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?

▷ Let \( X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B} \otimes \mathbb{R}^d) \), \(|\.|\) Euclidean norm,

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

▷ 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 $\Gamma$-quantization

$$X \mapsto \hat{X}^\Gamma: \Omega \rightarrow \Gamma := \{x^1, \ldots, x^N\}.$$  

$$\hat{X}^\Gamma := \operatorname{Proj}_\Gamma(X)$$

where

$\operatorname{Proj}_\Gamma$ denotes the projection on $\Gamma$ 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 $X - \hat{X}^\Gamma$ and $\hat{X}^\Gamma$?

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

$$|X(\omega) - \hat{X}^\Gamma(\omega)| = \text{dist}(X(\omega), \Gamma) = \min_{1 \leq i \leq N} |X(\omega) - x^i|.$$  

▸ **Mean quadratic induced error** (or quadratic quantization error):

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

▸ **Distribution of $\hat{X}^\Gamma$**: weights associated to each $x^i$:

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

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

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

\[ D_X^N : (\mathbb{R}^d)^N \rightarrow \mathbb{R}_+ \]

\[ \Gamma = (x_1, \ldots, x^N) \mapsto \|X - \hat{X} \Gamma\|_2^2 = \mathbb{E} \left( \min_{1 \leq i \leq N} |X - x^i|^2 \right) \]

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 \(D^X_N\) reaches a minimum at an (optimal) quantizer \(\Gamma^{(N,*)}\) of full size \(N\) (if \(\text{card(supp(\mathbb{P}))} \geq N\)). One derives

\[ e_N(X, \mathbb{R}^d) := \inf\{\|X - \hat{X} \Gamma\|_2, \text{card(\Gamma)} \leq N, \Gamma \subset H\} = \|X - \hat{X} \Gamma^{(N,*)}\|_2 \]
\[ \|X - \hat{X}^{\Gamma(N,*)}\|_2 = \min \{\|X - Y\|_2, \ Y : \Omega \to H, \ \text{card}(Y(\Omega)) \leq N\}. \]

**Example** \(N = 1\) :

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

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

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

\[ \mathbb{E}|X|^r < +\infty \quad (0 < r < +\infty). \]

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

\[ e_{r,N}(X, E) := \inf \left\{ \|X - \hat{X}^\Gamma\|_r, \ \Gamma \subset E, \ \text{card}(\Gamma) \leq N \right\}. \]

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

- Non-Euclidean norms on $E = \mathbb{R}^d$ like $\ell^p$-norms, $1 \leq p \leq \infty$, etc.
- dispersion of compactly supported distribution: $r = \infty$
### 1.5 Stationary Quantizers

- **Distortion** $D^X_N$ is $\cdot$-differentiable at $N$-quantizers $\Gamma \in (\mathbb{R}^d)^N$ of full size:

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

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

    \[
    \nabla D^X_N(\Gamma) = 0 \iff \hat{X}^\Gamma = \mathbb{E} \left( X \mid \hat{X}^\Gamma \right)
    \]

    since

    \[
    \sigma(\hat{X}^\Gamma) = \sigma(\{X \in C_i(\Gamma)\}, i = 1, \ldots, N).
    \]

- **An optimal** quadratic quantizer $\Gamma$ is stationary

First by-product:

$$\mathbb{E} X = \mathbb{E} \hat{X}^\Gamma.$$
1.6 Numerical Integration and conditional expectation (I) : cubature formulae

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

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

▷ If $F$ is Lipschitz continuous, then

$$\left| \mathbb{E}F(X) - \mathbb{E}F(\hat{X}^\Gamma) \right| \leq [F]_{\text{Lip}} \|X - \hat{X}^\Gamma\|_1 \leq [F]_{\text{Lip}} \|X - \hat{X}^\Gamma\|_2$$

in fact

$$\|X - \hat{X}^\Gamma\|_1 = \sup_{[F]_{\text{Lip}} \leq 1} \left| \mathbb{E}F(X) - \mathbb{E}F(\hat{X}^\Gamma) \right| .$$

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

$$\left\| \mathbb{E}(F(X) \mid \hat{X}^\Gamma) - F(\hat{X}^\Gamma) \right\|_r \leq [F]_{\text{Lip}} \|X - \hat{X}^\Gamma\|_r$$
Assume $F$ is $C^1$ on $H$, $DF$ is Lipschitz continuous and the quantizer $\Gamma$ is a stationary.

