Optimal quantization for the pricing of swing options

Olivier Bardou* Sandrine Bouthemy† and Gilles Pagès ‡

November 6, 2007

Abstract

In this paper, we investigate a numerical algorithm for the pricing of swing options, relying on the so-called optimal quantization method.

The numerical procedure is described in details and numerous simulations are provided to assert its efficiency. In particular, we carry out a comparison with the Longstaff-Schwartz algorithm.

Key words: Swing options, stochastic control, optimal quantization, energy.

Introduction

In deregulated energy markets, swing options arise as powerful tools for modeling supply contracts [21]. In such an agreement between a buyer and a seller, the buyer has to purchase, say daily, some quantity of a product at a given unitary price. The daily and cumulated purchased quantities are subjects to constraints. Therefore, the buyer has to manage his contract by constantly swinging from one state to other. This is the kind of agreement that usually links an energy producer to a trader. Numerous other examples of energy contracts can be modeled as swing options. From storage [8, 11] to electricity supply [25, 10], this kind of financial device is now widely used. And it has to be noticed that its field of application has recently been extended to the IT domain [14].

Nevertheless, the pricing of swing options remains a real challenge. Closely related to a multiple stopping problem [13, 12], swing options require the use of high level numerical schemes. Moreover, both the high dimensionality of the underlying price processes and the various constraints to be integrated in a model of contracts based on physical assets such as storage or gas fired power plants increase the difficulty of the problem.

Thus, the most recent techniques of mathematical finance have been applied in this context; from trees (see [38, 24, 26]) to Least Squares Monte Carlo based methodology (see [8, 17, 19]), finite elements [39] and duality approximation [30]. But none of these algorithms gives a totally satisfactory solution to the valuation and sensitivity analysis of swing contracts.

The aim of this paper is then to introduce and study a recent pricing method that seems very well suited to the question. Optimal Vector Quantization has already been successfully applied to the valuation of multi-asset American Options [3, 2, 4]. It turns out that this numerical technique

*Corresponding author. Gaz de France, Research and Development Division, 361 Avenue du Président Wilson - B.P. 33, 93211 Saint-Denis La Plaine cedex. E-mail: olivier-aj.bardou@gazdefrance.com, Phone: +33 1 49 22 54 58, Fax: +33 1 49 22 57 10

†Gaz de France, Research and Development Division, 361 Avenue du Président Wilson - B.P. 33, 93211 Saint-Denis La Plaine cedex. E-mail: sandrine.bouthemy@gazdefrance.com

‡Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail:gpa@ccr.jussieu.fr

is also very efficient in taking into account the physical constraints of swing contracts. For sake of simplicity we shall focus on gas supply contracts. After a brief presentation of such agreements and some background on optimal quantization methods (see also [33]), we describe the structure of an optimal quantization tree and its related quantized backward dynamic programming algorithm, including some error bounds. We show that a careful examination of the multi-factor underlying price process dynamics can dramatically improve the efficiency of the estimation procedure of the transition probability weights of the tree. Then, we carry out several numerical experiments on two "toy-dynamics" of our family of models.

The paper is organized as follows: in the first section, we describe in details the technical features of the supply contracts (with firm or penalized constraints) with an emphasis on the features of interest in view of a numerical implementation: the canonical decomposition and normal form, the stochastic control problem and the resulting backward dynamic programming, the existence of bang-bang strategies for some appropriate choices of local and global purchased volume constraints, following the result established in the companion paper [5]. Section 2 is devoted to some background on optimal vector quantization. In Section 3, our quantization based algorithm is briefly analyzed and the *a priori* error bound established in [5] is stated. A convergence result is derived for the "quantized price" toward the true price, including a rate when the quantization tree is made of optimal grids. In Section 4, various simulations are carried out. Our quantization tree method is compared with the well-known least squares regression algorithm "à la Longstaff-Schwartz" (*L-S*) from several points of view: modelling, design and performances. Finally, one key of the performances of the quantization trees in this framework is the Fast Transition Weight Estimation procedure that we describe in full details in the Annex.

1 Introduction to swing options

1.1 Description of the contract

A typical example of swing option is an energy (usually gas or electricity) supply contract with optional clauses on price and volume. The owner of such a contract is allowed to purchase some amount of energy q_{t_k} at times t_k, $k = 0, \ldots, n-1$, until the contract $t_n = T$, usually one year. The purchase price K_k – called strike price – may be constant or indexed to past values of crude oil. Throughout the paper we will consider that the strike prices are constant and equal to K over the term of the contract. The volume of gas q_{t_k} purchased at time t_i is subject to the local constraint

$$q_{\min} \leq q_{t_k} \leq q_{\max}.$$

The cumulative volume purchased prior to time t_k (*i.e.* up to t_{k-1}) is defined by $\bar{q}_{t_k} = \sum_{\ell=0}^{k-1} q_{t_\ell}$. It must satisfy the following global constraint (at maturity):

$$\bar{q}_T = \sum_{k=0}^{n-1} q_{t_k} \in [Q_{\min}, Q_{\max}].$$

Two approaches can be considered:

– The constraints on the global purchased volumes are firm.

– A penalty is applied if the constraints are not satisfied.

The price at time t of the forward contract delivered at time T is denoted by $F_{t,T}$. Then $(F_{0,t_k})_{0\leq k\leq n}$ is a deterministic process (the future prices at time 0) observable on the organised markets where they are traded.

Let $(S_{t_k})_{0\le k\le n}$ be the underlying Markov price process defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Note that S_{t_k} can be the observation at time t_k, $k = 0, \dots, n$ of a continuous-time process. Ideally S_t should be the spot price process of the gas *i.e.* $S_t = F_{t,t}$. However it does not correspond to a tradable instrument which often leads to consider in practice the day-ahead contract $F_{t,t+1}$. For more details, we refer to [23] where these issues are discussed.

We consider the natural filtration $\mathcal{F}^S = (\mathcal{F}^S_{t_k})_{0\le k\le n}$ of the process $(S_{t_k})_{0\le k\le n}$. The decision sequence $(q_{t_k})_{0\le k\le n-1}$ is defined on $(\Omega, \mathcal{A}, \mathbb{P})$ as well and is $\mathcal{F}^S$-adapted, *i.e.* q_{t_k} is $\mathcal{F}_{t_k} = \sigma(S_{t_0}, \dots, S_{t_k})$-measurable, $k = 0, \dots, n$. At time t_k the owner of the contract gets $q_{t_k}(S_{t_k} - K)$.

Remark 1.1. The results of this paper can also be applied to every physical asset or contract where the owner reward for a decision q_{t_k} is a function $\psi(t_k, q_{t_k}, S_{t_k})$. In the case of supply contracts, $\psi(t_k, q_{t_k}, S_{t_k}) = q_{t_k}(S_{t_k} - K)$. As for a storage, q_{t_k} represents the amount of gas the owner of the contract decides to inject or withdraw and the profit at each date is then

$$\psi(t_k, q_{t_k}, S_{t_k}) = \begin{cases} -q_{t_k}(S_{t_k} + c_I) & if \quad q_{t_k} \ge 0 \quad (Injection) \\ -q_{t_k}(S_{t_k} - c_W) & if \quad q_{t_k} \le 0 \quad (Withdrawal) \\ 0 & if \quad q_{t_k} = 0 \quad (Same\ level\ in\ the\ storage) \end{cases}$$

where c_I (resp. c_W) denotes the injection (resp. withdrawal) cost [8].

Remark 1.2. When $q_{\min} = Q_{\min} = 0$ and $q_{\max} = Q_{\max} = 1$, one easily shows that the above swing contract reduces to a Bermuda Call option (discrete time American option).

1.1.1 Case with penalties

We first consider that the penalties are applied at time T if the terminal constraint is violated. For a given consumption strategy $(q_{t_k})_{0\le k<n}$, the price is given at time 0 by

$$P(0, S_0, 0) = \mathbb{E}\left(\sum_{k=0}^{n-1} e^{-rt_k} q_{t_k}(S_{t_k} - K) + e^{-rT} Pen_{_T}(S_T, Q_T)|\mathcal{F}_0\right)$$

where r is the interest rate. The function $(x, \bar{q}) \mapsto Pen_{_T}(x, \bar{q})$ is the penalization: $Pen_{_T}(x, Q) \le 0$ and $|Pen_{_T}(x, Q)|$ represents the sum that the buyer has to pay if the global purchased volume constraints, say $Q_{\min}$ and $Q_{\max}$, are violated. [8] have already investigated this kind of contract.

Then for every non negative $\mathcal{F}_{t_{k-1}}$ measurable random variable $\bar{q}_{t_k}$ (representing the cumulated purchased volume until t_{k-1}), the price of the contract at time $t_k, k = 0, \dots, n-1$, is given by

$$P(t_k, S_{t_k}, \bar{q}_{t_k}) = \operatorname*{ess\,sup}_{(q_{t_\ell})_{k\le\ell<n}} \mathbb{E}\left(\sum_{\ell=k}^{n-1} e^{-r(t_\ell - t_k)} q_{t_\ell}(S_{t_\ell} - K) + e^{-r(T-t_k)} Pen_{_T}(S_T, Q_T)|S_{t_k}\right). \quad (1)$$

The standard penalization function is as follows:

$$Pen_{_T}(x, Q) = -\left(A\,x\,(Q - Q_{\min})_- + B\,x\,(Q - Q_{\max})_+\right) \quad (2)$$

where A and B are large enough – often equal – positive real constants.

1.1.2 Case with firm constraints

If we consider that constraints cannot be violated, then for every non negative $\mathcal{F}_{t_{k-1}}$-measurable random variable $\bar{q}_{t_k}$ defined on $(\Omega, \mathcal{A}, \mathbb{P})$, the price of the contract at time $t_k, k = 0, \dots, n-1$ is given by:

$$P(t_k, S_{t_k}, \bar{q}_{t_k}) = \operatorname*{ess\,sup}_{(q_{t_\ell})_{k\le\ell\le n-1}\in\mathcal{A}^{Q_{\min},Q_{\max}}_{[q_{\min},q_{\max}]}(k,Q)} \mathbb{E}\left(\sum_{\ell=k}^{n-1} e^{-r(t_\ell - t_k)} q_{t_\ell}(S_{t_\ell} - K)|S_{t_k}\right) \quad (3)$$

where

$$\mathcal{A}^{Q_{\min},Q_{\max}}_{[q_{\min},q_{\max}]}(k,Q)=\left\{(q_{t_\ell})_{k\le\ell\le n-1},\, q_{t_\ell}:(\Omega,\mathcal{F}_{t_\ell},\mathbb{P})\mapsto[q_{\min},q_{\max}],\,\sum_{\ell=k}^{n-1}q_{t_\ell}\in\left[(Q_{\min}-Q)_+,Q_{\max}-Q\right]\right\}.$$

At time 0, we have

$$P(0,S_0,0)=\sup_{(q_{t_k})_{0\le k\le n-1}\in\mathcal{A}^{Q_{\min},Q_{\max}}_{[q_{\min},q_{\max}]}(0,0)}\mathbb{E}\left(\sum_{k=0}^{n-1}e^{-rt_k}q_{t_k}(S_{t_k}-K)\right).$$

Note that this formally corresponds to the limit case of the contract with penalized constraints when $A=B=+\infty$. Furthermore, when the penalties $A,B\to+\infty$ in (2), the "penalized" price converges to the "firm" price (see [31]). This has been confirmed by extensive numerical implementations of both methods. In practice when A, $B\approx 10\,000$ both methods become indistinguishable for usual values of the volume constraints.

