# Multi-asset American options and parallel quantization

Anne Laure Bronstein, Gilles Pagès, Jacques Portès

# Multi-asset American options and parallel quantization

Anne Laure Bronstein* Gilles pagès† and Jacques Portès‡

September 10, 2008

**Abstract**

We present a parallel implementation of the optimal quantization method on a grid computing. Its purpose is to price instantaneously multidimensional American options. Numerical tests are proceeded with variable number of processors, from 4 to 128. Finally a spatial extrapolation of Richardson-Romberg is introduced to speed up the convergence rate and stabilize the results.

*Keywords: American options, optimal quantization, Parallel computing, Romberg extrapolation.*

## 1 Introduction

This paper is devoted to a numerical study of a probabilistic method for options pricing. We focus on multi-asset American options. We assume that the underlying assets dynamics follow a classical Black and Scholes model. Numerical methods for multidimensional American options have attracted significant interest in the literature. Important contributions include Longstaff and Schwartz [LS01], Tsitsiklis and Van Roy [TVR99], Broadie and Glasserman [BG07], Fournié al [FLLLT99], [FLLL01], Lions and Régnier [LS01].

We consider a probabilistic approach designed by Bally, Pagès and Printems, (see [BP03] and [BPP05]). This method is based on a spatial discretization of processes on optimal grids. In the fifties, optimal quantization has emerged in the fields of Signal processing and Information Theory, see Gersho and Gray [GG92] and Graf and Luschgy [GL00].

In financial institutions, quickness of execution as well as high accuracy are important criteria in the choice of a pricing method. With this observation in mind, we suggest some

*Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5. E-mail: albronstein@hotmail.com

†Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5. E-mail: gilles.pages@upmc.fr

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

improvements to the original quantization method. The quantization tree algorithm (or pricing procedure) is divided into three parts: the computation of the quantization grids, the estimation of the transition probabilities and the premium evaluation. As the first two tasks are time consuming, one usually proceeds these estimations off-line. To reduce drastically the computation time and proceed on-line, we suggest the application of a fast weight estimation method, which allows for a parallel implementation of our procedure on a grid.

The fast weight estimation method was first introduced in the paper by Bardou, Bouthemy and Pagès, (see [BBP07a]). It is based on centered Gaussian first order auto-regressive processes and yield time independent transition probabilities. The computation of these transition probabilities consists on large Monte Carlo simulations of couples of independent Gaussian random vectors. To improve the execution time of this procedure, we split into several processes the Monte Carlo simulations. The number of processes used to do the computations is function of the number of Monte Carlo simulations, the option dimension and the required accuracy. This parallel implementation reduces from several minutes to several seconds the execution time of this fast weight estimation procedure and allows for an almost instantaneously pricing algorithm.

With regard to computations accuracy, we suggest the application of a spatial Richardson-Romberg extrapolation method. This extrapolation is a powerful technique for improving the convergence rate of the algorithm. As in the American case the spatial term has a strong weight in the error bound, we consider a spatial extrapolation. Furthermore, the Richardson-Romberg extrapolation strongly improves the stability of the pricing method. We observe a decrease of the premium sensitivity with regard to the time discretization parameter, and therefore we get more accurate premium.

Section 2 is devoted to Bermuda options pricing. Section 3 recalls some basic notions on optimal quantization. In Section 4, following some results by Bally, Pagès and Printems, (see [BPP05]), some error bounds are estimated. In Section 5, we present a numerical study of American exchange options with a parallel implementation of the pricing algorithm on the CCR grid of Jussieu. A spatial Richardson-Romberg extrapolation method is introduced into the pricing procedure.

# 2 Bermuda option: an optimal stopping problem

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a given probability space and let define on this probability space a Markov structure process $(X_k)_{0\leq k\leq n} = (X_k^1, \ldots, X_k^d)_{0\leq k\leq n}$. This process will denote the discrete price of $d$-traded assets at time $t_k$, $k = 0, \ldots, n$ on a given market. Let $(\mathcal{F}_k)_{0\leq k\leq n}$ be the natural completed filtration generated by $(X_k)_{0\leq k\leq n}$ on $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that $\mathbb{P}$ is the risk-neutral probability. The holder of a Bermuda option has the right to receive at a time $t_k \in \{t_0, \ldots, t_n\}$ a flow of payoff $h(t_k, X_k)$. The Bermuda option pricing problem is associated to the following optimal stopping problem

$$V_k := \operatorname{ess\,sup} \left\{ \mathbb{E}\left(h(\tau, X_\tau)|\mathcal{F}_k\right), \tau \ \{t_k, \ldots, t_n\}\text{-valued stopping times} \right\},$$

where $V_k$ is identified to the option price at time $t_k$. To solve this problem, we rely on the dynamic programming formula:

$$\begin{aligned} V_n &:= h(t_n, X_n), \\ V_k &:= \max\{h(t_k, X_k), \mathbb{E}(V_{k+1}|X_k)\},\ k = 0, \dots, n-1. \end{aligned} \tag{1}$$

Observation of this formula leads to the following question: how to approximate the conditional expectations $\mathbb{E}(V_{k+1}|X_k)$ at the exercise dates $t_k$, $k = 0, \dots, n$. This paper is based on a spatial discretization method of $X$ initiated by Bally, Pagès and Printems, (see [BPP05]).

