Examples of applications of Optimal Quantization

Gilles Pagès

gpa@ccr.jussieu.fr

www.proba.jussieu.fr/pageperso/pages

Univ. PARIS 6 (Labo. Proba. et Modèles Aléatoires, UMR 7599)

American and Bermuda options

Vienna 17.10.2008

1 Introduction to optimal quadratic Vector Quantization ?

1.1 What is (quadratic) Vector Quantization ?

$\triangleright$ Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (\mathbb{R}^d, \mathcal{R}^{\otimes d})$, $|\,.\,|$ Euclidean norm,

$$\mathbb{E}|X|^2 < +\infty.$$

$\triangleright$ When

$\mathbb{R}^d \leftarrow (H, <.|.>)$ separable Hilbert space $\equiv$ Functional Quantization.

Example : If $H = L^2_T := L^2([0,T], dt)$ a process $X = (X_t)_{t \in [0,T]}$.

Discretization of the state/path space $H = \mathbb{R}^d$ or $L^2([0,T], dt)$

using

▷ N-quantizer (or N-codebook) :

$$\Gamma := \{x^1, \ldots, x^N\} \subset \mathbb{R}^d.$$

▷ Discretization by Γ-quantization

$$X \rightsquigarrow \widehat{X}^{\Gamma} : \Omega \to \Gamma := \{x^1, \ldots, x^N\}.$$

$$\widehat{X}^{\Gamma} := \mathrm{Proj}_{\Gamma}(X)$$

where

Proj_{Γ} denotes the projection on Γ following the nearest neighbour rule.

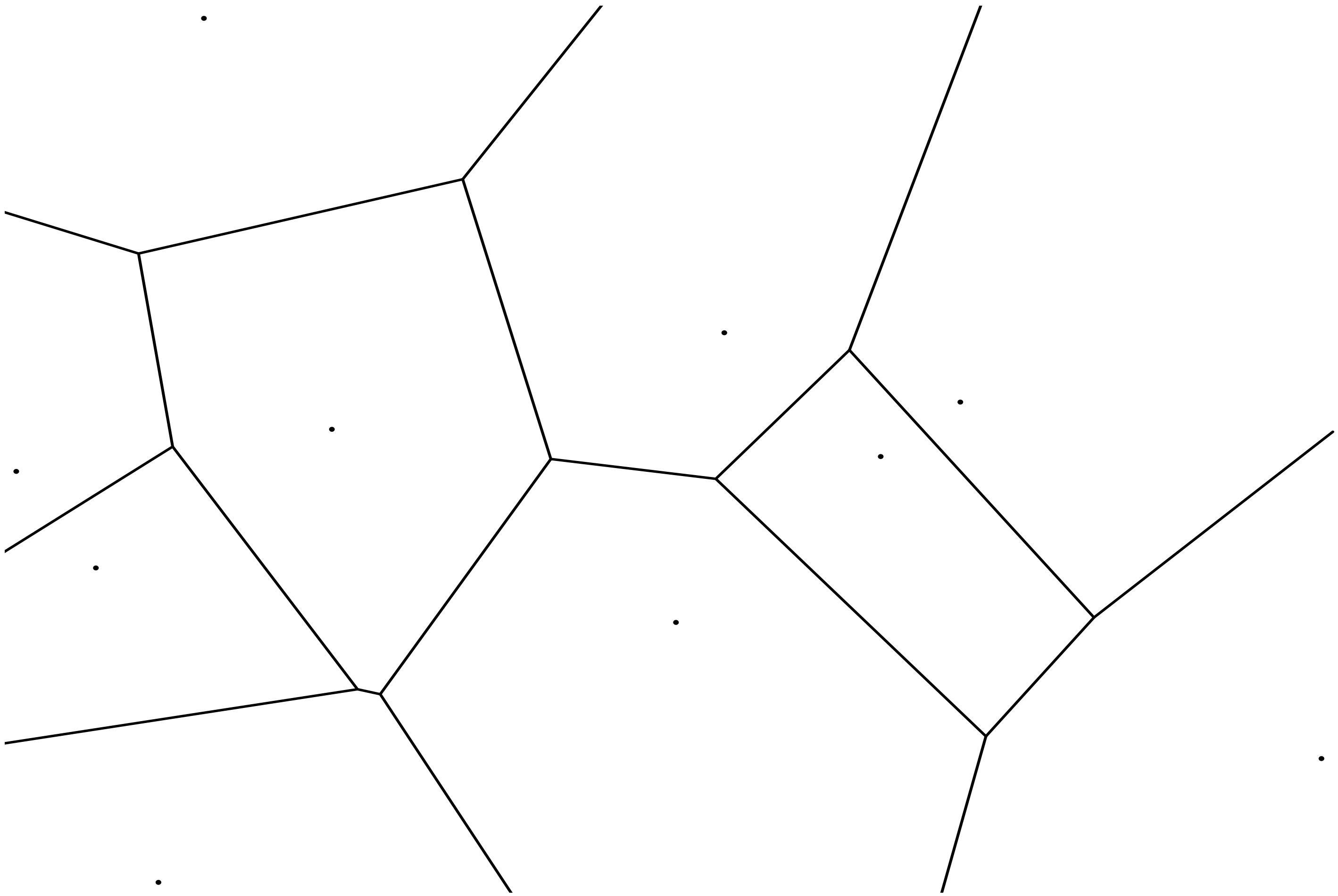


FIG. 1: A 2-dimensional 10-quantizer $\Gamma = \{x^1, \ldots, x^{10}\}$ and its Voronoi diagram. . .

1.2 What do we know about $X - \widehat{X}^{\Gamma}$ and $\widehat{X}^{\Gamma}$?

▷ Pointwise induced error : for every $\omega \in \Omega$,

$$|X(\omega) - \widehat{X}^{\Gamma}(\omega)| = \mathrm{dist}(X(\omega), \Gamma) = \min_{1\le i\le N} |X(\omega) - x^i|.$$

▷ Mean quadratic induced error (or quadratic quantization error) :

$$e_N(X, \Gamma) := \|X - \widehat{X}^{\Gamma}\|_2 = \sqrt{\mathbb{E}\left(\min_{1\le i\le N} |X - x^i|^2\right)}.$$

▷ Distribution of $\widehat{X}^{\Gamma}$: weights associated to each x^i :

$$\mathbb{P}(\widehat{X}^{\Gamma} = x^i) = \mathbb{P}(X \in C_i(\Gamma)), \qquad i = 1, \ldots, N$$

where $C_i(\Gamma)$ denotes the Voronoi cell of x^i (w.r.t. Γ) defined by

$$C_i(\Gamma) := \left\{\xi \in \mathbb{R}^d : |\xi - x^i| = \min_{1\le j\le N} |\xi - x^j|\right\}.$$

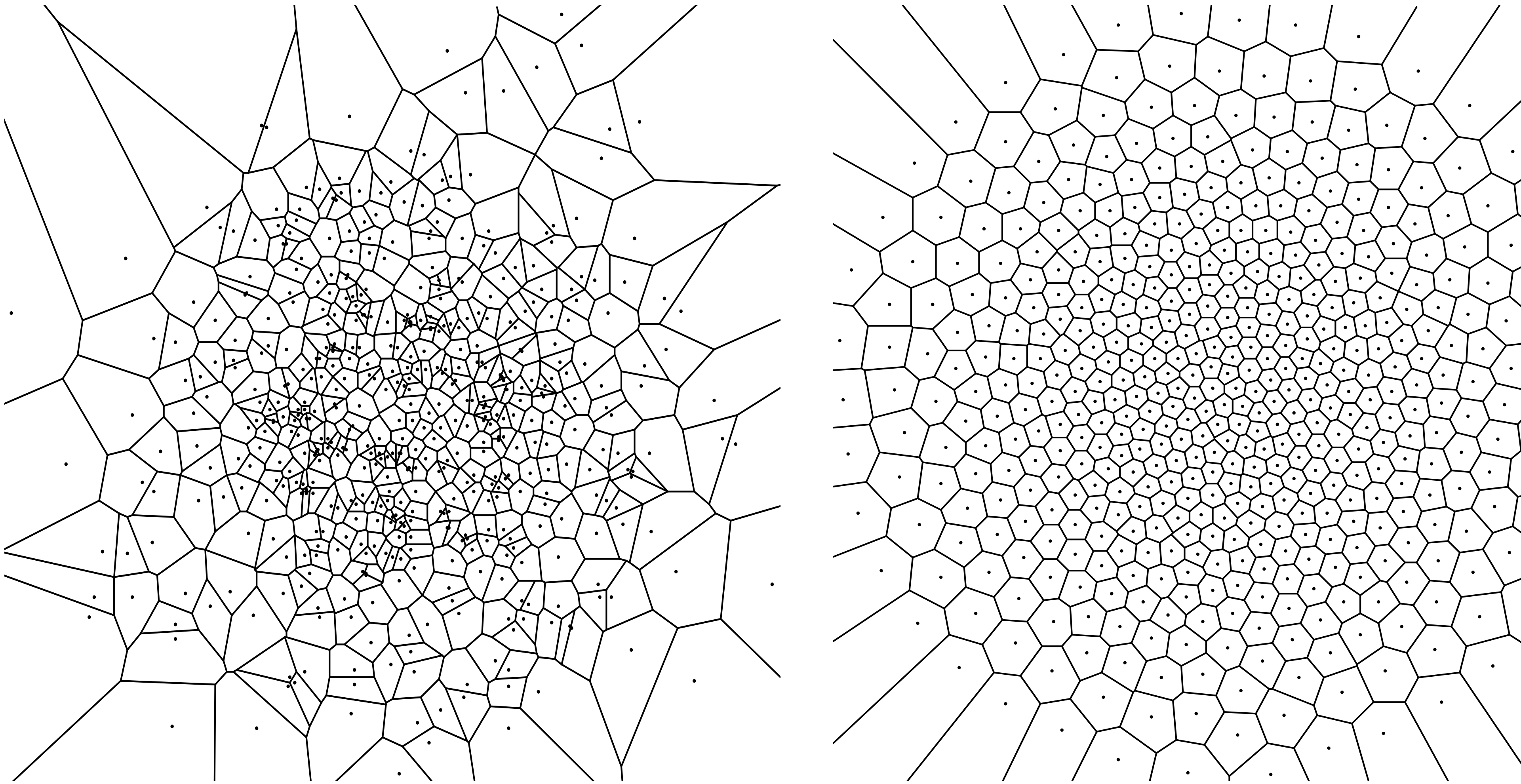


FIG. 2: Two N-quantizers related to $\mathcal{N}(0; I_2)$ of size $N = 500\ldots$

Which one is the best ?

1.3 Optimal (Quadratic) Quantization

The quadratic distortion (squared quadratic quantization error)

$$D_N^X \quad : \quad (\mathbb{R}^d)^N \longrightarrow \mathbb{R}_+$$

$$\Gamma = (x^1, \ldots, x^N) \longmapsto \|X - \widehat{X}^\Gamma\|_2^2 = \mathbb{E}\left(\min_{1\le i\le N} |X - x^i|^2\right)$$

is continuous [the quantization error is Lipschitz continuous !] for the (product topology on $(\mathbb{R}^d)^N$).

One derives (Cuesta-Albertos & Matran (88), Pärna (90), P. (93)) *by induction on* N that

D_N^X reaches a minimum at an (optimal) quantizer $\Gamma^{(N,*)}$

of full size N (if $\mathrm{card}(\mathrm{supp}(\mathbb{P})) \ge N$). One derives

$$e_N(X, \mathbb{R}^d) := \inf\{\|X - \widehat{X}^\Gamma\|_2,\ \mathrm{card}(\Gamma) \le N, \Gamma \subset H\} = \|X - \widehat{X}^{\Gamma^{(N,*)}}\|_2$$

$$\boxed{\|X - \widehat{X}^{\Gamma^{(N,*)}}\|_2 = \min\{\|X - Y\|_2,\ Y : \Omega \to H,\ \mathrm{card}(Y(\Omega)) \le N\}.}$$

Example $(N = 1)$:

Optimal 1-quantizer $\Gamma = \{\mathbb{E}\,X\}$ and $e_1(X, H) = \mathrm{stdev}(|X|_H)$.

1.4 Extensions to the $L^r(\mathbb{P})$-quantization of random variables $0 < r \le \infty$

▷ $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (\mathbb{R}^d, |\,.\,|)$

$$\mathbb{E}|X|^r < +\infty \qquad\qquad (0 < r < +\infty).$$

▷ The N-level $(L^r(\mathbb{P}), |\,.\,|)$-quantization problem for $X \in L^r_E(\mathbb{P})$

$$e_{r,N}(X, E) \;:=\; \inf\left\{\|X - \widehat{X}^{\Gamma}\|_r,\ \Gamma \subset E,\ \mathrm{card}(\Gamma) \le N\right\}.$$

Example $(N = 1,\ r = 1)$: Optimal 1-quantizer $\Gamma = \{\mathrm{med}(X)\}$ and $e_1(X, H) = \|X - \mathrm{med}(X)\|_1$.

