prices.

Key-words: sensitivity, stochastic programming, Danskin theorem, dual dynamic programming

^{*} Total, Paris La Defense, COMMANDS team, INRIA-Saclay and CMAP, Ecole Polytechnique, Palaiseau, France (zhihao.cen@polytechnique.edu).

[†] COMMANDS team, INRIA-Saclay, CMAP, Ecole Polytechnique, Palaiseau, and Laboratoire de Finance des Marchés de l'Energie, Paris, France (Frederic.Bonnans@inria.fr).

[‡] Total, Paris La Defense, France (thibault.christel@total.com).

Analyse de sensibilité de problèmes de gestion de contrats d'énergie par des techniques de programmation stochastique

Résumé : Nous considérons un modèle de gestion de contrats de produit de base en moyen terme. Les aléas ne portent que sur le dynamique de prix. Dans [8], nous avons proposé un algorithme pour ce type de problème, basé sur la méthode de quantification de processus aléatoire, et l'approche de programmation dynamique duale. Nous avons obtenu un estimateur précis de la valeur optimale et une stratégie sous-optimale par cet algorithme. Dans cet article, nous analysons la sensibilité par rapport aux paramètres dans le processus aléatoire (modèle de prix). Nous discutons un estimateur de coût marginal basé sur le théorème de Danskin. A la fin, des tests numériques appliqués à des problèmes réels dans le marché l'énergie ont été réalisés. Les comparaisons du résultat par la méthode dans [8] et celui par des autres méthodes, donnent des estimateurs précis sur le coût marginal.

Mots-clés : sensibilité, programmation stochastique, théorème Danskin, programmation dynamique duale

1 Motivation

We study a class of problems in the following settings. First of all, we consider a uniform discretization of time horizon $[0, T]$, $0 = t_0 < t_1 < \dots < t_K = T$, and a discrete time Markov process $(\xi_k) := (\xi_{t_k})$ (in the following, we will replace in index t_k by k for sake of clarity) in the probability space $L^2(\Omega, (\mathcal{F}_k), \mathbf{P}; \mathbb{R}^d)$. The canonic filtration associated with (ξ_k) is denoted by $\mathcal{F}_k := \sigma(\xi_s, 0 \leq s \leq k)$. We may write its dynamics as

$$\xi_{k+1} = f(W_k, \xi_k, \alpha_k) \quad 0 \leq k \leq K-2, \quad (1)$$

where ξ_0 is deterministic, (α_k) is a sequence of deterministic parameters, and (W_k) is a sequence of independent squared integrable random variables, independent of \mathcal{F}_k , finally f is a measurable mapping.

The stochastic dynamic decision problem has the following expression:

$$\begin{aligned} \inf_{(u_k) \in U} \quad & \mathbb{E} \left[\sum_{k=0}^{K-1} c_k(\xi_k) \cdot u_k \right] \\ \text{subject to} \quad & u_k \in \mathcal{U}_k, \quad \text{almost surely,} \\ & x_{k+1} = x_k + A_k u_k, \quad x_0 = 0, \\ & x_T \in \mathcal{X}_T \quad \text{almost surely.} \end{aligned} \quad (2)$$

where $u_k \in \mathbb{R}^n$ is the control variable, $U := L^2(\Omega, (\mathcal{F}_k), \mathbf{P}; \mathbb{R}^n)$ is the functional space of control variable u_k , \mathcal{U}_k is a compact nonempty polyhedral set in \mathbb{R}^n , $x_k \in \mathbb{R}^m$ represents the state variable, \mathcal{X}_T is the set of admissible final state assumed to be a nonempty polyhedral set in \mathbb{R}^m taking place in the final stage constraint over state variable x_T , $A_k \in \mathbb{R}^{m \times n}$ is the dynamic matrix, $c_k(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}$ is the running cost unit, assumed to be Lipschitz.

Such formulation (2) is motivated by the fact that most medium and long-term energy contracts can be modelled as multi stage stochastic dynamic programs. Indeed, most of the contracts define the terms on how the commodity will be exchanged throughout the duration of the contract. These terms generally state that the quantity to be taken at time t_k (modelled by the control u_k) is bounded by minimal and maximal amounts. They also specify the price at which the commodity will be exchanged at time t_k ; this price is generally a function of the spot price (modelled by $c_k(\xi_k)$) observed at that period. Since the spot price is unknown at the beginning, the problem becomes stochastic. The aim of the holder of such contract is to optimize its decisions with respect to the price information available to him at that time. Here, we choose to maximize the expectation of the revenue. In order to be compliant with the formulation of (2), we will flip the sign of the real cost-revenue functions to transform a maximizing problem into a standard minimization problem.

Most of the articles in the literature focus on the procedures on how to solve problem (2) and hence how to get the contract's price (see Carmona and Touzi [12], Carmona and Ludkovski [11], and Barrera-Esteve et al [6]). Although essential, this information needs to be coupled with sensitivity information in a trading perspective. Indeed, the values of sensitivities (called the greeks) allow traders to replicate the price variations originating from the contracts they have to manage on a regular basis. See [20, Chapter 17] for further discussions on hedging strategies using the greeks applied to vanilla options.

A lot of efforts have been made in order to assess proper sensitivities from discontinuous payoffs such as digital or barrier options. However, only a handful of papers focus on optimal exercise type options and most of them concern Bermudian options which are common in the banking industry. Recently, Cont and Fournié have developed functional Itô formula in order to study the sensitivities of path dependent options without control [14, 13]. Our purpose in this paper is to provide an alternative method to estimate sensitivities for problem (2) based on the Danskin's theorem which is well known among the optimization practitioners. The advantage of the Danskin's theorem is that it can be associated to any discretization type method such as PDE, Monte Carlo or optimal quantization. Also, it does not require to reestimate the conditional expectations once they have been computed during the price estimation procedure.

This paper is organised as follows: section 2 presents the resolution framework of problem (2) articulated around optimal quantization and stochastic dual dynamic programming (SDDP). Section 3 introduces the model driving the prices and section 4 focuses on the Danskin's theorem through the computation of sensitivities. Convergence results are also provided in section 4. Finally, numerical tests are performed in section 5.

2 A review of quantization discretization and stochastic dual dynamic programming approach

In this section, we briefly present the algorithm which was introduced in our previous article [8].

2.1 Discretization

In many real life problems, the discrete time random process (ξ_k) may possibly take value within an infinite set. In order to make the problem numerically tractable, we start by discretizing the random process by a tree. Because of the Markov property of random process, a recombining tree is the method of choice for reducing complexity. Bally et al [3, 2] introduced the vectorial quantization tree method as a new discretization method. The main idea is to replace ξ_k by the approximation $\hat{\xi}_k$ whose support is a finite set $\Gamma_k = \{\xi_k^i, 1 \leq i \leq N_k\}$. The quantization $\hat{\xi}$ is said to be optimal if it minimizes the error associated with the L^2 norm:

$$\begin{aligned} \hat{\xi} \in \operatorname{argmin} \left\{ \|\xi - \hat{\xi}\|_2 := \left(\sum_{k=0}^{K-1} \|\xi_k - \hat{\xi}_k\|_2^2 \right)^{1/2} \right. \\ \left. : \hat{\xi}_k \in \Gamma_k, \sum_{k=0}^{K-1} \#(\Gamma_k) = \sum_{k=0}^{K-1} N_k \leq N \right\} \end{aligned} \quad (3)$$

where N is the number of quantized points in the tree. Then, we define a *Voronoi tessellation* associated with the quantized points $\xi_k^i \in \Gamma_k$ by

$$C_k^i := \left\{ \xi_k : |\xi_k - \xi_k^i| \leq \min_{j \neq i} |\xi_k - \xi_k^j| \right\}.$$

The final step consists in building the recombining tree by computing probability transitions:

$$p_k^{ij} := \mathbf{P} \left[\hat{\xi}_{k+1} = \xi_{k+1}^j \mid \hat{\xi}_k = \xi_k^i \right] = \mathbf{P} \left[\xi_{k+1} \in C_{k+1}^j \mid \xi_k \in C_k^i \right]$$

based on the law of (ξ_k) . Bally and Pagès also proposed in [3] the competitive learning vector quantization (CLVQ) algorithm aimed at building a vectorial quantization tree based on the stochastic gradient method.

