Quantification fonctionnelle de processus stochastiques et applications

Gilles Pagès

gilles.pages@upmc.fr

Université PARIS 6

Laboratoire de Probabilités et Modèles aléatoires
UMR-CNRS 7599

Séminaire de mathématiques appliqées

Collège de France

13 Novembre 2009

Functional Quantization of stochastic processes and applications

Gilles Pagès

gilles.pages@upmc.fr

UPMC

Laboratoire de Probabilités et Modèles aléatoires
UMR-CNRS 7599

Séminaire de mathématiques appliquées

Collège de France

13 Novembre 2009

What is (quadratic) Functional Quantization ?

$\triangleright$ $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow H$, $(H, (\,.\,|\,.\,))$ separable Hilbert space

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

$\triangleright$ When $H = \mathbb{R}$, $\mathbb{R}^d$ $\equiv$ Vector Quantization of a random vector X.

[Old story stting in the the 1950's with many contributors, see *IEEE on Inf. Theory*, 1982, Gersho-Gray eds]

$\triangleright$ When $H = L^2_T := L^2([0,T], dt)$ $\equiv$ Functional Quantization of a process $X = (X_t)_{t\in[0,T]}$. [Not so old story]

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

using

$\triangleright$ N-quantizer (or N-codebook) :

$$\alpha := \{\alpha_1, \ldots, \alpha_N\} \subset H.$$

– When $H = \mathbb{R}^d$, each α_i is a vector of $\mathbb{R}^d$.

– When $H = L^2_T$, each $\alpha_i = (t \in [0, T] \mapsto \alpha_i(t))$ is a (class) of functions.

$\triangleright$ Discretization by α-quantization

$$X \rightsquigarrow \widehat{X}^{\alpha} : \Omega \to \alpha := \{\alpha_1, \ldots, \alpha_N\}.$$

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

where

$$\boxed{\mathrm{Proj}_{\alpha} \text{ denotes the projection on } \alpha \text{ following the nearest neighbour rule.}}$$

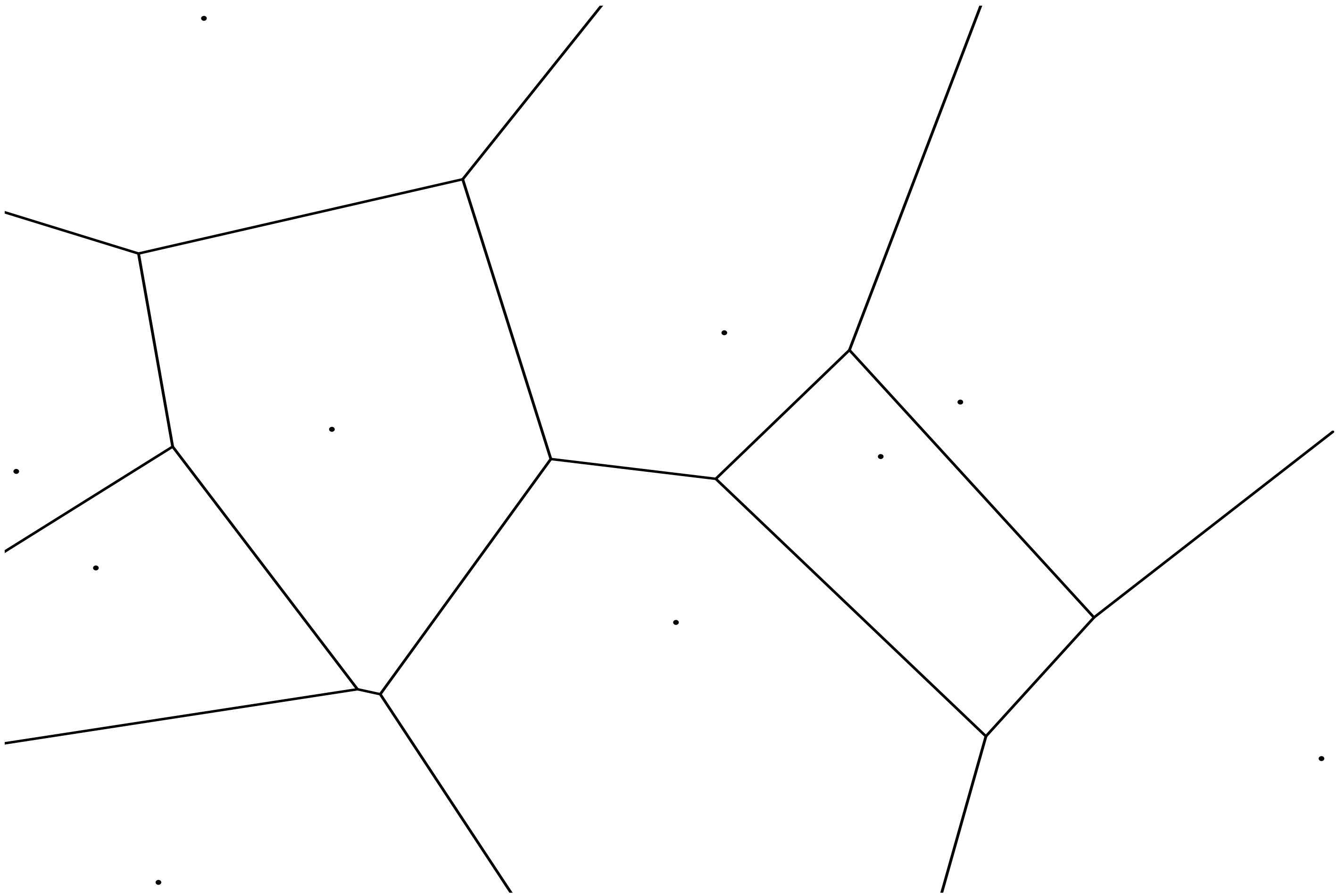


FIG. 1: A 2-dimensional 10-quantizer $\alpha = \{\alpha_1, \ldots, \alpha_{10}\}$ and its Voronoi diagram. . .