# 3 Quantization of the underlying structure process

## 3.1 Optimal quantization

Given $k = 0, \dots, n$, $X_k$ is approximated by a random vector $\hat{X}_k$ taking valued in a finite subspace of $\mathbb{R}^d$, named quantization grid, i.e.

$$\hat{X}_k = q_k(X_k), \quad q_k : \mathbb{R}^d \to \Gamma_k := \{x_1^k, \dots, x_{N_k}^k\}.$$

Let $x^k = (x_1^k, \dots, x_{N_k}^k)$ denotes the $N_k$-tuple induced by $\Gamma_k$.

Let $X_k \in L^p(\Omega, \mathcal{A}, \mathbb{P})$, $p \in [1, \infty[$. The $L^p$-mean error induced by replacing $X_k$ by $\hat{X}_k$ is called the $L^p$-mean quantization error and is given by $|| X_k - q_k(X_k)||_p$. To optimize this spatial discretization method, one hopes to minimize in $q_k$ and in $x$ the quantization error. Recall that a partition $(C_i(x^k))_{1 \le i \le N_k}$ of $\mathbb{R}^d$ is a Voronoi tessellation of the $N_k$-quantizer $x^k$, if for every $i \in \{1, \dots, N_k\}$, $C_i(x^k)$ is a Borel set satisfying

$$C_i(x^k) \subset \{u \in \mathbb{R}^d \mid |u - x_i^k| = \min_{1 \le j \le N_k} |u - x_j^k|\}.$$

Then, the solution of this minimization problem is given by the nearest neighbor projection induced by the Voronoi tessellation $(C_i(x^k))_{1 \le i \le N_k}$, on an optimal grid. Several procedures are available to design an optimal grid, e.g. the Competitive Learning Vector Quantization algorithm, the Lloyd's procedure. For a detailed version of this method, see Pagès and Printems [PP03].

## 3.2 The quantization tree

This spatial dicretization is applied to the backward dynamic programming formula (1). However it results a loss of the Markov property for $\hat{X}_k$. We force the Markov property and define by induction the following quantized dynamic programming formula given by

$$\begin{aligned} \hat{V}_n &:= h(t_n, \hat{X}_n), \\ \hat{V}_k &:= \max\{h(t_k, \hat{X}_k), \mathbb{E}(\hat{V}_{k+1}|\hat{X}_k)\},\ k = 0, \dots, n-1, \end{aligned}$$

and the quantization tree algorithm given by

$$\hat{v}_n(x_i^n) := h_n(x_i^n),\ i = 1, \dots, N_n,$$
$$\hat{v}_k(x_i^n) := \max\{h_k(x_i^k), \sum_{j=1}^{N_{k+1}} \hat{v}_{k+1}(x_j^{k+1})\pi_{ij}^k\},\ i = 1, \dots, N_k, k = 0, \dots, n-1, \tag{2}$$

where

$$\pi_{ij}^k = \mathbb{P}(\hat{X}_{k+1} = x_j^{k+1} | \hat{X}_k = x_i^k) = \mathbb{P}(X_{k+1} \in C_j(x^{k+1}) | X_k \in C_i(x^k)). \tag{3}$$

### 3.3 Complexity

Operation cost for a quantization tree descent is proportional to

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

## 4 From Bermuda option to American option

Set $t_k = \frac{kT}{n}$, $k = 0, \dots, n$ and let $n$ goes to $\infty$. By taking this limit, we get an American option analogue to the Bermuda one. Now, the derivative holder has the right to receive once at a time $t \in [0, T]$ a flow of payoff $h(t, X_t)$. Also, the American option price is identified to the solution of a continuous optimal stopping problem given by

$$\vartheta_t := \operatorname{ess\,sup} \{\mathbb{E}\left(h(\tau, X_\tau)|\mathcal{F}_t\right), \tau \text{ [t,T]-valued stopping times}\}.$$

Hence, it seems natural to approximate the price of an American option by the price of its Bermuda counterpart. A premium error bound follows from theorems 2.1 and 2.2 in Bally, Pagès and Printems, (see [BPP05]) and is given for semi-convex payoff by

$$\|\vartheta_0 - \hat{V}_0\|_p \leq \frac{C_1}{n} + C_2 \sum_{k=0}^{n} \|X_k - \hat{X}_k\|_p,$$

for $C_1$ and $C_2$ positive constants.