2.2 Stochastic dual dynamic programming algorithm

After having discretized the random process (ξ_k) with the help of vectorial quantization, we now focus on the resolution of problem (2). This problem can be naturally composited in time by the dynamic programming principle. We denote by $Q(k, x_k, \xi_k)$ the Bellman value:

$$\begin{aligned} Q(k, x_k, \xi_k) = \inf_{(u_s) \in U} \mathbb{E} \left[\sum_{s=k}^{K-1} c_s(\xi_s) \cdot u_s \right] \\ \text{subject to } u_s \in \mathcal{U}_s \text{ almost surely,} \\ x_{s+1} = x_s + A_s u_s, x_0 = 0, \\ x_T \in \mathcal{X}_T \text{ almost surely.} \end{aligned} \quad (4)$$

Then, the dynamic programming principle can be written as

$$Q(k, x_k, \xi_k) = \inf_{u_k \in \mathcal{U}_k} c_k(\xi_k) \cdot u_k + Q(k+1, x_{k+1}, \xi_k) \quad (5)$$

where

$$Q(k+1, x_{k+1}, \xi_k) = \mathbb{E}[Q(k+1, x_{k+1}, \xi_{k+1}) \mid \mathcal{F}_k]$$

and final time condition:

$$Q(T, x_T, \xi_{T-1}) = \chi_{\mathcal{X}_T}(x_T) = \begin{cases} 0 & \text{if } x_T \in \mathcal{X}_T \\ \infty & \text{otherwise} \end{cases}$$

In that perspective, we have opted for the stochastic dual dynamic programming (SDDP) type algorithm which is well suited to our problem. The main idea of SDDP algorithm is to approximate the Bellman value function $Q(k, x_k, \xi_{k-1})$ by the supremum of a family, depending on t_k and ξ_k , of affine function of x_k , denoted by $\vartheta(k+1, x_{k+1}, \hat{\xi}_k, \mathcal{M}_{k+1}^{\hat{\xi}_k})$. This method is based on successive iterations of forward and backward passes.

The forward pass generates a set of random trajectories following the dynamic of (ξ_k) and runs the optimization on the fly over these trajectories using the following approximation of dynamic programming:

$$\begin{aligned} \inf_{u_k \in \mathcal{U}_k} c_k(\xi_k) \cdot u_k + \vartheta(k+1, x_{k+1}, \hat{\xi}_k, \mathcal{M}_{k+1}^{\hat{\xi}_k}) \\ \text{subject to } x_{k+1} = x_k + A_k u_k, \\ \vartheta(k+1, x_{k+1}, \xi_k^i, \mathcal{M}_{k+1}^{\hat{\xi}_k}) \geq x^* \cdot x_{k+1} - e; \quad \forall (x^*, e) \in \mathcal{M}_{k+1}^{\hat{\xi}_k}, \end{aligned} \quad (6)$$

where $\mathcal{M}_{k+1}^{\hat{\xi}_k}$ is the set of optimality cuts cumulated at vertex $\hat{\xi}_k \in \Gamma_k$. The statistical average performed over the value associated to each forward pass provides an upper bound of the optimal value of (2).

The backward pass updates the optimality cuts $\mathcal{M}_{k+1}^{\hat{\xi}_k}$ on each vertex in the tree using:

$$\begin{aligned} \inf_{u_k \in \mathcal{U}_k} \quad & c_k(\hat{\xi}_k) \cdot u_k + \vartheta(k+1, x_{k+1}, \hat{\xi}_k, \mathcal{M}_{k+1}^{\hat{\xi}_k}) \\ \text{subject to} \quad & x_{k+1} = x_k + A_k u_k, \\ & \vartheta(k+1, x_{k+1}, \xi_k^i, \mathcal{M}_{k+1}^{\hat{\xi}_k}) \geq x^* \cdot x_{k+1} - e; \quad \forall (x^*, e) \in \mathcal{M}_{k+1}^{\hat{\xi}_k}. \end{aligned} \quad (7)$$

Therefore the backward pass generates a lower bound of the optimal value at first stage.

Note that we are using the original continuous random process (ξ_k) in the forward pass, whereas the discretized random process $(\hat{\xi}_k)$ is used in the backward pass.

3 Price model

In this paper, we assume that the random process, taking part in the modeling of futures commodity price evolution, follows the celebrated Black model [7]. Although simplistic, this model is equivalent to the one factor Gaussian HJM model (see [15]) which is well adapted for forward curve modeling. This model is discussed in the practice-based literature, Eydeland and Wolyniec[17, Chapter 5], and Lai et al [18]. Since we are using the vectorial quantization method to discretize the underlying random process, we are not limited to that model. Other financial price models can also be applied.

The price model takes as parameters the market volatilities σ , colleration ρ as well as the forward curve F_0 . The Black model states that future prices of commodities are martingales whose expectations correspond to the original forward curve level F_0 due to lack of arbitrage opportunities. The forward contract price F_s^t at time s , with maturity (also called tenor) t , follows under the equivalent martingale measure (EMM) \mathbf{P} the following dynamic:

$$\ln \frac{(F_{k+1}^t)^i}{(F_k^t)^i} = \sigma^i W_k^i - \frac{1}{2}(\sigma^i)^2 \quad i = 1, \dots, d; \quad (8)$$

where W_k^i is a standard normal distribution $\mathcal{N}(0, 1)$, with correlation

$$\text{corr}(W_s^i, W_s^j) = \rho_{ij}.$$

Then, the spot prices at time t_k are simply obtained by taking the t_k -expiring forward contracts at time t_k : $\xi_k = F_k^k$,

$$\xi_k^i := (F_k^k)^i = (F_0^k)^i \exp \left(\sum_{s=0}^{k-1} \sigma^i W_s^i - \frac{1}{2}(\sigma^i)^2 k \right) \quad i = 1, \dots, d; \quad (9)$$

where (F_0^k) is the original forward curve observed. Setting

$$S_k^i := \exp \left(\sum_{s=0}^{k-1} \sigma^i W_s^i - (\sigma^i)^2 k/2 \right),$$

we then have $\xi_k^i = (F_0^k)^i S_k^i$.

4 Sensitivity analysis

The principal contribution of this paper lies in the application of the Danskin's theorem for computing certain sensitivities of problem (2). In this section, we begin by stating this theorem and apply it to our sensitivity analysis. Then, we provide a result on the convergence properties of the sensitivity estimate of our approximation algorithm in section 2 to the actual sensitivities of the original problem (2).

In the subsequent analysis, given two Banach spaces X and Y , we say that $\phi : X \rightarrow Y$ is directionally differentiable at a point $x \in X$ in the direction $h \in X$ if the following limit exists:

$$\phi'(x; h) := \lim_{t \downarrow 0} \frac{\phi(x + th) - \phi(x)}{t}. \quad (10)$$

4.1 Danskin's theorem and its application

Let us denote by $v(F, \sigma)$ (resp. $U^*(F, \sigma)$) the optimal value function (resp. the optimal solution set) of the original problem (2). In this section, we show that $v(F, \sigma)$ is directionally differentiable at every $F \in \mathbb{R}^{d \times K}$ and $\sigma \in \mathbb{R}_+^d$, and Fréchet differentiable almost everywhere.