1.2 Canonical decomposition and normalized contract

In this section we detail the decomposition of the payoff of the swing contract (with firm constraints) into a swap part having a closed form expression and a contract with a minimal local constraint set to 0. This decomposition is mentioned in the literature (see *e.g.* [6, 26]). The second term can be reduced by homogeneity to a normalized contract. The swap part can be seen as a pseudo-control variate since it reduces the relative part of the global price that needs to be computed by some numerical methods. Practically, it may lead to an impressive increase of the precision of the price computation. The normalization of the resulting swing with lower local constraints induces a significant simplification by reducing the number of parameters to 2 (global constraints) instead of 4. Overall, its straightforward financial interpretation leads to a better understanding of the swing contract. Formally, one has

$$\begin{aligned}P(0,S_0) &= \underbrace{\mathbb{E}\left(\sum_{k=0}^{n-1}q_{\min}e^{-rt_k}(S_{t_k}-K)\right)}_{Swap} \\ &\quad+(q_{\max}-q_{\min})\underbrace{\sup_{(q_{t_k})\in\mathcal{A}^{\tilde{Q}_{\min},\tilde{Q}_{\max}}_{[0,1]}(0,0)}\mathbb{E}\left(\sum_{k=0}^{n-1}e^{-rt_k}q_{t_k}(S_{t_k}-K)\right)}_{Normalized\ Contract}\end{aligned} \tag{4}$$

where

$$\tilde{Q}_{\min}=\frac{(Q_{\min}-nq_{\min})_+}{q_{\max}-q_{\min}},\qquad \tilde{Q}_{\max}=\frac{(Q_{\max}-nq_{\min})_+}{q_{\max}-q_{\min}}. \tag{5}$$

The price models investigated in the following sections define the spot price as a process centered around the forward curve, and so $\mathbb{E}(S_t)=F_{0,t}$ is known for every $t\in[0,T]$. Thus, the swap part has a closed form given by

$$Swap_0=q_{\min}\sum_{i=0}^{n-1}e^{-rt_k}(F_{0,t_k}-K).$$

The adaptation to contracts with penalized constraints is straightforward and amounts to modifying the penalization function in an appropriate way.

1.3 Dynamic programming equation

It is shown in [8], that, in the penalized problem, the optimal consumption is the solution of a dynamic programming equation.

Proposition 1.1. *Assume that for some positive constants p and C, the following inequality holds for any $x > 0$, and $\bar{q} \in [n\, q_{\min}, n\, q_{\max}]$:*

$$\forall\, x \geq 0, \qquad |Pen_{_T}(x, \bar{q})| \leq C(1 + x^p).$$

Then there exists an optimal Markovian consumption $q^*(t_k, S_{t_k}, \bar{q}_{t_k})$ given by the maximum argument in the following dynamic programming equation:

$$\begin{cases} P(t_k, S_{t_k}, \bar{q}_{t_k}) = \max\limits_{q \in [q_{\min}, q_{\max}]} \Big\{ q(S_{t_k} - K) + e^{-r(t_{k+1} - t_k)} \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, \bar{q}_{t_k} + q) | S_{t_k}) \Big\}, \\ P(T, S_T, Q_T) = Pen_{_T}(S_T, Q_T). \end{cases} \tag{6}$$

Usually, the function $Pen_{_T}(x, \bar{q})$ is given by (2). Then, the case with firm constraints corresponds to the limit case where $Pen_{_T}(x, \bar{q}) = (-\infty)\mathbf{1}_{\{x \notin [Q_{\min}, Q_{\max}]\}}$.

When considering a contract with firm constraints, a more tractable form (see [5]) can be the following: $P(T, S_{_T}, \bar{q}_{_T}) = 0$ and for every k=0,. . . ,n-1,

$$\begin{aligned} P(t_k, S_{t_k}, \bar{q}_{t_k}) = \max \Big\{ & q(S_{t_k} - K) + e^{-r(t_{k+1} - t_k)} \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, \bar{q}_{t_k} + q) | S_{t_k}), \\ & q \in [q_{\min}, q_{\max}],\ \bar{q}_{t_k} + q \in [(Q_{\min} - (n-k) q_{\max})_+, (Q_{\max} - (n-k) q_{\min})_+] \Big\}. \end{aligned} \tag{7}$$

1.4 Bang-Bang consumption

1.4.1 Case with penalties on purchased volumes

In [8] is shown the following theoretical result.

Theorem 1.2. *Consider the penalized problem (6) and $Pen_{_T}(x, Q) = -xPen(Q)$ where Pen is a continuously differentiable function. If the following condition holds*

$$\mathbb{P}\left(e^{-rt_k}(S_{t_k} - K) + \mathbb{E}(e^{-rT} S_T Pen'(\bar{q}_T^*) | S_{t_k}, Q_{t_k}^*) = 0\right) = 0,$$

the optimal consumption at time t_k is necessarily of bang-bang type given by

$$\begin{aligned} q^*(t_k, S_{t_k}, \bar{q}_{t_k}^*) \quad = \quad & q_{\max} \mathbf{1}_{\{e^{-rt_k}(S_{t_k} - K) + \mathbb{E}(e^{-rT} S_T P'(\bar{q}_T^*) | S_{t_k}, \bar{q}_{t_k}^*) > 0\}} \\ & + q_{\min} \mathbf{1}_{\{e^{-rt_k}(S_{t_k} - K) + \mathbb{E}(e^{-rT} S_T P'(\bar{q}_T^*) | S_{t_k}, \bar{q}_{t_k}^*) < 0\}}. \end{aligned}$$

The above assumption seems difficult to check since it involves the unknown optimal consumption at time T. However, the condition could be satisfied provided one shows that the random variable $e^{-rt_k}(S_{t_k} - K) + \mathbb{E}(e^{-rT} S_T Pen'(\bar{q}_T^*) | S_{t_k}, \bar{q}_{t_k}^*)$ is absolutely continuous as noticed in [8].

1.4.2 Case with firm constraints

In a companion paper [5], we establish some properties of the value function of the swing options viewed *as a function of the global volume constraints* $(Q_{\min}, Q_{\max})$. Thanks to (4) one may assume without loss of generality that the contract is normalized, *i.e.* $q_{\min} = 0$ and $q_{\max} = 1$. We consider the following value function:

$$P(Q_{\min}, Q_{\max}) = \sup_{(q_{t_k})_{0 \leq k \leq n-1} \in \mathcal{A}^{Q_{\min}, Q_{\max}}(0,0)} \mathbb{E}\left(\sum_{k=0}^{n-1} e^{-rt_k} q_{t_k} (S_{t_k} - K) \right)$$

defined on the unit (upper) simplex $\{(u,v) \in \mathbb{R}^2,\ 0 \leq u \leq v \leq n\}$. (From now on, we drop the dependence of the control set in $[q_{\min}, q_{\max}] = [0,1]$).

Proposition 1.3. *The premium function* $(Q_{\min}, Q_{\max}) \mapsto P(Q_{\min}, Q_{\max})$ *is a concave, piecewise affine function of the global purchased volume constraints, affine on elementary triangles* $(m,M) + \{(u,v), 0 \leq u \leq v \leq 1\}$, $(m,M) \in \mathbb{N}^2$, $m \leq M \leq n$ *and* $(m,M) + \{(u,v), 0 \leq v \leq u \leq 1\}$, $(m,M) \in \mathbb{N}^2$, $m \leq M-1 \leq n-1$ *which tile the unit (upper) simplex.*

Theorem 1.4. *For integral valued global constraints, i.e.* $(Q_{\min}, Q_{\max}) \in \mathbb{N}^2$, *there always exists a bang-bang optimal strategy i.e. the* a priori $[0,1]$*-valued optimal purchased quantities* $q^*_{t_k}$ *are in fact always equal to* 0 *or* 1.

Remark 1.3. This result can be extended in some way to any couple of global constraints when all the payoffs are nonnegative (see [5]). Furthermore, it has nothing to do with the Markov dynamics of the underlying asset and holds in a quite general abstract setting.

An example of the premium function $(Q_{\min}, Q_{\max}) \mapsto P(Q_{\min}, Q_{\max})$ is depicted in Figure 1. The model parameters are those settled in Equation (15), Section 4.2.1. Now we turn to the problem

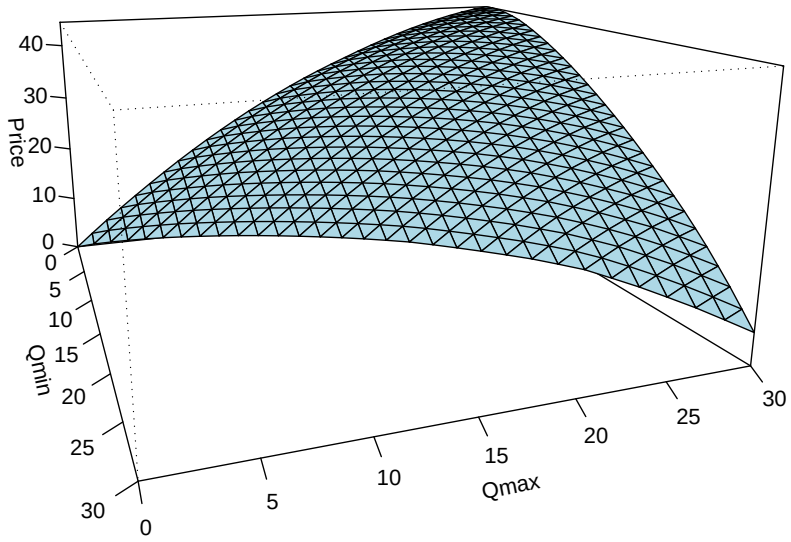

Figure 1: *Value function* $(Q_{\min}, Q_{\max}) \mapsto P(0, Q_{\min}, Q_{\max})$ *with penalized constraints,* $n = 30$, $q_{\min=0}$, $q_{\max=1}$, *two-factor model (see equation (15) in Section 4.2.1) and strike* $K = 18$

of the numerical evaluation of such contracts. As announced, we focus on an optimal quantization algorithm.

2 Optimal quantization

Optimal Quantization [32, 2, 3, 4] is a method coming from Signal Processing devised to approximate a continuous signal by a discrete one in an optimal way. Originally developed in the 1950's, it was introduced as a quadrature formula for numerical integration in the late 1990's, and for conditional expectation approximations in the early 2000's, in order to price multi-asset American style options.

Let X be an $\mathbb{R}^d$-valued random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Quantization consists in studying the best approximation of X by random vectors taking at most N fixed values $x^1, \ldots, x^N \in \mathbb{R}^d$.

Definition 2.1. Let $x = (x^1, \ldots, x^N) \in (\mathbb{R}^d)^N$. A partition $(C_i(x))_{i=1,\ldots,N}$ of $\mathbb{R}^d$ is a *Voronoi tessellation* of the N-quantizer x (or *codebook*; the term *grid* being used for $\{x^1, \ldots, x^N\}$) if, for every $i \in \{1, \ldots, N\}$, $C_i(x)$ is a Borel set satisfying

$$C_i(x) \subset \{\xi \in \mathbb{R}^d, |\xi - x^i| \leq \min_{i \neq j} |\xi - x^j|\}$$

where $|\,.\,|$ denotes the canonical Euclidean norm on $\mathbb{R}^d$.

The nearest neighbour projection on x induced by a Voronoi partition is defined by

$$\mathrm{Proj}_x : y \in \mathbb{R}^d \mapsto x^i \;\text{ if }\; y \in C_i(x).$$

Then, we define an *x-quantization* of X by

$$\hat{X}^x = \mathrm{Proj}_x(X).$$

The *pointwise error* induced when replacing X by $\widehat{X}^x$ is given by $|X - \widehat{X}^x| = d(X, \{x^1, \ldots, x^N\}) = \min_{1 \leq i \leq N} |X - x^i|$. When X has an absolutely continuous distribution, any two x-quantizations are $\mathbb{P}$-*a.s.* equal.

The quadratic *mean quantization error* induced by the the N-tuple $x \in \mathbb{R}^d$ is defined as the quadratic norm of the pointwise error *i.e.* $\|X - \hat{X}^x\|_2$ (where $\|Y\|_2 = \sqrt{\mathbb{E}Y^2}$ denotes the $L^2(\mathbb{P})$-norm on $(\Omega, \mathcal{F}, \mathbb{P})$).

We briefly recall some classical facts about theoretical and numerical aspects of Optimal Quantization. For details we refer *e.g.* to [22, 33].