$\triangleright$ **Other examples** :

– Non-Euclidean norms on $E = \mathbb{R}^d$ like ℓ^p-norms, $1 \leq p \leq \infty$, etc.

– dispersion of compactly supported distribution : $r = \infty$

1.5 Stationary Quantizers

▷ Distortion D_N^X is $|.|$-differentiable at N-quantizers $\Gamma \in (\mathbb{R}^d)^N$ of full size :

$$\nabla D_N^X(\Gamma) = 2\left(\int_{C_i(\Gamma)} (x^i - \xi)\mathbb{P}_X(d\xi)\right)_{1\leq i\leq N} = 2\left(\mathbb{E}(x^i - X)\mathbf{1}_{\{\widehat{X}^\Gamma = x^i\}}\right)_{1\leq i\leq N}$$

▷ **Definition :** If $\Gamma \subset (\mathbb{R}^d)^N$ is a zero of $\nabla D_N^X(\Gamma)$, then Γ is called a *stationary quantizer* (or self-consistent quantizer).

$$\boxed{\nabla D_N^X(\Gamma) = 0 \quad \Longleftrightarrow \quad \widehat{X}^\Gamma = \mathbb{E}\left(X \,|\, \widehat{X}^\Gamma\right)}$$

since

$$\sigma(\widehat{X}^\Gamma) = \sigma(\{X \in C_i(\Gamma)\}, i = 1, \ldots, N).$$

▷ An optimal quadratic quantizer Γ is stationary

First by-product :

$$\mathbb{E}X = \mathbb{E}\widehat{X}^\Gamma.$$

1.6 Numerical Integration and conditonal expectaion (I) : cubature formulae

Let $F : (\mathbb{R}^d)^N \longrightarrow \mathbb{R}$ be a functional and let $\Gamma \subset \mathbb{R}^d$ be an N-quantizer.

$$\mathbb{E}\left(F(\widehat{X}^{\Gamma})\right) = \sum_{i=1}^{N} F(x^i)\mathbb{P}(\widehat{X} = x^i)$$

▷ If F is Lipschitz continuous, then

$$\boxed{\left|\mathbb{E}F(X) - \mathbb{E}F(\widehat{X}^{\Gamma})\right| \le [F]_{\mathrm{Lip}}\|X - \widehat{X}^{\Gamma}\|_{1} \le [F]_{\mathrm{Lip}}\|X - \widehat{X}^{\Gamma}\|_{2}}$$

in fact

$$\|X - \widehat{X}^{\Gamma}\|_{1} = \sup_{[F]_{\mathrm{Lip}} \le 1} \left|\mathbb{E}\,F(X) - \mathbb{E}\,F(\widehat{X}^{\Gamma})\right|.$$

▷ If F is Lipschitz continuous, then for every $r \in [1, +\infty)$,

$$\boxed{\left\|\mathbb{E}(F(X)\,|\,\widehat{X}^{\Gamma}) - F(\widehat{X}^{\Gamma})\right\|_{r} \le [F]_{\mathrm{Lip}}\|X - \widehat{X}^{\Gamma}\|_{r}}$$

▷ Assume F is $\mathcal{C}^1$ on H, DF is Lipschitz continuous and the quantizer Γ is a stationary.

Taylor expansion yields

$$\left|\mathbb{E}\,F(X) - \mathbb{E}\,F(\widehat{X}^{\Gamma}) - \mathbb{E}\left(DF(\widehat{X}^{\Gamma}).(X - \widehat{X}^{\Gamma})\right)\right| \quad \leq \quad [DF]_{\rm Lip}\mathbb{E}\left|X - \widehat{X}^{\Gamma}\right|^2$$

▷ Assume F is $\mathcal{C}^1$ on H, DF is Lipschitz continuous and the quantizer Γ is a stationary. Taylor expansion $\Longrightarrow$

$$\left|\mathbb{E}F(X) - \mathbb{E}F(\widehat{X}^{\Gamma}) - \underbrace{\mathbb{E}\left(DF(\widehat{X}^{\Gamma}).(X - \widehat{X}^{\Gamma})\right)}_{=0}\right| \quad \leq \quad [DF]_{\text{Lip}}\mathbb{E}\left|X - \widehat{X}^{\Gamma}\right|^2$$

since

$$\mathbb{E}\left(DF(\widehat{X}^{\Gamma}).(X - \widehat{X}^{\Gamma})\right) = \mathbb{E}\left(DF(\widehat{X}^{\Gamma}).\mathbb{E}(X - \widehat{X}^{\Gamma} \mid \widehat{X}^{\Gamma})\right) = 0.$$

so that

$$\boxed{\left|\mathbb{E}F(X) - \mathbb{E}F(\widehat{X}^{\Gamma})\right| \leq [DF]_{\text{Lip}}\|X - \widehat{X}^{\Gamma}\|_2^2}$$

and

$$\|X - \widehat{X}^{\Gamma}\|_2^2 = \sup_{[DF]_{\text{Lip}} \leq 1} \left|\mathbb{E}F(X) - \mathbb{E}F(\widehat{X}^{\Gamma})\right|.$$

▷ Similarly

$$\boxed{\left\|\mathbb{E}(F(X)\,|\,\widehat{X}^{\Gamma}) - F(\widehat{X}^{\Gamma})\right\|_r \le [DF]_{\rm Lip}\|X-\widehat{X}^{\Gamma}\|_{2r}^2}$$

since

$$\left|\mathbb{E}(F(X)\,|\,\widehat{X}^{\Gamma}) - F(\widehat{X}^{\Gamma})\right| \le [DF]_{\rm Lip}\mathbb{E}(|X-\widehat{X}^{\Gamma}|^2\,|\,\widehat{X}^{\Gamma})$$

1.7 Quantized approximation of $\mathbb{E}(F(X)\,|\,Y)$

▷ Let X, Y $(\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow H$ and $F : H \to \mathbb{R}$ a Borel functional.

Let $\widehat{X} = \widehat{X}^{\Gamma}$ and $\widehat{Y} = \widehat{Y}^{\Gamma'}$ are (Voronoi) quantizations .

▷ Natural idea $\mathbb{E}(F(X)\,|\,Y) \approx \mathbb{E}(F(\widehat{X})\,|\,\widehat{Y})$.To what extend ?

$$\mathbb{E}(F(X)\,|\,Y) = \varphi_{F}(Y).$$

▷ In a Feller Markovian framework : regularity of $F \rightsquigarrow$ regularity φ_F

$$\mathbb{E}(F(X)\,|\,Y) - \mathbb{E}(F(\widehat{X})\,|\,\widehat{Y}) = \mathbb{E}(F(X)\,|\,Y) - \mathbb{E}(F(X)\,|\,\widehat{Y}) + \mathbb{E}(F(X) - F(\widehat{X})\,|\,\widehat{Y})$$

so that, using that conditional expectation is an L^2-contraction and $\widehat{Y}$ is $\sigma(Y)$-measurable,

$$\begin{aligned}\|\mathbb{E}(F(X)|Y) - \mathbb{E}(\mathbb{E}(F(\widehat{X})|Y)|\widehat{Y})\|_2 &\leq \|\varphi_F(Y) - \mathbb{E}(F(X)|\widehat{Y})\|_2 + \|F(X) - F(\\ &= \|\varphi_F(Y) - \mathbb{E}(\varphi_F(Y)|\widehat{Y})\|_2 + \|F(X) - F(\\ &\leq \|\varphi_F(Y) - \varphi_F(\widehat{Y})\|_2 + \|F(X) - F(\widehat{X})\|_2\end{aligned}$$

The last inequality follows from the very definition of conditional expectation given $\widehat{Y}$

$$\boxed{\|\mathbb{E}(F(X)\,|\,Y)-\mathbb{E}(F(\widehat{X})\,|\,\widehat{Y})\|_2 \le [F]_{\rm Lip}\|X-\widehat{X}\|_2+[\varphi_F]_{\rm Lip}\|Y-\widehat{Y}\|_2.}$$

▷ Non-quadratic case the above inequality remains valid provided $[\varphi_F]_{\rm Lip}$ is replaced by $2[\varphi_F]_{\rm Lip}$.

▷ These are the ingredients for the proofs of both theorems for

– Bermuda options (orders 0 & 1).

– Swing options

1.8 Vector Quantization rate ($H = \mathbb{R}^d$)

▷ THEOREM (*a*) ASYMPTOTIC (Zador, Kiefer, Bucklew & Wise, Graf & Luschgy al., from 1963 to 2000).

Let $X \in L^{r+}(\mathbb{P})$ and $\mathbb{P}_X(d\xi) = \varphi(\xi)\,d\xi \overset{\perp}{+} \nu(d\xi)$. Then

$$e_{N,r}(X, \mathbb{R}^d) \sim \widetilde{J}_{2,d} \times \left(\int_{\mathbb{R}^d} \varphi^{\frac{d}{d+2}}(u)\,du\right)^{\frac{1}{d}+\frac{1}{r}} \times N^{-\frac{1}{d}} \quad \text{as} \quad N \to +\infty.$$

(*b*) NON ASYMPTOTIC (Luschgy-P., 2006). Let $r, \delta > 0$. There exists a universal constant $C_{r,\delta} \in (0, \infty)$

$$\forall\, N \geq 1, \qquad e_{N,r}(X, \mathbb{R}^d) \leq C_{r,\delta} \|X\|_{r+\delta} N^{-\frac{1}{d}}$$

▷ The true value of $\widetilde{J}_{r,d}$ is unknown for $d \geq 3$ but (Euclidean norm)

$$\widetilde{J}_{r,d} \sim \sqrt{\frac{d}{2\pi e}} \approx \sqrt{\frac{d}{17,08}} \quad \text{as} \quad d \to +\infty.$$

Conclusions : • For every N the same rate as with "naive" product-grids for the $U([0,1]^d)$ distribution with $N = m^d$ points + the best constant

• No escape from "The curse of dimensionality" ...

• Equalization of local inertia (see *Comm. in Statist.*, S.Delattre-J.C. Fort-G. P., 2004)

2 Numerical optimization of the grids : Gaussian and non-Gaussian vectors

2.1 The case of normal distribution $\mathcal{N}(0; I_d)$ on $\mathbb{R}^d$

▷ As concerns Gaussian $\mathcal{N}(0, I_d)$

Already quantized for you

(see J. Printems-G.P., *MCMA* 2003).

$\triangleright$ For $d = 1$ up to 10 and $N = 1 \leq N \leq 5\,000$, new grid files available including(L^1&L^2-distortion, local L^1&L^2-pseudo-inertia, etc).

on

Download at our WEBSITE :

www.quantize.maths-fi.com

2.2 The 1-dimension...

▷ THEOREM (Kiefer (82), LLoyd (82), Lamberton-P. (90)) $H = \mathbb{R}$. If $\mathbb{P}_X(d\xi) = \varphi(\xi)\, d\xi$ with $\log\varphi$ concave, then there is exactly one stationary quantizer. Hence

$$\forall\, N \geq 1, \qquad \operatorname{argmin} D_N^X = \{\Gamma^{(N)}\}.$$

EXAMPLES : The normal distribution, the gamma distributions, etc.

▷ Voronoi cells : $C_i(\Gamma) = [x^{i-\frac{1}{2}}, x^{i+\frac{1}{2}}[, \;\; x^{i+\frac{1}{2}} = \frac{x^{i+1}+x^i}{2}$.

▷ Gradient : $\nabla D_N^X(\Gamma) = 2\left(\displaystyle\int_{x^{i-\frac{1}{2}}}^{x^{i+\frac{1}{2}}} (x^i - \xi)\varphi(\xi)d\xi\right)_{1\leq i\leq N}$

Hessian : $D^2(D_N^X)(\Gamma) = \quad \ldots\ldots \quad$ only involves $\int_0^x \varphi(\xi)d\xi$ and $\int_0^x \xi\varphi(\xi)d\xi$

$\triangleright$ Thus if $X \sim \mathcal{N}(0;1)$: only $\mathrm{erf}(x)$ and $e^{-\frac{x^2}{2}}$ are needed.

$\triangleright$ Instant search for the unique optimal quantizer using a Newton-Raphson descent on $\mathbb{R}^N$ ...with an arbitrary accuracy.