We first study the objective function of problem (2) which can be written as:

$$f(F, \sigma, u) = \mathbb{E} \left[\sum_{k=0}^{K-1} c_k(\xi_k(F_0^k, \sigma)) \cdot u_k \right]. \quad (11)$$

$f(F, \sigma, u)$ is a continuous function on $\mathbb{R}^{d \times K} \times \mathbb{R}_+^d \times U$ according to the continuity of c_k and the continuity of (ξ_k) in the price model (9). Then, we show the Fréchet differentiability by the following lemma.

Lemma 4.1. *The objection function $f(F, \sigma, u)$ is Fréchet differentiable with respect to $F \in \mathbb{R}^{d \times K}$ and $\sigma \in \mathbb{R}^d$, and its derivative is*

$$D_{F, \sigma} f(F, \sigma, u) = \mathbb{E} \left[\sum_{k=0}^{K-1} D_{F, \sigma} c_k \cdot u_k \right]. \quad (12)$$

Proof. In view of the Lipschitz continuity of c_k , Rademacher's theorem [21, Theorem 9.60] states that it is almost everywhere Fréchet differentiable, and its derivative is bounded. Note that the derivative of $f(F, \sigma, u)$ will appear the derivative of c_k as an integrand against measure of ξ_k . Furthermore, at time t_k , the random variable ξ_k follows a log-normal distribution according to our price model, whose density function is absolutely continuous with respect to the Lebesgue measure. Thus, using the chain rule of derivative on the composite mapping as well as the dominated convergence theorem ends the proof. \square

By this lemma, we immediately deduce another property of $f(F, \sigma, u)$: its derivative $D_{F, \sigma} f$ is also a continuous function on $\mathbb{R}^{d \times K} \times \mathbb{R}_+^d \times U$ and is weakly continuous with respect to u .

Next, we recall Danskin's Theorem:

Theorem 4.2. [9, Theorem 4.13] *Let V be a nonempty, compact topological space and $\psi : \mathbb{R}^n \times V \rightarrow \mathbb{R}$. Let $\phi(x)$ be the optimal value function and $V^*(x)$ be the optimal solution set:*

$$\phi(x) := \min_{v \in V} \psi(x, v) \quad V^*(x) := \operatorname{argmin}_{v \in V} \psi(x, v)$$

If $\psi(\cdot, v)$ is differentiable for every $v \in V$ and $D_x \psi(x, v)$ is continuous on $\mathbb{R}^n \times V$, then $\phi(x)$ is locally Lipschitz continuous, directionally differentiable, and

$$\phi'(x; h) = \inf_{v \in V^*(x)} D_x \psi(x, v) h. \quad (13)$$

In particular, if for some $x \in \mathbb{R}^n$, the set $V^(x) = \{v^*\}$ is a singleton, then the min-function is differentiable at x and*

$$\phi'(x) = D_x \psi(x, v^*). \quad (14)$$

In order to show the directional differentiability of $v(F, \sigma)$, we only need to verify the compactness assumption of feasible solution space in the Danskin's theorem 4.2. In our framework, the feasible solution set U taking value in (\mathfrak{U}_k) is a bounded closed subset of $L^2 \cap L^\infty$. From the Banach Alaoglu Bourbaki theorem [10, theorem 3.16], it is a weak compact subset in L^2 .

Corollary 4.3. *The optimal function $v(F, \sigma)$ is Fréchet differentiable at almost every $F \in \mathbb{R}^{d \times K}$ and $\sigma \in \mathbb{R}_+^d$, and at the point where $v(F, \sigma)$ is differentiable, we have*

$$Dv(F, \sigma) = D_{F, \sigma} f(F, \sigma, u), \quad \text{for all } u \in U^*. \quad (15)$$

Proof. The almost every Fréchet differentiability is a direct consequence of Rademacher's theorem [21, Theorem 9.60] and the fact that $v(F, \sigma)$ is locally Lipschitz by Danskin's theorem. And if u_1 and u_2 are two optimal solutions such that $D_{F, \sigma} f(F, \sigma, u_1) \neq D_{F, \sigma} f(F, \sigma, u_2)$, then (13) states that f is not Fréchet differentiable at (F, σ) . Then, we have a contradiction. \square

The direct application of Danskin's theorem leads to one method to compute the value of sensitivities. We will focus on the sensitivities $\delta_F(dF)$ and $\delta_\sigma(d\sigma)$ and give an expression the directional derivative of the optimal value $v(F, \sigma)$ with respect to the forward curve F and volatility vector σ in the directions dF and $d\sigma$:

$$\delta_F(dF) := v'(F, \sigma; dF, 0) = \inf_{(u_k^*) \in U^*} \mathbb{E} \left[\sum_{k=0}^{K-1} D_F c_k dF_0^k \cdot u_k^* \right] \quad (16)$$

$$\delta_\sigma(d\sigma) := v'(F, \sigma; 0, d\sigma) = \inf_{(u_k^*) \in U^*} \mathbb{E} \left[\sum_{k=0}^{K-1} D_\sigma c_k d\sigma \cdot u_k^* \right] \quad (17)$$

Danskin's theorem requires all the optimal solutions, which are generally difficult to obtain in stochastic programmings, which are not strongly convex in general. In most cases, we can only get *one* optimal policy (u_k^*) , and the whole optimal policy set is hard to compute. Therefore, we can only obtain an upper

bound of the sensitivity value:

$$\delta_F(dF) \leq \bar{\delta}_F(dF, u^*) := \mathbb{E} \left[\sum_{k=0}^{K-1} D_F c_k dF_0^k \cdot u_k^* \right] \quad (18)$$

$$\delta_\sigma(d\sigma) \leq \bar{\delta}_\sigma(d\sigma, u^*) := \mathbb{E} \left[\sum_{k=0}^{K-1} D_\sigma c_k d\sigma \cdot u_k^* \right] \quad (19)$$

However, by the corollary 4.3, $v(F, \sigma)$ is Fréchet differentiable almost everywhere. Thus, the inequalities in (18) and (19) become equality at almost every F and σ . So, knowing one optimal policy $u^* \in U^*$, allows us to get the derivative value almost everywhere.

Furthermore, we can study the derivatives associated to the standard base of F and σ . At almost every F and σ , we have:

$$(\delta_F)_k^i := \delta_F(e_k^i) = \mathbb{E} \left[\frac{\partial c_k(\xi_k)}{\partial (F_0^k)^i} \cdot u_k^* \right] \quad i = 1, \dots, d \quad (20)$$

$$(\delta_\sigma)^i := \delta_\sigma(e^i) = \mathbb{E} \left[\sum_{k=0}^{K-1} \frac{\partial c(\xi_k)}{\partial \sigma^i} \cdot u_k^* \right] \quad i = 1, \dots, d \quad (21)$$

Then, at almost every F and σ :

$$\delta_F(dF) = \sum_{k=0}^{K-1} \sum_{i=1}^d (dF)_k^i (\delta_F)_k^i \quad \delta_\sigma(d\sigma) = \sum_{i=1}^d (d\sigma)^i (\delta_\sigma)^i. \quad (22)$$

Finally, a simple computation on the price model (9) shows that:

$$\frac{\partial \xi_k^i}{\partial (F_0^k)^i} = S_t^i, \quad (23)$$

$$\frac{\partial \xi_k^i}{\partial \sigma^i} = \xi_k^i \left(\sum_{s=0}^{k-1} W_s^i - \sigma^i k \right). \quad (24)$$

4.2 Convergence of sensitivity estimate

In this section, we start by presenting two discrete approximations of problem (2) helping us describe the forward and backward approximations made in our algorithm in section 2. Then we show some properties of the solutions of the discrete approximation problems and present their convergence to the solution of problem (2). Eventually, we deduce the convergence of the sensitivity estimates of the discrete approximation problems to the sensitivities of continuous problem.