Theorem 2.1. *[22] Let $X \in L^2(\mathbb{R}^d, \mathbb{P})$. The quadratic distortion function (squared quantization error)*

$$x = (x^1, \ldots, x^N) \longmapsto \mathbb{E}(\min_{1 \leq i \leq N} |X - x^i|^2) = \|X - \hat{X}^x\|_2^2$$

reaches a minimum at some quantizer x^*. Furthermore, if the distribution $\mathbb{P}_X$ has an infinite support then $x^{*,(N)} = (x^{*,1}, \ldots, x^{*,N})$ has pairwise distinct components and $N \mapsto \min_{x \in (\mathbb{R}^d)^N} \|X - \hat{X}^x\|_2$ is decreasing to 0 as $N \uparrow +\infty$.

Figure 2 shows a quadratic optimal – or at least close to optimality – quantization grid for a bivariate normal distribution $\mathcal{N}(0, I_2)$. The rate of convergence to 0 of the optimal quantization error is ruled by the so-called Zador Theorem.

Theorem 2.2. *(a)* Sharp rate. *(see [22]) Let $X \in L^{2+\delta}(\mathbb{P}), \delta > 0$, with $\mathbb{P}_X(d\xi) = \varphi(\xi)\lambda_d(d\xi) + \nu(d\xi)$, $\nu \perp \lambda_d$ (λ_d Lebesgue measure on $\mathbb{R}^d$). Then, there is a constant $\widetilde{J}_{2,d} \in (0, \infty)$, such that*

$$\lim_{N \to +\infty} (N^{\frac{1}{d}} \min_{x \in (\mathbb{R}^d)^N} \|X - \hat{X}^x\|_2) = \widetilde{J}_{2,d} \left(\int_{\mathbb{R}^d} \varphi^{\frac{d}{d+2}} d\lambda_d \right)^{\frac{1}{2} + \frac{1}{d}}.$$

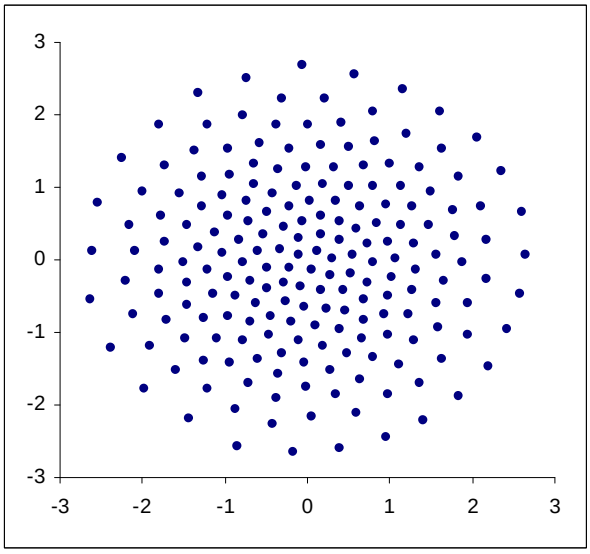

Figure 2: *Optimal quadratic quantization of size* $N = 200$ *of the bi-variate normal distribution* $\mathcal{N}(0, I_2)$ *with* $N = 200$.

(*b*) NON ASYMPTOTIC RATE. *(see [29]) Let* $\eta > 0$. *There exists a real constant* $C_{d,\eta} > 0$ *and an integer* $N_{d,p,\eta} \geq 1$ *such that or any* $\mathbb{R}^d$*-valued random vector* X, *for* $N \geq N_{d,\eta}$,

$$\min_{x \in (\mathbb{R}^d)^N} \|X - \hat{X}^x\|_2 \leq C_{d,\eta} \|X\|_{2+\eta} N^{-\frac{1}{d}}.$$

The true value of $\widetilde{J}_{2,d}$ is unknown as soon as $d \geq 3$. One only knows that $\widetilde{J}_{2,d} = \sqrt{\frac{d}{2\pi e}} + o(\sqrt{d})$.

The following result can be derived from some differentiability properties of the distortion function.

Proposition 2.3. *[32, 33] Any* L^2*-optimal quantizer* $x \in \mathbb{R}^d$ *is stationary in the following sense*

$$\mathbb{E}(X|\hat{X}^x) = \hat{X}^x.$$

Remarks. • Note that for any stationary quantizer, $\mathbb{E}(X) = \mathbb{E}(\hat{X}^x)$.

• Except in one dimension for distributions with log-concave probability densities, there are several stationary quantizers, not all of them being minima (even local).

The random vector $\hat{X}^x$ takes its values in a finite space $\{x^1, \ldots, x^N\}$, so for every continuous functional $f : \mathbb{R}^d \to \mathbb{R}$ with $f(X) \in L^2(\mathbb{P})$, we have

$$\mathbb{E}(f(\hat{X}^x)) = \sum_{i=1}^{N} f(x^i) \mathbb{P}(X \in C_i(x))$$

which is the quantization-based quadrature formula to approximate $\mathbb{E}\,(f(X))$ [32, 33]. As $\hat{X}^x$ is close to X, it is natural to estimate $\mathbb{E}(f(X))$ by $\mathbb{E}(f(\hat{X}^x))$ when f is continuous. Furthermore, when f is smooth enough, one can provide an upper bound for the resulting error using $\|X - \hat{X}^x\|_2$, or even $\|X - \hat{X}^x\|_2^2$ (when the quantizer x is stationary).

The same idea can be used to approximate the conditional expectation $\mathbb{E}(f(X)|Y)$ by $\mathbb{E}(f(\hat{X})|\hat{Y})$, but one also needs the transition probabilities:

$$\mathbb{P}(X \in C_j(x)|Y \in C_i(y)).$$

The application of this technique to the quantization of spot price processes is discussed in detail in the Annex, page 24.

3 Pricing swing contracts with a quantization tree

3.1 Description of the algorithm (general setting)

In this section we assume that $(S_{t_k})_{0\le k\le n}$ is a Markov process. For the sake of simplicity, we consider that there is no interest rate. We also consider a normalized contract, as defined in Section 1.2.

In the penalized problem, the price of the swing option is given by the following dynamic programming equation (see Equation (6)):

$$\begin{cases} P(t_k, S_{t_k}, \bar{q}_{t_k}) = \max\limits_{q\in\{0,1\}} [q(S_{t_k} - K) + \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, \bar{q}_{t_k} + q)|S_{t_k})],\ k = 0, \dots, n-1, \\ \\ P(T, S_T, \bar{q}_T) = Pen_T(S_T, \bar{q}_T) \end{cases}$$

where $t_k = k\Delta, k = 0, \dots, n$, $\Delta = \frac{T}{n}$.

We assume the bang-bang feature of the optimal consumption (owing to Theorem 1.2 in Section 1.4.1 or as a close approximation of the firm constraint case when A and B are large enough). It allows us to limit the possible values of q in the dynamic programming equation to $q \in \{0, 1\}$. The possible values of the cumulative consumption at time t_k are

$$\bar{q}^{\ell}_{t_k} = \ell, \quad 0 \le \ell \le k. \tag{8}$$

At every time t_k we consider an (optimized) N_k-quantization $\widehat{S}_{t_k} = \widehat{S}^{x_k^{(N)}}_{t_k}$, $k = 0, \dots, n$, based on an optimized quantization N_k-tuple (or grid) $x_k^{(N_k)} := \left(s_k^1, \dots, s_k^{N_k}\right)$ of the spot S_{t_k}.

The modeling of the yields of the future prices by multi-factor Gaussian processes with memory (see Section 4 for a toy example) implies that $(S_t)_{t\in[0,T]}$ is a log-normal process. Then the quantization of S_{t_k} can be obtained by a simple dilatation-contraction (by a factor $\mathrm{StD}(S_{t_k})$) from optimal quantization grids of the (possibly multivariate) normal distribution, to be downloaded on the website [35]

`www.quantize.maths-fi.com`.

Then we compute the price at each time t_k, for all points on the corresponding grid, and for all the possible cumulative consumptions:

$$\begin{cases} P(t_k, s_k^i, \bar{\hat{q}}_{t_k}) = \max\limits_{q\in\{0,1\}} [q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, \bar{\hat{q}}_{t_k} + q)|\hat{S}_{t_k} = s_k^i)],\ k = 0, \dots, n-1, \\ \qquad\qquad i = 1, \dots, N_k, \\ P(T, s_T^i, \bar{\hat{q}}_T) = Pen_T(s_T^i, \bar{\hat{q}}_T),\ i = 1, \dots, N_n. \end{cases} \tag{9}$$

▷ *Firm constraints.* When considering a contract with firm constraints, we need to compute the price at each time t_k, for all the points of the quantization grid of the spot price, and for all the admissible cumulative consumptions (See Figure 3)

$$\bar{\hat{q}}^{\ell}_{t_k} = \ell + (Q_{\min} - (n-k))_+, \ell = 0, \dots, \min(k, Q_{\max}) - (Q_{\min} - (n-k))_+. \tag{10}$$

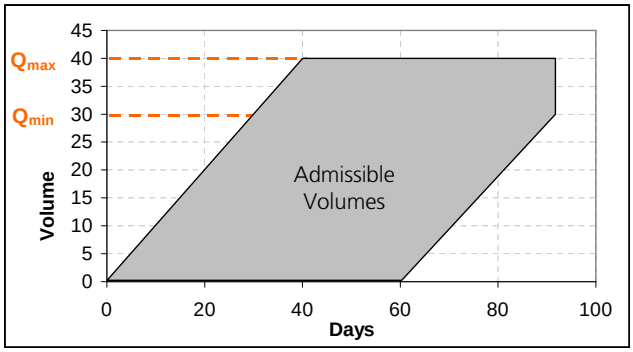

Figure 3: *Volume constraints: possible partial cumulative consumptions lie in the grey area.*

Using the bang-bang feature (See Section 1.4.2) and the dynamic programming principle (7), this price is given by $P(T,.,.) \equiv 0$ and for every $k = 0, \ldots, n-1$,

$$P(t_k, s_k^i, \bar{\hat{q}}_{t_k}) = \max\Big\{q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, \bar{\hat{q}}_{t_k} + q)|\hat{S}_{t_k} = s_k^i), \qquad (11)$$
$$q \in \{0,1\},\ \bar{\hat{q}}_{t_k} + q \in [(Q_{\min} - (n-k))_+, Q_{\max}]\Big\}.$$

Since $\hat{S}_{t_k}$ takes its values in a finite space, we can rewrite the conditional expectation as:

$$\mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, Q)|\hat{S}_{t_k} = s_k^i) = \sum_{j=1}^{N_{k+1}} P(t_{k+1}, s_{k+1}^j, Q)\pi_k^{ij}$$

where

$$\pi_k^{ij} = \mathbb{P}(\hat{S}_{t_{k+1}} = s_{k+1}^j | \hat{S}_{t_k} = s_k^i)$$

is the *quantized transition probability* between times t_k and t_{k+1}. The whole set of quantization grids equipped with the transition matrices make up the so-called "quantization tree". The transition weights (π_k^{ij}) matrices are the second quantity needed to process the quantized dynamic programming principle (9 or 11). A specific fast parallel quantization procedure has been developed in our multi-factor Gaussian framework to speed up (and parallelize) the computation of these weights (see Annex). In a more general framework, one can follow the usual Monte Carlo approach described in [2] to compute the grids and the transition matrices of the global quantization tree.

When (S_{t_k}) no longer has a Markov dynamics but appears as a function of a Markov chain (X_{t_k}): $S_{t_k} = f(X_{t_k})$, one can proceed as above, except that one has to quantize (X_{t_k}). Then the dimension of the problem (in terms of quantization) is that of the structure process (X_{t_k}).

Of course one can always implement the above procedure formally regardless of the intrinsic dynamics of (S_{t_k}). This yields to a drastic dimension reduction (from that of X down to that of (S_{t_k})). Doing so, we cannot apply the convergence theorem (see Section 3.3) which says that in a Markovian framework the premium resulting from (9) or (11) converges toward the true one as the size of the quantization grid goes to infinity.