$\triangleright$ For $\mathcal{N}(0;1)$ and $N = 1, \ldots, 500$, tabulation within 10^{-14} accuracy of both optimal N-quantizers and companion parameters :

$$\Gamma^{(N)} = (x^{(N),1}, \ldots, x^{(N),N})$$

and

$$\mathbb{P}(X \in C_i(\Gamma^{(N)})),\ i = 1, \ldots N, \qquad \text{and} \qquad \|X - \widehat{X}^{\Gamma^{(N)}}\|_2.$$

Download at our WEBSITE :

www.quantize.math-fi.com

2.3 Optimal quantization by simulation or general distribution

2.3.1 Competitive Learning Vector Quantization

- Grid $\Gamma := \{x^1, \ldots, x^N\} \longleftrightarrow (x^1, \ldots, x^N)$

$$D_N^X(\Gamma) := \|X - \widehat{X}^{\Gamma}\|_2^2 = \mathbb{E}\left(d_N(\Gamma, X)\right)$$

with $(\Gamma, \xi) \mapsto d_N(\Gamma, \xi)$ is a *local potential* defined by

$$d_N(\Gamma, \xi) = \min_{1 \le i \le N} |\xi - x^i|^2.$$

- D_N^X is continuously differentiable at grids Γ of full size N and

$$\frac{\partial D_N^X}{\partial x^i}(\Gamma) \; := \; \mathbb{E}\,\frac{\partial d_N}{\partial x^i}(\Gamma, X) = \int_{\mathbb{R}^d} \frac{\partial d_N}{\partial x^i}(\Gamma, \xi)\mathbb{P}_X(d\xi),$$

with a *local gradient* $\displaystyle \frac{\partial d_N}{\partial x^i}(\Gamma, \xi) \; := \; 2(x^i - \xi)\mathbf{1}_{\{\mathrm{Proj}_{\Gamma}(x) = x^i\}}, \qquad 1 \le i \le N.$

- ∇D_N^X has an integral representation

||

Minimization of D^X_N using a stochastic gradient descent

- INGREDIENTS : – $\xi^1, \ldots, \xi^t, \ldots$ *simulated independent copies of* X,

  – Step sequence $\delta_1, \ldots, \delta_t \ldots$.

Usually : $\delta_t = \dfrac{A}{B+t} \searrow 0$ or $\delta_t = \eta \approx 0$.

- STOCHASTIC GRADIENT DESCENT Formally reads

$$\Gamma(t) = \Gamma(t-1) - \delta_t \nabla d_N(\Gamma(t-1), \xi^t), \qquad |\Gamma^0| = N.$$

- GRID UPDATING : $(t \rightsquigarrow t+1) : \Gamma(t) := \{x^{1,t}, \ldots, x^{N,t}\}$,

COMPETITION : winner selection $i(t+1) \in \mathrm{argmin}_i |x^{i,t} - \xi^{t+1}|$

LEARNING :
$$\begin{cases} x^{i(t+1),t+1} := \mathrm{Homothety}(\xi^{t+1}, 1-\delta_{t+1})(x^{i(t+1),t}) \\ x^{i,t+1} := x^{i,t},\ i \neq i(t+1). \end{cases}$$

- HEURISTICS : $\Gamma^t \longrightarrow \Gamma^* \in \text{argmin(loc)}_\Gamma \mathrm{D}^{\mathrm{X}}_{\mathrm{N}}(\Gamma)$ as $t \to \infty$.

- COMPUTATION OF THE "COMPANION PARAMETERS" :

  – Weights $\pi^{i,*} = \mathbb{P}(\widehat{X}^{\Gamma^*} = x^{i,*})$, $i = 1, \ldots, N$:

$$\pi^{i,t+1} := (1 - \delta_{t+1})\pi^{i,t} + \delta_{t+1}\mathbf{1}_{\{i=i(t+1)\}} \xrightarrow{a.s.} \pi^{i,*} = \mathbb{P}(\widehat{X}^{\Gamma^*} = x^{i,*}).$$

  – (Quadratic) Quantization error $D^X_N(\Gamma^*) = \|X - \widehat{X}^{\Gamma^*}\|_2$:

$$D^{X,t+1}_N := (1 - \delta_{t+1})D^{X,t}_N + \delta_{t+1}|x^{i(t+1),t} - \xi^{t+1}|^2 \xrightarrow{a.s.} D^X_N(\Gamma^*).$$

Extra C.P.U. time cost ≈ 0 !

CLVQ $\equiv$ NON LINEAR MONTE CARLO SIMULATION

- Extension to the whole quantization tree

2.3.2 Randomized Lloyd's I procedure

▷ Randomized fixed point procedure based on the stationarity equality :

$$\widehat{X}^{\Gamma(t+1)} = \mathbb{E}(X \,|\, \widehat{X}^{\Gamma(t)}), \qquad \Gamma(0) \subset \mathbb{R}^d, \quad |\Gamma| = N.$$

▷ $\Gamma(\ell) = \{x_1^{(\ell)}, \ldots, x_N^{(\ell)}\}$ being computed,

$$x_i^{(\ell+1)} := \mathbb{E}(X^{\Gamma(\ell)} \,|\, X^{\Gamma(\ell)} \in C_i(\Gamma(\ell))) \quad = \quad \lim_{M \to \infty} \frac{\sum_{m=1}^{M} X_m \mathbf{1}_{\{X_m \in C_i(\Gamma(\ell))\}}}{|\{1 \leq m \leq M,\ X_m \in C_i(\Gamma(\ell))\}|}$$

based on repeated nearest neighbour searches.

▷ Improvements : splitting method.

$$\Gamma_{N+1}(0) = \Gamma_N(\infty) \cup \{X(\omega)\}$$

▷ Alternative based on minimum local inertia search (A. Sagna in progress).

2.3.3 Fast nearest neighbour procedure in $\mathbb{R}^d$

▷ The Partial Distance Search paradigm (Chen, 1970) : Target $= 0$!!

Running record dist to $0 :=$ Rec.

Let $x = (x^1, \dots, x^d) \in \mathbb{R}^d$

$$(x^1)^2 \geq Rec^2 \quad \Longrightarrow \quad |x| \geq Rec$$
$$\vdots$$
$$(x^1)^2 + \cdots + (x^\ell)^2 \geq Rec^2 \quad \Longrightarrow \quad |x| \geq Rec$$
$$\vdots$$

▷ The K-d tree (Friedmann, Bentley, Finkel , 1977) : store the N points of $\mathbb{R}^d$ in a tree of depth $O(\log(N))$ ...

▷ Further recent improvements (Mc Names) : K-d-tree $+CPA$.

3 Multi-asset American/Bermuda Options

▷ *d Traded risky assets :* $S_t = (S_t^1, \ldots, S_t^d) \qquad t \in [0, T]$.

with natural (augmented...) filtation $\mathcal{F}^S = \mathcal{F}_t^S)_{t\in[0,T]}$.

▷ *Discounted price :* $\widetilde{S}_t^i = \frac{S_t^i}{S_t^0} = e^{-rt} S_t^i, \qquad i = 1, \ldots, d.$

is a $(\mathbb{P}, \mathcal{F}^S)$-martingale under the risk-neutral probability (if AOA holds) where r is a *riskless asset and Mathematical interest rate.*

▷ *American Payoff process :* $(h_t)_{t\in[0,T]}$ is a≥ 0, $\mathcal{F}^S$-adapted process.

▷ *American option on* $(h_t)_{t\in[0,T]}$ *:*

Choose to receive h_t once within 0 and T

▷ *Bermuda option on* $(h_t)_{t\in[0,T]}$ *:*

Choose to receive h_{t_k} once, $k = 0, \ldots, n$.

usually with $t_k = \dfrac{kT}{n}$, $k = 0, \ldots, n$.

Examples :

$\triangleright$ *Call/Put Option* :

Right to buy/sell once the asset S at *the strike price* K

American : once at $t \in [0, T]$ *vs* Bermuda : once at a time $t = t_k = \dfrac{kT}{n}$, $k = 0, \ldots, n$.

$h_t = (S_t^1 - K)^+$ or $h_t = (K - S_t^1)^+$.

$\triangleright$ *"Vanilla" American Options* :

Right to receive once $h_t = h(t, S_t) \geq 0$ within time 0 and T

vs Bermuda : once at a time $t = t_k = \dfrac{kT}{n}$, $k = 0, \ldots, n$.

Example : Exchange American/Bermuda options (Villeneuve) :

$$h_t = (S_t^1 - \lambda\, S_t^2)^+.$$

▷ *"Exotic" American/Bermuda Options* : $h_t \neq h(t, S_t)$.

Example : American/Bermuda Asian options :

$h_t = \left(\frac{1}{T-T_0}\int_{T_0}^{T} S_s ds - K\right)^+$.

American/Bermuda Lookback options, etc.

▷ *"Shout" Options* :

Right to "shout" once within time 0 and T

vs Bermuda : once at a time $t = t_k = \dfrac{kT}{n}$, $k = 0, \dots, n$.

to receive (a non adapted) h_t at T.

3.1 Pricing Bermuda options : the dynamical programming principle

3.2 Markov structure process

(Replace $t_k = \frac{kT}{n}$ by k) Let $(X_k)_{0\le k\le n}$ be a Markov structure process.

with transition $P_{k-1,k}(g)(x) = \mathbb{E}\,(g(X_{k+1}\,|\,X_k = x)$ such that

- $\mathcal{F}^X_k = \mathcal{F}^S_{t_k}$
- Risky asset vector satisfies

$$S_{t_k} = (S^1_{t_k}, \ldots, S^d_{t_k}) = G(X_k)$$

- Payoff process satisfies

$$h_{t_k} = h(k, X_k).$$

- Simulability : $(X_k)_{0\le k\le n}$ can be simulated (at a reasonable cost).

- Typical structure processes (for American/Bermuda "Vanilla" options) :

$$X_k := \begin{cases} S_{t_k} & (Ex : X_k = W_{t_k} \text{the multi-dim } B\text{-}S \text{ model}) \\ \log(S_{t_k}) & \\ \bar{S}_{t_k} & \text{(Euler scheme)} \end{cases}$$

- For path-dependent options (Asian, lookback, etc)

$$X_k := \begin{cases} (S_{t_k}, \frac{1}{t_k}(S_0 + \cdots + S_{t_k})) \\ (\bar{S}_{t_k}, \frac{1}{t_k}(\bar{S}_0 + \cdots + \bar{S}_{t_k})), \\ (S_{t_k}, \max_{0 \le i \le k} S_{t_i}), \\ \text{etc.} \end{cases}$$

3.3 Arbitrage and value function

STEP 1

$$\begin{cases} \mathcal{V}_n & := & h(n, X_n) \\ & & \\ \mathcal{V}_k & := & \max\left(h(k, X_k), \mathbb{E}(\mathcal{V}_{k+1}|\mathcal{F}_k^X)\right). \end{cases}$$

STEP 2 Backward induction based on the Markov property

Markov $\Longrightarrow$ Conditioning given $\mathcal{F}_k^X$ = Conditioning given X_k.

$$\boxed{\mathcal{V}_k = v_k(X_k), \qquad k = 0, \ldots, n.}$$

3.4 Vector Quantization approach (Bally-P.-Printems, from 2000 to 2005)

Based on the value function.

APPROXIMATION 1 : QUANTIZATION

Substitution by nearest neighbour projection on grids Γ_k :

$$\widehat{X}_k = \pi_k(X_k) \longleftarrow X_k$$

where $\pi_k : \mathbb{R}^d \to \Gamma_k$, Γ_k is a grid of size N_k, $\Gamma_k = \{x_k^1, \ldots, x_k^{N_k}\} \subset \mathbb{R}^d$.

But loss of the Markov property...

Approximation 2 : Markov approximation

$$\text{Quantized obstacle} : h(k, \widehat{X}_k),\ k = 0 \dots, n.$$

The Markov property is forced : one defines $\widehat{V}_k$ by a backward induction

$$(\text{QDPP-I}) \equiv \begin{cases} \widehat{V}_n & := & h(n, \widehat{X}_n) \\ \widehat{V}_k & := & \max(h(k, \widehat{X}_k), \mathbb{E}(\widehat{V}_{k+1} \,|\, \widehat{X}_k)), \quad k = 0, \dots, n-1. \end{cases}$$

Again a Backward induction

$$\widehat{V}_k = \widehat{v}_k(\widehat{X}_k), \quad k = 0, \dots, n.$$

where

$$
(\text{QDPP-II}) \equiv \left\{ \begin{array}{rcl} \widehat{v}_n(x_n^i) & = & h(n, x_n^i), \qquad i = 1, \ldots, N_n \\ \widehat{v}_k(x_k^i) & = & \max\left(h(k, x_k^i), \sum_{j=1}^{N_k} \widehat{p}_k^{ij}\, \widehat{v}_{k+1}(x_{k+1}^j)\right), \quad i = 1, \ldots, N \\ & & \qquad\qquad k = 1, \ldots, n - \end{array} \right.
$$

NUMERICAL TASK(S) Optimize and Compute *off-line*

– Task 1 : (good) grids Γ_k including the quantization error.

and

– Task 2 : (accurate) quantized transitions $\widehat{p}_k^{ij} := \dfrac{\mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j, \widehat{X}_k = x_k^i)}{\mathbb{P}(\widehat{X}_k = x_k^i)}$.

Conclusion

(QDPP-II) is instantaneous for the *on line* computation of any portfolio of options.