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

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

$$|X(\omega) - \widehat{X}^{\alpha}(\omega)|_{_H} = \mathrm{dist}_{_H}(X(\omega), \alpha) = \min_{1 \le i \le N} |X(\omega) - \alpha_i|_{_H}.$$

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

$$e_{_N}(X, H, \alpha) = \|X - \widehat{X}^{\alpha}\|_{_2} = \sqrt{\mathbb{E}\left(\min_{1 \le i \le N} |X - \alpha_i|^2_{_H}\right)}.$$

▷ Distribution of $\widehat{X}^{\alpha}$: weights associated to each α_i :

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

where $C_i(\alpha)$ denotes the Voronoi cell of α_i (w.r.t. α) defined by

$$C_i(\alpha) := \left\{\xi \in H : |\xi - \alpha_i|_{_H} = \min_{1 \le j \le N} |\xi - \alpha_j|_{_H}\right\}.$$

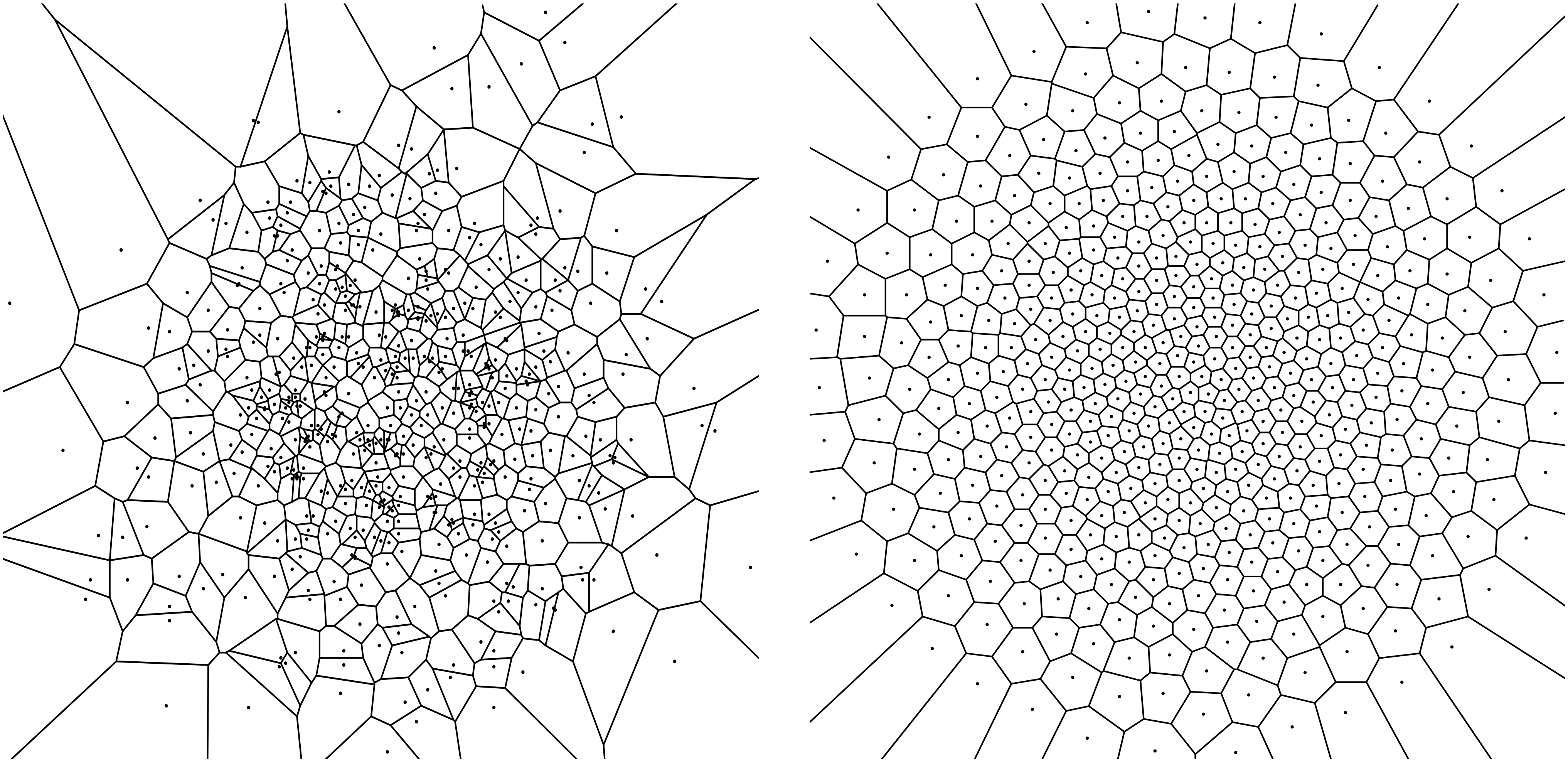


Fig. 2: Two N-quantizers related to $\mathcal{N}(0; I_2)$ of size $N = 500\ldots$

(with J. Printems)

Which one is the best ?

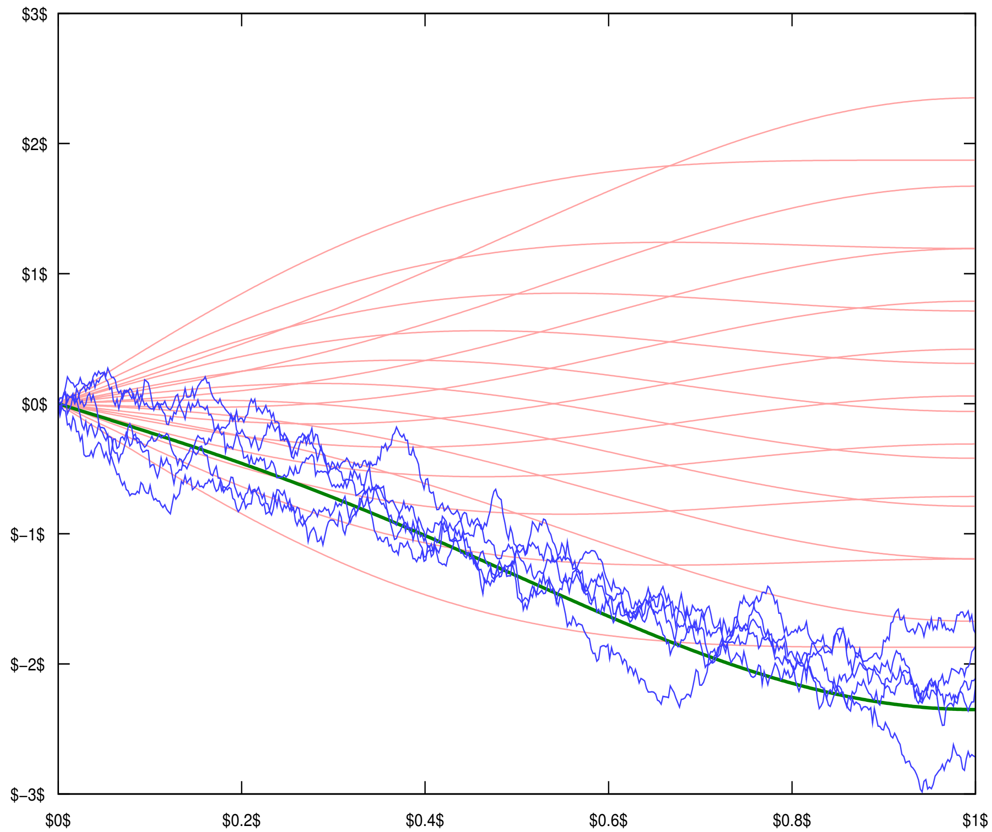


FIG. 3: A $N = 20$-quantizers of Brownian motion *vs* some Brownian paths. . . . . .