This introduces a methodological residual error that can be compared to that appearing in the algorithm for American option pricing (see [7]). However, one checks on simulations that this residual error turns out to be often negligible (see Section 4.3).

3.2 Complexity

The first part of the algorithm consists in designing the quantization tree and the corresponding weights. The complexity of this step is directly connected to the size of the quantization grids

chosen for the transitions computation in 1-dimension, or to the number of Monte Carlo simulations otherwise. However those probabilities have to be calculated once for a given price model, and then several contracts can be priced using the same quantization tree. So we will first focus on the complexity of the quantized dynamic programming algorithm.

We first consider a penalized normalized contract, *i.e.* $q_{\min} = 0$ and $q_{\max} = 1$. The implementation of the dynamic programming principle requires three interlinked loops. For each time step k (going backward from $n-1$ to 0), one needs to compute at all the points s_k^i, $i = 1, \dots, N_k$ of the quantization grid and for every possible cumulative consumption $\bar{\hat{q}}_{t_k}^{\ell}$ $(0 \le \ell \le k)$ (see (8)) the functional

$$\max_{q \in \{0,1\}} [q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, \bar{\hat{q}}_{t_k} + q) | \hat{S}_{t_k} = s_k^i)], \ k = 0, \dots, n-1,$$

which means computing twice a sum of N_{k+1} terms. The computation of the penalization function has a complexity of nN_n.

Hence, the global complexity in the penalized setting is proportional to

$$\sum_{k=0}^{n} (k+1) N_k N_{k+1}.$$

(with the convention $N_n = 1$). If all layers in the quantization tree have the same size, *i.e.* $N_k, = \bar{N}$, $\forall k = 1, \dots, n$, the complexity is proportional to $n^2 \bar{N}^2$. This is not an optimal design but on the other hand *only one grid needs to be stored*. It is possible to reduce the algorithm complexity by optimizing the grid sizes N_i (subject to the constraint $\sum_k N_k = n\bar{N}$) [1], but it costs more memory space.

In the case of firm constraints, the dynamic programming principle (11) has to be computed at every time $k = 0, \dots, n-1$, at admissible cumulative consumption $\bar{\hat{q}}_{t_k}^{\ell}$ defined by (10)). The resulting complexity is strictly lower, namely proportional to

$$\sum_{k=0}^{n-1} (\min(k, Q_{\max}) + 1 - (Q_{\min} - (n-k))_+) N_k N_{k+1}$$

Of course, the complexity in the case of firm constraints is lower than that for a penalized problem, and depends on the global constraints $(Q_{\min}, Q_{\max})$. But the implementation is easier in the case of a penalized problem, because one does not need to check if the cumulative consumption volume is admissible. Both approaches have been numerically tested and results are indistinguishable for large enough penalties. For the ease of implementation, the approach with penalties has been adopted (although it is not optimal in terms of computational complexity).

In order to reduce the complexity of the algorithm, one usually prunes the quantization tree. In most examples, at each layer k, many terms of the transition matrix $(\pi_{ij}^k)_{i,j}$ are equal to 0 or negligible. So while the transition probabilities are estimated, all the transitions that are not visited are deleted. This step is important because it allows to reduce the algorithm complexity significantly.

In practice we can even neglect transitions whose probability is very low, say less than 10^{-5}.

Remark 3.1. In this section we focused of the complexity of the dynamic programming formula, assuming that the grid computation and the transition weight estimation has been performed *off-line*. A new approach to transition weight estimation in the Gaussian framework is investigated

[1] To minimize the complexity, set $N_k \approx \frac{2n\bar{N}}{(k+1)\log(n)}$, $k = 0, \dots, n-1$, which leads to a global complexity proportional to $\frac{4n^2\bar{N}^2}{\log(n)}$

in the Appendix, which makes possible to parallelize the standard Monte Carlo procedure. (For theoretical and numerical aspects of the standard Monte Carlo procedure and its performances we refer *e.g.* to [33] or [1].) Whatsoever, all these procedures heavily depend on repeated nearest neighbour searches. This problem is known to become very challenging as the dimension of the state space increases. Several fast nearest neighbour search procedures (among, say, $\bar{N}$ points) have been devised since the early 1970's. The most efficient ones prescribe to organize the data set as a tree (which requires a $O(\bar{N}\log\bar{N})$ complexity), making then possible a nearest neighbour search with a $O(\log\bar{N})$ complexity: the K-d-tree algorithm is the most commonly implemented (see [20]). This procedure is implemented in our weight estimation procedure to reduce the computation time: its complexity for a Monte Carlo sample of size M is proportional to $n(\bar{N}+M)\log\bar{N}$.

As concerns the grid optimization, all algorithms are also based in higher dimensions on repeated nearest neighbour searches: however a stochastic gradient like the Competitive Learning Vector Quantization algorithm proceeds on a moving set whereas the randomized fixed point procedure like Lloyd's I procedure acts most of the time on a fixed set, see [34]. Fast nearest neighbour search procedures can be implemented straightforwardly in the second case. The K-d-tree itself can be speeded up when working on distributions with correlated marginals by adding a PCA pre-processing phase (see [28] or recent advances in that direction).

3.3 Convergence

In [5] an error bound for the pricing of swing options by optimal quantization is proved.

Let $P_0^n(Q)$ denote the price of the *swing firm contract* at time 0 where n is the number of possible gas delivery instants and $Q=(Q_{\min},Q_{\max})$ is the global constraint. We consider a contract with normalized local constraints $q_{\min}=0$ and $q_{\max}=1$. The "quantized" price $\hat{P}_0^n(Q)$ is the approximation of the price obtained using optimal quantization.

Proposition 3.1. *Assume there is a real exponent $p\in[1,+\infty)$ such that the (d-dimensional) Markov structure process $(X_k)_{0\le k\le n-1}$ satisfies*

$$\max_{0\le k\le n-1}|X_k|\in L^{p+\eta}(\mathbb{P}),\eta>0.$$

At each time $k\in\{0,\dots,n-1\}$, we implement a (quadratic) optimal quantization grid $x^{\bar{N}}$ of size $\bar{N}$ of X_k. Then

$$\|\sup_{Q\in T^+(n)}|P_0^n(Q)-\hat{P}_0^n(Q)|\,\|_p\le C\frac{n}{\bar{N}^{\frac{1}{d}}}$$

where $T^+(n):=\{(u,v),0\le u\le v\le n\}$ is the set of admissible global constraints (at time 0).

In fact this error bound turns out to be conservative and several numerical experiments, such as those presented in Section 4, suggest that in fact the true rate (for a fixed number n of purchase instants) behaves like $O(\bar{N}^{-\frac{2}{d}})$.

4 Numerical experiments

In this section the same grid size has been used at each time step, *i.e.* we always have $N_k=\bar{N},k=0,\dots,n-1$. The results have been obtained by implementing the penalized problem and using the canonical decomposition (see Section 1.2).

4.1 The one-factor model

Swing options are often priced using the least squares regression method "à la Longstaff-Schwartz" [27]. This section aims to compare our numerical results to those obtained with Longstaff-Schwartz method. We consider a one factor model, which corresponds to a one dimensional Markov structure process.

4.1.1 Quantization tree for a one dimensional structure process

We consider the following diffusion model for the forward contracts $(F_{t,T})_{0\leq t\leq T}$:

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma e^{-\alpha(T-t)} dW_t$$

where W is a standard Brownian motion. It yields:

$$S_t = F_{0,t} \exp\left(\sigma \int_0^t e^{-\alpha(t-s)} dW_s - \frac{1}{2}\Lambda_t^2\right)$$

where

$$\Lambda_t^2 = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}).$$

Denote $X_k = \int_0^{k\Delta} e^{-\alpha(t-s)} dW_s$. The structure process $(X_k)_{k\geq 0}$ can be quantized using the fast parallel quantization method described in the Annex (page 24). Let $x_k^{(\bar{N})}$ denote an (optimal) quantization grid of X_k of size $\bar{N}$. We have to compute for every $k \in \{0, \ldots, n-1\}$, and every $(i,j) \in \{1, \ldots, \bar{N}\}^2$, the following (quantized) transition probability weights:

$$p_k^{ij} = \mathbb{P}(\eta_1 \in C_i(x_k^{(\bar{N})}); \alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(x_{k+1}^{(\bar{N})})) \tag{12}$$

where $(\eta_1, \eta_2) \sim \mathcal{N}(0, I_2)$, and α_k and β_k are scalar coefficients for which closed expressions are available.

This computation can be achieved can be done by using quantization again as presented in the Annex, page 24 (see (21)).

4.1.2 Comparison with the regression method

We first use the following parameters for the one factor model:

$$\sigma = 70\%,\ \alpha = 4,\ F_{0,t_k} = 20,\ k = 0, \ldots, n. \tag{13}$$

The following Tables 1 and 2 present the results obtained with Longstaff Schwartz and optimal quantization, for different strike values. The L-S algorithm has been implemented with 4 polynomial regression functions $f_k(x) = x^k$, $k = 0, 1, \ldots, 3$ following [8] and $M_{LS} = 1000$ Monte-Carlo sample paths have been used. The 95% confidence interval of the Monte Carlo estimate is given in the table. Different grid sizes $\bar{N}$ have been used to quantize the spot price process whereas the transitions have been computed with a 500-point grid. The local volume constraints $q_{\min}$ and $q_{\max}$ are set to 0 and 6 respectively.

We first consider in Table 1 a case without constraints which means that the swing option is a strip of calls, whose price can easily be computed with the Black-Scholes formula.

Table 2 presents the results obtained with the global constraints presented on Figure 4.

The results seem consistent for both methods, the price given by quantization always belongs to the 95% confidence interval of the Longstaff-Schwartz method. One can note that it is true even for small grids, which means that quantization gives quickly a good price approximation. Moreover, the price given by quantization is very close of the theoretical price in the case of a call strip.

	$K=5$	$K=10$	$K=15$	$K=20$
Longstaff-Schwartz	[32518,33282]	[21572,22335]	[11113,11815]	[3758,4227]
10 point grid	32815	21865	11342	3917
20 point grid	32841	21894	11399	3956
50 point grid	32849	21903	11411	3976
100 point grid	32849	21905	11411	3976
200 point grid	32851	21905	11412	3979
Theoretical price	32850	21904	11413	3977

Table 1: *Comparison between Longstaff-Schwartz (with 1000 sample paths) and Quantization for different grid sizes $\bar{N}$ for a contract with no global constraints (the contract becomes a strip of call options, whose theoretical price can be obtained by the Black-Scholes formula).*

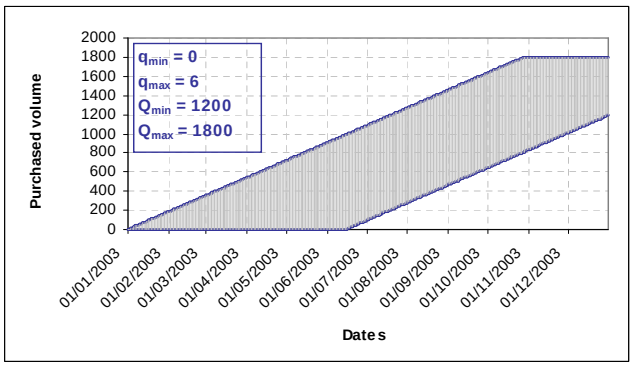

Figure 4: *Consumption constraints: the admissible cumulated volumes are represented in the grey area.*

	$K=5$	$K=10$	$K=15$	$K=20$
Longstaff-Schwartz	[27853,28503]	[18858,19507]	[10174,10797]	[2581,3133]
10 point grid	28588	19588	10796	3236
20 point grid	28338	19341	10612	3026
50 point grid	28185	19188	10475	2873
100 point grid	28156	19159	10446	2840
200 point grid	28148	19151	10438	2831

Table 2: *Comparison between Longstaff-Schwartz (with 1000 sample paths) and Quantization for different grid sizes $\bar{N}$ for the contract represented on Figure 4.*

4.1.3 Execution *CPU* time

In this section are compared the execution CPU times to price swing options using optimal quantization and *L-S* method.