Interpretation Global Transition operators approximation

Grids Γ_k + quantized transitions $\widehat{p}_k^{ij}$

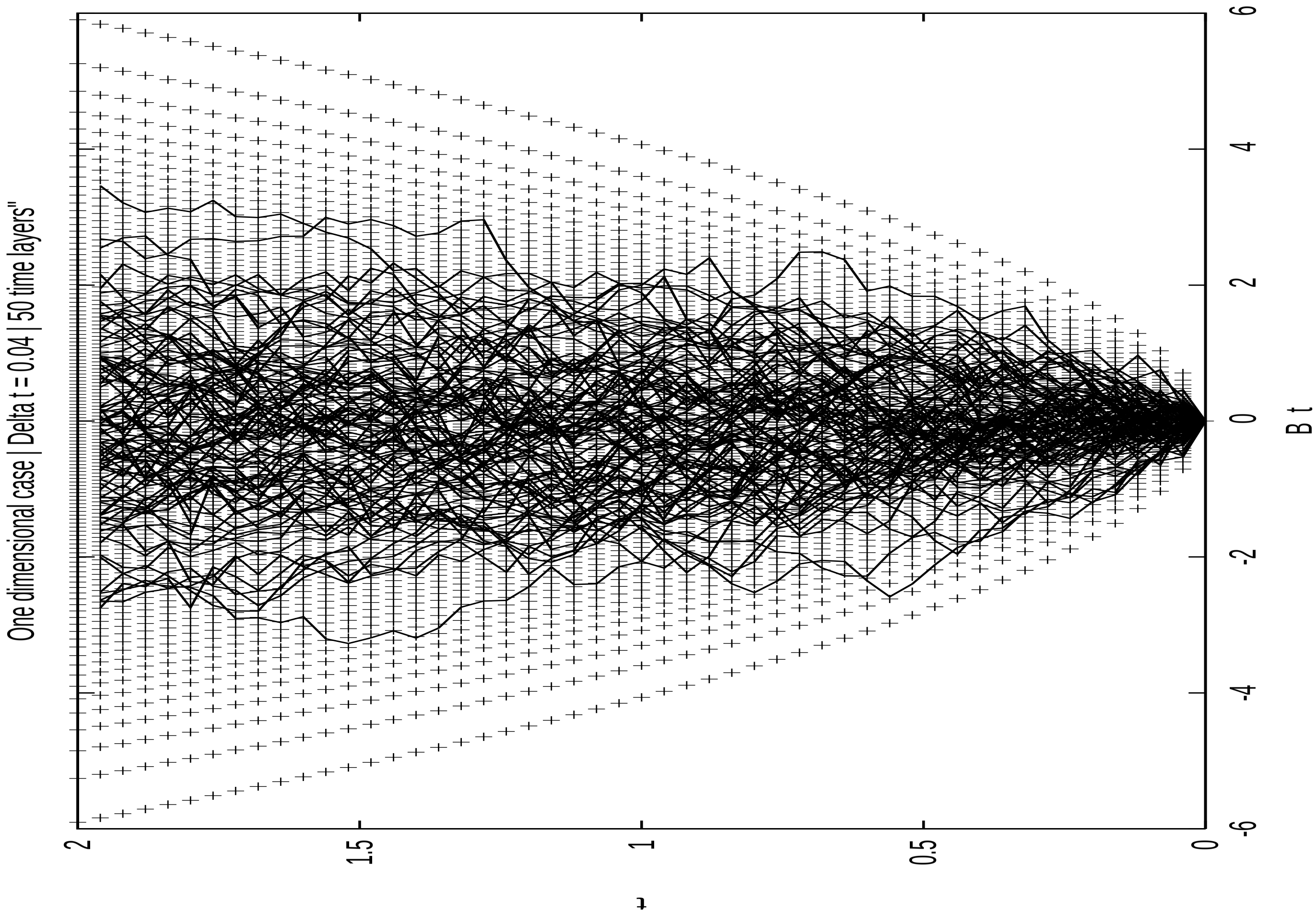
$$\Downarrow$$

$$\boxed{\widehat{P}_{k-1,k}(x_k^i, dy) = \sum_j \widehat{p}_k^{ij} \delta_{x_k^j}}$$

with

$$\widehat{P}_{k-1,k}(x_k^i, dy) \approx P_{k-1,k}(x, dy), \qquad k = 1, \dots, n.$$

3.5 Quantization tree (I)



- For every $k \in \{0, \ldots, n\}$, $|\Gamma_k| = N_k$.
- Theoretical complexity of a tree descent : $\kappa \sum_{k=0}^{n-1} N_k \, N_{k+1}$.
- Global size of the tree (constraint) : $\sum_{k=0}^{n} N_k = N$.

The theoretical complexity is minimal when (Schwarz Inequality)

$$N_k = \frac{N}{n+1}$$

with complexity $\frac{n}{(n+1)^2} N^2$. Not so important in practise since

Most connections $\widehat{p}_k^{ij}$ are negligible $\Longrightarrow$ pruning. . .

Display a pruned quantization tree

THEOREM (a) (Bally-Pagès, 2001 ($MCMA$) to 2005 ($Math.Fin.$))
Scheme of order 0 (described above, to be compared to *non conformal finite elements of order* 0). If $h(k,.)$ are Lipschitz, the transitions $P_{k,k-1}$ are Lipschitz, the

$$\|\mathcal{V}_0 - \widehat{v}_0(\widehat{X}_0)\|_2 \le C_{X,\varphi} \sum_{k=0}^{n} \|X_k - \widehat{X}_k^{\Gamma_k}\|_2.$$

(b) (Bally-Pagès-Printems, ($Math.Fin.$), 2003) Scheme of order 1 (to be compared to *non conformal finite elements of order* 1). If $(\ldots)$

$$\|\mathcal{V}_0 - \widehat{v}_0(\widehat{X}_0)\|_2 \le C_{X,\varphi} \sum_{k=0}^{n} \|X_k - \widehat{X}_k^{\Gamma_k}\|_2^2.$$

3.6 Optimal design of the quantization tree

IDEA : optimal integral allocation problem

Item (a) of the theorem & Zador's Theorem (non asympotic version)

$$\begin{aligned}
\|\mathcal{V}_0 - \widehat{v}_0(\widehat{X}_0)\|_2 &\leq C_{X,\varphi} \sum_{k=0}^{n} \|X_k - \widehat{X}_k^{\Gamma_k}\|_2 \\
&\leq C_{X,\varphi} C_\delta \sum_{k=0}^{n} \|X_k\|_{2+\delta} |\Gamma_k|^{-\frac{1}{d}} \\
&= C_{X,\varphi} C_\delta \sum_{k=0}^{n} \|X_k\|_{2+\delta} N_k^{-\frac{1}{d}}
\end{aligned}$$

Amounts to solving the

$$\min_{N_0+\cdots+N_n=N} \sum_{k=0}^{n} \|X_k\|_{2+\delta} N_k^{-\frac{1}{d}}$$

i.e. denoting the (upper) integral part of x by $\lceil x \rceil$,

$$N_k = \left\lceil \frac{\left(\|X_k\|_{2+\delta}\right)^{\frac{d}{d+1}}}{\sum_{0\le\ell\le n}\left(\|X_\ell\|_{2+\delta}\right)^{\frac{d}{d+1}}} N \right\rceil, \qquad k = 0, \ldots, n$$

so that

$$\boxed{\left\|\mathcal{V}_0 - \widehat{v}_0(\widehat{X}_0)\right\|_2 \le C_{X,\varphi} C_\delta \left(\sum_{k=0}^{n} \left(\|X_k\|_{2+\delta}\right)^{\frac{d}{d+1}} \right)^{1-\frac{1}{d}} \widetilde{N}^{-\frac{1}{d}}.}$$

with $\widetilde{N} = N_0 + \cdots + N_n$ (usually $> N$).

EXAMPLES :

• Brownian motion $X_k = W_{t_k}$: Then $\widehat{W}_0 = 0$ and

$$\|W_{t_k}\|_{2+\delta} = C_\delta \sqrt{t_k}, \qquad k = 0, \ldots, n.$$

Hence $N_0 = 1$ and

$$N_k \approx \frac{2(d+1)}{d+2} \left(\frac{k}{n}\right)^{\frac{d}{2(d+1)}} N, \qquad k = 1, \ldots, n.$$

$$|\mathcal{V}_0 - \widehat{v}_0(0)| \le C_{W,\delta} \left(\frac{2(d+1)}{d+2}\right)^{1-\frac{1}{d}} \frac{n^{1+\frac{1}{d}}}{N^{\frac{1}{d}}} = O\big(\frac{n}{\bar{N}^{\frac{1}{d}}}\big), \quad \bar{N} = \frac{N}{n}.$$

Theoretically not crucial. Numerically it is...

• Stationary process (ex : $X_k = OU_{t_k}$) : Only needs

ONE OPTIMAL GRID ...and ONE QUANTIZED TRANSITION MATRIX

since $\|X_k\|_{2+\delta} = \|X_0\|_{2+\delta}$.

Hence

$$N_k = \left\lceil \frac{N}{n+1} \right\rceil, \quad k = 0, \ldots, n.$$

$$\|\mathcal{V}_0 - \widehat{v}_0(\widehat{X}_0)\|_2 \le C_{X,\delta} \frac{n^{1+\frac{1}{d}}}{N^{\frac{1}{d}}} = C_{X,\delta} \frac{n}{\bar{N}^{\frac{1}{d}}} \qquad \bar{N} = \frac{N}{n}.$$

3.7 Computing the quantized transitions $\widehat{p}_k^{ij}$

3.7.1 Standard Monte Carlo estimation

• As a companion procedure of grid updating :

– Nearest neighbour search at every time step to update the grid $\Gamma_k \subset \mathbb{R}^d$ *via* $CLVQ$ and the transition frequency estimators

– or "batch" estimation *via* randomized Lloyd's I procedure

• Freeze the grids and carry on the MC estimation of the transitions.

– M independent copies $X^m = (X_0^m, X_1^m, \ldots, X_n^m)$, $m = 1, \ldots, M$ "launched" in the quantization tree

3.7.2 Alternative methods

• Fast tree quantization for Gaussian structure processes (Bardou-Bouthemy-P. (2006) for swing options[...]).

• The "spray" method ("gerbes" in French) (P.-Pham-Printems (2005) for filtering by optimal quantization)

3.8 δ-Hedging, higher order schemes...

3.8.1 Computing the δ-hedge, $X_k = S_{t_k}$ (B-S) or $\bar{S}_{t_k}$ (local vol).

- Quantized δ-Hedging :
- $\widehat{\zeta}_k^n := \frac{n}{Tc^2(\widehat{S}_{t_k})} \widehat{\mathbb{E}}_k \Big((\widehat{v}_{k+1}^n(\widehat{S}_{t_{k+1}}) - \widehat{v}_k^n(\widehat{S}_{t_k}))(\widehat{S}_{t_{k+1}} - \widehat{S}_{t_k}) \Big)$.
- Similar formulae for the Euler scheme...

$(\mathcal{H}) \equiv (i)\ \sigma \in C_b^\infty(\mathbb{R}^d), \quad (ii)\ \sigma\sigma^* \geq \varepsilon_0 I_d, \quad (iii)\ \|x\sigma'(x)\|_\infty < +\infty.$

- Bermuda Error :

$$\mathbb{E}\int_0^T |c^*(S_u)(Z_u - \zeta_u^n)|^2\, ds \leq C_{h,\sigma} \frac{(1+|s_0|)^q}{\varepsilon_0} \frac{1}{n^{\frac{1}{6}}}.$$

- Quantization Error :

$$\mathbb{E}\int_0^T |\zeta_u^n - \widehat{\zeta}_u^n| du \leq C(1+|s_0|)| \frac{n^{\frac{3}{2}}}{(N/n)^{\frac{1}{d}}}.$$

4 Numerical experiments

4.1 Numerical experiments I : Exchange geometric options

- Exchange American options on geometric assets.
- Reference : Villeneuve-Zanette, 1998 *Finite differences for 2-Dim exchange American options with dividends.*
- Model : Standard $2d$-dim (B & S) model with *non correlated* Brownian Motions (The most "hostile" to quantization...).
- Maturity : $T = 1$ year. Volatility : $\sigma_i = \dfrac{20\%}{\sqrt{d}},\ i = 1, \ldots, d.$
  - $2d$-dim pay-off : $$h(t,x) = \left(\prod_{i=1}^{d} e^{-\mu_i t} S_t^i - \prod_{i=d+1}^{2d} e^{-\mu_i t} S_t^i\right)^+.$$
- Initial values :
  $\prod_{i=1}^{d} S_0^i = 40,\ \prod_{i=d+1}^{2d} S_0^i = 36$ (in-the money), $\mu_1 := 5\,\%,\ \mu_2 = 0, \ldots$
  $\prod_{i=1}^{d} S_0^i = 36,\ \prod_{i=d+1}^{2d} S_0^i = 40$ (out-of-the money), $\mu_{d+1} := 0\,\%, \ldots$

4.1.1 Results : Premium and δ-hedge : 0-order scheme

Maturity	3 months		6 months		9 months		12 months	
AM_{ref}	**4.4110**		**4.8969**		**5.2823**		**5.6501**	
	Price	Error (%)	Price	Error (%)	Price	Error (%)	Price	Error (%)
$d = 2$	4.4111	0.0023	4.8971	0.0041	5.2826	0.0057	5.6505	0.0071
$d = 4$	4.4076	0.08	4.9169	0.34	5.3284	0.82	5.7366	1.39
$d = 6$	4.4156	0.1	4.9276	0.63	5.3550	1.38	5.7834	2.20
$d = 10$	4.4317	0.47	4.9945	2.00	5.4350	2.89	5.8496	3.53

TAB. 1: AMERICAN PREMIUM & RELATIVE ERROR. Different maturities and dimensions.