Combining Zador's Theorem [Z82] with an optimal dispatching rule of the elementary quantizers among $n$ time discretization steps, i.e., for $n \geq 1$ and $N \geq n + 1$, we have

$$N_k := \left\lceil \frac{t_k^{\frac{d}{2(d+1)}}(N-1)}{t_1^{\frac{d}{2(d+1)}} + \ldots + t_n^{\frac{d}{2(d+1)}}} \right\rceil, \tag{4}$$

where $\lceil x \rceil := \min\{k \in \mathbb{N} : k \geq x\}$, yields an a priori error bound in time and space for semi-convex payoff given by

$$\|\vartheta_{t_k} - \hat{V}_k(\hat{X}_k)\|_p \leq C_p e^{C_p T} \left( \frac{1 + |s_0|}{n} + \frac{n^{1+1/d}}{N^{1/d}} \right).$$

# 5 Numerical study of American exchange option premium

In this Section, we proceed to numerical tests on American exchange options.

## 5.1 The framework

We consider a $d$-dimensional Black-Scholes model for the underlying assets dynamics given by

$$dS_t^l = (r - \mu_l)S_t^l dt + \sigma_l S_t^l dW_t^l, \quad t \in [0,T],\, l = 1, \ldots, d,$$

where $(W_t)_{0 \le t \le T}$ is a $d$-dimensional standard Brownian motion.

If the derivative holder exercises its contract, he'll receive a reward based on the value of the underlying assets at the exercise date and defined by

$$h(y) := \max(y^1 \cdot \ldots \cdot y^p - y^{p+1} \cdot \ldots \cdot y^{2p}, 0), \quad \text{with } d := 2p.$$

To simplify the computations, w.l.o.g. we set the interest rate to 0. (Indeed $h$ doesn't depend on $r$).

### 5.1.1 A more accurate computation: the introduction of a European control variable

To reduce the price estimation variance, we introduce a sequence of control variate variables in the quantization tree algorithm, (see [BPP05], Section 5). In the exchange case, the variable considered is the European exchange option with similar maturity as the American one. The European premium has a closed form solution given by

$$\begin{aligned} \mathrm{Ex}_{\mathrm{BS}}(\Theta, y, y', \tilde{\sigma}, \mu) &:= \mathrm{erf}(d_1)\exp(-\mu\Theta)y - \mathrm{erf}(d_1 - \tilde{\sigma}\sqrt{\Theta})y', \\ d_1(y, y', \tilde{\sigma}, \Theta, \mu) &:= \frac{\log(y/y') + (\tilde{\sigma}^2/2 - \mu)\Theta}{\tilde{\sigma}\sqrt{\Theta}} \\ \text{and} \quad \mathrm{erf}(y) &= \int_{-\infty}^{y} e^{-u^2/2} du/\sqrt{2\pi}, \end{aligned}$$

with

$$\Theta := T - t,\ \tilde{\sigma} := \left(\sum_{l=1}^{d} \sigma_l^2\right)^{1/2},\ \mu := \sum_{l=1}^{p} \mu_l - \sum_{l=p+1}^{d} \mu_l,\ y := \prod_{l=1}^{p} S_t^l,\ y' := \prod_{l=p+1}^{d} S_t^l.$$

### 5.1.2 Exchange option parameters

We focus on American exchange options in dimensions 2, 4 and 6 in and out of the money. The model parameters are the following:

- a maturity $T$ of one year,
- a dividend rate $\mu$ of 5%,
- a volatility $\sigma$ of 20%.

## 5.2 A new quantization tree

With the introduction of a sequence of European control variables, the quantization tree given by (2) turns into the following pricing algorithm.

$$\begin{aligned} v_i^{h,n} &:= h_i^n \quad i = 1, \dots, N_n, \\ v_i^{h,k} &:= M_i^k + \max\left\{ (h_i^k - M_i^k), \sum_{j=1}^{N_{k+1}} \pi_{ij}^k v_j^{(h-M),k+1} \right\}, \quad i = 1, \dots, N_k, k = 0, \dots, n-1. \end{aligned} \tag{5}$$

where the obstacle is given by

$$h_i^k := h(s_i^{k,1}, \dots, s_i^{k,d}),$$

the European control variable by

$$M_i^k := \mathrm{Ex}_{\mathrm{BS}}(T - t_k, \prod_{l=1}^{p} s_i^{k,l}, \prod_{l=p+1}^{d} s_i^{k,l}, \tilde{\sigma}, \mu),$$

with

$$s_i^{k,l} := s_0^l \exp\left( -(\mu_l + \frac{\sigma_l^2}{2}) k\Delta t + \sigma_l x_i^{k,l} \right), \quad l = 1, \dots, d,$$

and the transition probabilities by

$$\pi_{ij}^k = \mathbb{P}(\hat{X}_{k+1} = x_j^{k+1} | \, \hat{X}_k = x_i^k). \tag{6}$$

Here $x^k$ is an optimal $N_k$-quantizer of $X_k$.