The (common) size of the grids of the quantization tree is $\bar{N} = 100$. The transition weights computation is made by quantization with a grid of size $N_{trans} := 500$. And $M_{LS} = 1000$ Monte Carlo simulations are used. The maturity of the contract is one year ($n = 365$). The *L-S* algorithm was implemented with $M_{LS} = 1000$ Monte Carlo sample paths.

The computer that has been used has the following characteristics.

Processor: Celeron; *CPU* 2,4 Ghz; 1,5 Go of RAM; Microsoft Windows 2000.

The execution CPU times given in Table 3 concern the pricing of one contract: design of the quantization tree and pricing using the quantized dynamic programming as concerns the quantization approach, forward MC simulation and backward regressions as concerns the the L-S method.

L-S	Quantization: Transition weight estimation + Pricing	Quantization: Pricing only
160 s	38.5 s	2.5 s

Table 3: *Execution time for the pricing of one contract.*

There is no need for re-computing the quantization tree if the underlying price model has not changed. This is why quantization is really faster than L-S in this case, as one can note from the results reported in Table 4.

L-S	Quantization Transition weight estimation + 10 Pricing
$\approx$ 1600 s	61 s

Table 4: *Execution time for the pricing of 10 contracts.*

4.1.4 Robustness Analysis

When contracts such as swing options ought to be signed, negotiations usually concern the volume constraints. This is why the valuation technique has to be very sensitive and coherent to constraint variations. In this section we will compare the robustness to global constraints for L-S method and the optimal quantization tree.

Figure 5 represents the price of the contract as a function of the global constraints $Q_{\min}$ and $Q_{\max}$. The parameters of the one factor model are those given in (13), $q_{\min} = 0$, and $q_{\max} = 6$. The size of the quantization grids at each time step of the tree is $\bar{N} = 100$ and $N_{trans} = 500$ for the transition weight computation. The number of sample paths in the L-S algorithm is $M_{LS} = 1000$. The L-S regression method has been implemented so that the same Monte Carlo sample is used to compute all the contracts as $Q_{\min}$ and $Q_{\max}$ vary. Figure 5 emphasizes the fact that *two different samples (of the same size) may produce two significantly distant price surfaces.*

One can notice that the two surfaces obtained from L-S method are quite distant, and we observe that if we use many Monte Carlo samples, the price surface obtained by quantization appears as the mean of the resulting Monte Carlo computed price surfaces.

In order to improve the performance of the L-S method, one should use more Monte Carlo simulations. Nevertheless, an increasing number of sample paths implies greater issues in the management of the RAM. Indeed, it should be noticed that the practical implementation of regression methods relies on costly matrices handling algorithms, SVN decompositions for example. So, regression methods require a very careful management of the RAM, adding further difficulties to their efficient implementation. Such an issue has not to be addressed in quantization based methods.

4.1.5 Convergence

In this section we will study the convergence of the quantization method. We focus on the convergence of the pricing part of the algorithm.

We consider a one year maturity contract ($n = 365$) with the following volume constraints: $q_{\min} = 0, q_{\max} = 6, Q_{\min} = 1302, Q_{\max} = 1902$, and the first year of the daily forward curve depicted on Figure 6. The parameters of the one factor model are those given in (13).

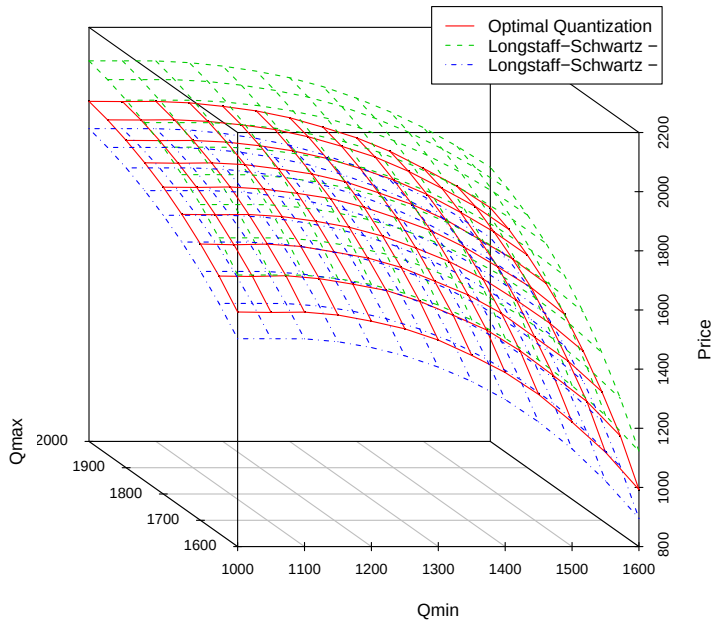

Figure 5: *Sensitivity to global constraints. L-S algorithm: the same Monte Carlo sample has been used to price all the contracts of each price surface. Two different samples (both of size* $M_{LS} = 1000$*) produce two different price surfaces.*

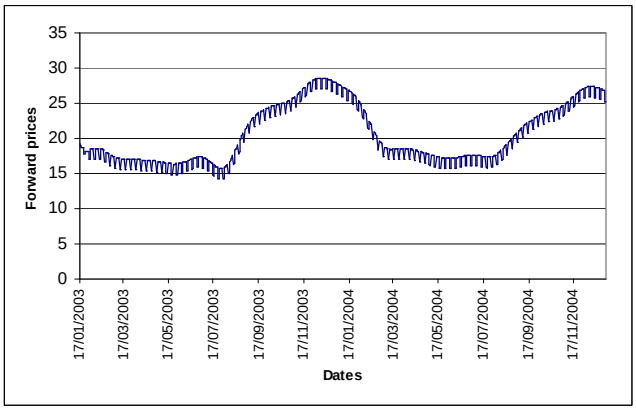

Figure 6: *Daily forward curve.*

One checks from Figure 7 that the true value of the premium is reached at $\bar{N} = 400$. Let $P(\bar{N})$ be the price obtained for a quantization grid of size $\bar{N} < 400$, the error has been computed as $|P(\bar{N}) - P(400)|, \bar{N} \leq 400$. We assume that the error can be written as a functional of the grid size $\bar{N}$ with the following shape:

$$\bar{N} \mapsto \frac{C}{\bar{N}^{\vartheta}}.$$

A linear regression in a logarithmic scale is done to find the functional that best fits the empirical

error. The ϑ coefficient obtained is 1.96.

Figures 7 and 8 show the obtained numerical convergence and the corresponding fitted functional.

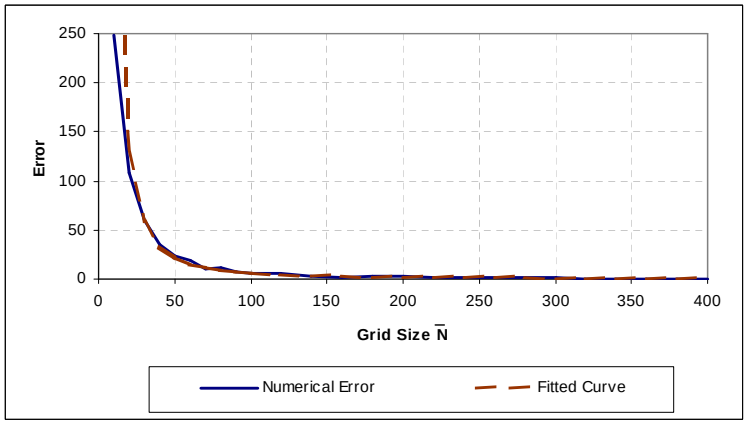

Figure 7: *Convergence rate: Numerical error as a function of the size $\bar{N}$ of the quantization grid.*

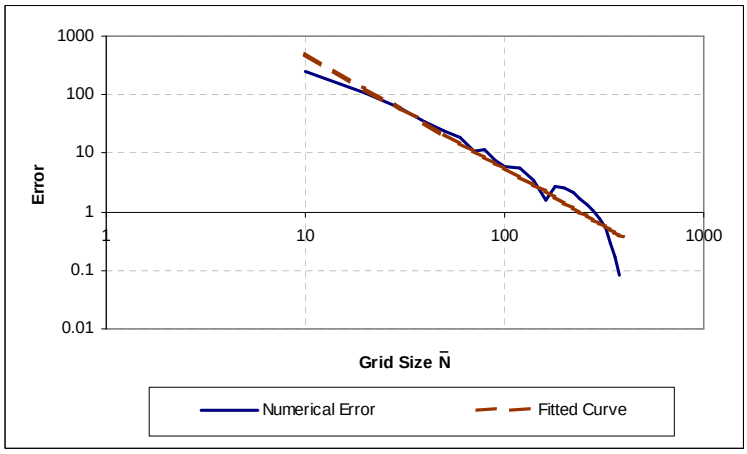

Figure 8: *Numerical Convergence in a logarithmic scale.*

The same experiments have been done for other contracts, results are presented in Table 5. $q_{\min}$ and $q_{\max}$ are set to 0 and 6.

Forward Curve	Strike	Constraints ($Q_{\min}$-$Q_{\max}$)	Estimated ϑ
Figure 6	20	1300-1900	1.96
Flat (20)	20	1302-1902	2.07
Flat (20)	10	1302-1902	2.32
Flat (20)	20	1002-2004	1.95
Flat (20)	20	1602-1800	2.26

Table 5: *Estimation of the convergence rate for different contracts: $\bar{N}^{-\vartheta}$ is the convergence rate that best fits the numerical error.*

We can conclude that the convergence rate of the quantization algorithm for pricing swing options is close to $O(\frac{1}{N^2})$. It is much better than Monte-Carlo, and leads to think that optimal quantization is an efficient alternative to L-S method for this problem.

4.2 Two factor model

We consider the following diffusion model for the forward contracts $(F_{t,T})_{0\leq t\leq T}$:

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma_1 e^{-\alpha_1(T-t)}dW_t^1 + \sigma_2 e^{-\alpha_2(T-t)}dW_t^2$$

where W^1 and W^2 are two Brownian motions with correlation coefficient ρ.

Standard computations based on Itô formula yield

$$S_t = F_{0,t}\exp\left(\sigma_1\int_0^t e^{-\alpha_1(t-s)}dW_s^1 + \sigma_2\int_0^t e^{-\alpha_2(t-s)}dW_s^2 - \frac{1}{2}\Lambda_t^2\right)$$

where

$$\Lambda_t^2 = \frac{\sigma_1^2}{2\alpha_1}(1-e^{-2\alpha_1 t}) + \frac{\sigma_2^2}{2\alpha_2}(1-e^{-2\alpha_2 t}) + 2\rho\frac{\sigma_1\sigma_2}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)t}).$$

Unlike the one factor model, the spot price process obtained from the two factor model is not a Markov process. Hence the dynamic programming equation (6) cannot be used directly. However, the structure process of the two factor model (See Annex, page 27)

$$X_t = \left(\int_0^t e^{-\alpha_1(t-s)}dW_s^1, \int_0^t e^{-\alpha_2(t-s)}dW_s^2\right)$$

is a Markov process, and $S_t = f(X_t)$ where $f : \mathbb{R}^2 \mapsto \mathbb{R}$ is a continuous function. So we can rewrite the dynamic programming equation as follows:

$$\begin{cases} P(t_k, X_{t_k}, Q_{t_k}) = \max_{q\in[q_{\min},q_{\max}]}\{q(f(X_{t_k})-K) + \mathbb{E}(P(t_{k+1}, X_{t_{k+1}}, Q_{t_k}+q)|X_{t_k})\}, \\ \\ P(T, X_T, Q_T) = Pen_T(f(X_T), Q_T). \end{cases} \tag{14}$$

To build the quantization tree of the $\mathbb{R}^2$-valued structure process $(X_{t_k})_{k\geq 0}$, we use some two-dimensional pre-computed Gaussian grids. The transition weights are computed using the "Fast Weight Estimation" procedure (that can be parallelized) described and analyzed in the Annex.