(a) (b)

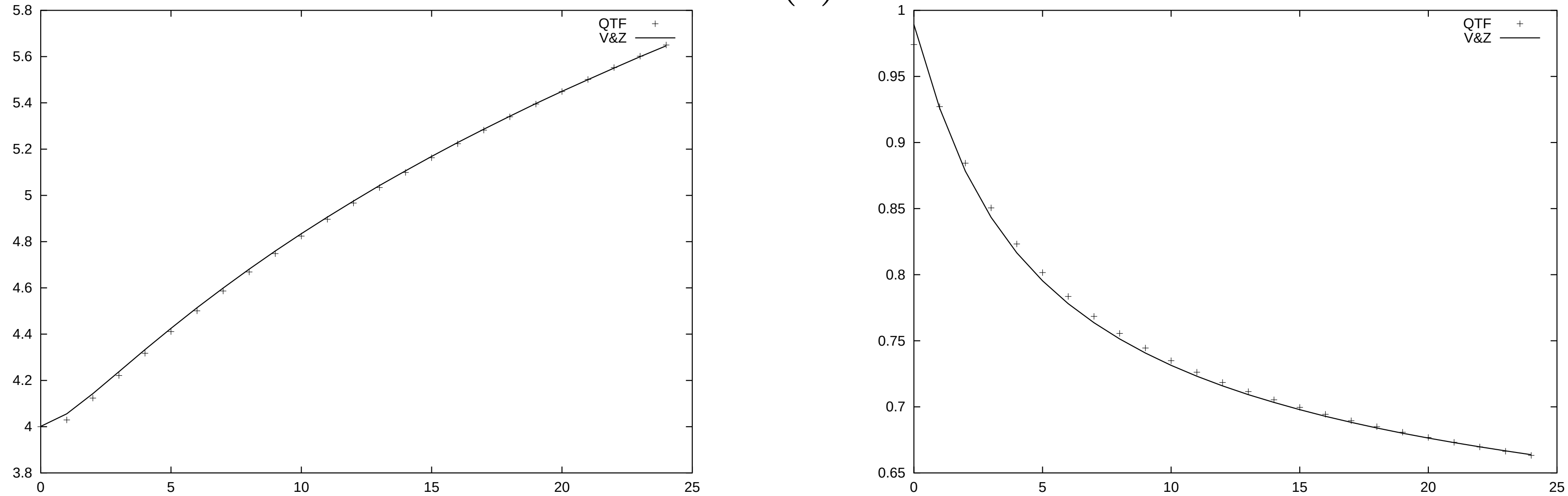


FIG. 3: $d = 2$, $n = 25$ and $\bar{N} = 300$. (a) American premium as a function of the maturity. (b) Hedging strategy on the first asset. The cross + depicts the premium obtained with the method of quantization and – depicts the reference premium (V & Z).

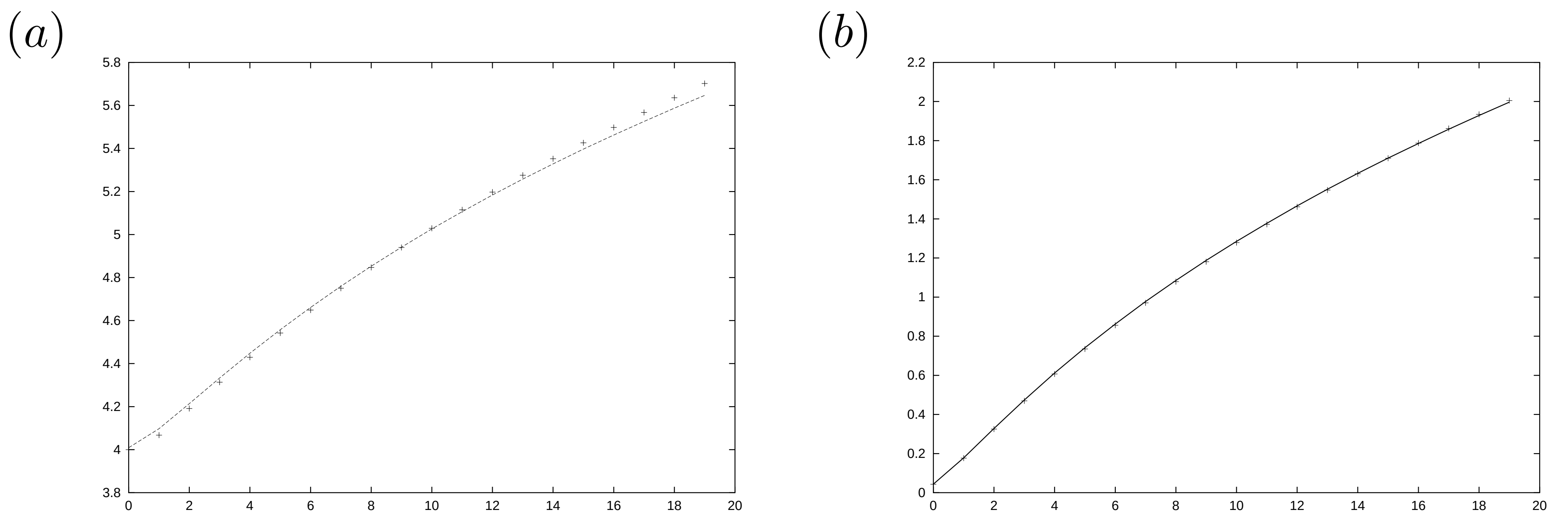


FIG. 4: $d = 4$. AMERICAN PREMIUM AS A FUNCTION OF THE MATURITY. (a) *In-the-money.* (b) *Out-of-the-money. + depicts the premium obtained with the method of quantization and – depicts the reference premium (V & Z).*

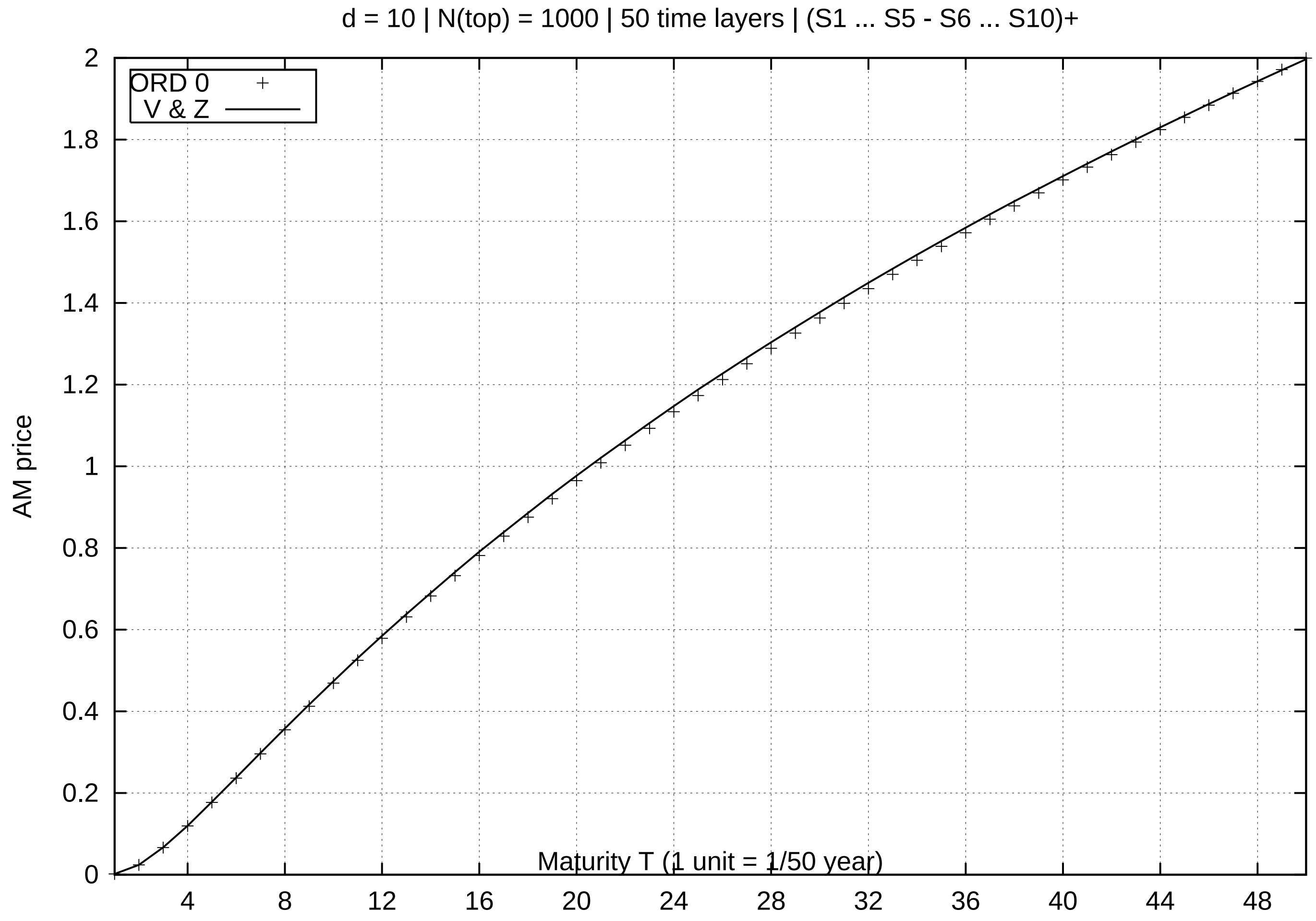


FIG. 5: Exchange option 10D $(S^1 \cdots S^5 - S^6 \cdots S^{10})_+$: out-of-the-money

4.1.2 0-order scheme *vs* 1-order scheme

Maturity	3 months		6 months		9 months		12 months	
AM_{ref}	**4.4110**		**4.8969**		**5.2823**		**5.6501**	
	Price	Error (%)	Price	Error (%)	Price	Error (%)	Price	Error (%)
$d = 4$								
AM_0	4.4076	0.08	4.9169	0.34	5.3284	0.82	5.7366	1.39
AM_1	4.4058	0.1	4.8991	0.04	5.2881	0.08	5.6592	0.13
$d = 6$								
AM_0	4.4156	0.1	4.9276	0.63	5.3550	1.38	5.7834	2.20
AM_1	4.4099	0.02	4.8975	0.01	5.3004	0.34	5.6557	0.10
$d = 10$								
AM_0	4.4317	0.47	4.9945	2.00	5.4350	2.89	5.8496	3.53
AM_1	4.4194	0.19	4.8936	0.07	5.1990	1.58	5.4486	3.56

TAB. 2: Relative errors of AM_0 and AM_1 with respect to a reference price for different maturities and dimensions.