(with S. Corlay)

W is Gaussian process with independent increments

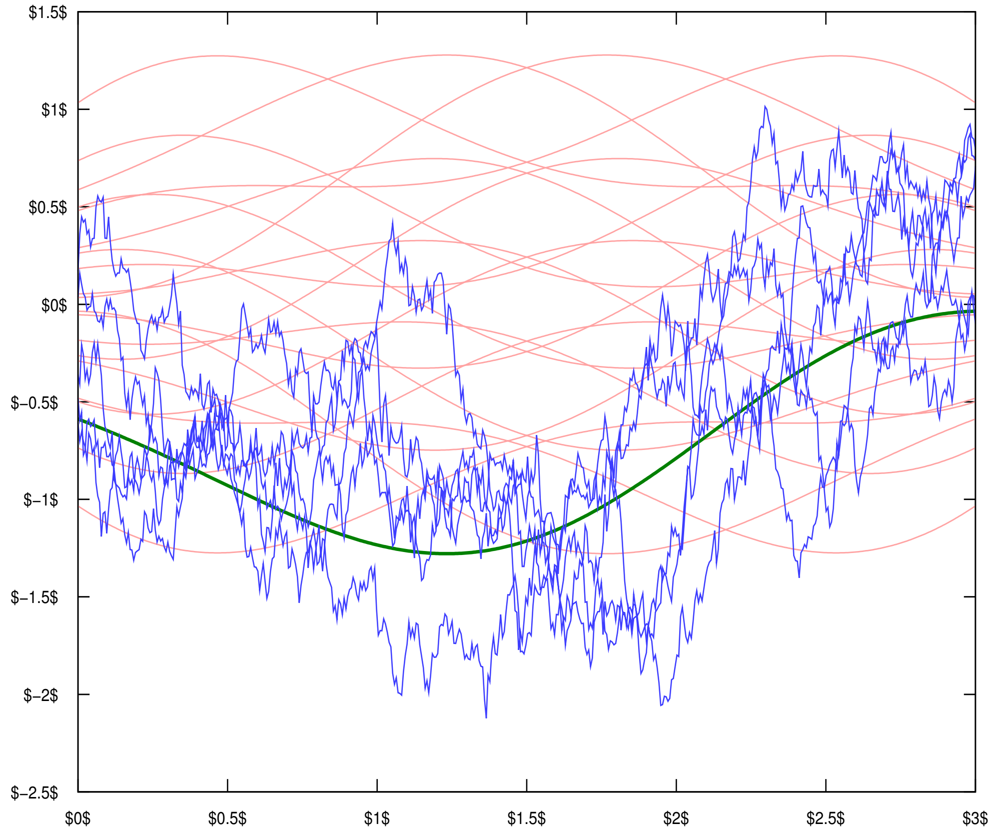


FIG. 4: A $N = 20$-quantizers of a stationary Ornstein-Uhlenbeck process *vs* some paths……

(with S. Corlay)

$$X_t = \int_{-\infty}^{t} e^{-(t-s)} dW_s \quad || \quad dX_t = -X_t dt + dW_t, \; X_0 \sim \mathcal{N}(0; \frac{1}{2})$$

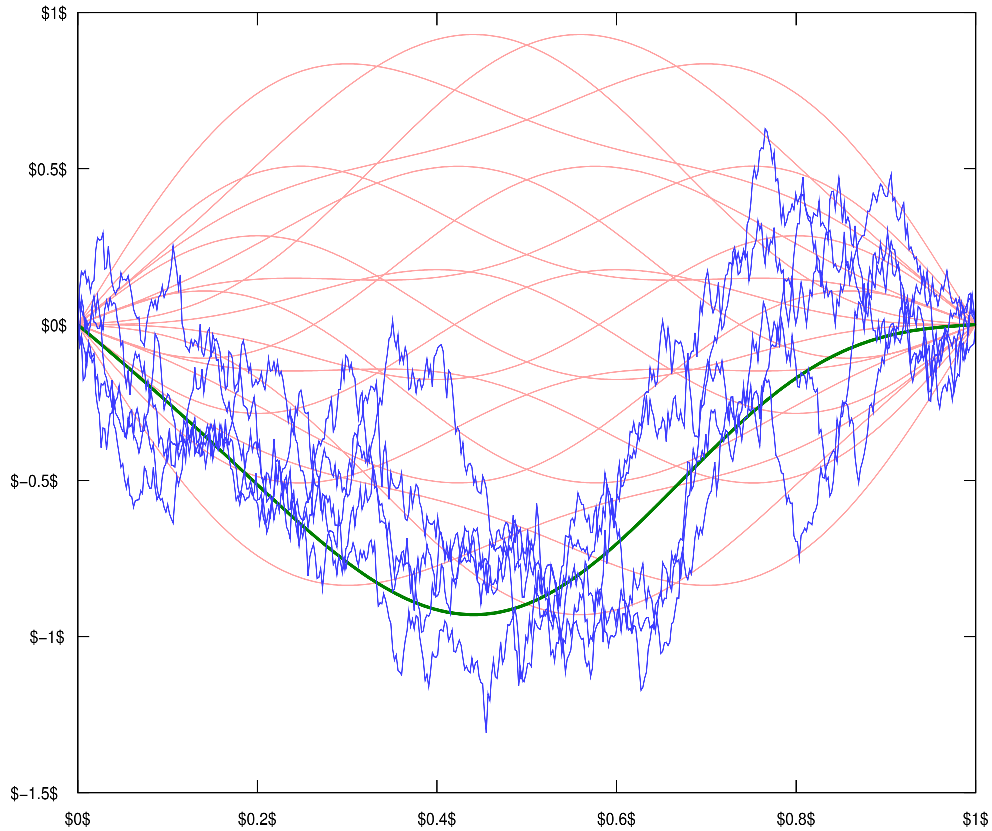


FIG. 5: A $N = 20$-quantizers of Brownian bridge *vs* some paths......

(with S. Corlay)

$$X_t = W_t - tW_1, \ t \in [0, 1]$$

non Gaussian diffusion processes ? etc.

Some questions

$\triangleright$ What is the connection between blue chaotic lines and pink smooth lines ?

$\triangleright$ How to get the pink smooth lines from the blue chaotic lines ?

$\triangleright$ Can we replace the blue chaotic lines by the pink smooth lines (for numerics, in a SDE or in a $SPDE$) ?

$\triangleright$ Can we take advantage of the pink smooth lines to simulate the blue chaotic lines ?

Optimal (Quadratic) Quantization

The quadratic distortion (squared quadratic quantization error)

$$\begin{aligned} D^X_N \quad &: \quad H^N \longrightarrow \mathbb{R}_+ \\ \alpha = (\alpha_1, \ldots, \alpha_N) &\longmapsto \|X - \widehat{X}^{\alpha}\|_2^2 = \mathbb{E}\left(\min_{1\le i\le N} |X - \alpha_i|^2_H\right) \end{aligned}$$

is lower semi-continuous for the (product) weak topology on H^N.