Taylor expansion yields

$$\left| \mathbb{E} F(X) - \mathbb{E} F(\hat{X}^\Gamma) - \mathbb{E} \left( DF(\hat{X}^\Gamma).\left( X - \hat{X}^\Gamma \right) \right) \right| \leq [DF]_{\text{Lip}} \mathbb{E} \left| X - \hat{X}^\Gamma \right|^2$$
Assume $F$ is $C^1$ on $H$, $DF$ is Lipschitz continuous and the quantizer $\Gamma$ is a stationary. Taylor expansion $\implies$

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

since

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

so that

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

and

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

\[
\left\| \mathbb{E}(F(X) \mid \hat{X}^\Gamma) - F(\hat{X}^\Gamma) \right\|_r \leq [DF]_{\text{Lip}} \|X - \hat{X}^\Gamma\|_{2r}^2
\]

since

\[
\left| \mathbb{E}(F(X) \mid \hat{X}^\Gamma) - F(\hat{X}^\Gamma) \right| \leq [DF]_{\text{Lip}} \mathbb{E}(\|X - \hat{X}^\Gamma\|^2 \mid \hat{X}^\Gamma)
\]
1.7 Quantized approximation of $\mathbb{E}(F(X) \mid Y)$

Let $X, Y (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow H$ and $F : H \rightarrow \mathbb{R}$ a Borel functional. Let $\hat{X} = \hat{X}^\Gamma$ and $\hat{Y} = \hat{Y}^{\Gamma'}$ are (Voronoi) quantizations.

Natural idea $\mathbb{E}(F(X) \mid Y) \approx \mathbb{E}(F(\hat{X}) \mid \hat{Y})$. To what extend?

$$\mathbb{E}(F(X) \mid Y) = \varphi_F(Y).$$

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

$$\mathbb{E}(F(X) \mid Y) - \mathbb{E}(F(\hat{X}) \mid \hat{Y}) = \mathbb{E}(F(X) \mid Y) - \mathbb{E}(F(X) \mid \hat{Y}) + \mathbb{E}(F(X) - F(\hat{X}) \mid \hat{Y})$$

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

$$\| \mathbb{E}(F(X) \mid Y) - \mathbb{E}(\mathbb{E}(F(\hat{X}) \mid Y) \mid \hat{Y}) \|_2 \leq \| \varphi_F(Y) - \mathbb{E}(F(X) \mid \hat{Y}) \|_2 + \| F(X) - F(\hat{X}) \|_2 \leq \| \varphi_F(Y) - \varphi_F(\hat{Y}) \|_2 + \| F(X) - F(\hat{X}) \|_2$$

The last inequality follows from the very definition of conditional expectation given $\hat{Y}$. 
\[ \| \mathbb{E}(F(X) \mid Y) - \mathbb{E}(F(\hat{X}) \mid \hat{Y}) \|_2 \leq [F]_{\text{Lip}} \| X - \hat{X} \|_2 + [\varphi_F]_{\text{Lip}} \| Y - \hat{Y} \|_2. \]

- Non-quadratic case the above inequality remains valid provided \([\varphi_F]_{\text{Lip}}\) is replaced by \(2[\varphi_F]_{\text{Lip}}\).

- These are the ingredients for the proofs of both theorems for
  - Bermuda options (orders 0 & 1).
  - Swing options
1.8 **Vector Quantization rate** \( (H = \mathbb{R}^d) \)

\[ \nabla \textbf{Theorem (a) Asymptotic} \quad (\text{Zador, Kiefer, Bucklew & Wise, Graf & Luschgy al., from 1963 to 2000}). \]

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

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

\[ \nabla \text{(b) Non Asymptotic} \quad (\text{Luschgy-P., 2006}). \quad \text{Let} \quad r, \delta > 0. \quad \text{There exists a} \quad \text{universal constant} \quad C_{r,\delta} \in (0, \infty) \]

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

\[ \nabla \quad \text{The true value of} \quad \tilde{J}_{r,d} \quad \text{is unknown for} \quad d \geq 3 \quad \text{but} \quad (\text{Euclidean norm}) \]

\[
\tilde{J}_{r,d} \sim \sqrt{\frac{d}{2\pi e}} \approx \sqrt{\frac{d}{17,08}} \quad \text{as} \quad d \to +\infty.
\]
**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” . . .