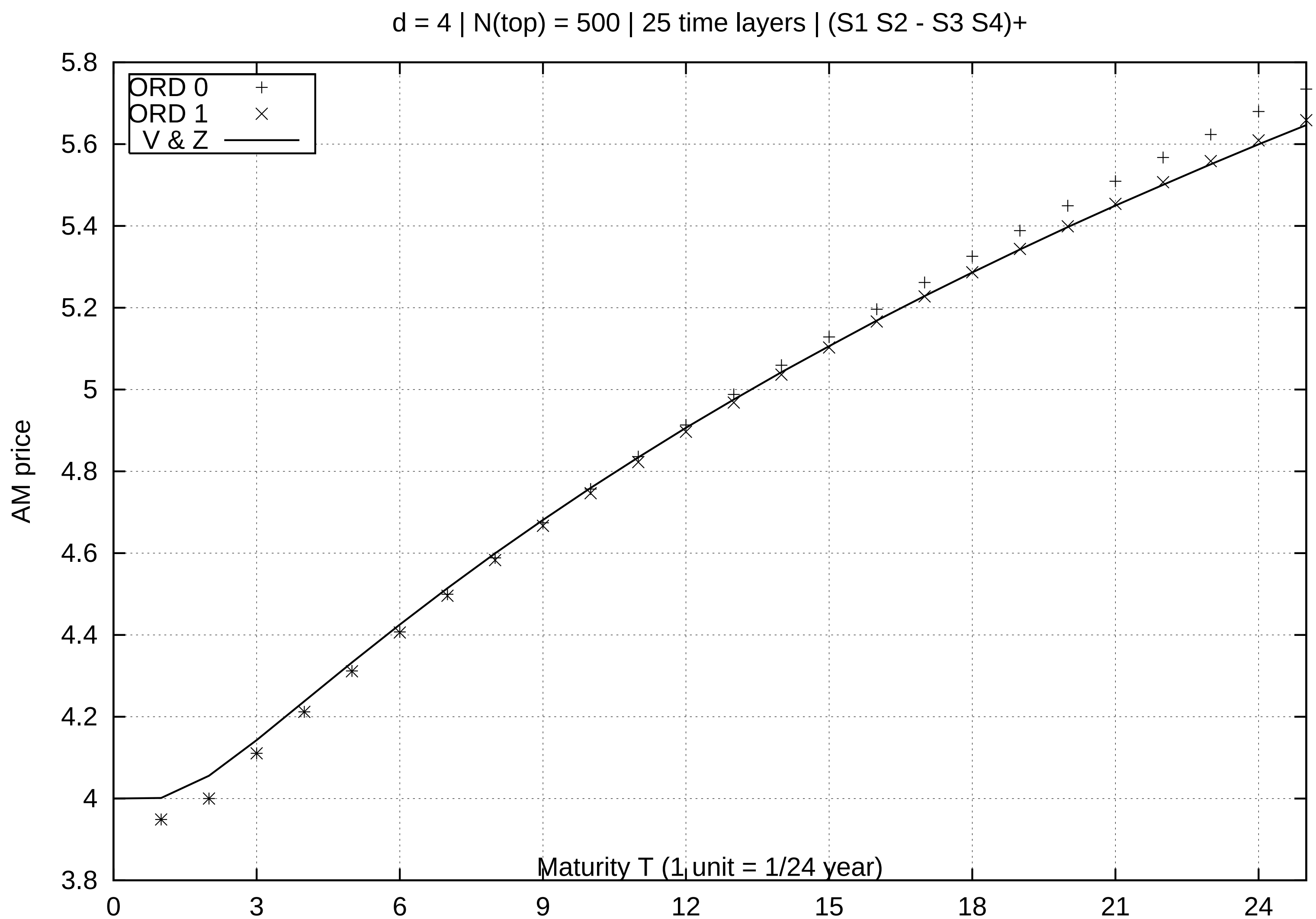


Fig. 6: Exchange option $4D$ $(S^1S^2 - S^3S^4)_+$: In-the-money. Dimension $d = 4$, $n = 25$ and $N_{25} = 500$. American option function of the maturity T. The crosses denote the quantized version with order 0 (+) and order 1 (×)

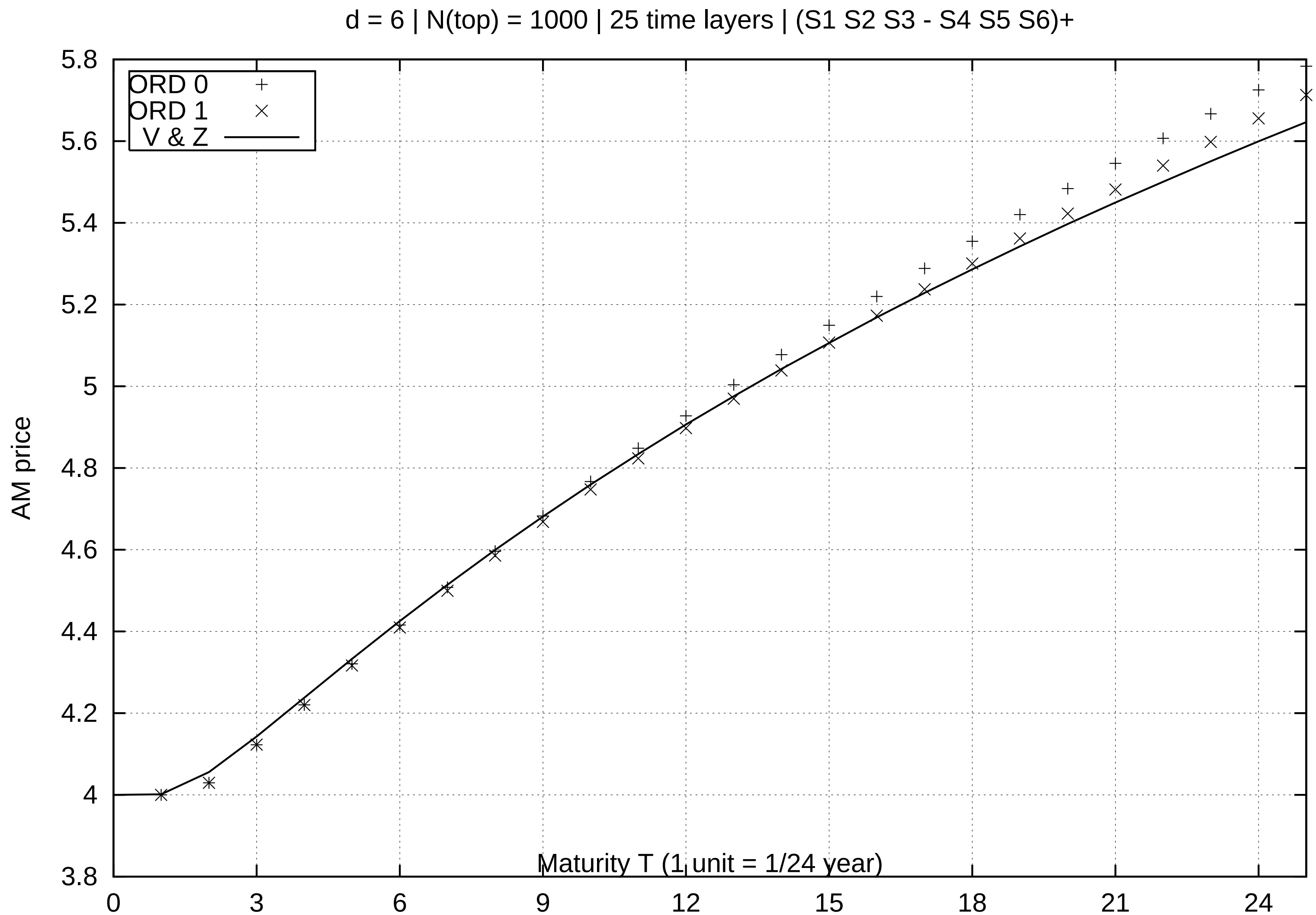


FIG. 7: Quantized version order 0 (+), order 1 (×). (a) Dimension $d = 6$, $n = 25$, $N_{25} = 1000$, In-the-money case. Value of the American option function of the maturity T.

Computation velocity : Pentium II, 800 MHz, 500 MO RAM [2003. . .]

$$d = 5 \qquad N = 2.10^4 \qquad n = 10$$

- Design of the quantization tree (grid/weights) : 3 seconds ;
- (Premium+ δ-Hedge) (QBDPP) : 3 per second.

4.2 Swing Options

▷ Take or Pay contract on gas (with firm constraints)

– Spot or day-ahead delivery contract S_{t_k} *assumed to Markov (for convenience) i.e.*

$$X_k = S_{t_k}$$

– *Local volume constraints* : Buy daily $q_{t_k} \in [q_{\min}, q_{\max}]$ m^3 of natural gas at price K_k

– *Global volume constraints* $Q_{\min} \leq q_0 + q_{t_1} + \cdots + q_{t_{n-1}} \leq Q_{\max}$.

$$P(Q_{\min}, Q_{\max}, s_0) = \sup_{(q_{t_k})_{0\leq k\leq n-1}\in \mathcal{A}_{Q_{\min},Q_{\max}}} \mathbb{E}\left(\sum_{k=0}^{n-1} q_{t_k} e^{-r(T-t_k)}(S_{t_k} - K_k)\right)$$

where the set of admissible daily purchased quantities is given by

$$\mathcal{A}_{Q_{\min},Q_{\max}} = \left\{(q_{t_k})_{0\leq k\leq n-1},\ q_{t_k} \underline{\in} \mathcal{F}^S_{t_k},\ q_{\min} \leq q_{t_k} \leq q_{\max},\ Q_{\min} \leq \sum_{0\leq k\leq n-1} q_{t_k} \leq Q_{\text{ma}}\right.$$

4.3 Pricing swing by (optimal) Quantization (2006)

(Bardou-Bouthemy-P. 2007).

Supply contracts and swing options

▷ Typical derivative products on energy markets :

Strip of Calls options with global physical constraints (volumes)

▷ Example : Used to model "reactive storage" and "supply contracts" for gas.

▷ We will focus on swing options for gas supply contracts :

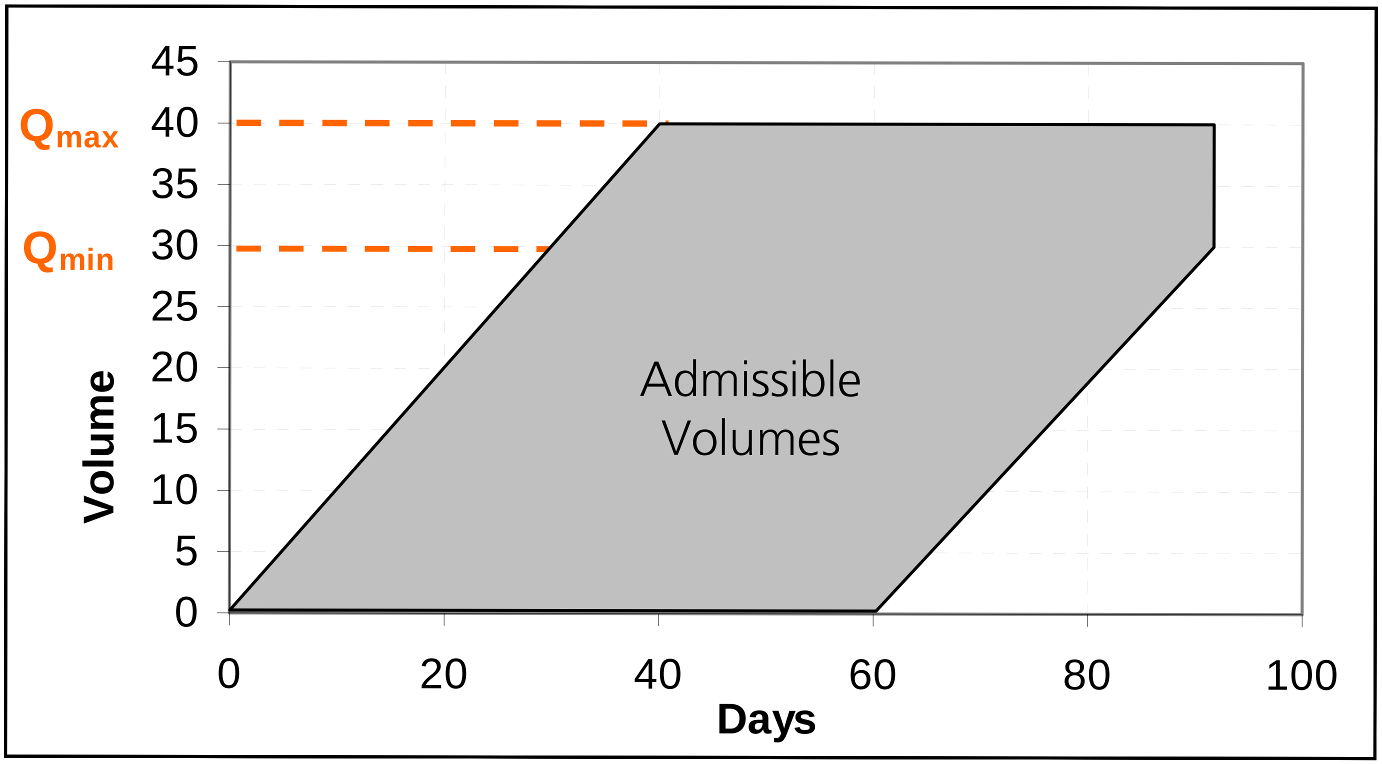
– Right to buy daily some gas at a strike price