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

$$\boxed{D^X_N \text{ reaches a minimum at an (optimal) quantizer } \alpha^{(N,*)}}$$

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

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

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

Example $(N = 1)$:

Optimal 1-quantizer $\alpha = \{\mathbb{E}\, X\}$ and $e_1(X, H) = \sqrt{\mathbb{E}\,|X|^2 - |\mathbb{E}\, X|^2}$.

Extensions to the $L^r(\mathbb{P})$-quantization of Radon random variables

$\rhd$ $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (E, \|\ \ \|_E)$ separable Banach space

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

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

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

$\triangleright$ **Examples** : Non-Euclidean norms on $E = \mathbb{R}^d$, $E = L^p_T := L^p([0,T], dt)$, $1 \le p < \infty$, $E = \mathcal{C}([0,T]), \|\,.\,\|_{\sup}$, etc.

$\triangleright$ Existence of an optimal quantizer holds true for reflexive Banach spaces (see Pärna (90)) and $E = L^1_T$, but may fail even when $N = 1\ldots$

$\triangleright$ Recent existence results, see Graf-Luschgy-P. (2006, *J. of Approx.*).

Stationary Quantizers

▷ Distorsion D_N^X is $|.|_H$-differentiable at N-quantizers $\alpha \in H^N$ of full size :

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

▷ **Definition :** If $\alpha \subset H^N$ is a zero of $\nabla D_N^X(\alpha)$, then α is called a *stationary quantizer* (or self-consistent quantizer).

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

since

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

▷ An optimal quantizer α is stationary

(First by-product : $\mathbb{E}X = \mathbb{E}\widehat{X}^\alpha$).

Numerical Integration/Conditional expectation (I) : cubature formulae

Let $F : H \longrightarrow \mathbb{R}$ be a functional and let $\alpha \subset H$ be an N-quantizer.

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

▷ If F is Lipshitz continuous,then

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

in fact

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

Likewise

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

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

Taylor expansion yields

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

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

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

since

$$\mathbb{E}\left(DF(\widehat{X}^{\alpha}).(X - \widehat{X}^{\alpha})\right) = \mathbb{E}\left(DF(\widehat{X}^{\alpha}).\mathbb{E}(X - \widehat{X}^{\alpha} \,|\, \widehat{X}^{\alpha})\right) = 0.$$

so that

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

Likewise

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

$\triangleright$ The key for numerical applications : F Lipschitz continuous

$$\mathbb{E}(F(X)\,|\,Y) = \varphi_{_F}(Y) \quad \varphi \text{ Lipschitz continuous.}$$

Then, if $\widehat{X}$ and $\widehat{Y}$ are quantizations of X and Y

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

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

$\triangleright$ THEOREM (Zador et al., from 1963 to 2000) Let $X \in L^{2+}(\mathbb{P})$ and $\mathbb{P}_X(d\xi) = \varphi(\xi)\,d\xi \overset{\perp}{+} \nu(d\xi)$. Then

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

$\triangleright$ The true value of $\widetilde{J}_{2,d}$ is unknown for $d \geq 3$ but (Euclidean norm)

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

CONCLUSIONS : • The curse of dimensionality of course...

• The same result holds with any $L^r(\mathbb{P})$-quantization with $r \in (0, \infty)$ replacing 2 (including $\widetilde{J}_{r,d} \sim \widetilde{J}_{2,d}$ as $d \to \infty$).

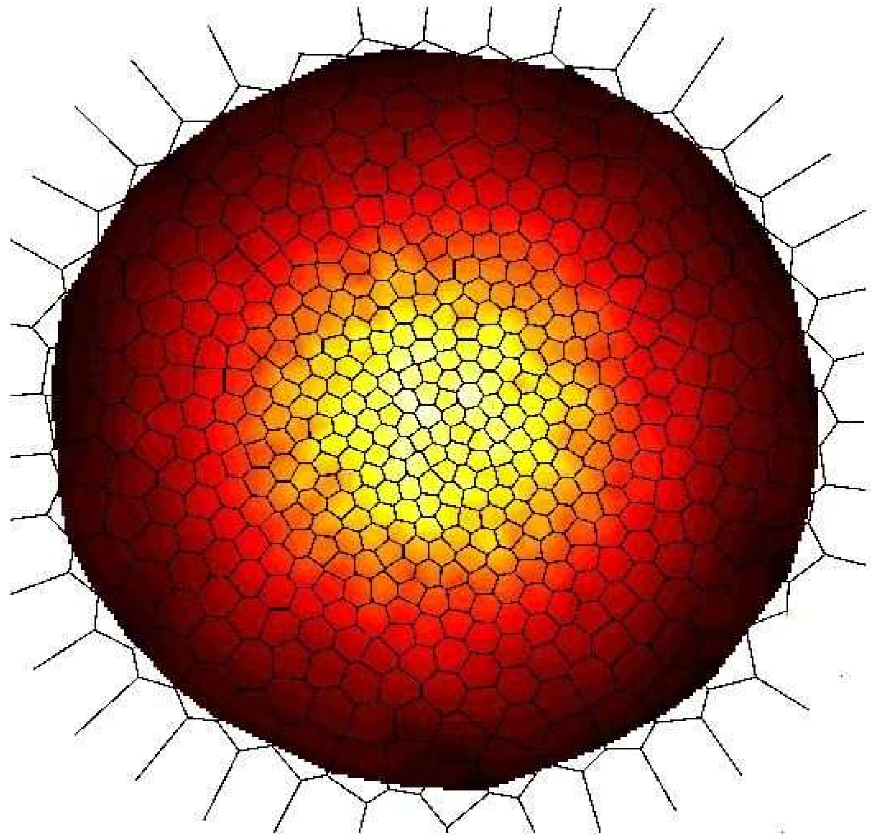


FIG. 6: An N-quantization of $X \sim \mathcal{N}(0; I_2)$ with coloured weights : $\mathbb{P}(X \in C_a(\alpha))$, $a \in \alpha$

(with J.Printems)

$\triangleright$ Local inertia : $a \longmapsto \mathbb{E}|X-a|^2 \mathbf{1}_{X \in C_a(\alpha)} \approx$ Constant

The 1-dimension...

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

$$\forall\, N \geq 1, \qquad \mathrm{argmin} D_N^X = \{\alpha^{(N)}\}$$

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

▷ Voronoi cells : $C_i(\alpha) = [\alpha_{i-\frac{1}{2}}, \alpha_{i+\frac{1}{2}}[$, $\alpha_{i+\frac{1}{2}} = \frac{\alpha_{i+1}+\alpha_i}{2}$.

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

Hessian : $D^2(D_N^X)(\alpha) =$ ...... 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 $\mathrm{erf}(x)$ 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 optimal N-quantizers and companion parameters :

$$\alpha^{(N)} = (\alpha_1^{(N)}, \ldots, \alpha_N^{(N)})$$

and

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

Download at our WEBSITE :

$\boxed{\text{www.quantize.maths-fi.com}}$

$\triangleright$ For $d = 1$ up to 10? Also available for Gaussian $\mathcal{N}(0, I_d)$ $(1 \leq N \leq 4\,000)$.