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

For $d = 1$ up to 10 and $N = 1 \leq N \leq 5000$, 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 \( P_X(d\xi) = \varphi(\xi) d\xi \) with \( \log \varphi \) concave, then there is exactly one stationary quantizer. Hence

\[
\forall N \geq 1, \quad \text{argmin} D^X_N = \{ \Gamma^{(N)} \}.
\]

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

\begin{itemize}
  \item Voronoi cells: \( C_i(\Gamma) = [x^i - \frac{1}{2}, x^i + \frac{1}{2}] \), \( x^i + \frac{1}{2} = \frac{x^{i+1} + x^i}{2} \).
  \item Gradient: \( \nabla D^X_N(\Gamma) = 2 \left( \int_{x^i - \frac{1}{2}}^{x^i + \frac{1}{2}} (x^i - \xi) \varphi(\xi) d\xi \right) \)
    \quad \text{1} \leq i \leq N
\end{itemize}

Hessian: \( D^2(D^X_N)(\Gamma) = \ldots \ldots \) only involves \( \int_0^x \varphi(\xi) d\xi \) and \( \int_0^x \xi \varphi(\xi) d\xi \)
Thus if \( X \sim \mathcal{N}(0; 1) \) : only \( \text{erf}(x) \) and \( e^{-\frac{x^2}{2}} \) are needed.

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

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

\[
\mathbb{P}(X \in C_i(\Gamma^{(N)})), \ i = 1, \ldots N, \quad \text{and} \quad \|X - \hat{X}^{\Gamma^{(N)}}\|_2.
\]

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

\[ D^X_N(\Gamma) := \|X - \hat{X}^\Gamma\|_2^2 = \mathbb{E}(d_N(\Gamma, X)) \]

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

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

- $D^X_N$ is continuously differentiable at grids $\Gamma$ of full size $N$ and

\[ \frac{\partial D^X_N}{\partial x^i}(\Gamma) := \mathbb{E} \frac{\partial d_N}{\partial x^i}(\Gamma, X) = \int_{\mathbb{R}^d} \frac{\partial d_N}{\partial x^i}(\Gamma, \xi) \mathbb{P}_X(d\xi), \]

with a local gradient

\[ \frac{\partial d_N}{\partial x^i}(\Gamma, \xi) := 2(x^i - \xi)1_{\{\text{Proj}_\Gamma(x) = x^i\}}, \quad 1 \leq i \leq N. \]

- $\nabla D^X_N$ 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 $\delta_1, \ldots, \delta_t$ . . . .

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

**Stochastic Gradient Descent** Formally reads

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

**Grid updating:** $(t \rightsquigarrow t + 1): \Gamma(t) := \{x^{1,t}, \ldots, x^{N,t}\}$,

**Competition:** winner selection $i(t + 1) \in \arg\min_i |x^{i,t} - \xi^{t+1}|$

**Learning:**
\[
\begin{align*}
    x^{i(t+1),t+1} &:= \text{Homothety}(\xi^{t+1}, 1 - \delta_{t+1})(x^{i(t+1),t}) \\
    x^{i,t+1} &:= x^{i,t}, \ i \neq i(t + 1).
\end{align*}
\]
• **Heuristics** : $\Gamma^t \rightarrow \Gamma^* \in \text{argmin}(\text{loc}) \Gamma D_N^X(\Gamma)$ as $t \rightarrow \infty$.

• **Computation of the “companion parameters”** :
  
  - Weights $\pi_i^*, t = \mathbb{P}(\hat{X}^{\Gamma^*} = x^{i,*}), i = 1, \ldots, N$ :
    $$\pi_i^{t+1} := (1 - \delta_{t+1}) \pi_i^t + \delta_{t+1} 1_{\{i=i(t+1)\}} \text{ a.s. } \pi_i^* = \mathbb{P}(\hat{X}^{\Gamma^*} = x^{i,*}).$$

  - (Quadratic) Quantization error $D_N^X(\Gamma^*) = \|X - \hat{X}^{\Gamma^*}\|_2$ :
    $$D_N^{X,t+1} := (1 - \delta_{t+1}) D_N^{X,t} + \delta_{t+1} |x^{i(t+1),t} - \xi^{t+1}|^2 \text{ a.s. } D_N^X(\Gamma^*).$$

  Extra C.P.U. time cost $\approx 0$!