4.2.1 Call strip

We first consider a case without global constraints ($Q_{\min} = 0$, $Q_{\max} = nq_{\max}$). The swing option becomes a strip of calls for which a closed form "à la Black-Scholes" is available. We compare the results with the theoretical price, for several values of the strike K. The maturity of the contract is one month ($n = 30$). The daily constraints are set to

$$q_{\min} = 0,\ q_{\max} = 6.$$

The transitions weights have been computed with $M = 100\,000$ Monte-Carlo simulation. The parameters of the two factor model are:

$$\sigma_1 = 36\%, \alpha_1 = 0.21, \sigma_2 = 111\%, \alpha_2 = 5.4, \rho = -0.11 \quad \text{and } F_{0,t_k} = 20,\ k = 0,\ldots,n-1. \tag{15}$$

Table 6 presents the results obtained for these strips of calls. Although the quantized process is now taking values in $\mathbb{R}^2$, prices are close to the theoretical price even for small grids.

	$K=5$	$K=10$	$K=15$	$K=20$
Theoretical Price	2700	1800	924	269
$\bar{N}=50$	2690	1790	914	260
$\bar{N}=100$	2696	1796	920	264
$\bar{N}=200$	2704	1804	929	272
$\bar{N}=300$	2702	1802	927	273

Table 6: *Results obtained with a bivariate structure process for different grid sizes $\bar{N}$, for a contract with no global constraints (the contract becomes a strip of call options, whose theoretical premium is obtained by the Black-Scholes formula).*

4.2.2 Execution CPU time

The size of the quantization grids in the quantization tree is $\bar{N} = 200$ and the transition weights have been computed using $M = 100\,000$ Monte Carlo simulations. The maturity of the contract is one year ($n = 365$). The execution time in this two-dimensional setting is almost twice that of the 1-dimensional one. It remains twice shorter than the execution time of the regression method in the 1-dimensional setting.

Transition weight estimation + Pricing	Pricing only
75 s	4 s

Table 7: *Execution time for the pricing of one contract using quantization of a bivariate structure process*

As concerns the *L-S* algorithm, its computation time for the same *size* of the regression function set contract is a bit less than doubled ($\approx$ 300 seconds). As expected, there is a loss of accuracy (with monomial functions $f(x,y) = x^k y^\ell$, $0 \leq k$, $\ell \leq 1$). To reach an equivalent confidence interval as in 1-dimension requires at least 12 functions, with the induced increase of the computation CPU time. .

4.2.3 Convergence

We use the same procedure as in section 4.1.5 to find the functional $\bar{N} \mapsto C\bar{N}^{-\vartheta}$ that best fits the empirical error.

The model parameters are those given in (15), and $F_{0,t_k} = 20$, $k = 0, \ldots, n-1$. The contract parameters are set as follows: the maturity is set to one month ($n = 30$), $q_{\min} = 0$, $q_{\max} = 6$, $Q_{\min} = 78$, $Q_{\max} = 144$,

Figure 9 shows an example of the empirical error and Table 8 gathers the values of ϑ obtained for different contracts. The transition weights of the quantization tree have been computed with 3 000 000 of Monte Carlo sample paths. This number is deliberately very large, *in order to cancel out the error charged to the transitions computation.*

The convergence of the quantization algorithm is close to $O(\frac{1}{\bar{N}})$. The convergence rate is linked to the dimension d of the structure process, and from the results obtained in section 4.1.5 and in this section, we can conjecture that the convergence rate is close to $O(\bar{N}^{-2/d})$ in d-dimensions. This is better than the error bound theoretically established in Section 3.3. Moreover, this convergence rate corresponds more or less to that of a finite element method.

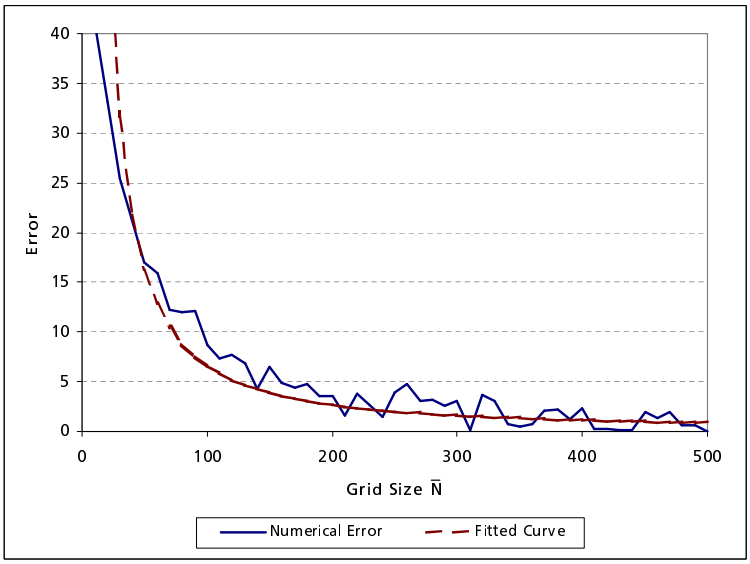

Figure 9: *Numerical rate of convergence as a function of the size* $\bar{N}$ *of the quantization grid.* $n = 30$, $F_{0,t_k} = 20$, $K = 10$, $q_{\min} = 0$, $q_{\max} = 6$, $Q_{\min} = 78$, $Q_{\max} = 144$.

Forward Curve	Strike K	Constraints ($Q_{\min}$-$Q_{\max}$)	Estimated ϑ
Flat (20)	10	78-144	1.30
Flat (20)	20	78-144	1.44
Flat (20)	20	42-156	0.92
Flat (20)	20	102-120	1.34

Table 8: *Estimation of the convergence rate for different contracts:* $\bar{N} \mapsto \bar{N}^{-\vartheta}$ *is the convergence rate that best fits the numerical error.*

4.2.4 Some elements of comparison

Gas future contracts are often modeled by multi-factor (Gaussian) processes like the family of processes investigated in the Annex (which includes the two dynamics tested in the previous sections). The number d of these factors defines the dimension of the state space of the underlying Markov structure process (the virtual spot price is a functional of this process). In such a framework, the use of multinomial trees becomes unrealistic since, for example, it is basically impossible to "catch" the correlation structure of the factors correctly.

Optimized quantization trees overcome this major drawback: they are in some way "model-driven" (when viewed as the result of a stochastic optimization procedure implemented by simulation). So they naturally fit the dynamics of the underlying asset and need no sharp *a priori* knowledge from the user. Moreover, the main advantage of tree methods, which is to be an approximation the Markov dynamics of a process (its transitions), still holds. Quantization based methods are not payoff dependent like regression methods which depend simultaneously on both the payoff function and the asset dynamics.

As the dimension increases, optimal vector quantization suffers from the so-called *curse of dimensionality* owing to its $O(\bar{N}^{-\frac{1}{d}})$ rate of convergence. This is true but can be partially overcome by using some Romberg extrapolation methods, at least up to $d = 10$ or $d = 12$ as recently pointed out in [36] (dealing with numerical integration problems). As concerns the L-S regression methods, the Monte Carlo component of the global rate of convergence is as expected dimension

free (see [15] for American options) but the choice of the regression function set quickly becomes a quite challenging problem as d increases, in particular for stochastic control problems. Once again, quantization approaches seem to have a better flexibility in multiple (medium) dimensions since in some way it appears as a model-driven non parametric selection procedure (among stepwise functions). This feature explains why this approach produces some simple global *a priori* error bounds.

Finally, the successful first implementations of the method to jump dynamics in view of applications to the electricity market (a multi-factor models based on a NIG processes instead of Brownian motion) confirm that the performances of quantization methods are not connected with Gaussian processes (see [9]): in fact no significant difference was observed in terms of performances.

4.3 Dimension reduction

In the case of multi-factorial models, we need to quantize the structure process $(X_{t_k})_k$ instead of the spot process $S_{t_k} = f(X_{t_k})$ in order to work with a Markov process (See section 4.2).

From an operational point of view, it is interesting to study the results obtained by formally quantizing the spot process (S_{t_k}), regardless to its dynamics, and using the approximation

$$\mathbb{E}(X|\mathcal{F}_{t_k}) \simeq \mathbb{E}(X|S_{t_k})$$

for any random variable X, although the spot price (S_{t_k}) is not a Markov process. Similar approximation has already been proposed by Barraquand-Martineau in [7]. Numerical tests have shown that the resulting prices remain very close to those obtained by quantizing the structure process in the case of a two factor model. Execution time and convergence rate are significantly faster (they are those of a one dimensional model). The quantization tree can be computed as presented in the Annex page 24 using Equation (18).

Although there is no theoretical evidence on the resulting systematic error, this approach seems useful to get quick results. Table 9 presents some results. The parameters of the two factor model are those given in (15) and the volume constraints are represented on Figure 4. The forward curve is flat: $F_{0,t_k} = 20,\ k = 0, \ldots, n,\ n = 365$. The convergence rate obtained in this case is always

Quantized Process	$K = 5$	$K = 10$	$K = 15$	$K = 20$
Spot Price	29837	19848	11732	4467
Structure Process	29045	20177	12012	4795

Table 9: *Quantization of the spot price vs Quantization of the bivariate structure process*

$O(\bar{N}^{-2})$, which is consistent with the general rate $O(\bar{N}^{-2/d})$, because the quantized process is $\mathbb{R}$-valued.

Therefore, when the spot process is not Markov, the quantization tree method can be performed all the same way as if it were, with small damage in practice. This is of course not a general consideration but rather an observation over the considered problem. The dramatic increase in the computation effort that can be spared from this observation can justify in this case a small loss of accuracy. Moreover, this approach is the one that is operational and seems really efficient.

Conclusion

In this article, we have introduced an optimal quantization tree method for the valuation of swing options. These options are of great interest in numerous modeling issues of the energy markets and their accurate pricing is a real challenge.

Our method has been compared to the L-S least squares algorithm and seems to perform much better on various examples. In fact, the optimal quantization method shares the good properties of the so-called multinomial tree methods but is not limited by the dimension of the underlying. Moreover, some very general theoretical *a priori* estimates for the global error have been established for this quantization based method which seems not to be the case for other approaches.

Thus, optimal quantization methods suit very well to the valuation of complex derivatives and further studies can be done in order to extend the present results to other structured products arising in the energy sector.

References

[1] Bally V. (2004). The central limit theorem for a nonlinear algorithm based on quantization. Stochastic analysis with applications to mathematical finance, *Proc. R. Soc. Lond.*, Ser. A Math. Phys. Eng. Sci., **460**(2041), 221-241.

[2] Bally V., Pagès G. (2001). A quantization algorithm for solving multi-dimensional optimal stopping problems, *Bernoulli*, **6**(9), 1-47.

[3] Bally V., Pagès G. (2003). Error analysis of the quantization algorithm for obstacle problems, *Stochastic Processes & Their Applications*, **106**(1), 1-40.

[4] Bally V., Pagès G., Printems J. (2005). A quantization tree method for pricing and hedging multi-dimensional American options, *Mathematical Finance*, **15**(1), 119-168.

[5] Bardou O., Bouthemy S., Pagès G. (2007). When are swing options bang-bang and how to use it?, pre-print LPMA-1141.

[6] Barbieri A., Garman M. B. (1996). Understanding the valuation of swing contracts, *Energy and Power Risk Management*, **1**.

[7] Barraquand J. and Martineau D. (1995). Numerical valuation of high dimensional multivariate American securities, *Journal of Financial and Quantitative Analysis* **303**, 383-405.

[8] Barrera-Esteve C., Bergeret F., Dossal C., Gobet E., Meziou A., Munos R., Reboul-Salze D. (2006). Numerical methods for the pricing of swing options: a stochastic control approach, *Methodology and Computing in Applied Probability.*

[9] Bronstein A.-L., Wilbertz B. (2007). A quantization tree algorithm: improvements and financial applications for swing options. Preprint, Univ. Trier (Germany)

[10] Carmona R. and Ludkovski M. (2007). Optimal Switching with Applications to Energy Tolling Agreements, pre-print.