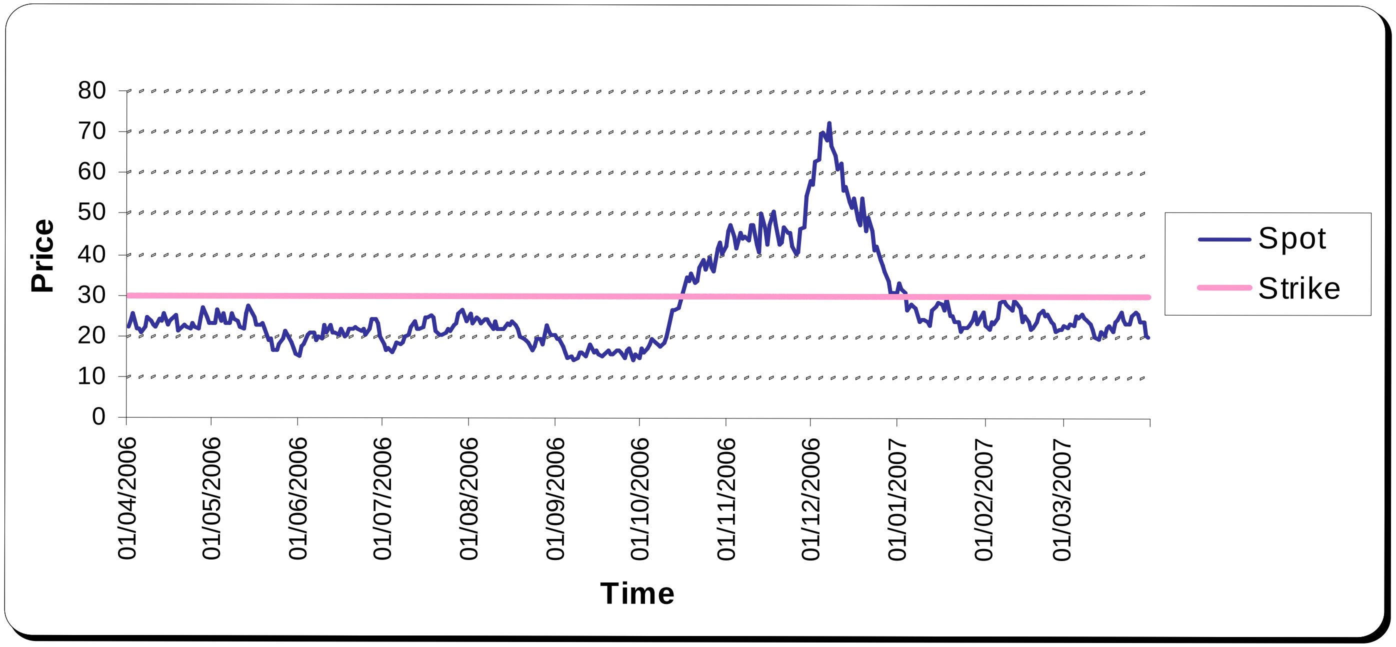
– Daily ("local") min-max constraints on the purchased volumes

– Annual ("global") min-max constraints on the purchased volumes

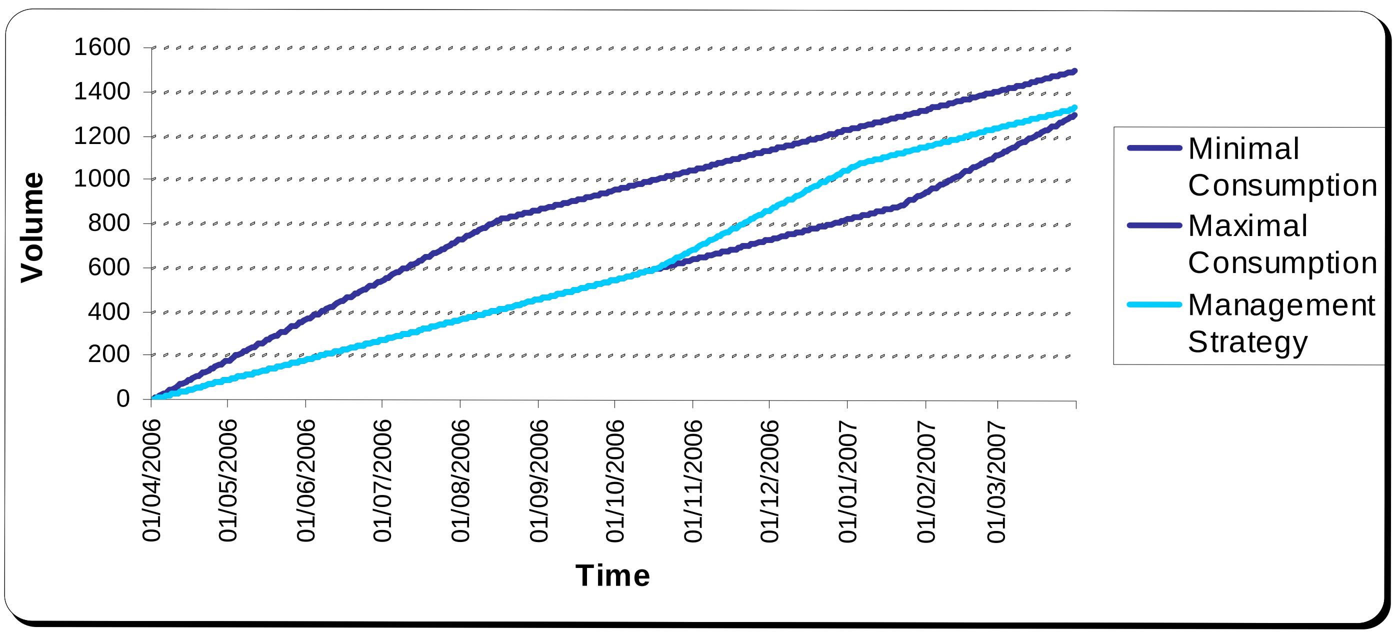
– Strike prices are possibly indexed on a basket of underlyings (petroleum products)

$$q_{min} = 0,\ q_{max} = 1,\ Q_{min} = 30,\ Q_{max} = 40$$





Price
80
70
60
50
40
30
20
10
0
01/04/2006
01/05/2006
01/06/2006
01/07/2006
01/08/2006
01/09/2006
01/10/2006
01/11/2006
01/12/2006
01/01/2007
01/02/2007
01/03/2007
Time
Spot
Strike



Volume
1600
1400
1200
1000
800
600
400
200
0
01/04/2006
01/05/2006
01/06/2006
01/07/2006
01/08/2006
01/09/2006
01/10/2006
01/11/2006
01/12/2006
01/01/2007
01/02/2007
01/03/2007
Time
Minimal Consumption
Maximal Consumption
Management Strategy

It is a stochastic control problem ($r = 0$)

▷ Dynamic programming principle on the price $P(t_k, S_{t_k}, Q_{t_k})$

$$P(t_k, S_{t_k}, Q_{t_k}) = \max\Big\{q(S_{t_k} - K) + \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, Q_{t_k} + q)|S_{t_k}), \tag{1}$$
$$q \in [q_{\min}, q_{\max}],\ Q_{t_k} + q \in [(Q_{\min} - (n-k)q_{\max})_+, (Q_{\max} - (n-k)q_{\min})_+]\Big\}.$$

▷ Bang-bang control (Bardou-Bouthemy-P. (2007)).

If $\left(\dfrac{Q_{\max} - nq_{\min}}{q_{\max} - q_{\min}}, \dfrac{Q_{\min} - nq_{\min}}{q_{\max} - q_{\min}}\right) \in \mathbb{N} \times \mathbb{N}$, then

the optimal control is bang-bang *i.e.* $\{q_{\min}, q_{\max}\}$-valued

▷ Quantized Dynamic programming principle Let $\widehat{S}_{t_k}$ be an (optimal) quantization of S_{t_k} taking values in $\Gamma_k := \{s_k^1, \ldots, s_k^{N_k}\}$, $k = 0, \ldots, n$.

$$\begin{cases} P(t_k, s_k^i, \hat{Q}_{t_k}) = \max_{q \in \mathcal{A}_k^{\hat{Q}_{t_k}}} [q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, \hat{Q}_{t_k} + q) | \hat{S}_{t_k} = s_k^i)] \\ \qquad i = 1, \ldots, N_k, \\ \mathcal{A}_k^{\hat{Q}_{t_k}} = \Big\{ q \in \{q_{\min}, q_{\max}\}, \hat{Q}_{t_k} + q \in [(Q_{\min} - (n-k) q_{\max})_+, (Q_{\max} - (n-k) q_{\min})_+ \\ P(T, s_T^i, \hat{Q}_T) = P_T(s_T^i, \hat{Q}_T), \; i = 1, \ldots, N_n. \end{cases} \tag{2}$$

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

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

where

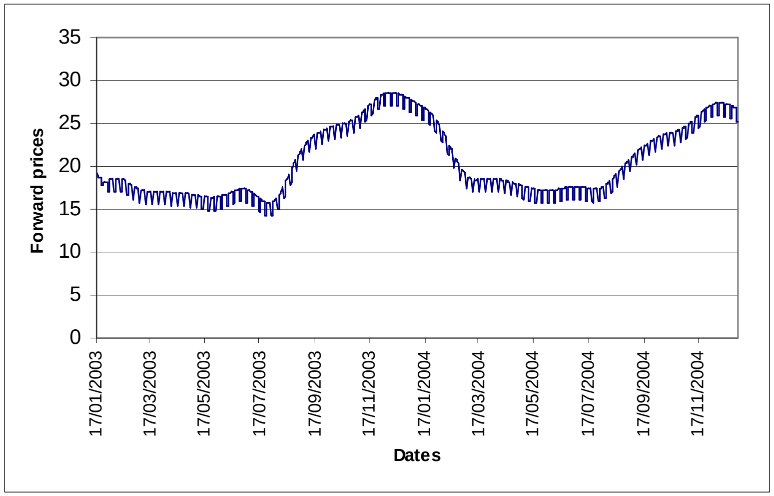
$$\widehat{p}_k^{ij} = \mathbb{P}(\hat{S}_{t_{k+1}} = s_{k+1}^j | \hat{S}_{t_k} = s_k^i)$$

- Dynamics : We consider the one factor toy-model given by

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

where $\sigma = 70\%$, $\alpha = 4$ and $t_k = k/n$.

- Future prices Real data (day 17/01/2003)



The contract parameters are $q_{\min} = 0$, $q_{\max} = 6$, $K_{t_k} = K = 20$ and $n = 30$ (1 year).

- Technical Parameters :

  - Quantization approach $n = 30$ (1 year), $N_k = \bar{N} = 100$

- Processor : Céléron, CPU 2,4 GHz. RAM 1,5 Go

▷ The function $(Q_{\min}, Q_{\max}) \mapsto P(s_0, (Q_{\min}, Q_{\max})$ is concave, piecewise affine on small triangles with integer vertices.

▷

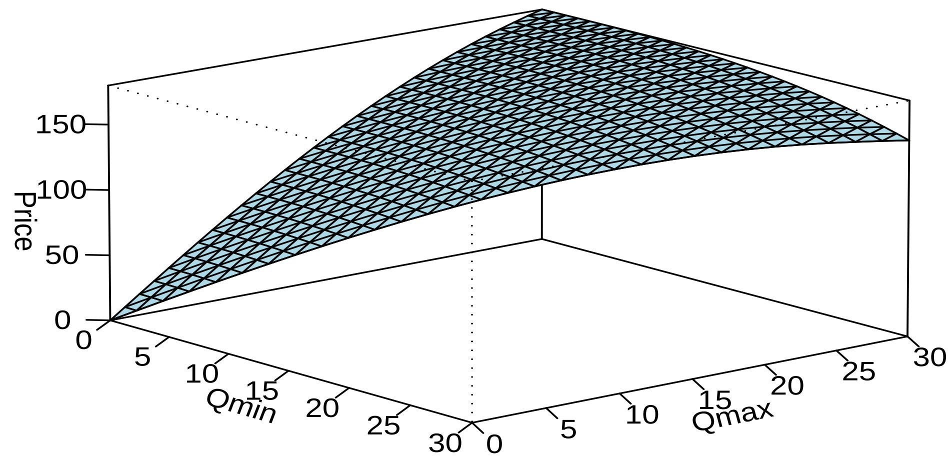


FIG. 8: PRICE SURFACE BY OPTIMAL QUANTIZATION AS A FUNCTION OF THE GLOBAL CONSTRAINTS, $n = 30$

4.4 Quantization *vs* L-S for Swing options (2006).

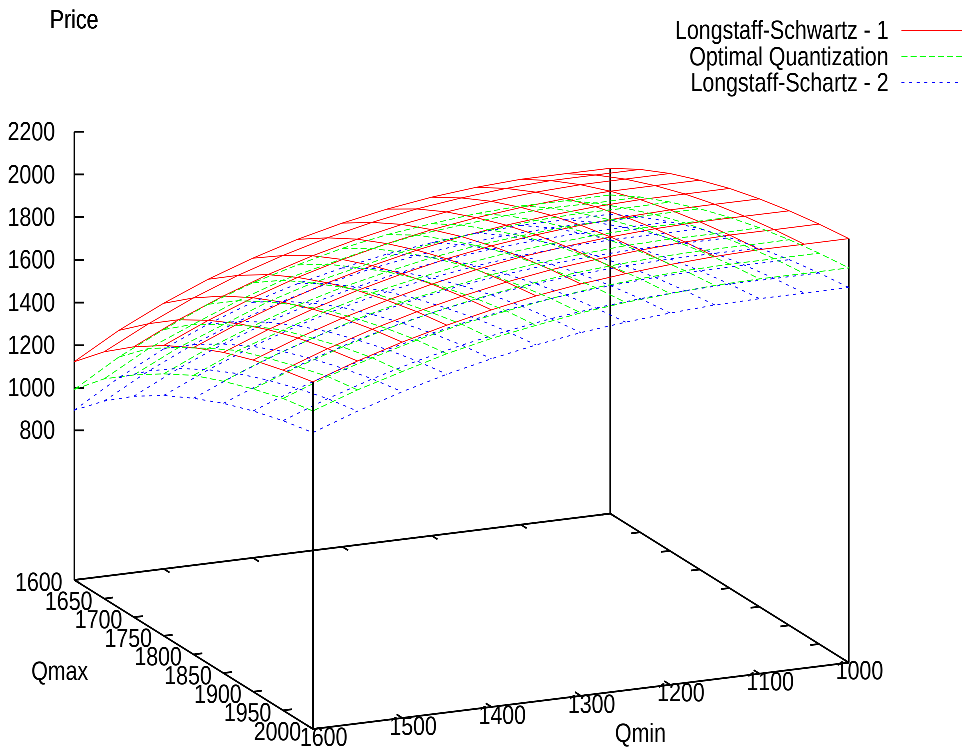


Fig. 9: Price Surface by L-S (dotted lines) and by Optimal Quantization (solid lines)

- First results :
  - 1 contract :

L-S	Quantization : Transitions + pricing	Quantization : Pricing alone
160 sec	38.5 sec	2.5 sec

  - 10 contracts :

L-S	Quantization
1600 sec	61 sec

- If less RAM available :
  - Quantization is unchanged
  - L-S slows down because the computers “swaps”. . .