How? Stochastic optimization methods, see further on...

Optimal Functional Quantization (of the Brownian motion)

▷ $H = L^2_T := L^2([0,T], dt)$, $(f|g) = \int_0^T f(t)g(t)dt$, $|f|_{L^2_T} = \sqrt{(f|f)}$.

▷ The Brownian motion W : centered Gaussian process with covariance operator $C_W(f) : f \longmapsto (t \mapsto \int_{[0,T]^2} (s \wedge t) f(s) ds)$.

▷ Diagonalization of C_W yields the Karhunen-Loève system ($\equiv$ CPA of W)

$$\boxed{e_n^W(t) = \sqrt{2T}\sin\left((n-\frac{1}{2})\pi\frac{t}{T}\right), \qquad \lambda_n = \left(\frac{T}{\pi(n-\frac{1}{2})}\right)^2, \; n \geq 1}$$

$$W_t \overset{L^2_T}{=} \sum_{n\geq 1}(W|e_n^W)_2\, e_n^W(t) = \sum_{n\geq 1}\sqrt{\lambda_n}\,\xi_n\, e_n^W(t)$$

$$\xi_n \sim \mathcal{N}(0;1), \quad n \geq 1, \quad \text{i.i.d.}$$

$\triangleright$THEOREM (Luschgy-P., JFA (2002) and AP (2003)) Let α^N, $N \geq 1$, be a sequence of optimal N-quantizers.

$\triangleright$ $\alpha^N = (\alpha_1^N, \cdots, \alpha_N^N) \subset \operatorname{span}\{e_1^W, \ldots, e_{d(N)}^W\}$ with

$$d(N) \gtrsim \log N/2 \quad [\text{Conjecture}: d(N) \sim \log N].$$

$\triangleright$ $e_{_N}(W, L_{_T}^2) = \|W - \widehat{W}^{\alpha^N}\|_{_2} \sim \dfrac{\sqrt{2}}{\pi}\dfrac{1}{\sqrt{\log N}}$. $\quad (\frac{\sqrt{2}}{\pi} = \sqrt{0.2026...})$

$\triangleright$ Reduction to finite dimension (Pythagore)

$$(\mathcal{O}_N)\left\{\begin{array}{rcl} \|W - \widehat{W}^{\alpha^N}\|_2^2 & = & \|Z - \widehat{Z}^{\beta(N)}\|_2^2 + \sum_{k \geq d(N)+1} \lambda_k \\ Z \sim \bigotimes_{k=1}^{d(N)} \mathcal{N}(0, \lambda_k) & \& & \|Z - \widehat{Z}^{\beta(N)}\|_{_2} = e_N(Z, \mathbb{R}^{d(N)}) \end{array}\right.$$

Then $$\boxed{\widehat{W}^{\alpha^N} = \sum_{k=1}^{d(N)} (\widehat{Z}^{\beta(N)})_k e_k^W.}$$

Optimal Quadratic Functional Quantization of Gaussian processes

THEOREM (Luschgy-P., *JFA* (2002) and *AP* (2003)) Let $X = (X_t)_{t\in[0,1]}$ be a Gaussian process with K-L eigensystem $(\lambda_n^X, e_n^X)_{n\geq 1}$. Let α^N, $N \geq 1$, be a sequence of quadratic optimal N-quantizers for X. If

$$\lambda_n^X \sim \frac{\kappa}{n^b} \quad \text{as } n \to \infty \qquad (b > 1).$$

$\triangleright$ $\alpha^N = (\alpha_1^N, \cdots, \alpha_N^N) \subset \text{span}\{e_1^X, \ldots, e_{d^X(N)}^X\}$ with

$$d^X(N) \gtrsim \frac{1}{b^{1/(b-1)}} \frac{2}{b} \log N \quad [\text{Conjecture}: d^X(N) \sim \frac{2}{b} \log N].$$

$\triangleright$ $e_N(X, L^2_{[0,1]}) = \|X - \widehat{X}^{\alpha^N}\|_2 \sim \sqrt{\kappa}\left(\frac{b^b}{(b-1)^{b-1}}\right)^{\frac{1}{2}} \frac{1}{(2\log N)^{\frac{b-1}{2}}}$.

$\triangleright$ Extensions to $\lambda_n^X \begin{pmatrix} \leq \\ \geq \end{pmatrix} \varphi(n)$, φ regularly varying, index $-b \leq -1$.

Applications to classical (centered) Gaussian processes

Sharp rates for $e_N(X, L^2_T)$ available for

- Brownian bridge, Ornstein-Uhlenbeck process, Gaussian diffusions (same rate).

– Fractional Brownian motion with Hurst constant $H \in (0, 1)$

$$e_N(W^H, L^2_T) \sim \frac{c_2}{(\log N)^H}.$$

– Brownian sheet, m-fold integrated Brownian motion, etc.

Extensions to $p \neq 2$ (methods are different)

– Brownian motion and fractional Brownian motion : Dereich-Scheutzow (2005) based on self-similarity properties, random quantization, small balls

$$e_{N,r}(W^H, L^p_T) \sim \frac{c_p}{(\log N)^H}.$$

Optimal quadratic Functional Quantization (of W) : numerical aspects $(T = 1)$

▷ Good news : $(\mathcal{O}_N)$ is a finite dimensional optimization problem.

▷ Bad news : $\lambda_1 = 0.40528...$ and $\lambda_2 = 0.04503... \approx \lambda_1/10\,!!!$

▷ A way out :

$$(\mathcal{O}_N) \equiv \begin{cases} N\text{-optimal quantization of } \displaystyle\bigotimes_{k=1}^{d(N)} \mathcal{N}(0,1) \\ \text{for the covariance norm } |(z_1, \ldots, z_{d(N)})|^2 = \sum_{k=1}^{d(N)} \lambda_k z_k^2. \end{cases}$$

$\triangleright$ A toolbox (see *e.g.* P.-Printems, *MCMA*, 2003, book by Gersho & Gray (97), Mrad & Ben Hamida (04), etc) :

– *Competitive Learning Vector Quantization* :

Recursive stochastic approximation gradient descent

based on the representation of the gradient of the distortion *i.e.*

$$\nabla D_N^Z(\alpha) = \mathbb{E}(\nabla D_N^Z(\alpha, \zeta)),\ \zeta \sim \mathcal{N}(0, I_d), \quad \zeta_t \sim \zeta, \quad i.i.d.$$

so that

$$\begin{aligned}(\alpha^N)(t+1) &= (\alpha^N)(t) - \frac{c}{t+1}\nabla D_N^Z((\alpha^N)(k), \zeta_{t+1}), \quad (\alpha^N)(0) \subset \mathbb{R}^d \\ &= \text{nearest neighbor search} + \text{Dilatation}_{\zeta_{t+1}, 1-\frac{c}{t+1}}\text{ (winner)}\end{aligned}$$

– *“Lloyd I procedure”* : randomized fixed point procedure based on the stationarity equality :

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

▷ $\alpha(t) = \{x_1^{(t)}, \ldots, x_N^{(t)}\}$ being computed,

$$\begin{aligned} x_i^{(t+1)} &:= \mathbb{E}(X \,|\, X^{\alpha(t)} \in C_i(\Gamma(\ell))), \quad i = 1, \ldots, N \\ &= \lim_{M \to \infty} \frac{\sum_{m=1}^{M} X_m \mathbf{1}_{\{X_m \in C_i(\alpha(t))\}}}{|\{1 \leq m \leq M,\ X_m \in C_i(\alpha(t))\}|} \end{aligned}$$

based on repeated nearest neighbour searches.