### 5.2.1 An algorithm in three steps

To compute the quantization tree algorithm, one proceeds step by step.

- Step 1: the computation of optimal $N_k$-quantizer of $S_{t_k}$.

  In the Black-Scholes model, the asset price can be seen as function of the Brownian motion, i.e. $S_{t_k} = \psi(t_k, X_k)$, with $X_k := W_{t_k}$. Thus, optimal quantizers of $S_{t_k}$ can be computed as function of optimal quantizers of $W_{t_k}$. Recalling Section 3.1, there exist

several algorithms to compute the $L^p$-optimal $N_k$-quantizer $x^k$ of $W_{t_k}$. However, one notices that it also could be obtained by a dilatation of the optimal $N_k$-quantizer $\bar{x}^k$ of the normal distribution. That is for $k \in \{0, \ldots, n\}$,

$$x^k = \sqrt{t_k}\bar{x}^k$$

where $\bar{x}^k$ is already known. Indeed, quantizers of the normal law have already been computed for several dimensions and several grids sizes and are downlable at the url:

http://quantification.finance-mathematique.com

In view of these observations, one could execute step one instantaneously.

- Step 2: the computation of the transition probabilities (6).

  To proceed this step, one appeals to Monte Carlo simulations. The execution time depends on the option dimension and variates between 15 minutes and several hours. Computations are usually proceeded off-line during the night.

- Step 3: the option premium computation.

  Once the two first steps are executed, the algorithm execution time is very quick and we instantaneously get option prices.

The objective of this numerical work is to study several methods to improve drastically the execution time of step 2.

### 5.2.2 The transition probabilities computation

- The diffusion method: accurate but too long.

  We simulate some standard Brownian motion trajectories from $t_0 = 0$ to the maturity of the exchange option $t_n = T$. This simulation is based on the independence and stationary properties of the Brownian motion increments. Indeed, one will notice that the law of the family $(W_{t_{k+1}} - W_{t_k})_{0 \leq k \leq n-1}$ , with $t_k = \frac{kT}{n}, k = 0, \ldots, n$, is similar to a family of i.i.d. random vectors with $\mathcal{N}(0, \frac{T}{n} I_d)$ distribution. Therefore for $k \in \{1, \ldots, n\}$, $W_{t_k}$ is simulated by

$$\sqrt{\frac{T}{n}} \sum_{i=0}^{k-1} \epsilon_{i+1}, \tag{7}$$

  where $\epsilon_i$, $i \in \{1, \ldots, k\}$ are i.i.d random variables, with normal distribution. The Monte Carlo proxies of the theorical transitions (6) are then simulated for $k \in \{0, \ldots, n-1\}$, $i \in \{1, \ldots, N_k\}$ and $j \in \{1, \ldots, N_{k+1}\}$ by

$$\tilde{\pi}_{ij}^k = \frac{\frac{1}{M} \sum_{m=1}^{M} 1_{C_j(x^{k+1})}(\tilde{W}_{t_{k+1}}^m) 1_{C_i(x^k)}(\tilde{W}_{t_k}^m)}{\frac{1}{M} \sum_{m=1}^{M} 1_{C_i(x^k)}(\tilde{W}_{t_k}^m)} \tag{8}$$

and by

$$\tilde{\pi}^0_{1j} = \frac{1}{M}\sum_{m=1}^{M} 1_{C_j(x^1)}(\tilde{W}^m_{t_1}), \tag{9}$$

where for $k \in \{0,\dots,n\}$, $(\tilde{W}^m_{t_k})_{1\le m\le M}$ are $M$ copies of the random vector $W_{t_k}$ given by (7). This method is very consistent as the quantization tree is built step by step and gives very accurate results. However, this advantage is also a problem as it implies lots of simulations and doesn't allow a split of the transition probabilities computations since the dynamics plays an important role here.

- The fast weight estimation method (FWE): accurate and quick!

This method, introduced in the paper of Bardou, Bouthemy and Pagès, (see [BBP07a]) appeals to centered Gaussian first order auto-regressive processes. In order to apply this method to the American exchange option, we consider the centered Gaussian first order auto-regressive Brownian motion process in $\mathbb{R}^d$ given by $W_{t_{k+1}} = W_{t_k} + \sqrt{\frac{T}{n}}\epsilon_{k+1}$, for $k \in \{0,\dots,n-1\}$ where $(\epsilon_k)_{1\le k\le n}$ are i.i.d random vectors with $\mathcal{N}(0, I_d)$ distribution.