[11] Carmona R. and Ludkovski M. (2005). Gas Storage and Supply Guarantees: An Optimal Switching Approach, pre-print.

[12] Carmona R. and Dayanik S. (2003). Optimal Multiple-Stopping of Linear Diffusions and Swing Options, pre-print.

[13] Carmona R. and Touzi N. (2004). Optimal multiple stopping and valuation of swing options, *Mathematical Finance*, to appear.

[14] Clearwater S. H. and Huberman B. A. (2005). Swing Options: a Mechanism for Pricing It Peak Demand.

[15] É. Clément, P. Protter, D. Lamberton (2002). An analysis of a least squares regression method for American option pricing, *Finance & Stochastics*, **6**, n^0 2, 449-47.

[16] Clewlow L., Strickland C. and Kaminski V. (2002). Risk Analysis of Swing Contracts, *Energy and power risk management.*

[17] Dörr U. (2003). Valuation of swing options and examination of exercise strategies by Monte Carlo techniques. Master's thesis, University of Oxford.

[18] Eydeland A. and Wolyniec K. (2003). Energy and Power Risk Management, *Wiley Finance.*

[19] Figueroa M. G., Pricing multiple interruptible-swing contracts. BWPEF 0606, Birbeck School of Economics, Mathematic and Statistics.

[20] Friedman, J. H., Bentley, J.L. and Finkel R.A. (1977). An Algorithm for Finding Best Matches in Logarithmic Expected Time, *ACM Transactions on Mathematical Software*, **3**(3), 209-226.

[21] Geman H. (2005). *Commodities and commodity derivatives - Modeling and Pricing for Agriculturals, Metals and Energy*, Wiley Finance.

[22] Graf S., Luschgy H. (2000). *Foundations of Quantization for Probability Distributions.* Lect. Notes in Math. 1730, Springer, Berlin.

[23] Hinz J. (2006). Valuing production capacities on flow commodities, *Mathematical Methods of Operations Research*, **64**(2), 187-209.

[24] Jaillet P., Ronn E.I., Tompaidis S. (2004). Valuation of Commodity-Based Swing Options, *Management Science*, **50**, 909-921.

[25] Keppo J. (2004). Pricing of Electricity Swing Options, *Journal of Derivatives*, **11**, 26-43.

[26] Lari-Lavassani A., Si*MC*hi M., Ware A. (2001). A discrete valuation of Swing options, *Canadian Applied Mathematics Quarterly*, **9**(1), 35-74.

[27] Longstaff F. and Schwartz E.S. (2001). Valuing American Options by Simulation: A Simple Least Squares Approach, *The Review of Financial Studies*, **14**, 113-147.

[28] *MC*Names J. (2001). A Fast Nearest-Neighbor Algorithm Based on a Principal Axis Search Tree, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(9), 964-976.

[29] Luschgy H., Pagès G., (2006). Functional quantization rate and mean regularity of processes with an application to Lévy processes, to appear in *Annals of Applied Probability.*

[30] Meinshausen N. and Hambly B.M. (2004). Monte Carlo methods for the valuation of multiple exercise options, *Mathematical Finance* , **14**,(4).

[31] Pagès G. (2007). From penalized to firm constraints for swing options. In progress.

[32] Pagès G. (1998). A space vector quantization method for numerical integration, *J. Computational and Applied Mathematics*, 89, 1-38.

[33] Pagès G., Pham H., Printems J. (2004). Optimal quantization methods and applications to numerical problems in finance, *Handbook of Numerical Methods in Finance*, ed. S. Rachev, Birkhauser, 253-298.

[34] Pagès G. and Printems J. (2003). Optimal quadratic quantization for numerics: the Gaussian case, *Monte Carlo Methods and Appl.* 9(2), 135-165.

[35] Pagès G. and Printems J. (2005). `www.quantize.maths-fi.fr`, website devoted to quantization, *maths-fi.com*.

[36] Pagès G. and Printems J. (2007). Optimal quadratic quantization in finance: from random vectors to processes, prepub. LPMA, Univ. Paris 6, to appear in *Mathematical Handbook for finance*, Elsevier.

[37] Ross S.M. and Zhu Z. (2006). Structure of swing contract value and optimal strategy, *The 2006 Stochastic Modeling Symposium and Investment Seminar, Canadian Institute of Actuaries*

[38] Thompson A.C. (1995). Valuation of path-dependent contingent claims with multiple exercise decisions over time: the case of Take or Pay. *Journal of Financial and Quantitative Analysis*, **30**, 271-293.

[39] Winter C. and Wilhelm M. (2006). Finite Element Valuation of Swing Options, pre-print.

[40] Zador P.L. (1982). Asymptotic quantization error of continuous signals and the quantization dimension, *IEEE Trans. Inform. Theory*, **28**, Special issue on quantization, A. Gersho & R.M. Grey eds, 139-149.

Annex: Fast Parallel Weight Estimation

In this annex we propose an efficient method to quantize a wide family of spot price dynamics (S_{t_k}). To be precise we will assume that a time discretization Δ being fixed,

$$S_{k\Delta} = f_k(X_k), k \geq 0 \tag{16}$$

where $(X_k)_{k\geq 0}$ is a $\mathbb{R}^m$-valued Gaussian auto-regressive process and $(f_k)_{0\leq k\leq n}$ a family of continuous functions. The fast quantization method applies to the Gaussian process $(X_k)_{0\leq k\leq n}$. We will apply it to a scalar two factor model in full details. As a conclusion to this section we will sketch the approach to a multi-factor model.

Quantization of the Gaussian structure process

We consider a centered Gaussian first order auto-regressive process in $\mathbb{R}^m$:

$$X_{k+1} = AX_k + T\varepsilon_{k+1} \tag{17}$$

where $A \in \mathcal{M}(m\times m, \mathbb{R})$, $T \in \mathcal{M}(m\times m, \mathbb{R})$ lower triangular, and (ε_k) i.i.d. with $\mathcal{N}(0, I_m)$ distribution.

Denote by $D(Z) = [\mathbb{E}(Z_iZ_j)]_{1\leq i,j\leq m}$ the covariance matrix of Z. We have $\forall k \in \mathbb{N}$:

$$D(X_{k+1}) = AD(X_k)A^* + TT^*.$$

Denote Σ_k the lower triangular matrix such that $D(X_k) = \Sigma_k\Sigma_k^*$.

We consider for every $k = 0, \ldots, n-1$, an *optimal (quadratic) quantizer* $x_k^{(N_k)}$ of size N_k, for the $\mathcal{N}(0, I_m)$ distribution. The quantization grid of the random variable X_k is taken as a dilatation of $x_k^{(N_k)}$,*i.e.*

$$\bar{x}_k = \Sigma_k \, x_k^{(N_k)} := (\Sigma_k x^{(N_k),i})_{1 \le i \le N_k}.$$

To compute the conditional expectations in the dynamic programming equation, we need to get the following transition probabilities:

$$\pi_k^{ij} = \mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}) | X_k \in C_i(\bar{x}_k))$$

where $C_i(x)$ denotes the i-the Voronoi cell of the generic quantizer $x \in (\mathbb{R}^d)^N$. Then

$$\mathbb{P}(X_k \in C_i(\bar{x}_k)) = \mathbb{P}(Z \in C_i(x_k^{(N_k)})),$$

with $Z \sim \mathcal{N}(0, I_m)$. This probability is provided as a companion parameter with the normal distribution grid files (available on [35]).

To get the transition probability π_k^{ij} we need to compute

$$p_k^{ij} = \mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}), X_k \in C_i(\bar{x}_k)).$$

Proposition 4.1. *Let X be a discrete time process described as above. Let U, V be two gaussian random variables $\mathcal{N}(0, I_m)$. Then we have for every $k \in \{0, \ldots, n-1\}$, every $i \in \{1, \ldots, N_k\}$, every $j \in \{1, \ldots, N_{k+1}\}$,*

$$\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}), X_k \in C_i(\bar{x}_k)) = \mathbb{P}(U \in C_i(x_k^{(N_k)}), A_{k+1}U + B_{k+1}V \in C_j(x_{k+1}^{N_{k+1}})) \tag{18}$$

where A_k and B_k are $m \times m$ matrices whose coefficients depend on k, and on the matrices A, T. If $k = 0$,

$$\mathbb{P}(X_1 \in C_j(\bar{x}_1)) = \mathbb{P}(V \in C_j(x_1^{(N_1)})). \tag{19}$$

Proof. We have:

$$\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)) = \mathbb{P}(AX_k + T\varepsilon_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)).$$

We consider the couple $(X_k, T\varepsilon_{k+1})$. $T\varepsilon_{k+1}$ is independent of X_k. Let $\eta = (\eta_1, \eta_2)$ a couple of independent Gaussian random vectors: $\eta_i \sim N(0, I_m), i = 1, 2$. Then $(X_k, T\varepsilon_{k+1}) \sim (\Sigma_k \eta_1, T\eta_2)$ and

$$\begin{aligned}
\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)) &= \mathbb{P}(A\Sigma_k\eta_1 + T\eta_2 \in C_j(\bar{x}_{k+1}); \Sigma_k\eta_1 \in C_i(\bar{x}_k)) \\
&= \mathbb{P}(\Sigma_{k+1}^{-1}\,(A\Sigma_k\eta_1 + T\eta_2) \in C_j(x_{k+1}^{(N_{k+1})}); \eta_1 \in C_i(x_k^{(N_k)})).
\end{aligned}$$

Setting

$$A_{k+1} = \Sigma_{k+1}^{-1} A \Sigma_k, \quad B_{k+1} = \Sigma_{k+1}^{-1} T$$

we get

$$\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)) = \mathbb{P}(A_{k+1}\eta_1 + B_{k+1}\eta_2 \in C_j(x_{k+1}^{(N_{k+1})}); \eta_1 \in C_i(x_k^{(N_k)})).$$

If $k = 0$ and $\Sigma_0 \equiv 0$, the quantity

$$\mathbb{P}(X_k \in C_j(\bar{x}_k)) = \mathbb{P}(\eta_1 \in C_j(x_k^{(N_k)}))$$

is given as a companion parameter with the quantization grids of the normal distribution. □

Remark 4.1. Equation (18) emphasizes the fact that the transitions can be computed in parallel.

Remark 4.2. To simplify the structure of the quantization tree we propose to consider the same normalized grid of size $N_k = N$ at each step k but other choices are possible like those recommended in [2].

Numerical methods

Hereafter we will focus on the numerical computation of these transitions.

• The standard Monte Carlo approach The simplest way is to use a Monte Carlo method ([2]). One just needs to simulate trajectories of the process $(X_k)_{k=0,\dots,n-1}$ and to compute the quantities p_k^{ij}. This approach can be used whatever the dimension m of the structure process X is.

• Fast Parallel Weight Estimation An other Monte Carlo based approach can be settled using Proposition 4.1. Only one sample of N_{MC} of $\mathcal{N}(0; I_m)$-distributed numbers (bivariate normal distribution for $m = 2$) need to be simulated. The main point is that the weight estimation procedure can be parallelized by computing the transitions $(p_k^{ij})_{ij}$ of every time step simultaneously.

To be precise, the following procedure has to be done N_{MC} times in order to compute the quantities p_k^{ij}:

1. Simulation of a $\mathcal{N}(0; I_m)$-distributed random vector (U, V).
2. Simultaneously for each k, $k = 0, \dots, n-1$, computation of the matrices A_{k+1} and B_{k+1}
3. Search for the nearest neighbor of U in the grid $x_k^{(N_k)}$ and of $A_{k+1}U + B_{k+1}V$ in the grid $x_{k+1}^{(N_{k+1})}$,
4. Compute $p_k^{ij} = p_k^{ij} + \frac{1}{N_{MC}}$ where i, j are the results of the nearest neighbor search

The more efficient the nearest neighbor search is, the quicker the method will be. Numerous numerical methods have been designed to compute this search, in particular the K-d-tree based approach presented in [20] (including a Partial Distance Search).