We consider a sequence of optimal vector quantization tree $\hat{\xi}^m \in \Gamma^m$, such that the total number of quantized points is N^m and $\lim_{m \rightarrow \infty} N^m = \infty$. Denote by $(\hat{\mathcal{F}}_k^m)$ the canonic filtration associated to the discretized random process $(\hat{\xi}_k^m)$: $\hat{\mathcal{F}}_k^m := \sigma(\hat{\xi}_s^m, 0 \leq s \leq k)$.

Let us consider two discrete approximations of problem (2). The first one enforces a constant strategy in each Voronoi cell. We call it as the *quantization approximation* of problem (2):

$$\begin{aligned}
v_Q^m(F, \sigma) &= \inf_{(u_k^m) \in U} \mathbb{E} \left[\sum_{k=0}^{K-1} c_k(\xi_k) \cdot u_k^m \right] \\
\text{subject to } u_k^m &\in \mathfrak{U}_k, \quad \text{almost surely} \\
u_k^m(\xi_k, x_k^m) &= \sum_{\xi_k^{i,m} \in \Gamma_k^m} u_k^{i,m}(x_k^m) \mathbb{1}_{C_k^{i,m}}(\xi_k) \\
x_{k+1}^m &= x_k^m + A_k u_k^m, \quad x_0^m = 0 \\
x_T^m &\in \mathfrak{X}_T^m \quad \text{almost surely}
\end{aligned} \tag{25}$$

By forward induction, we can deduce that u_k is $\hat{\mathcal{F}}_k^m$ adapted and that x_k is $\hat{\mathcal{F}}_{k-1}^m$ measurable. This quantization approximation describes the approximation made in the forward pass of our algorithm.

Since u_k is $\hat{\mathcal{F}}_k^m$ adapted, we may write the cost function as function of (25) as

$$\begin{aligned}
\mathbb{E} \left[\sum_{k=0}^{K-1} c_k(\xi_k) \cdot u_k^m \right] &= \mathbb{E} \left[\sum_{k=0}^{K-1} \mathbb{E} \left[c_k(\xi_k) \cdot u_k^m | \hat{\mathcal{F}}_k^m \right] \right] \\
&= \mathbb{E} \left[\sum_{k=0}^{K-1} \mathbb{E} \left[c_k(\xi_k) | \hat{\mathcal{F}}_k^m \right] \cdot u_k^m \right].
\end{aligned} \tag{26}$$

In this formulation the decision $u_k^{i,m}$ to be taken when the state belongs to the cell C_k^i at step t_k actually depends on the previous cells along the realization of ξ . We may replace in each Voronoi cell $\mathbb{E} \left[c_k(\xi_k) | \hat{\mathcal{F}}_k^m \right]$ by $c_k(\mathbb{E}[\xi_k | \xi_k \in C_k^i])$ in the previous formulation. It can be shown, see [22, Chapter 3], that it is equivalent to one large size formulation on scenarios. We call it as the *scenario approximation* of problem (2):

$$\begin{aligned}
v_S^m(F, \sigma) &= \inf_{(u_k^m) \in U} \sum_{(\xi_k^{i,m})_k} p^{i,m} \left(\sum_{k=0}^{K-1} c_k(\xi_k^{i,m}) \cdot u_k^{i,m} \right) \\
\text{subject to } u_k^{i,m} &\in \mathfrak{U}_k \\
\text{(non-anticipativity)} \quad u_k^{i,m} &= u_k^{j,m}, \quad \text{if } \xi_{[k]}^{i,m} = \xi_{[k]}^{j,m} \\
x_{k+1}^{i,m} &= x_k^{i,m} + A_k u_k^{i,m}, \quad x_0^{i,m} = 0 \\
x_K^{i,m} &\in \mathfrak{X}_K^{i,m}
\end{aligned} \tag{27}$$

where $\xi_{[k]}^{i,m} := (\xi_0^{i_0,m}, \dots, \xi_k^{i_k,m}) \in (\Gamma_0^m, \dots, \Gamma_k^m)$ is a sub-trajectory in the quantization tree until time t_k , $p^{i,m} := \prod_{k=0}^{K-2} p_k^{i_k i_{k+1}, m}$ is the probability associated to one trajectory $\xi_{[K-1]}^{i,m}$. This scenario approximation describes the approximation made in the backward pass in our algorithm.

Denote by $v_Q^m(F, \sigma)$ and $u_Q^{*,m}(F, \sigma)$ (resp. $v_S^m(F, \sigma)$ and $u_S^{*,m}(F, \sigma)$) the optimal value function and one optimal solution of the quantization approximation problem (25) (resp. of the scenario approximation problem (27)).

In the same way as (20) and (21), we obtain the sensitivity estimates corresponding to the above discretized approximation problems by Danskin's theorem:

$$(\delta_{Q,F}^m)_k^i := (v_Q^m)'(F, \sigma; e_k^i, 0) \leq (\bar{\delta}_{Q,F}^m)_k^i := \mathbb{E} \left[\frac{\partial c_k(\xi_k)}{\partial (F_0^k)^i} \cdot (u_Q^{*,m})_k \right] \quad i = 1, \dots, d \quad (28)$$

$$(\delta_{Q,\sigma}^m)_k^i := (v_Q^m)'(F, \sigma; 0, e^i) \leq (\bar{\delta}_{Q,\sigma}^m)_k^i := \mathbb{E} \left[\sum_{k=0}^{K-1} \frac{\partial c(\xi_k)}{\partial \sigma^i} \cdot (u_Q^{*,m})_k \right] \quad i = 1, \dots, d \quad (29)$$

$$(\delta_{S,F}^m)_k^i := (v_S^m)'(F, \sigma; e_k^i, 0) \leq (\bar{\delta}_{S,F}^m)_k^i := \mathbb{E} \left[\frac{\partial c_k(\hat{\xi}_k)}{\partial (F_0^k)^i} \cdot (u_S^{*,m})_k \right] \quad i = 1, \dots, d \quad (30)$$

$$(\delta_{S,\sigma}^m)_k^i := (v_S^m)'(F, \sigma; 0, e^i) \leq (\bar{\delta}_{S,\sigma}^m)_k^i := \mathbb{E} \left[\sum_{k=0}^{K-1} \frac{\partial c(\hat{\xi}_k)}{\partial \sigma^i} \cdot (u_S^{*,m})_k \right] \quad i = 1, \dots, d \quad (31)$$

Lemma 4.4. *If the distribution of ξ_k is absolutely continuous with respect to the Lebesgue measure, then the optimal value functions of the two discrete approximation problems $v_S^m(F, \sigma)$ and $v_Q^m(F, \sigma)$ are Fréchet differentiable at almost every $F \in \mathbb{R}^{d \times K}$ and $\sigma \in \mathbb{R}_+^d$.*

Proof. The scenario approximation problem (27) can be viewed as a large linear program. According to the linear programming theory, v_S^m is piecewise linear, concave with respect to the coefficient of the objective function defined in whole space, so v_S^m is locally Lipschitz and then almost everywhere Fréchet differentiable by Rademacher's theorem [21, Theorem 9.60]. Under the assumption on the distribution of ξ_k , and the fact that $(F, \sigma) \rightarrow \hat{\xi}^m(F, \sigma)$ is injective and Fréchet differentiable, chain rule of Fréchet derivative of composite mapping holds almost everywhere and we deduce the lemma for scenario approximation problem (27).