Let $W$ be the auto-regressive process described above and let $(\eta_1, \eta_2)$ be a couple of independent random vectors normally distributed. Then, the transition probabilities (6) satisfy for $k \in \{1,\dots,n-1\}$, $i \in \{1,\dots,N_k\}$ and $j \in \{1,\dots,N_{k+1}\}$

$$\pi^k_{ij} = \frac{\mathbb{P}(\alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(\bar{x}^{k+1}), \eta_1 \in C_i(\bar{x}^k))}{\mathbb{P}(\eta_1 \in C_i(\bar{x}^k))} \tag{10}$$

and for $j \in \{1,\dots,N_1\}$

$$\pi^0_{1j} = \mathbb{P}(\eta_2 \in C_j(\bar{x}^1)), \tag{11}$$

where for $k \in \{0,\dots,n-1\}$,

$$\alpha_{k+1} = \sqrt{\frac{k}{k+1}} \quad \text{and} \quad \beta_{k+1} = \frac{1}{\sqrt{k+1}},$$

and $\bar{x}^k$ is the $L^p$-optimal $N_k$-quantizer of the normal distribution. Therefore each transition probability can be computed independently and all the computations are executable separately with an unique sample of $(\eta_1, \eta_2)$. However, to keep a certain consistency in the tree foundation, we simulate at each time step $t_k$ a sample $(\tilde{\eta}^m_1, \tilde{\eta}^m_2)_{1\le m\le M}$ and we compute the transition probabilities estimations for $i \in \{1,\dots,N_k\}$ and $j \in \{1,\dots,N_{k+1}\}$ by

$$\tilde{\pi}^k_{ij} = \frac{\frac{1}{M}\sum_{m=1}^{M} 1_{C_j(\bar{x}^{k+1})}(\alpha_{k+1}\tilde{\eta}^m_1 + \beta_{k+1}\tilde{\eta}^m_2)1_{C_i(\bar{x}^k)}(\tilde{\eta}^m_1)}{\frac{1}{M}\sum_{m=1}^{M} 1_{C_i(\bar{x}^k)}(\tilde{\eta}^m_1)}.$$

Since these computations are time independent, the computations $(\pi^0_{ij})_{i,j}$, $(\pi^1_{ij})_{i,j}$,... $(\pi^{n-1}_{ij})_{i,j}$ could be done simultaneously on different processors.

| | $d=2$ | $d=4$ | $d=6$ |
|---|---|---|---|
| $n$ | 25 | 11 | 9 |
| $\bar{N}$ | 335 | 750 | 1000 |

Table 1: Quantization parameters.

## 5.3 A parallel implementation of the transition probabilities: an incredible time reward!

We will study a parallel implementation of the fast weight estimation method on the CCR grid of Jussieu.

### 5.3.1 Parallel computing

In computer science, parallel computing consists in the partitioning of a computation in elementary tasks that will be simultaneously proceeded by several processors. This method is used to speed up algorithm execution time and to increase the number of tasks executed. Today's computer have between one and four processors. To optimize computing resources, numerous processors are coupled together through a support named grid computing. These grids offer appropriate support to do parallel computing. Computations are proceeded on these grids through a Message Passing Interface (MPI). This interface defines a library of functions used to send messages on the grids. These functions are compatible with the C language.

### 5.3.2 From 14 minutes to 14 seconds

We have proceeded several tests on the American exchange options described above in dimensions 2, 4 and 6. The model parameters are similar as in the Section 5.1.2 for maturity, drift and volatility and we consider $S_0$ =(40,36). The quantization parameters (number of layers $n$, average number of points per layer $\bar{N} := N/n$), are given by Table 1 and satisfy the optimal dispatching rule (4). The reference price amounts to 5.6468 and was computed by a two dimensional finite difference algorithm devised by Villeneuve et Zanette [VZ02]. The control variate variable (European Black and Scholes premium) is equal to 5.2674. These tests have been done on the CCR grid of Jussieu. First, following the FWE method, we implement a sequential procedure to estimate the transition probabilities (10)-(11) and insert these results in the quantization tree algorithm to check the quantized premium accuracy. Then, following a similar procedure with exactly the same parameters, we split the Monte Carlo simulations between several processes. For $M$ Monte Carlo simulations and $x$ processes, denoted between 0 and $x-1$, each process, (from 1 to $x-1$), receive $\lceil \frac{M}{x-1} \rceil$ Monte Carlo simulations to execute. MPI functions are used to send the data and receive the results. The process 0 plays the role of an orchestra conductor. It coordinates these transfers and proceeds to final computation.

An incredible result in dimension 2: if one wishes to get a very accurate estimation of the transition probabilities, one could execute 50 million trials on 128 processors in 14 seconds. A similar procedure, with a sequential implementation lasts 14 minutes and 18 seconds. I.e., the parallel procedure is 61 times quicker than the sequential one. The quantized premium is equal to 5,6484, i.e., we have an error of 0,026% with regard to the reference price of 5,6468.