**CLVQ \equiv \text{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:
  \[
  \hat{X}^{\Gamma(t+1)} = \mathbb{E}(X | \hat{X}^{\Gamma(t)}), \quad \Gamma(0) \subset \mathbb{R}^d, \quad |\Gamma| = N.
  \]

- \( \Gamma(\ell) = \{x_1^{(\ell)}, \ldots, x_N^{(\ell)}\} \) being computed,
  \[
  x_i^{(\ell+1)} := \mathbb{E}(X^{\Gamma(\ell)} | X^{\Gamma(\ell)} \in C_i(\Gamma(\ell))) = \lim_{M \to \infty} \frac{\sum_{m=1}^{M} X_m \mathbf{1}_{\{X_m \in C_i(\Gamma(\ell))\}}}{|\{1 \leq m \leq M, X_m \in C_i(\Gamma(\ell))\}|}
  \]
  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 \( x = (x^1, \ldots, x^d) \in \mathbb{R}^d \)

\[
(x^1)^2 \geq \text{Rec}^2 \implies |x| \geq \text{Rec}
\]

\[
\vdots
\]

\[
(x^1)^2 + \cdots + (x^\ell)^2 \geq \text{Rec}^2 \implies |x| \geq \text{Rec}
\]

\[
\vdots
\]

▷ 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

- Traded risky assets: \( S_t = (S^1_t, \ldots, S^d_t) \quad t \in [0,T] \).

  with natural (augmented...) filtration \( \mathcal{F}^S = \mathcal{F}^S_t \) \( t \in [0,T] \).

- Discounted price: \( \tilde{S}^i_t = \frac{S^i_t}{S^0_t} = e^{-rt}S^i_t, \quad i = 1, \ldots, d \).

  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 \( \geq 0, \mathcal{F}^S \)-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 \( h_{tk} \) once, \( k = 0, \ldots, n \).

  usually with \( t_k = \frac{kT}{n}, \quad k = 0, \ldots, n \).
Examples:

▷ **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 $t = t_k = \frac{kT}{n}$, $k = 0, \ldots, n$.

  $$h_t = (S^1_t - K)^+ \text{ or } h_t = (K - S^1_t)^+.$$  

▷ **“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 $t = t_k = \frac{kT}{n}$, $k = 0, \ldots, n$.

- Example: Exchange American/Bermuda options (Villeneuve):

  $$h_t = (S^1_t - \lambda S^2_t)^+.$$
“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 $t = t_k = \frac{kT}{n}$, $k = 0, \ldots, n$.

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 \((X_k)_{0 \leq k \leq n}\) be a Markov structure process. with transition \( P_{k-1,k}(g)(x) = \mathbb{E}(g(X_{k+1} | X_k = x) \) such that

- \( \mathcal{F}_k^X = \mathcal{F}_{t_k}^S \)
- Risky asset vector satisfies
  \[ S_{t_k} = (S_{t_k}^1, \ldots, S_{t_k}^d) = G(X_k) \]
- Payoff process satisfies
  \[ h_{t_k} = h(k, X_k). \]
- Simulability: \((X_k)_{0 \leq k \leq n}\) can be simulated (at a reasonable cost).
• Typical structure processes (for American/Bermuda “Vanilla” options): 

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

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

\[ X_k := \begin{cases} 
(S_{tk}, \frac{1}{t_k}(S_0 + \cdots + S_{tk})) \\
(\bar{S}_{tk}, \frac{1}{t_k}(\bar{S}_0 + \cdots + \bar{S}_{tk})), \\
(S_{tk}, \max_{0 \leq i \leq k} S_{t_i}), \\
\text{etc}.
\end{cases} \]
3.3 Arbitrage and value function

**Step 1**

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

**Step 2** Backward induction based on the Markov property

Markov \(\Rightarrow\) Conditioning given \(\mathcal{F}_k^X = \text{Conditioning given } X_k\).

\[\mathcal{V}_k = v_k(X_k), \quad k = 0, \ldots, n.\]
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 $\Gamma_k$:

$$\hat{X}_k = \pi_k(X_k) \leftarrow X_k$$

where $\pi_k : \mathbb{R}^d \rightarrow \Gamma_k$, $\Gamma_k$ is a grid of size $N_k$, $\Gamma_k = \{x_{k1}^1, \ldots, x_{N_k}^k\} \subset \mathbb{R}^d$.