The quantization approximation problem (25) only differs from (27) by replacing the coefficient in objective function $\mathbb{E}[c(\xi_t) | \xi_t \in C_t^i]$ by $c(\mathbb{E}[\xi_t | \xi_t \in C_t^i])$. Thus, concavity of value function v_Q^m with respect to the coefficient in objective function still holds. With the same arguments, we get the almost everywhere Fréchet differentiability for the optimal value function of the quantization approximation problem (25). \square

Therefore, we are able to obtain exact sensitivity values for both discretize approximation problems following Danskin's theorem almost everywhere, i.e. the inequalities (28)- (31) become equalities almost everywhere.

Lemma 4.5. *A subsequence of the optimal strategy $u_Q^{*,m}$ of the quantization approximation problem (25) converges weakly in $L^2(\Omega, (\mathcal{F}_t), \mathbf{P}; \mathbb{R}^{n \times K})$ to one optimal strategy u^* of continuous problem (2).*

Proof. We denote by v_Q^* the limit inferior of the associated value function $v_Q^* := \liminf_{m \rightarrow \infty} v_Q^{*,m}$, and the subsequence associated to this limit inferior

$u_Q^{*,m_{n1}}$. The feasible set of (25) is a subset of U taking value in \mathfrak{U} which is non empty, convex and bounded. The optimal solution $u_Q^{*,m_{n1}}$ is therefore a bounded sequence in a Hilbert space U . Therefore, we deduce the existence of a weakly convergent subsequence $u_Q^{*,m_{n2}}$ that converges in the weak topology to u_Q^* following [10, Corollary 3.30]. Since the feasible set is closed and convex, it is weakly closed and hence u_Q^* is feasible. Since $u_Q^{*,m}$ is always feasible for the original problem (2), we get $v \leq v_Q^{*,m}$ for every m . Let m go to ∞ and we obtain the first inequality $v \leq v_Q^*$.

In order to obtain the reverse inequality, let us build a sequence which is feasible for the quantization approximation problem (25) and which converges to an optimal solution of problem (2). The feasible solution set U is included in L^2 , where C^∞ is dense. Then for one optimal solution $(u_k^*) \in U^*$ of the continuous problem (2), there exists one sequence $(u_k^{n,c}) \in C^\infty$ converging to (u_k^*) such that $\sum_{k=0}^{K-1} \|u_k^{n,c} - u_k^*\| \leq 1/n$. We build

$$u_k^{n,m}(\xi_{[k]}) = \mathbb{E} \left[u_k^{n,c}(\xi_{[k]}) \mid \hat{\mathcal{F}}_k^m \right],$$

which is one feasible solution of the quantization approximation problem (25). The convergence of $u_k^{n,m}$ to $u_k^{n,c}$ in L^2 is due to the convergence $\|\xi - \hat{\xi}^m\| \rightarrow 0$ and the dominated convergence theorem. We note $v_Q^{n,m}$ the value associated with strategy $u_k^{n,m}$. Therefore, $v_Q^m \leq v_Q^{n,m}$. Let us consider the sequence $v_Q^{n,n}$, and let $n \rightarrow \infty$, we obtain the second inequality $v_Q^* \leq v$. Thus, we get $v_Q^* = v$ and hence prove the lemma. \square

One immediate corollary is

Corollary 4.6. *The sensitivity estimates of the quantization approximation problem $\bar{\delta}_{Q,F}^m$ and $\bar{\delta}_{Q,\sigma}^m$ defined in (28) and (29) converge to sensitivities value of continuous problem δ_F and δ_σ (defined in (16) and (17)) at almost every $F \in \mathbb{R}^{d \times K}$ and $\sigma \in \mathbb{R}_+^d$.*

Lemma 4.7. *A subsequence of the optimal strategy u_S^* of scenario approximation problem (27) converges weakly in $L^2(\Omega, (\mathcal{F}_t), \mathbf{P}; \mathbb{R}^{n \times K})$ to an optimal strategy u^* of continuous problem (2).*

Proof. In view of lemma 4.5, we just need to compare the optimal value function v_S^m and v_Q^m . Both discretized approximation problems have the same feasible solution space $L^2(\Omega, (\hat{\mathcal{F}}_k^m), \mathbf{P}) \rightarrow (\mathfrak{U}_k)$. By a straightforward computation, we can obtain that

$$|v_Q^m - v_S^m| \leq \sum_{k=0}^{K-1} [c_k]_{lip} \|\mathfrak{U}_k\|_\infty \|\xi_k - \hat{\xi}_k\|_2 \leq C_{QS} \sum_{k=0}^{K-1} \|\xi_k - \hat{\xi}_k\|_2 \leq 2C_{QS} \|\xi - \hat{\xi}\|_2$$

where $\|\mathfrak{U}\|_\infty = \sup \{|u| \in \mathfrak{U}\}$, and C_{QS} is a constant which only depends on $[c_k]_{lip}$ and $\|\mathfrak{U}\|_\infty$.

Following Zador's theorem (see below theorem 4.9), we get that $\lim_{m \rightarrow \infty} v_Q^m = \lim_{m \rightarrow \infty} v_S^m$. Then, the lemma can be proved by following the same argument as in the proof of 4.5. \square

Then, we have a similar corollary as corollary 4.6 for the scenario approximation problem:

Corollary 4.8. *The sensitivity estimates of the scenario approximation problem $\bar{\delta}_{S,F}^m$ and $\bar{\delta}_{S,\sigma}^m$ defined in (30) (31) converge to the sensitivities value of continuous problem δ_F and δ_σ defined in (16) and (17), at almost every $F \in \mathbb{R}^{d \times K}$ and $\sigma \in \mathbb{R}_+^d$.*

Proof. This corollary is obtained following the fact that $(\hat{\xi}_k^m)$ converges to (ξ_k) and that $u_S^{*,m}$ converges weakly in U to u^* and the [10, Proposition 3.13]. \square

Finally, we recall the Zador's theorem:

Theorem 4.9. [19, Theorem 6.2] *If $\mathbb{E}[|\xi|^{p+\eta}] < \infty$ for some $\eta > 0$, then,*

$$\lim_N \left(N^{p/d} \min_{|\Gamma| \leq N} \|\xi - \hat{\xi}\|_p^p \right) = J_{p,d} \left(\int |g|^{d/(d+p)}(u) du \right)^{1+p/d}$$

where $\mathbf{P}(du) = g(u)\lambda_d(du) + \nu, \nu \perp \lambda_d$ (λ_d Lebesgue measure on \mathbb{R}^d). The constant $J_{p,d}$ corresponds to the case of the uniform distribution on $[0, 1]^d$.

This theorem implies that $\min_{|\Gamma| \leq N} \|\xi - \hat{\xi}\|_p = O(N^{1/d})$.

5 Algorithm and numerical tests

5.1 Algorithm

After getting one optimal (ϵ -optimal) strategy $(u_k^{*,tr-sddp})$ thanks to our discretization and SDDP algorithm described in section 2, we use the Monte Carlo method to compute the expectations in (28)-(29).