However, accurate estimations are obtained with fewer trials. Several tests have been done for American exchange in dimension 6. For example, transition probabilities were estimated with 900 000 trials split on 8 processors (2 computers). Parallel computing lasts 8 seconds, whereas the sequential one lasts 3 minutes 35 seconds. I.e., parallel computing is almost 27 times quicker than sequential computing in this case. The quantized premium is here equal to 5,6634, i.e., we have an error of 0,29% in dimension 6.

To get an understanding of the grid capacity, several tests have been done for 32 processors and 900 000 Monte Carlo trials. In dimension 2, for an accuracy of 0,026% the parallel procedure lasts 6 seconds, whereas the sequential one lasts 44 seconds. I.e., the parallel procedure is 7 times quicker than the sequential one. In dimension 4, for an accuracy of 0,15% the parallel procedure lasts 10 seconds whereas the sequential one 1 minute 37 seconds. So here, the parallel procedure is almost 10 times quicker than the sequential one. In dimension 6, for an accuracy of 0,18% we have an execution time of 11 seconds in parallel computing and 3 minutes 35 seconds in the sequential one. That is, parallel computing is 20 times quicker than sequential computing.

**Remark 1** *Here, as the number of processors is large and the number of Monte Carlo low, sending message and final computation take longer with 32 processors than with 8 processors in dimension 6.*

## 5.4 Pricing accuracy

Recalling Section 4 , one observes, for semi convex payoffs, an error bound given by

$$|\, p(\bar{N}, n) - p_{VZ}| \leq \frac{C_1}{n} + C_2 \frac{n}{\bar{N}^{\frac{1}{d}}}. \tag{12}$$

Here $p(\bar{N}, n)$ represents the quantized premium of the American exchange option computed with $n$ time steps and $\bar{N} := \frac{N}{n}$ average elementary quantizers per layer. And $p_{VZ}$ is the Villeneuve and Zanette reference premium. Several tests are executed to observe the dependency of the error bound in $N$ and $n$. first, we note the ascendency of the spatial term on the temporal one. Then, we observe a convergence rate closer to $2/d$ in the spatial term which is stronger than the theoretical rate of $1/d$. These tests suggest that heuristically we have

$$|\, p(\bar{N}, n) - p_{VZ}| \sim \frac{C_1}{n} + C_2 \frac{n}{\bar{N}^{\frac{2}{d}}}. \tag{13}$$

In view of these observations, a spatial Richardson-Romberg extrapolation is applied on quantized results.

## 5.5 A more stable method: a spatial Richardson-Romberg extrapolation

### 5.5.1 A spatial Richardson-Romberg extrapolation

Let $F : \mathbb{R}^d \to \mathbb{R}$ be a 2-times differentiable functional with Lipschitz Hessian $D^2F$. Let $(\hat{X}^{(N)})_{N\geq 1}$ be a sequence of quatratic optimal quantizations. Then a Taylor expansion yields

$$\mathbb{E}(F(X)) = \mathbb{E}(F(\hat{X}^{(N)})) + \frac{1}{2}\mathbb{E}\left(D^2F(\hat{X}^{(N)})(X-\hat{X}^{(N)})^{\otimes 2}\right) + \mathcal{O}\left(\mathbb{E}(|X-\hat{X}^{(N)}|^3)\right).$$

We can see in [GLP06] that for the normal distribution we have,

$$\mathbb{E}\left(|X-\hat{X}^{(N)}|^3\right) = \mathcal{O}(N^{-\frac{3-\epsilon}{d}}), \epsilon > 0,$$

if we suppose that

$$\mathbb{E}(D^2F(\hat{X}^{(N)})(X-\hat{X}^{(N)})^{\otimes 2}) = \frac{c_{F,X}}{N^{\frac{2}{d}}} + \mathcal{O}\left(\mathbb{E}(|X-\hat{X}^{(N)}|^3)\right),$$

then,

$$\mathbb{E}(F(X)) = \mathbb{E}(F(\hat{X}^{(N)})) + \frac{c_{F,X}}{2N^{\frac{2}{d}}} + \mathcal{O}(N^{-\frac{3-\epsilon}{d}}).$$

So, one can appeal to a spatial Richardson-Romberg extrapolation to compute $\mathbb{E}(F(X))$. Let $\bar{N}_1$ and $\bar{N}_2$ be two optimal quantizer sizes, then we have

$$\mathbb{E}(F(X)) = \frac{\bar{N}_2^{\frac{2}{d}}\,\mathbb{E}(F(\hat{X}^{(N_2)})) - \bar{N}_1^{\frac{2}{d}}\,\mathbb{E}(F(\hat{X}^{(N_1)}))}{\bar{N}_2^{\frac{2}{d}} - \bar{N}_1^{\frac{2}{d}}} + \mathcal{O}(N^{-\frac{3-\epsilon}{d}}).$$

### 5.5.2 How to proceed?

One will compute two quantized prices following the quantization tree algorithm given by (5), that we denote by $p(\bar{N}_1, n)$ and $p(\bar{N}_2, n)$. The model parameters are similar as in the Section 5.3.2 for maturity, drift, volatility and initial condition. The average number of points per layer, $\bar{N}$, is fixed and the number of layers, $n$, is variable (see Table 2). The Villeneuve and Zanette reference price amounts to 5.6468 and the control variate variable (European Black and Scholes premium) is equal to 5.2674.