• Quantized Parallel Weight Estimation If $m = 1$, the transitions can be computed using again optimal quantization, because in low dimension (say $d \leq 4$), quantization converges faster than Monte Carlo method. In this case, we have to compute a two dimensional expectation.

We estimate for every $k \in \{0, \dots, n-1\}$, every $i \in \{1, \dots, N_k\}$ and every $j \in \{1, \dots, N_{k+1}\}$ the following probabilities:

$$p_k^{ij} = \mathbb{P}(\eta_1 \in C_i(x_k^{(N_k)}); \alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(x_{k+1}^{(N_{k+1})})) \tag{20}$$

where $(\eta_1, \eta_2) \sim \mathcal{N}(0, I_2)$, and α_k and β_k are scalar coefficients satisfying $\alpha_k^2 + \beta_k^2 = 1$:

$$\begin{aligned} \alpha_k &= \tfrac{\Sigma_k A}{\Sigma_{k+1}} \\ \beta_k &= \tfrac{T}{\Sigma_{k+1}}. \end{aligned}$$

To alleviate notations, we temporarily set $x = x_k^{(N_k)}$ and $y = x_{k+1}^{(N_{k+1})}$.

We define $\left[x^{i-\frac{1}{2}}, x^{i+\frac{1}{2}}\right] = \left[\frac{1}{2}(x^i + x^{i-1}), \frac{1}{2}(x^i + x^{i+1})\right] = C_i(x)$ (the same shortcut is implicitly defined for y).

In order to reduce the problem dimension, it is possible to write the probability p_{ij}^k as a double integral, and to integrate first with respect to the second variable by using Fubini theorem:

$$\mathbb{P}(\eta_1 \in C_i(x);\ \alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(y)) = \mathbb{E}\left(\mathbf{1}_{\{x^{i-\frac{1}{2}} \leq \eta_1 \leq x^{i+\frac{1}{2}}\}}\left(\mathcal{N}(\frac{y^{j-\frac{1}{2}} - \alpha_{k+1}\eta_1}{\beta_{k+1}}) - \mathcal{N}(\frac{y^{j+\frac{1}{2}} - \alpha_{k+1}\eta_1}{\beta_{k+1}})\right)\right) \tag{21}$$

where $\mathcal{N}(x)$ is the distribution function of the standard gaussian law.

For this one-dimensional expectation computation, quantization can be used again since it converges faster than Monte Carlo method. Parallelization can also be done as in larger dimension.

Example: Two factor model

We consider the following diffusion model for the forward contracts $(F_{t,T})_{0\leq t\leq T}$:

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma_1 e^{-\alpha_1(T-t)}dW_t^1 + \sigma_2 e^{-\alpha_2(T-t)}dW_t^2$$

where W^1 and W^2 are two Brownian motions with correlation coefficient ρ.

Standard computations based on Itô formula yield

$$S_t = F_{0,t}\exp\left(\sigma_1\int_0^t e^{-\alpha_1(t-s)}dW_s^1 + \sigma_2\int_0^t e^{-\alpha_2(t-s)}dW_s^2 - \frac{1}{2}\Lambda_t^2\right)$$

where

$$\Lambda_t^2 = \frac{\sigma_1^2}{2\alpha_1}(1-e^{-2\alpha_1 t}) + \frac{\sigma_2^2}{2\alpha_2}(1-e^{-2\alpha_2 t}) + 2\rho\frac{\sigma_1\sigma_2}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)t}).$$

We have $S_t = F_{0,t}\exp\left(\sigma_1 X_t^1 + \sigma_2 X_t^2 - \frac{1}{2}\Lambda_t^2\right)$, where X_t is the following structure process:

$$X_t = \left(\int_0^t e^{-\alpha_1(t-s)}dW_s^1, \int_0^t e^{-\alpha_2(t-s)}dW_s^2\right). \tag{22}$$

Proposition 4.2. *Let $Z = (Z_t)$ be an Ornstein-Uhlenbeck process. $Z_t = Z_0 + \int_0^t e^{-\alpha(t-s)}dB_s$ where B is a standard Brownian motion, and Z_0 is Gaussian and independent of B. Z can be written at discrete times $k\Delta$ as a first order auto-regressive process:*

$$Z_{k+1} = e^{-\alpha\Delta}Y_k + \sqrt{1-e^{-2\alpha\Delta}}\sqrt{\frac{1}{2\alpha}}\varepsilon_{k+1} \tag{23}$$

where (ε_k) is i.i.d and $\varepsilon_1 \sim \mathcal{N}(0,1)$.

X_t is made up with two Ornstein-Uhlenbeck processes. Using Proposition 4.2, it yields:

Proposition 4.3.

$$X_{k+1} = AX_k + T\varepsilon_{k+1}$$

with $\varepsilon_k \sim \mathcal{N}(0, I_2)$ i.i.d. and

$$A = \begin{bmatrix} e^{-\alpha_1\Delta} & 0 \\ 0 & e^{-\alpha_2\Delta} \end{bmatrix}$$

$$r = \rho\frac{\frac{1}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta})}{\sqrt{\frac{1}{4\alpha_1\alpha_2}(1-e^{-2\alpha_1\Delta})(1-e^{-2\alpha_2\Delta})}}$$

$$T = \begin{bmatrix} \frac{1}{2\alpha_1}(1-e^{-2\alpha_1\Delta}) & 0 \\ \frac{1}{2\alpha_2}(1-e^{-2\alpha_2\Delta})r & \frac{1}{2\alpha_2}(1-e^{-2\alpha_2\Delta})\sqrt{1-r^2} \end{bmatrix}.$$

Hence it is possible to use the fast parallel quantization method described in section 4.3.

General multi-factor Gaussian model

More generally, we consider a family of price dynamics that can be written as follows:

$$\frac{dF_{t,T}}{F_{t,T}} = \sum_{i=1}^{m} P_i(T-t)e^{-\alpha_i(T-t)}dW_t^i \tag{24}$$

where $P_i(x)$ is a polynomial function of degree d_i, for every $i = 1, \dots, m$, and W is a Brownian motion, with $d < W^i, W^j >_t = \rho_{ij}dt$.

The two factor model (Section 4.3) corresponds to $m = 2$, $P_i \equiv \sigma_i$, $i = 1, 2$.

In order to price a swing option with such a model, we first need to quantize it. Equation (24) yields:

$$F_{t,T} = F_{0,T}e^{\sum_{i=1}^{m}\int_0^t P_i(T-s)e^{-\alpha_i(T-s)}dW_s^i - \frac{1}{2}\phi(t,T)} \tag{25}$$

where

$$\phi(t,T) = \sum_{i=1}^{m}\int_0^t P_i^2(T-s)e^{-2\alpha_i(T-s)}ds + \sum_{i\neq j}\rho_{ij}\int_0^t P_i(T-s)P_j(T-s)e^{-(\alpha_i+\alpha_j)(T-s)}ds.$$

Practically we focus on the spot price $F_{t,t}$ or the day-ahead contract $F_{t,t+1}$. Unfortunately these processes are not Markovian in a general setting, except when $m = 1$ and $d_1 = 0$ (Ornstein-Uhlenbeck process).

We consider a discretization time step $\Delta > 0$, and we set, for all $i \in \{1, \dots, m\}$ and for all $l \in \{0, \dots, d_i\}$

$$X_k^{i,l} = \int_0^{k\Delta} P_i\left((k+l)\Delta - s\right)e^{-\alpha_i((k+l)-s)}dW_s^i.$$

Proposition 4.4. *$X_k = [X_k^{i,l}]_{1\leq i\leq m, 0\leq l\leq d_i}$ is a $\mathbb{R}^{d_1+\cdots+d_m+m}$-valued gaussian AR(1).*

Lemma 4.5. *Let $P \in \mathbb{R}[Z]$, $d^oP = d$ and $\theta \in \mathbb{R}^*$. Then $(P(Z + l\theta))_{0\leq l\leq d}$ is a basis of $\mathbb{R}^d$.*

Proof. Using a dimension argument, only the linear independence of the family has to be checked. And we have

$$\begin{aligned}\sum_{k=0}^{d}\lambda_k P(Z+k\theta) = 0 \quad &\Leftrightarrow \quad \sum_{k=0}^{d}\lambda_k\sum_{j=0}^{d}\frac{(k\theta)^j}{j!}P^{(j)}(Z) = 0\\ &\Leftrightarrow \quad \sum_{j=0}^{d}\frac{\theta^j}{j!}\left(\sum_{k=0}^{d}\lambda_k k^j\right)P^{(j)}(Z) = 0.\end{aligned}$$

Since $\left(P^{(j)}(Z)\right)_{0\leq j\leq d}$ is a basis of $\mathbb{R}_d[Z]$, it yields

$$\forall j \in \{0, \dots, d\}, \quad \sum_{k=0}^{d}\lambda_k k^j = 0$$

so that $\lambda_k = 0$, $0 \leq k \leq d$ since $\det[k^j]_{0\leq k,j\leq d} \neq 0$ (Vandermonde determinant). □

Proof. (of Proposition 4.4) We can extend the definition of $X_k^{i,l}$ to $l \in \mathbb{N}$. It is easy to check that

$$\begin{aligned}X_{k+1}^{i,l} &= \int_0^{(k+1)\Delta} P_i\left((k+1+l)\Delta - s\right)e^{-\alpha_i((k+1+l)\Delta-s)}dW_s^i\\ &= X_k^{i,l+1} + \varepsilon_{k+1}^{i,l}\end{aligned}$$

where $\varepsilon_{k+1}^{i,l} = \int_{k\Delta}^{(k+1)\Delta} P_i\left((k+1+l)\Delta - s\right) e^{-\alpha_i((k+1+l)\Delta - s)} dW_s^i$.

If $l = d_i$,

$$X_{k+1}^{i,d_i} = X_k^{i,d_i+1} + \varepsilon_{k+1}^{i,d_i}.$$

According to Lemma 4.5,

$$P_i\left(Z + (d_i+1)\Delta\right) = \sum_{l=0}^{d_i} \lambda^{i,l} P_i(Z + l\Delta).$$

Hence

$$\begin{aligned}
X_k^{i,d_i+1} &= \int_0^{k\Delta} P_i\left((k+d_i+1)\Delta - s\right) e^{-\alpha_i((k+d_i+1)\Delta - s)} dW_s^i \\
&= \sum_{l=0}^{d_i} \lambda^{i,l} \int_0^{k\Delta} P_i\left((k+l)\Delta - s\right) e^{-\alpha_i((k+l)\Delta - s)} dW_s^i e^{-\alpha_i(d_i+1-l)\Delta} \\
&= \sum_{l=0}^{d_i} \lambda^{i,l} e^{-\alpha_i(d_i+1-l)\Delta} X_k^{i,l} \\
&= \sum_{l=0}^{d_i} \tilde{\lambda}^{i,l} X_k^{i,l}.
\end{aligned}$$

Finally we have

$$X_{k+1}^{i\cdot} = A^i X_k^{i\cdot} + \varepsilon_{k+1}^{i\cdot}$$

where

$$A^{i\cdot} = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & 0 & 1 \\
\tilde{\lambda}^{i,0} & \cdots & \cdots & \cdots & \tilde{\lambda}^{i,d_i}
\end{pmatrix}$$

and

$$X_{k+1} = AX_k + \varepsilon_k, X_0 = 0$$

where $\varepsilon_k \in \sigma\left(W_u^i - W_{k\Delta}^i, k\Delta \le u \le (k+1)\Delta, i = 1, \ldots, m\right)$ is independent of $\mathcal{F}_{k\Delta}^W$. The process $(X_k)_k$ is thus a gaussian AR(1). □

X_k is the structure process for the spot price $S_{k\Delta} = F_{k\Delta,k\Delta}$. Its dimension is $\sum_{i=1}^m (d_i+1)$. For the two factor model, the structure process is $\mathbb{R}^2$-valued, because $m = 2$, and $d_i = 0, i = 1, 2$. This is coherent with (22).