Then $\alpha(t+1) = \{x_i(t+1)\},\ i = 1, \ldots, N\}$, etc.

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

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

Running record dist to 0 := Rec.

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

$$
\begin{aligned}
(x^1)^2 \geq Rec^2 \quad &\Longrightarrow \quad |x| \geq Rec \\
&\vdots \\
(x^1)^2 + \cdots + (x^\ell)^2 \geq Rec^2 \quad &\Longrightarrow \quad |x| \geq Rec \\
&\vdots
\end{aligned}
$$

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

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

Rough quantization based tree search method (S. Corlay, in progress).

▷ As a result : Computation of

- Optimal (optimized...) stationary codebooks $\beta(N)$ for W

$$N = 1 \text{ up to } 10\,000 \text{ with } d(N) = 1 \text{ up to } 9.$$

- the companion parameters : for every $N \geq 1$

– The weights = distribution of $\widehat{W}^{\alpha^N}$

$$\mathbb{P}(\widehat{W}^{\alpha^N} = \alpha^N{}_i) = \mathbb{P}(\widehat{Z}^{\beta^{(N)}} = \beta_i^{(N)}) \quad (\leftarrow \text{ in } \mathbb{R}^{d(N)}).$$

– The quadratic quantization error $\|W - \widehat{W}^{\alpha^N}\|_2$.

Download at our WEBSITE :

www.quantize.maths-fi.com

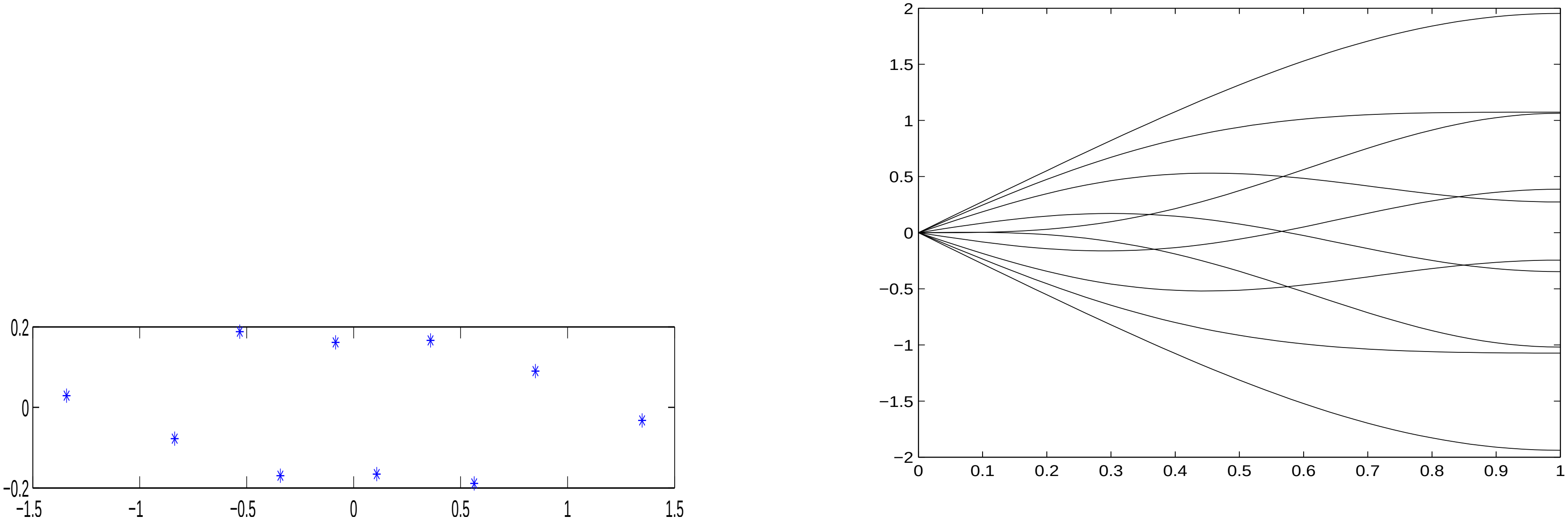


FIG. 7: Optimized FQ of the Brownian motion W for $N = 10 : \beta(10)$ depicted in $\mathbb{R}^2$ *vs* the paths of the 10-quantizer $\alpha^{(10)}$ in the K-L basis