Then applying Section 5.5.1, an estimation of the American exchange premium is given by

$$p_{VZ} \sim \frac{\bar{N}_2^{\frac{2}{d}} p(\bar{N}_2, n) - \bar{N}_1^{\frac{2}{d}} p(\bar{N}_1, n)}{\bar{N}_2^{2/d} - \bar{N}_1^{2/d}}. \tag{14}$$

One will observe that this extrapolation strongly stabilizes the premium estimations and speed up the convergence rate. An additional attribute of Richardson-Romberg method is

| | $d$ =2 | $d$ =4 | $d$ =6 |
|---|---|---|---|
| $\bar{N}_1$ | 100 | 750 | 1000 |
| $\bar{N}_2$ | 400 | 1000 | 1500 |
| $n$ | 5 to 65 | 4 to 32 | 4 to 16 |

Table 2: Quantization parameters.

| $S_0$ | Ref VZ | Diff | D Error | FWE | FWE Error |
|---|---|---|---|---|---|
| (80,40) | 40 | 40 | $< 10^{-4}\%$ | 40 | $< 10^{-4}\%$ |
| (60,40) | 20 | 20 | $< 10^{-4}\%$ | 20 | $< 10^{-4}\%$ |
| (44,40) | 5.9822 | 5.9827 | 0,008358% | 5.9813 | 0,01504 % |
| (36,40) | 1.9969 | 1.9969 | $< 10^{-4}\%$ | 1.9965 | 0,02003% |
| (40,44) | 2.3364 | 2.3375 | 0.04708% | 2.3370 | 0.02568% |
| (40,60) | 0.31339 | 0.31276 | 0.2010% | 0.31284 | 0.1755% |
| (40,80) | 0.021208 | 0.021064 | 0.6790% | 0.021076 | 0.6224% |

Table 3: American exchange option premiums in dimension 2: $\sigma$ =20, $\mu$ =5 and $T$ =1. All the prices have been computed with the same transitions given by (10)-(11).

the following: we observe in Figures 1-4, for $\bar{N}_1$ and $\bar{N}_2$ fixed, a translation to the right of the optimal number of time steps which should be used in the quantization tree. This feature has little impact in low dimension, but becomes crucial for high dimensions. Since at least 6 or 8 time steps are necessary to get accurate results and since the optimal number of time steps is a decreasing function of the dimension.

# 6 Conclusion

- Strength of the quantization method : once the transition probabilities computed, many options can be price simultaneously and instantaneously with these data.

- A drawback of this method was the off-line computation of the transition probabilities. Parallel implementation brings an answer to this problem and allow an on-line computation of these estimations.

- Then several American options with different structures and different parameters, (e.g. volatility, dividend rate, initial condition), can be prices on an excel file simultaneously and instantaneously, (see Table 3).

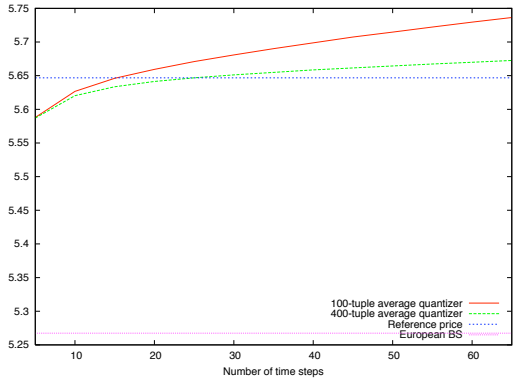


Figure 1: Quantized premium for $d$ =2.

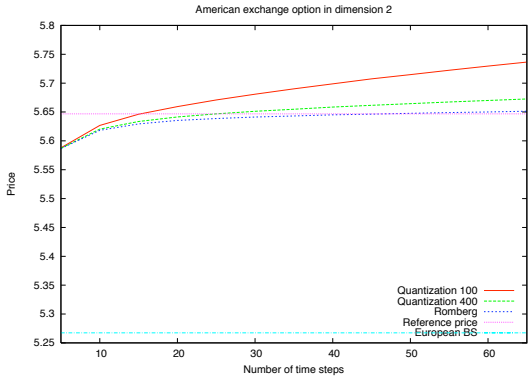


Figure 2: Richardson-Romberg for $d$ =2.