We simulate a sample of $\xi^s, s = 1, \dots, S$ the trajectories and compute the value of sensitivities for each trajectory:

$$^s(\delta_F)_k^i = \frac{\partial c_k(\xi_k^s)}{\partial (F_0^k)^i} \cdot {}^s u_k^{*,tr-sddp} \quad i = 1, \dots, d \quad (32)$$

$$^s(\delta_\sigma)^i = \sum_{k=0}^{K-1} \frac{\partial c_k(\xi_k^s)}{\partial \sigma^i} \cdot {}^s u_k^{*,tr-sddp} \quad i = 1, \dots, d \quad (33)$$

Therefore we obtain the mean and variance of this sample:

$$\mu_{(\delta_F)_k^i}^{tr-sddp} = \frac{1}{S} \sum_{s=1}^S {}^s(\delta_F)_k^i \quad \sigma_{(\delta_F)_k^i}^{tr-sddp} = \left(\frac{1}{S-1} \sum_{s=1}^S ({}^s(\delta_F)_k^i - \mu_{(\delta_F)_k^i}^{tr-sddp})^2 \right)^{1/2} \quad (34)$$

$$\mu_{(\delta_\sigma)^i}^{tr-sddp} = \frac{1}{S} \sum_{s=1}^S {}^s(\delta_\sigma)^i \quad \sigma_{(\delta_\sigma)^i}^{tr-sddp} = \left(\frac{1}{S-1} \sum_{s=1}^S ({}^s(\delta_\sigma)^i - \mu_{(\delta_\sigma)^i}^{tr-sddp})^2 \right)^{1/2} \quad (35)$$

5.2 Comparison methods

We shall now address the question of the accuracy of the results obtained from the application of Danskin's theorem. In that perspective, we benchmark them against the sensitivities obtained by the mean of other algorithms.

The first comparison lies in the discretization of state variable (x_k) . Indeed, many popular algorithms impose the discretization of (x_k) as a prerequisite whilst our SDDP based algorithm removes this constraint. Although more efficient than SDDP in low dimensions, the (x_k) discrete algorithms become numerically intractable when the dimension of (x_k) increase.

The second source of comparison concerns the discretization of the random process (ξ_k) . In our framework, we have chosen the optimal quantization tree method as a way to discretize the randomness. As long as (ξ_k) remains in low dimension, we may also implement a PDE based method.

Finally, the last and probably most important criteria of comparison resides in the way the sensitivities are computed. We challenge the numerical results obtained by the application of the Danskin's theorem against the finite difference method which is as simple as it is popular.

5.2.1 Danskin + quantization tree + state space discretization

The first comparison method is the algorithm proposed in [5] that combines quantization tree method and the dynamic programming principle on discretized state space (\hat{x}_k) . Let us stress again that this method proves to be very efficient as long as the state variable (x_k) remains in low dimension. Let $v^{tr-disc}$ stands for the Bellman value and $u^{*,tr-disc}$ for the optimal strategy:

$$v^{tr-disc}(k, \hat{x}_k, \hat{\xi}_k) = \min_{u_k \in \mathcal{U}_k} \left(c_k(\hat{\xi}_k)u_k + \mathbb{E}[v^{tr-disc}(k+1, \hat{x}_k + u_k, \hat{\xi}_{k+1}) | \hat{\xi}_k] \right) \quad (36)$$

$$u^{*,tr-disc}(k, \hat{x}_k, \hat{\xi}_k) = \operatorname{argmin}_{u_k \in \mathcal{U}_k} \left(c_k(\hat{\xi}_k)u_k + \mathbb{E}[v^{tr-disc}(k+1, \hat{x}_k + u_k, \hat{\xi}_{k+1}) | \hat{\xi}_k] \right) \quad (37)$$

where

$$\mathbb{E}[v^{tr-disc}(k+1, x, \hat{\xi}_{k+1}) | \hat{\xi}_k = \xi_k^i] = \sum_{\xi_{k+1}^j \in \Gamma_{k+1}} p_k^{ij} v^{tr-disc}(k+1, x, \xi_{k+1}^j)$$

We can reuse the same formula as (32)-(35) with the optimal policy $u^{*,tr-disc}$ to compute the sensitivities values that will be denoted by $(\delta_F)^{tr-disc}$ and $(\delta_\sigma)^{tr-disc}$. Thanks to this method, we are able to compare two methods to compute the optimal solutions obtained on the same quantization tree as the one used in our SDDP based algorithm.

5.2.2 Danskin + PDE + state space discretization

If the random process (ξ_k) is in low dimension, one natural idea is to use a PDE method to compute conditional expectations. Therefore, the second comparison method presented in this paragraph combines a PDE method with the dynamic programming principle on discretized state space (\hat{x}_k) . In order to use a PDE method, we need to build one continuous time process from the discrete time random process. To be consistent with our price model (8), we will use the continuous version of the Black model:

$$dF_t = \sigma F_t dW_t \quad (38)$$

and make interpolation of F_0^t between t_k and t_{k+1} :

$$F_0^t = \frac{t - t_k}{t_{k+1} - t_k} F_0^{k+1} + \frac{t - t_{k+1}}{t_k - t_{k+1}} F_0^k \quad t_k \leq t \leq t_{k+1}. \quad (39)$$

Let $v^{pde-disc}(k, x_k, \xi_k)$ stands for the Bellman value, and $u^{*,pde-disc}(k, x_k, \xi_k)$ for the optimal strategy:

$$v^{pde-disc}(k, \hat{x}_k, \xi_k) = \min_{u_k \in \mathcal{U}_k} (c_k(\xi_k)u_k + \mathbb{E}[v^{pde-disc}(k+1, \hat{x}_k + u_k, \xi_{k+1}) | \xi_k]) \quad (40)$$

$$u^{*,pde-disc}(k, \hat{x}_k, \xi_k) = \operatorname{argmin}_{u \in \mathcal{U}_k} (c_k(\xi_k)u_k + \mathbb{E}[v^{pde-disc}(k+1, \hat{x}_k + u_k, \xi_{k+1}) | \xi_k]) \quad (41)$$

Let $\tilde{v}(t, x, \xi_t) := \mathbb{E}[v^{pde-disc}(k+1, x, \xi_{k+1}) | \xi_t]$, $t_k \leq t \leq t_{k+1}$. Combining the Feynman-Kac formula and Itô's lemma on the continuous extension of our price model (38) (39), we compute the conditional expectations by solving the following PDE backwardly from t_{k+1} to t_k :

$$\begin{cases} \frac{\partial \tilde{v}}{\partial t} + \frac{1}{2} \sum_{i=1}^d (\sigma^i)^2 (\xi^i)^2 \frac{\partial^2 \tilde{v}}{\partial (\xi^i)^2} + \sum_{i \neq j} \rho_{ij} \sigma^i \sigma^j \xi^i \xi^j \frac{\partial^2 \tilde{v}}{\partial \xi^i \partial \xi^j} = 0 & t_k \leq t \leq t_{k+1} \\ \tilde{v}(t_{k+1}, \cdot, \xi_{t_{k+1}}) = v^{pde-disc}(t_{k+1}, \cdot, \xi_{t_{k+1}}) \end{cases} \quad (42)$$

This parabolic partial differential equation is straightforward to solve numerically in dimension 1. For example, we may solve it by the use of finite difference method with an implicit scheme, which is introduced in appendix A.

Once we have obtained the optimal solution $u^{*,pde-disc}$, we may still apply Danskin's theorem by using the same formula (32)-(35). We denote the sensitivity result by $(\delta_F)^{pde-disc}$ and $(\delta_\sigma)^{pde-disc}$.

Note that this method does not depend on the quantization tree. Hence, the results obtained from the method presented in the current paragraph evidence the accuracy of discretizing (ξ_k) by the quantization tree method.