d(N)=2

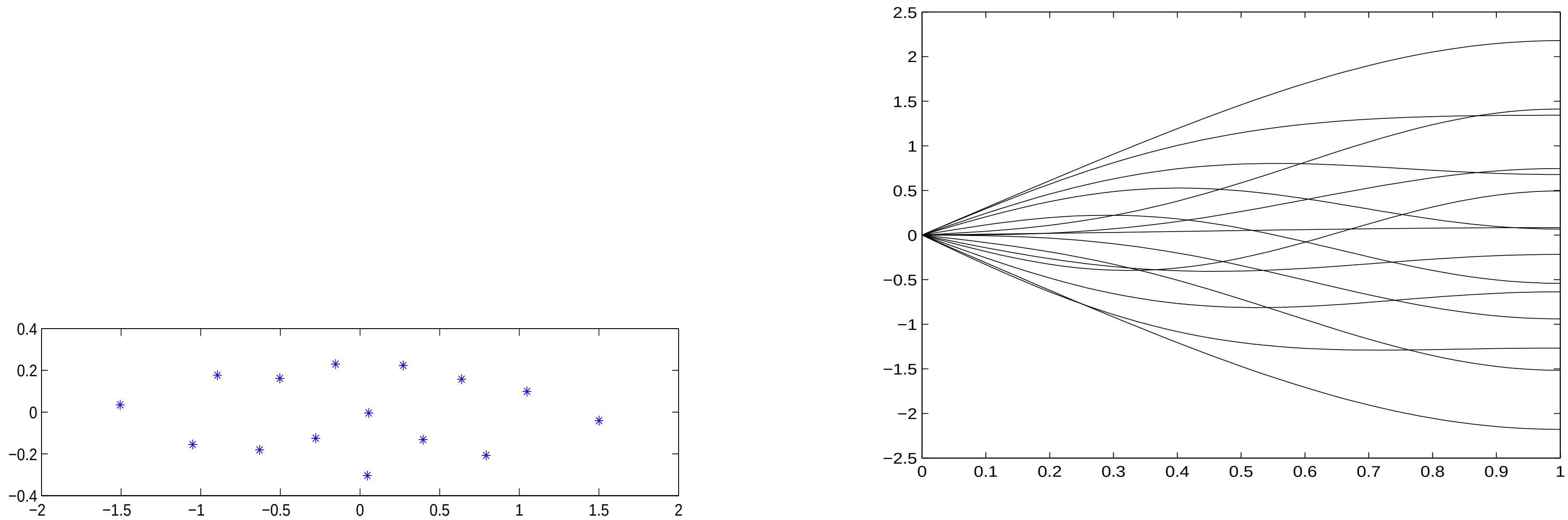


FIG. 8: Optimized FQ of the Brownian motion W for $N = 15 : \beta(15)$ depicted in $\mathbb{R}^2$ *vs* the paths of the 15-quantizer $\alpha^{(15)}$ paths

d(N)=2

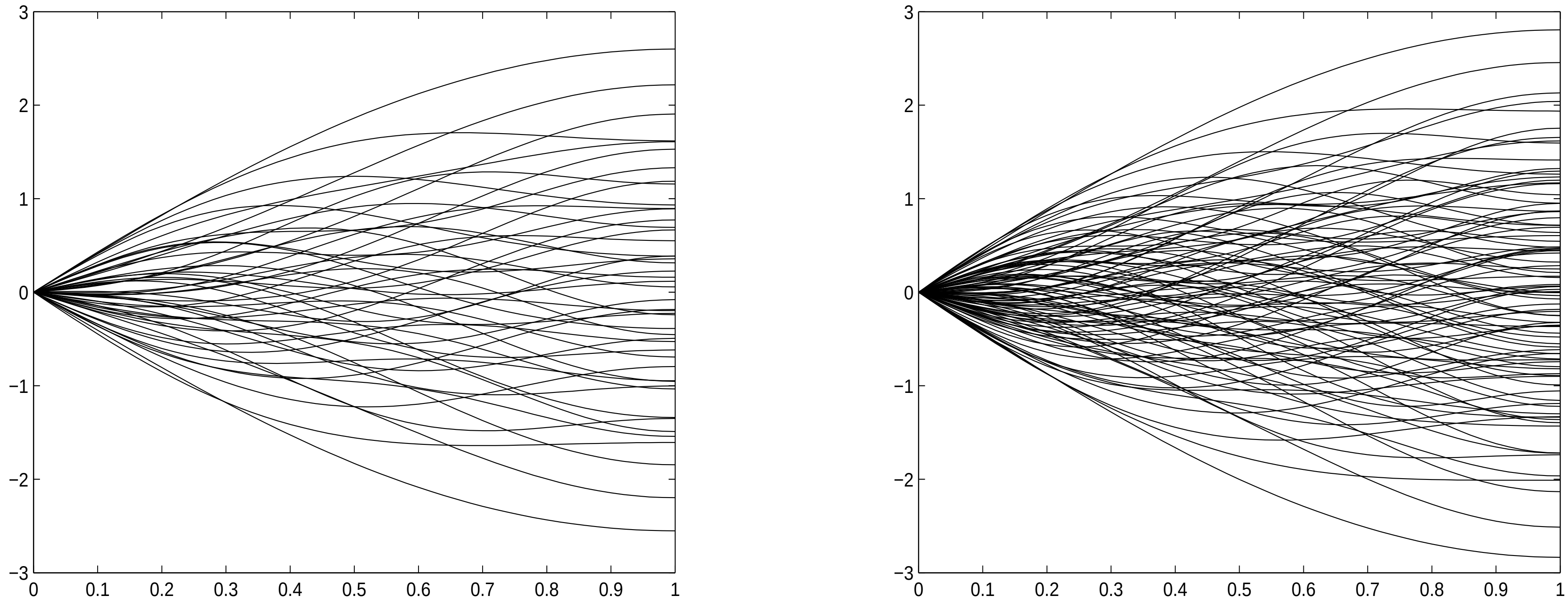
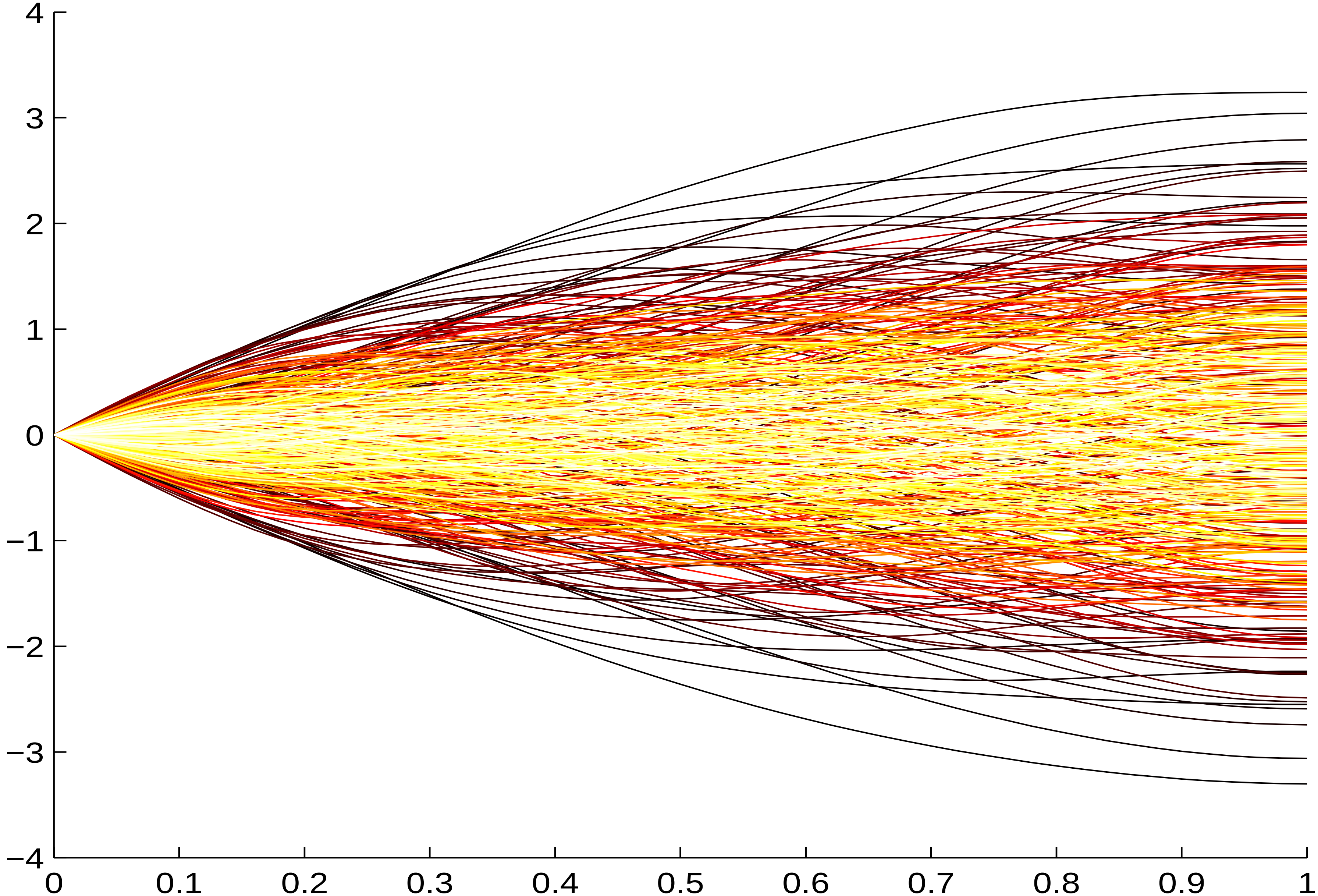