5 Empirical rate of convergence

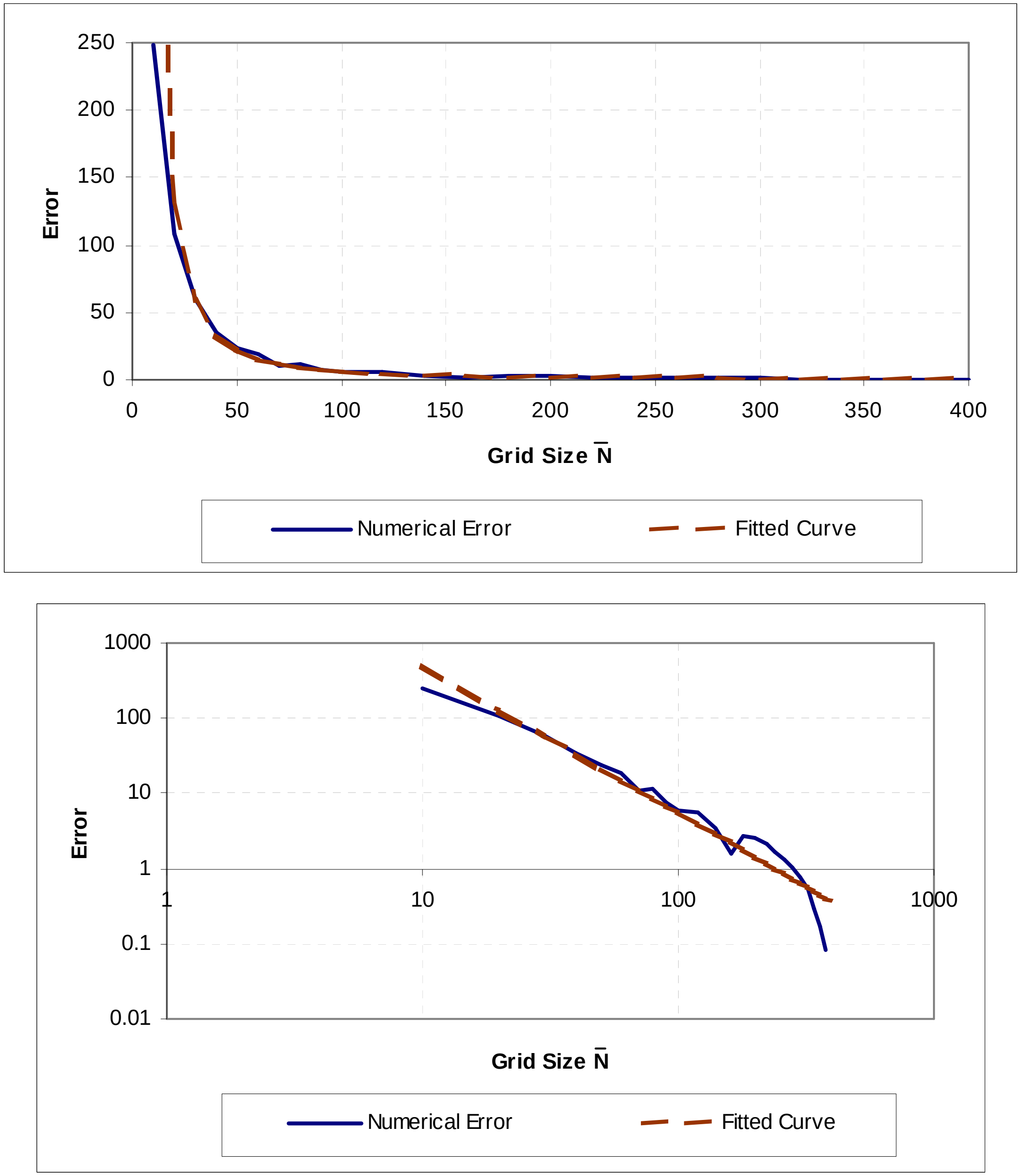


Fig. 10: Numerical convergence : the error as a function of $\bar{N}$

6 Numerical improvements

▷ Variance reduction (≈ “randomized quantization”, P.-Printems, *MCMA*, 2005) : X_k, $k \geq 1$, independent copies of X and $\widehat{X}_k$ (optimal) N-quantization of X_k.

$$\mathbb{E}F(X) \approx \mathbb{E}F(\widehat{X}) + \frac{1}{M}\sum_{k=1}^{M} X_k - \widehat{X}_k,$$

$$\begin{aligned} \mathrm{Var}\left(\frac{1}{M}\sum_{k=1}^{M} X_k - \widehat{X}_k\right) &= \frac{\|X-\widehat{X}\|_2^2 - (\mathbb{E}\,F(X) - \mathbb{E}\,F(\widehat{X}))^2}{M} \\ &\leq \frac{\|X-\widehat{X}\|_2^2}{M} \leq \frac{C_X}{MN^{\frac{1}{d}}}. \end{aligned}$$

QUESTION : Efficient simulation of $\widehat{X}$, given X ? Yes …

– in 1-dimension,

– for “product quantizers” in d-dimensions.

$\triangleright$ RICHARDSON-ROMBERG (R-R) EXTRAPOLATION.

– Let $F : \mathbb{R}^d \to \mathbb{R}$, twice differentiable functional with Lipschitz Hessian $D^2 F$.

– Let $(\widehat{X}^{(N)})_{N \geq 1}$ be a sequence of optimal quadratic quantizations. Then

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

– Under some assumptions [...]

$$\mathbb{E}|X - \widehat{X}|^3 = O(N^{-\frac{3}{d}}) \qquad \text{if } \ d \geq 2,$$

or $\mathbb{E}\,|X - \widehat{X}|^3 = O(N^{-\frac{3-\varepsilon}{d}})$, $\varepsilon > 0$, if $d = 2$.

– If furthermore, we make the conjecture that

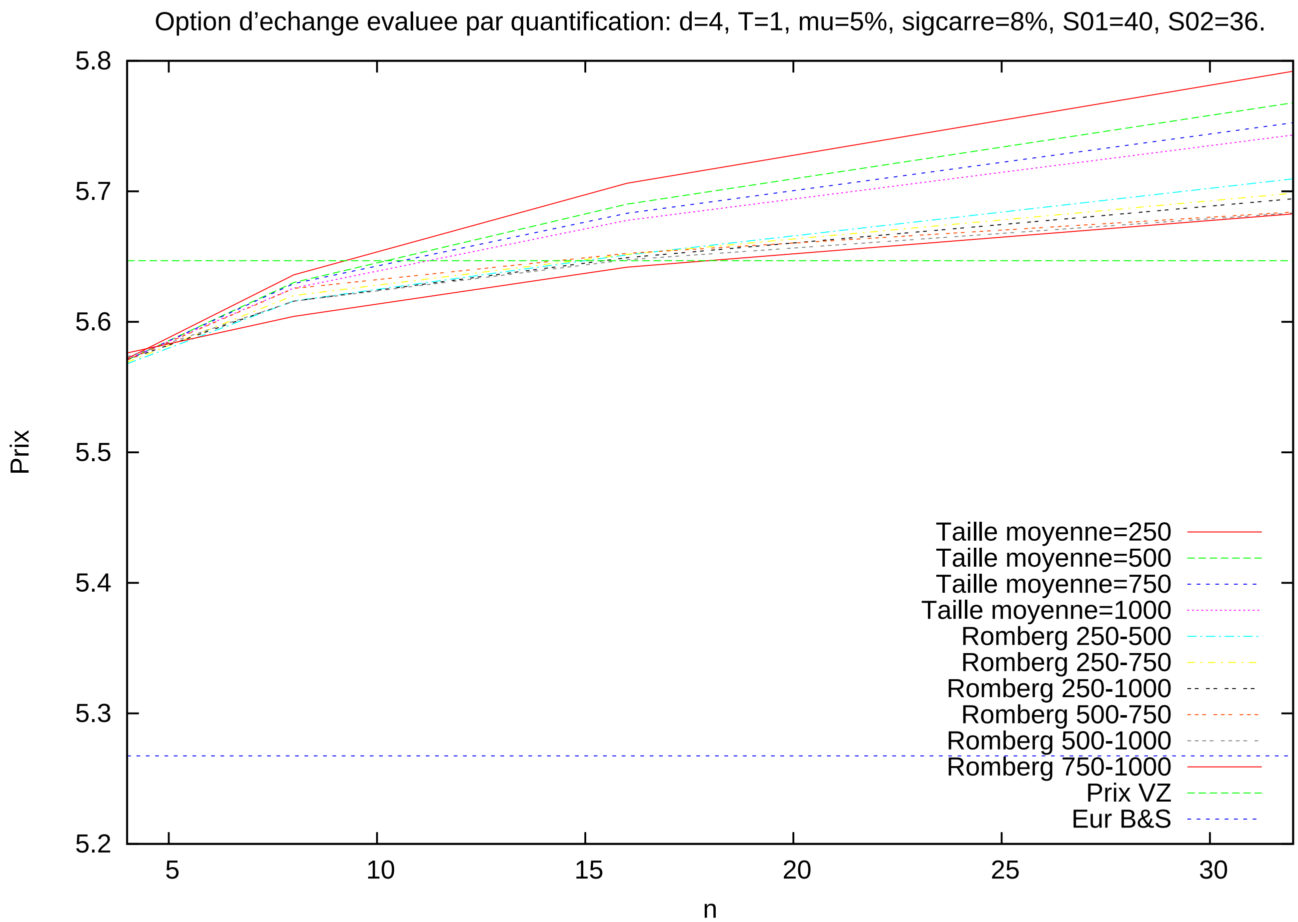
$$\mathbb{E}\left(D^2 F(\widehat{X}^{(N)}).(X - \widehat{X}^{(N)})^{\otimes 2}\right) = c_{F,X} N^{-\frac{2}{d}} + O(N^{-\frac{3}{d}})$$

It becomes possible to design an R-R extrapolation to compute $\mathbb{E}(F(X))$.

Let N_1 and N_2 be two sizes (*e.g.* $N_1 = N/2$ and $N_2 = N$).

Then linear combining (3) with N_1 and N_2,

$$\mathbb{E}(F(X)) = \frac{N_2^{\frac{2}{d}}\mathbb{E}(F(\widehat{X}^{(N_2)})) - N_1^{\frac{2}{d}}\mathbb{E}(F(\widehat{X}^{(N_1)}))}{N_2^{\frac{2}{d}} - N_1^{\frac{2}{d}}} + O\left(\frac{1}{(N_1 \wedge N_2)^{\frac{1}{d}}(N_2^{\frac{2}{d}} - N_1^{\frac{2}{d}})}\right)$$


Option d'echange evaluee par quantification: d=4, T=1, mu=5%, sigcarre=8%, S01=40, S02=36.
Prix
n
5.8
5.7
5.6
5.5
5.4
5.3
5.2
5
10
15
20
25
30
Taille moyenne=250
Taille moyenne=500
Taille moyenne=750
Taille moyenne=1000
Romberg 250-500
Romberg 250-750
Romberg 250-1000
Romberg 500-750
Romberg 500-1000
Romberg 750-1000
Prix VZ
Eur B&S