5.2.3 Finite difference (fd.) + PDE + state space discretization

The last method used to compare the accuracy of the sensitivities estimate is probably the most used in practice. Once an efficient numerical method to compute the optimal value $v(F, \sigma)$ is available, the finite difference method may be used to approximate the derivative value:

$$(\delta_F^{fd})_k^i = \frac{v(F + \epsilon e_k^i, \sigma) - v(F - \epsilon e_k^i, \sigma)}{2\epsilon} \quad (43)$$

$$(\delta_\sigma^{fd})^i = \frac{v(F, \sigma + \epsilon e^i) - v(F, \sigma - \epsilon e^i)}{2\epsilon} \quad (44)$$

In the following numerical tests, we use PDE and the state variable discretization to compute the optimal value $v(F, \sigma)$. We denote by $(\delta_F)^{fd-pde}$ and $(\delta_\sigma)^{fd-pde}$ the sensitivities estimates. This method allows us to compare the relevance of the values obtained from Danskin's theorem.

5.3 Swing option

The first test is performed on swing options which are fairly common in the energy markets. A swing option allows its holder to purchase a total amount of commodity during a predetermined period of time. A typical swing option contract has the following parameters: a feasible control set $\mathfrak{U}_k = [0, 1]$ for every time step t_k , a final state set $\mathfrak{X}_K = [L, U]$, and a time horizon $K = 50$ in our example. The running cost function $c(\xi_k)$ is equal to $\xi_k - K_k$, where the strike $K_k = F_0^k$.

Under certain circumstances, swing options featuring a one dimension control possesses the bang-bang property. Bardou et al [5, 4] show that there always exists a bang-bang optimal strategy when $(L, U) \in \mathbb{Z}^2$. Therefore, the discretization on x_t takes value in \mathbb{Z} only.

In the current test, we take the lower bound of exercise right L equal to 20, and upper bound of exercise right U equal to 30. In this numerical test and the next one, we set in price model forward curve F_0^k to $1.0 + 0.2 \sin(2\pi t_k/T)$, and volatility σ to 0.042. Concerning the parameters intervening in the PDE resolution (see section A), we set the lower bound $\underline{\xi}$ to 0.1, the upper bound $\bar{\xi}$ to 5.0, the number of discretization space steps J to 200, and finally the number of time steps L to 5. Finally, we take $\epsilon = 0.01$ in the finite difference estimation. The program is written in C++ with GNU Linear Programming Kit (GLPK) 4.44, and the test was performed on a PC with a 3.07GHz Intel Xeon CPU and 12GByte main memory. The results obtained for optimal values are given in table 1.

	tree+sddp.			tree+discret.	pde.+discret.
	ub.	$\sigma(\text{ub.})$	lb.		
optimal value	-1.4429	0.0125	-1.4375	-1.4361	-1.4244

Table 1: Optimal values for the 3 methods.

The sensitivity values with respect to each contract of the forward curve are displayed in figure 1.

Finally, the sensitivity values with respect to σ are given in table 2.

	tree+sddp.		tree+discret.		pde.+discret.		fd.+ pde.
	mean	std. dev.	mean	std. dev.	mean	std. dev.	
δ_σ	-32.958	0.158	-32.963	0.158	-32.993	0.160	-33.014

Table 2: Sensitivity value with respect to σ obtained by the 4 methods.

We remark that the sensitivities values obtained by all 4 methods are close to each other. This result proves the accuracy of the sensitivities obtained by our original algorithm (Danskin + quantization tree + SDDP). Furthermore, we also observe the closeness of the deltas obtained by finite difference and those obtained by Danskin for the case where the discretization methods are the same

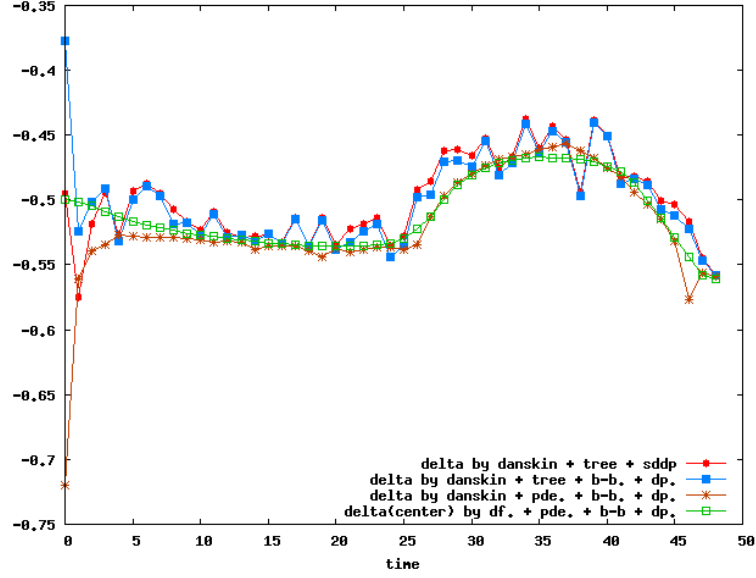


Figure 1: Sensitivity values with respect to F obtained by the 4 methods. Since the standard deviations of the first 3 methods represent less than 1% of the sensitivity values, we will not plot them for sake of clarity.

(i.e. PDE+discretization). However, we observe a significant deviation for δ_0 . One possible reason is that the optimal value function $v(F, \sigma)$ is not Fréchet differentiable at time t_0 .

Compared to other methods, our algorithm does not procure any advantage in CPU time for such a low dimension problem. In the next example, we will compare the performance of these 4 methods when we consider a problem with higher dimension control variable u_t and state variable x_t .

5.4 Small commodity portfolio case study

Let us consider a small network of distribution of natural gas. One company buys gas from production site A and delivers to two consuming sites B and C. The margin obtained is indexed on the spot commodity price. The contract is also subject to flow constraints (see table 3).

site	loc. quan. \mathcal{U}_t	glo. quan. \mathcal{X}_T	running cost $c_k(\xi_k)$
A	[5, 10]	[400, 400]	
B	[0, 7]	[200, 200]	$(A \rightarrow B) \begin{cases} (\xi - 1.1) + 0.1 & \text{if } \xi \leq 1.1 \\ 0.5(\xi - 1.1) + 0.1 & \text{otherwise} \end{cases}$
C	[2, 8]	[200, 200]	$(A \rightarrow C) \begin{cases} (\xi - 0.9) - 0.1 & \text{if } \xi \leq 0.9 \\ 0.5(\xi - 0.9) - 0.1 & \text{otherwise} \end{cases}$

Table 3: Small commodity portfolio setting: constraints and running cost.

Unfortunately, this problem does not possess the convenient bang-bang property of the optimal solution. Therefore, discretizing state variable $x_t \in \mathbb{R}^2$ will make the program both time and memory consuming.

During the following tests, we take the discretization steps on state variables $h_B = h_C = 0.5$. The results for optimal values are given in table 4.

	tree+sddp.			tree+discret.	pde.+discret.
	ub.	$\sigma(\text{ub.})$	lb.		
optimal value	-16.2809	0.8251	-16.0024	-13.861	-13.8361
CPU time (s)		319		207	297

Table 4: Optimal value and CPU time of 3 methods.

Our algorithm costs around 50% more CPU time than the combination of a quantization tree and a state space discretization and only around 10% more than the combination of a PDE method and a state space discretization. But the result of our algorithm is much better than the standard algorithm on optimal value (around 20% in this test). Furthermore, the discretized programs will become intractable when the state space variable is in dimension greater than 3. In fact, real life portfolio problems are mostly in dimension 10 and more (see second example in our previous article [8]). This remark also applies to problems involving several layers of randomness like swing options which often include a second stochastic component driving the strike price. Thus, our algorithm gives a good answer to such problems.

Despite the gap between the values obtained from the SDDP algorithm and the discretized methods, we still present the sensitivity values obtained by the different methods knowing in advance that such a comparison is skewed.

Hence, the sensitivities values with respect to each contract of the forward curve is given by figure 2.