But loss of the Markov property...
Approximation 2: Markov approximation

Quantized obstacle: \( h(k, \hat{X}_k) \), \( k = 0 \ldots, n \).

The Markov property is forced: one defines \( \hat{V}_k \) by a backward induction

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

Again a Backward induction

\[
\hat{V}_k = \hat{v}_k(\hat{X}_k), \quad k = 0, \ldots, n.
\]
where

\[
\begin{align*}
\hat{v}_n(x_n^i) & = h(n, x_n^i), \quad i = 1, \ldots, N_n \\
\hat{v}_k(x_k^i) & = \max \left( h(k, x_k^i), \sum_{j=1}^{N_k} \hat{p}_{kj} \hat{v}_{k+1}(x_{k+1}^j) \right), \quad i = 1, \ldots, N_k \\
& \quad k = 1, \ldots, n - 1
\end{align*}
\]

**Numerical Task(s)**

Optimize and Compute *off-line*

- Task 1: (good) grids $\Gamma_k$ including the quantization error.

and

- Task 2: (accurate) quantized transitions $\hat{p}_{kj} := \frac{\mathbb{P}(\hat{X}_{k+1} = x_{k+1}^j, \hat{X}_k = x_k^i)}{\mathbb{P}(\hat{X}_k = x_k^i)}$. 

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

**INTERPRETATION**  
Global Transition operators approximation

Grids $\Gamma_k$ + quantized transitions $\hat{p}_{ij}^k$

\[
\hat{P}_{k-1,k}(x_k^i, dy) = \sum_j \hat{p}_{ij}^k \delta_{x_k^j} \\
\]

with

\[
\hat{P}_{k-1,k}(x_k^i, dy) \approx P_{k-1,k}(x, dy), \quad k = 1, \ldots, n.
\]
One dimensional case | Δt = 0.04 | 50 time layers
• 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 $\hat{p}_{ij}^k$ 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 $h(k, \cdot)$ are Lipschitz, the transitions $P_{k,k-1}$ are Lipschitz, the

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

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

$$\| \mathcal{V}_0 - \hat{v}_0(\hat{X}_0) \|_2 \leq C_{X,\varphi} \sum_{k=0}^{n} \| X_k - \hat{X}^{\Gamma_k}_k \|_2^2.$$
3.6 Optimal design of the quantization tree

**Idea**: optimal integral allocation problem

Item (a) of the theorem & Zador’s Theorem (non asymptotic version)

\[
\|\mathcal{V}_0 - \hat{v}_0(\hat{X}_0)\|_2 \leq C_{X,\varphi} \sum_{k=0}^{n} \|X_k - \hat{X}_{\Gamma_k}^\Gamma\|_2 \\
\leq C_{X,\varphi} C_\delta \sum_{k=0}^{n} \|X_k\|_{2+\delta} |\Gamma_k|^{-\frac{1}{\alpha}} \\
= C_{X,\varphi} C_\delta \sum_{k=0}^{n} \|X_k\|_{2+\delta} N_k^{-\frac{1}{\alpha}}
\]

Amounts to solving the

\[
\min_{N_0 + \cdots + N_n = N} \sum_{k=0}^{n} \|X_k\|_{2+\delta} N_k^{-\frac{1}{\alpha}}
\]
i.e. denoting the (upper) integral part of \( x \) by \( \lceil x \rceil \),

\[
N_k = \left[ \frac{\left( \| X_k \|_{2+\delta} \right)^{\frac{d}{d+1}}}{\sum_{0 \leq \ell \leq n} \left( \| X_\ell \|_{2+\delta} \right)^{\frac{d}{d+1}}} N \right], \quad k = 0, \ldots, n
\]

so that

\[
\| \nu_0 - \hat{v}_0(\hat{X}_0) \|_2 \leq C_{X,\varphi} C_\delta \left( \sum_{k=0}^{n} \left( \| X_k \|_{2+\delta} \right)^{\frac{d}{d+1}} \right)^{1-\frac{1}{d}} \tilde{N}^{-\frac{1}{d}}.
\]

with \( \tilde{N} = N_0 + \cdots + N_n \) (usually \( > N \)).
Examples:

- Brownian motion $X_k = W_{t_k}$: Then $\hat{W}_0 = 0$ and
  \[
  \|W_{t_k}\|_{2+\delta} = C_\delta \sqrt{t_k}, \quad k = 0, \ldots, n.
  \]

Hence $N_0 = 1$ and
\[
N_k \approx \frac{2(d+1)}{d+2} \left( \frac{k}{n} \right)^{\frac{d}{2(d+1)}} N, \quad k = 1, \ldots, n.
\]

\[
|\nu_0 - \nu_0(0)| \leq C_{W,\delta} \left( \frac{2(d+1)}{d+2} \right)^{1-\frac{1}{d}} \frac{n^{1+\frac{1}{d}}}{N^{\frac{1}{d}}} = O\left( \frac{n}{N^{\frac{1}{d}}} \right), \quad \tilde{N} = \frac{N}{n}.
\]

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

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

$$\|\nu_0 - \hat{v}_0(\hat{X}_0)\|_2 \leq C_{X,\delta} \frac{n^{1 + \frac{1}{d}}}{N^{\frac{1}{d}}} = C_{X,\delta} \frac{n}{\bar{N}^{\frac{1}{d}}} \quad \bar{N} = \frac{N}{n}.$$
3.7 Computing the quantized transitions $\hat{p}_{ij}^k$

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 CLV Q 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 $X^m = (X^m_0, X^m_1, \ldots, X^m_n)$, $m = 1, \ldots, M$

“launched” in the quantization tree

3.7.2 Alternative methods

- Fast tree quantization for Gaussian structure processes
  (Bardou-Bouthemy-P. (2006) for swing options[...]).
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:
  \[ \hat{\zeta}_n^k := \frac{n}{Tc^2(\hat{S}_{t_k})} \hat{E}_k \left( (\hat{v}_{k+1}^n (\hat{S}_{t_{k+1}}) - \hat{v}_k^n (\hat{S}_{t_k})) (\hat{S}_{t_{k+1}} - \hat{S}_{t_k}) \right). \]

- Similar formulae for the Euler scheme...

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

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

- Quantization Error:
  \[ \mathbb{E} \int_0^T |\zeta_u^n - \hat{\zeta}_u^n| \, du \leq C (1 + |s_0|) \frac{n^{\frac{3}{4}}}{(N/n)^{\frac{1}{4}}}. \]
4 Numerical experiments

4.1 Numerical experiments I : Exchange geometric options

- Exchange American options on geometric assets.
- Model : Standard 2\textit{d}-dim (B & S) model with \textit{non correlated} Brownian Motions (The most “hostile” to quantization...).
- Maturity : $T = 1$ year. Volatility : $\sigma_i = \frac{20\%}{\sqrt{d}}$, $i = 1, \ldots, d$.

- \textbf{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 :
  \[ \prod_{i=1}^{d} S_0^i = 40, \prod_{i=d+1}^{2d} S_0^i = 36 \text{ (in-the money)}, \mu_1 := 5\%, \mu_2 = 0, \ldots \]
  \[ \prod_{i=1}^{d} S_0^i = 36, \prod_{i=d+1}^{2d} S_0^i = 40 \text{ (out-of-the money)}, \mu_{d+1} := 0\%, \ldots \]
4.1.1 Results: Premium and $\delta$-hedge: 0-order scheme