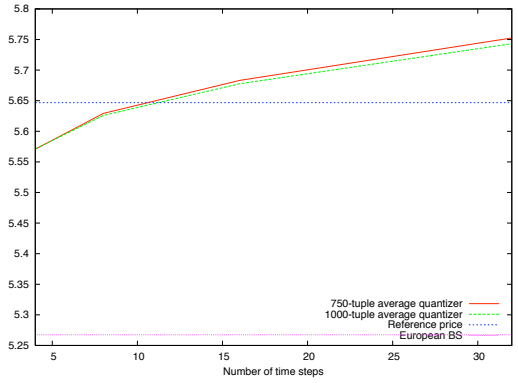


Figure 3: Quantized premium for $d$ =4.

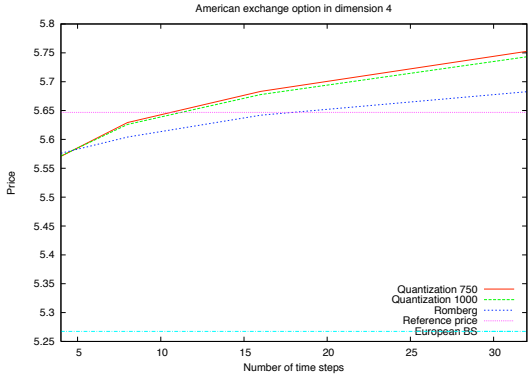


Figure 4: Richardson-Romberg for $d$ =4.

# References


[BP03] Bally V. and Pagès G. (2003), A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems, Bernoulli, vol **9**, $n^0$ 6, pp. 1003-1049.

[BPP05] Bally V., Pagès G. and Printems J. (2005), A quantization tree method for pricing and hedging multi-dimensional American options, Mathematical Finance, vol **15**, $n^0$ 1, pp. 119-168.

[BBP07a] Bardou O., Bouthemy S. and Pagès G., (2007), Optimal quantization for the pricing of swing options, to appear in Applied Mathematical Finance.

[BG07] Broadie M., and Glasserman P., (1997), Pricing American-style securities using simulation, J. Econ. Dynam. Control, vol **21**, $n^0$ 8-9, pp. 1323-1352.

[FLLL01] Fournié É., Lasry J.M., Lebouchoux J. and Lions P.L, (2001), Applications of Malliavin calculus to Monte Carlo methods in Finance II, Finance Stoch., vol **5**, pp. 201-236.

[FLLLT99] Fournié É., Lasry J.M., Lebouchoux J., Lions P.L and Touzi N, (1999), Applications of Malliavin calculus to Monte Carlo methods in Finance, Finance Stoch., vol **3**, pp. 391-412.

[GG92] Gersho A. and Gray R.M. (1992), Vector quantization and signal compression, Kluwer, Boston.

[GL00] Graf S. and Luschgy H. (2000), Foundations of quantization for probability distributions, Lecture Notes in Mathematics $n^0$1730, Springer, Berlin.

[GLP06] Graf S. and Luschgy H. and Pagès G. (2008), Distortion mismatch in the quantization of probability measures, ESAIM P&S, vol **12**, pp. 127-153 .

[K82] Kiefer J. (1982), Exponential rate of convergence for the Lloyd's method I, IEEE Trans. Info. Theory, Special issue on Quantization, vol **28**, $n^0$ 2 pp. 205-210.

[LS01] Lions P.L. and Régnier H. (2001), Calcul des prix et des sensibilités d'une option Américaine par une méthode de Monte Carlo, working paper coc ixis.

[LS01] Longstaff F.A. and Schwartz E.S. (2001), Valuing American options by simulation: a simple least-squares approach, Rev. Financial Stud., vol **14**, pp. 113-148.

[P97] Pagès G., (1997) A space vector quantization method for numerical integration, Journal of Applied and Computational Mathematics, vol **89**, pp. 1-38.

[PP03] Pagès G. and Printems J., (2003) Optimal quadratic quantization for numerics: the Gaussian case, Monte Carlo Methods and Applications, vol **9**, $n^0$ 2 pp. 135-166.

[TVR99] Tsitsiklis J.N. and Van Roy B. (1999), Optimal stopping of markov processes: Hilbert space theory, approximation, algorithms, and an application to pricing high-dimentional financial derivatives, IEEE Trans. Automat. Cont., vol **44**, $n^0$ 10, pp. 1840-1851.

[VZ02] Villeneuve S. and Zanette A. (2002), Parabolic A.D.I. methods for pricing American option on two stocks, Math. Oper. Res., vol **27**, $n^0$ 1, pp. 121-149.

[Z82] Zador P.L. (1982), Asymptotic quantization error of continuous signals and the quantization dimension, IEEE Trans. Info. Theory, Special issue on Quantization, vol **28**, $n^0$ 2, pp. 139-149.