Finally, the sensitivity values with respect to σ are given in table 5.

	tree+sddp.		tree+discret.		pde.+discret.		fd.+ pde.
	mean	std. dev.	mean	std. dev.	mean	std. dev.	
δ_σ	-63.951	8.368	-16.420	9.077	-16.537	8.077	-15.115

Table 5: Sensitivity value with respect to σ obtained by 4 methods.

As expected, we observe a large difference between the one obtained by our algorithm and the other methods, especially when it comes to the sensitivities δ_σ . However, we also remark a common pattern for the sensitivities δ_F in figure 2, evidencing that the results provided by our algorithm and Danskin theorem seem reasonable. Finally, the conclusion of the last paragraph remains valid: for similar discretization method, the Danskin theorem provides similar results as the finite difference method.

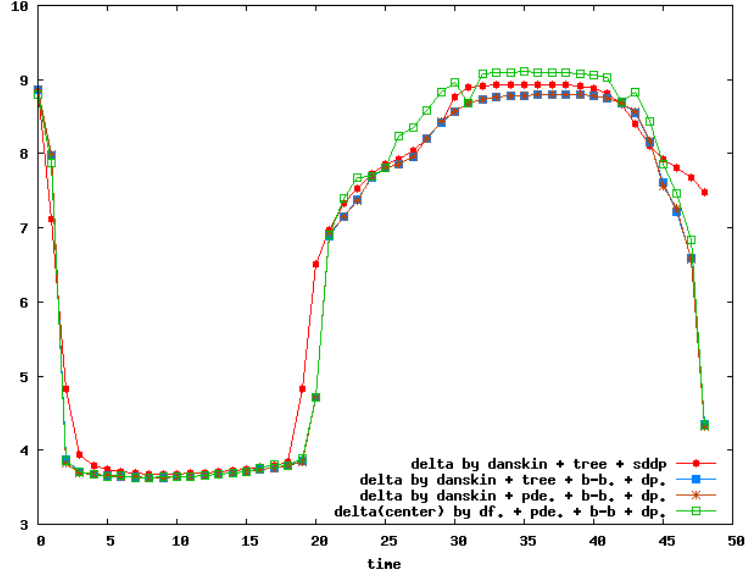


Figure 2: Sensitivity values with respect to F obtained by the 4 methods. Since the standard deviations of the first 3 methods represent less than 0.1% of the sensitivity values, we will not plot them for sake of clarity.

A Implicit scheme of finite difference for 1 dimension PDE

Generally, the parabolic partial differential equation is very difficult to solve numerically in high dimension. We refer to Allaire [1] and Ern and Guermand [16] for more details.

In the case that ξ_t is in dimension one, the equation (42) is written as

$$\begin{cases} \frac{\partial \tilde{v}}{\partial t} + \frac{1}{2}\sigma^2 \xi^2 \frac{\partial^2 \tilde{v}}{\partial \xi^2} = 0, & t_k \leq t \leq t_{k+1}, \xi \in [\underline{\xi}, \bar{\xi}], \\ \tilde{v}(t_{k+1}, \cdot, \xi_{t_{k+1}}) = g(t_{k+1}, \cdot, \xi_{t_{k+1}}). \end{cases} \quad (45)$$

To apply the finite difference scheme, we need to discretize both time t and space ξ . Let the time step be $h_t = (t_{k+1} - t_k)/L$, and the space step be $h_\xi = (\bar{\xi} - \underline{\xi})/J$, where L, J are positive integers. We denote by \tilde{v}_j^l an approximation of $\tilde{v}(t_k + lh_t, \underline{\xi} + jh_\xi)$ where $0 \leq l \leq L$ and $0 \leq j \leq J$. The standard implicit finite difference scheme is

$$\begin{cases} \frac{\tilde{v}_j^{l+1} - \tilde{v}_j^l}{h_t} + \frac{1}{2}\sigma^2 \xi_j^l \frac{\tilde{v}_{j+1}^l - 2\tilde{v}_j^l + \tilde{v}_{j-1}^l}{h_\xi^2} = 0, \\ \tilde{v}_j^L(\cdot) = g(t_{k+1}, \cdot, \underline{\xi} + jh_\xi). \end{cases} \quad (46)$$

Thus, we need the boundary condition at bounds of ξ . By the arguments of problem context, we take the following boundary conditions:

$$\begin{cases} \frac{\partial^2 \tilde{v}}{\partial \xi^2} = 0, \xi = \underline{\xi}, \\ \frac{\partial^2 \tilde{v}}{\partial \xi^2} = 0, \xi = \bar{\xi}. \end{cases} \quad (47)$$

Thus, the finite difference scheme at bounds are:

$$\begin{cases} \tilde{v}_0^{l+1} - \tilde{v}_0^l = 0, \\ \tilde{v}_J^{l+1} - \tilde{v}_J^l = 0. \end{cases} \quad (48)$$

References

- [1] G. Allaire. *Numerical analysis and optimization: An introduction to mathematical modelling and numerical simulation (Numerical mathematics and scientific computation)*. Oxford university Press, USA, 2007. Translated from the French by A. Craig.
- [2] V. Bally and P. Pagès. Error analysis of the quantization algorithm for obstacle problems. *Stochastic Processes & Their Applications*, 106(1):1–40, 2003.
- [3] V. Bally and P. Pagès. A quantization algorithm for solving discrete time multi-dimensional optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.
- [4] O. Bardou, S. Bouthemy, and G. Pagès. Pricing swing options using optimal quantization, 2007.
- [5] O. Bardou, S. Bouthemy, and G. Pagès. Optimal quantization for the pricing of swing options. *Applied Mathematical Finance*, 16(2):183–217, 2009.
- [6] C. Barrera-Esteve, C.F. Bergeret, C. Dossal, E. Gobet, A. Meziou, R. Munos, and D. Reboul-Salze. Numerical methods for the pricing of swing options: a stochastic control approach. *Methodology and computing in applied probability*, 8(4):517–540, 2006.
- [7] F. Black. The pricing of commodity contracts. *Journal of Financial Economy*, 3:167–179, 1976.
- [8] J.F. Bonnans, Z. Cen, and T. Christel. Energy contracts management by stochastic programming techniques. INRIA Research Report 7289, May, 2010.
- [9] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer-Verlag, New York, 2000.
- [10] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer-Verlag, New York, 2010.

- [11] R. Carmona and M. Ludkovski. Valuation of energy storage: an optimal switching approach. *Quantitative Finance*, 10:359–374, 2009.
- [12] R. Carmona and N. Touzi. Optimal multiple stopping and valuation of swing options. *Mathematical Finance*, 18:239–268, 2000.
- [13] R. Cont and D. Fournié. A functional extension of the Ito formula. *Comptes Rendus Mathématiques*, 348(1-2):57 – 61, 2010.
- [14] R. Cont and D.A. Fournié. Functional Itô calculus and stochastic integral representation of martingales.
- [15] R.A. Jarrow D.C Heath and A. Morton. Bond pricing and the term structure of interest rates : a new methodology for contingent claim valuation. *Econometrica*, 60:77–105, 1992.
- [16] A. Ern and J-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.
- [17] A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. John Wiley & Sons, Hoboken, NJ, 2003.
- [18] G.Lai, F. Margot, and N. Secomandi. An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation. *Oper. Res.*, 58:564–582, May 2010.
- [19] S. Graf and H. Luschgy. *Foundations of quantization for probability distributions*. Springer-Verlag, Berlin, 2000.
- [20] J.C. Hull. *Options, Futures and other derivatives*. Prentice-Hall, New Jersey, seventh edition, 2010.
- [21] R.T. Rockafellar and R. J-B. Wets. *Variational analysis*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1997.
- [22] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming : Modelling and Theory*. SIAM, Philadelphia, 2009.



Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399