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3 months</th>
<th>6 months</th>
<th>9 months</th>
<th>12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>$AM_{ref}$</td>
<td>4.4110</td>
<td>4.8969</td>
<td>5.2823</td>
<td>5.6501</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>Error (%)</td>
<td>Price</td>
<td>Error (%)</td>
</tr>
<tr>
<td>$d = 2$</td>
<td>4.4111</td>
<td>0.0023</td>
<td>4.8971</td>
<td>0.0041</td>
</tr>
<tr>
<td>$d = 4$</td>
<td>4.4076</td>
<td>0.08</td>
<td>4.9169</td>
<td>0.34</td>
</tr>
<tr>
<td>$d = 6$</td>
<td>4.4156</td>
<td>0.1</td>
<td>4.9276</td>
<td>0.63</td>
</tr>
<tr>
<td>$d = 10$</td>
<td>4.4317</td>
<td>0.47</td>
<td>4.9945</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Tab. 1: American Premium & Relative error. Different maturities and dimensions.
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

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3 months</th>
<th>6 months</th>
<th>9 months</th>
<th>12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>( AM_{ref} )</td>
<td>Price</td>
<td>Error (%)</td>
<td>Price</td>
<td>Error (%)</td>
</tr>
<tr>
<td>( d = 4 )</td>
<td>( AM_0 )</td>
<td>4.4076</td>
<td>0.08</td>
<td>4.9169</td>
</tr>
<tr>
<td></td>
<td>( AM_1 )</td>
<td>4.4058</td>
<td>0.1</td>
<td>4.8991</td>
</tr>
<tr>
<td>( d = 6 )</td>
<td>( AM_0 )</td>
<td>4.4156</td>
<td>0.1</td>
<td>4.9276</td>
</tr>
<tr>
<td></td>
<td>( AM_1 )</td>
<td>4.4099</td>
<td>0.02</td>
<td>4.8975</td>
</tr>
<tr>
<td>( d = 10 )</td>
<td>( AM_0 )</td>
<td>4.4317</td>
<td>0.47</td>
<td>4.9945</td>
</tr>
<tr>
<td></td>
<td>( AM_1 )</td>
<td>4.4194</td>
<td>0.19</td>
<td>4.8936</td>
</tr>
</tbody>
</table>

**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^1 S^2 - S^3 S^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 ($\times$)
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…]

\[ d = 5 \quad N = 2 \cdot 10^4 \quad n = 10 \]

- Design of the quantization tree (grid/weights): 3 seconds;
- (Premium+ \( \delta \)-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 $q_{t_k} \in [q_{\text{min}}, q_{\text{max}}]$ m$^3$ of natural gas at price $K_k$
- Global volume constraints $Q_{\text{min}} \leq q_0 + q_{t_1} + \cdots + q_{t_{n-1}} \leq Q_{\text{max}}$.

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

where the set of admissible daily purchased quantities is given by

\[
A_{Q_{\text{min}}, Q_{\text{max}}} = \left\{ (q_{t_k})_{0 \leq k \leq n-1}, q_{t_k} \in \mathcal{F}_{t_k}^S, q_{\text{min}} \leq q_{t_k} \leq q_{\text{max}}, Q_{\text{min}} \leq \sum_{0 \leq k \leq n-1} q_{t_k} \leq Q_{\text{max}} \right\}
\]

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)
\[ 0 \forall = \max Q, 30, 0, \forall 1 = \max b, 0 = \min b \]
It is a stochastic control problem ($r = 0$)

\[ P(t_k, S_{t_k}, Q_{t_k}) = \max\{q(S_{t_k} - K) + \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, Q_{t_k} + q)|S_{t_k})\}, \]

\[ q \in [q_{\min}, q_{\max}], Q_{t_k} + q \in [(Q_{\min} - (n - k)q_{\max})_+, (Q_{\max} - (n - k)q_{\min})_+] \} . \]

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

If \( \left( \frac{Q_{\max} - nq_{\min}}{q_{\max} - q_{\min}}, \frac{Q_{\min} - nq_{\min}}{q_{\max} - q_{\min}} \right) \in \mathbb{N} \times \mathbb{N} \), then the optimal control is bang-bang i.e. \( \{q_{\min}, q_{\max}\}\)-valued
Quantized Dynamic programming principle

Let $\hat{S}_{tk}$ be an (optimal) quantization of $S_{tk}$ taking values in $\Gamma_k := \{s_{k1}^1, \ldots, s_{kn}^{N_k}\}$, $k = 0, \ldots, n$.

\[
P(t_k, s_k^i, \hat{Q}_{tk}) = \max_{q \in A_k^{\hat{Q}_{tk}}} \left\{ q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{tk+1}, \hat{Q}_{tk} + q | \hat{S}_{tk} = s_k^i) \right\}
\]

\[
A_k^{\hat{Q}_{tk}} = \left\{ q \in \{q_{\text{min}}, q_{\text{max}}\}, \hat{Q}_{tk} + q \in [(Q_{\text{min}} - (n - k)q_{\text{max}})_+, (Q_{\text{max}} - (n - k)q_{\text{min}})_+] \right\}
\]

\[
P(T, s_T^i, \hat{Q}_T) = P_T(s_T^i, \hat{Q}_T), i = 1, \ldots, N_n.
\]

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

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

where

\[
\hat{p}_{k}^{ij} = \mathbb{P}(\hat{S}_{tk+1} = s_{k+1}^j | \hat{S}_{tk} = s_k^i)
\]
• **Dynamics:** We consider the one factor toy-model given by

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

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

• **Future prices** Real data (day 17/01/2003)

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

• **Technical Parameters:**

  - Quantization approach \( n = 30 \) (1 year), \( N_k = \bar{N} = 100 \)
• **Processor**: Céleron, CPU 2,4 GHz. RAM 1,5 Go
The function \((Q_{\text{min}}, Q_{\text{max}}) \mapsto P(s_0, (Q_{\text{min}}, Q_{\text{max}}))\) is concave, piecewise affine on small triangles with integer vertices.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{price_surface}
\caption{Price Surface by Optimal Quantization as a function of the global constraints, \(n = 30\)}
\end{figure}
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:

<table>
<thead>
<tr>
<th>L-S</th>
<th>Quantization</th>
<th>Quantization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transitions + pricing</td>
<td>Pricing alone</td>
</tr>
<tr>
<td>160 sec</td>
<td>38.5 sec</td>
<td>2.5 sec</td>
</tr>
</tbody>
</table>

• 10 contracts:

<table>
<thead>
<tr>
<th>L-S</th>
<th>Quantization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1600 sec</td>
<td>61 sec</td>
</tr>
</tbody>
</table>

• 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 $\tilde{N}$
6 Numerical improvements

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

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

\[
\text{Var} \left( \frac{1}{M} \sum_{k=1}^{M} X_k - \hat{X}_k \right) = \frac{\|X - \hat{X}\|_2^2 - (\mathbb{E} F(X) - \mathbb{E} F(\hat{X}))^2}{M}
\]

\[
\leq \frac{\|X - \hat{X}\|_2^2}{M} \leq \frac{C_X}{MN^{\frac{1}{d}}}
\]

QUESTION: Efficient simulation of \( \hat{X} \), given \( X \)? Yes . . .

– in 1-dimension,

– for “product quantizers” in \( d \)-dimensions.
Richardson-Romberg (R-R) extrapolation.

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

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

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

(3)

- Under some assumptions [...] \[
\mathbb{E}|X - \hat{X}|^3 = O(N^{-3 \over d}) \quad \text{if} \quad d \geq 2,
\]

or \[
\mathbb{E}|X - \hat{X}|^3 = O(N^{-3 - \varepsilon \over d}), \quad \varepsilon > 0, \quad \text{if} \quad d = 2.
\]

- If furthermore, we make the conjecture that

$$
\mathbb{E} \left( D^2 F(\hat{X}^{(N)}). (X - \hat{X}^{(N)}) \otimes^2 \right) = c_{F,X} N^{-2 \over d} + O(N^{-3 \over d})
$$
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 $N_2$,

$$
\mathbb{E}(F(X)) = \frac{N_2^2 \mathbb{E}(F(\hat{X}^{(N_2)})) - N_1^2 \mathbb{E}(F(\hat{X}^{(N_1)}))}{N_2^2 - N_1^2} + O\left(\frac{1}{(N_1 \land N_2)^{\frac{1}{d}}(N_2^\frac{2}{d} - N_1^\frac{2}{d})}\right)
$$
Option d'échange évaluee par quantification: $d=4$, $T=1$, $\mu=5\%$, $\sigma_{carre}=8\%$, $S_01=40$, $S_02=36$. 

**THE ROMBERG IMPACT ON THE PREMIUM COMPUTATION**