FIG. 9: Optimized Functional N-quantizers $\alpha^{(N)}$ of the Brownian motion W with $N = 48$ and $N = 96$

$d(48) = 3$ and $d(96) = 4$



Brownian motion on [0,1], N=400 points

Product Functional Quantization (of the Brownian motion, etc)

(Numerical aspects : P.-Printems, MCMA, 2006)

▷ Let $(e_n^W)_{n\geq 1}$ be the K-L o.n. basis

$$\forall\, t \in [0,T],\ W_t \overset{L^2_T}{=} \sum_{n\geq 1}(W|e_n^W)_2\, e_n(t) = \sum_{n\geq 1}\sqrt{\lambda_n}\,\xi_n\, e_n^W(t)$$

$$\xi_n \sim \mathcal{N}(0;1), \quad n\geq 1, i.i.d.$$

▷ Quantization by (infinite) product-quantizers

$$\widehat{W}_t^{(N)} \overset{\text{def}}{=} \sum_{n\geq 1}\sqrt{\lambda_n}\,\widehat{\xi}_n^{(N_n)}\, e_n^W(t) = \sum_{n=1}^{m}\sqrt{\lambda_n}\,\widehat{\xi}_n^{(N_n)}\, e_n^W(t)$$

where $\prod_{n=1}^{m} N_n \leq N$ and $\widehat{\xi}_n^{(N_n)} = \mathrm{Proj}_{\beta^{(N_n)}}(\xi_n)$ optimal N_n-quantization of ξ_n

$\triangleright$ Alternative expression : multi-index

$$\underline{i} := (i_1, \ldots, i_m, 1, 1, \ldots, 1, \ldots).$$

$$\widehat{W}_t^{(N)} = \sum_{1 \le i_1 \le N_1, \ldots, 1 \le i_m \le N_m} \underbrace{\mathbf{1}_{\{\widehat{\xi}_n^{(N_n)} = \beta_{i_n}^{(N_n)}, n=1,\ldots,m\}}}_{=\{W \in C_{\underline{i}}(\alpha^{(N)})\}} \underbrace{\sum_{n=1}^{m} \sqrt{\lambda_n}\, \beta_{i_n}^{(N_n)} e_n^W(t)}_{\text{elementary quantizer } \alpha_{\underline{i}}^{(N)}}$$

$\triangleright$ Elementary Quantizer $\alpha_{\underline{i}}^{(N)}$:

$$\boxed{\alpha_{\underline{i}}^{(N)}(t) := \sum_{n=1}^{m} \sqrt{\lambda_n}\, \beta_{i_n}^{(N_n)} e_n(t)}$$

$\triangleright$ Voronoï cell of $\alpha_{\underline{i}}^{(N)}$:

$$\boxed{C_{\underline{i}}(\alpha^{(N)}) = \prod_{n=1}^{m} [\beta_{i_n - \frac{1}{2}}^{(N_n)}, \beta_{i_n + \frac{1}{2}}^{(N_n)}[}$$

Quantization rate by product quantizers

▷ THEOREM (Luschgy-P., *JFA* (2002) and *AP* (2004))

$$\min\left\{\|\,|W-\widehat{W}|_{L^2_T}\|_2,\ 1\le N_1\cdots N_m\le N,\ m\ge 1\right\}\le \frac{c_W}{(\log N)^{\frac{1}{2}}}$$

▷ PROOF :

$$\begin{aligned}\|\,|W-\widehat{W}|_{L^2_T}\|_2^2 &= \sum_{n\ge 1}\lambda_n\|\widehat{\xi}_n^{(N_n)}-\xi_n\|_2^2\\ &\le C\left(\sum_{n=1}^{m}\frac{1}{n^2N_n^2}+\sum_{n\ge m+1}\lambda_n\right)\end{aligned}$$

with $\prod_n N_n \le N$. Set

$$m=[\log N]\,,\qquad N_k=\left[\frac{(m!N)^{\frac{1}{m}}}{k}\right],k=1,\dots,m.\qquad\diamondsuit$$

$\boxed{\text{Optimal scalar product quantizers are then rate optimal}}$

Using product quantizers for applications ?

• The N-quantizers $\alpha^{(N)}_{i_1,\ldots,i_{m(N)}}$ are explicit .

• The weights of Voronoi cells $\mathbb{P}(\widehat{\xi}_n^{(N_n)} = \beta^{(N_n)}_{i_n},\, n = 1,\ldots,m(N))$ are explicit too ...

since the normalized coordinates ξ_n are independent so that

$$\mathbb{P}(\widehat{\xi}_n^{(N_n)} = \beta^{(N_n)}_{i_n},\, n = 1,\ldots,m(N)) = \prod_{n=1}^{m(N)} \underbrace{\mathbb{P}(\widehat{\xi}_n^{(N_n)} = \beta^{(N_n)}_{i_n})}_{1D \Longrightarrow tabulated!}$$

$$\boxed{\text{The distribution of a } K\text{-}L \text{ product quantization } \widehat{W} \text{ is known.}}$$

- Numerical aspects : optimal "integer bit allocation" *i.e.* solving

$$\min\left\{\sum_{n=1}^{m}\lambda_n\|\widehat{\xi}_n^{(N_n)}-\xi_n\|_2^2+\sum_{n\geq m}\lambda_n,\ 1\leq N_1\cdots N_m\leq N,\ m\geq 1\right\}$$

It has already been computed (up to $N = 12\,000$) : a file including the optimal allocations is available on the website

www.quantize.maths-fi.com

N	N_{rec}	Quant. Error	Opti. Alloc.
1	1	0.7071	1
10	10	0.3138	5-2
100	96	0.2264	12-4-2
1 000	966	0.1881	23-7-3-2
10 000	9 984	0.1626	26-8-4-3-2-2
100 000	97 920	0.1461	34 − 10 − 6 − 4 − 3 − 2 − 2

Brownian product quantizations

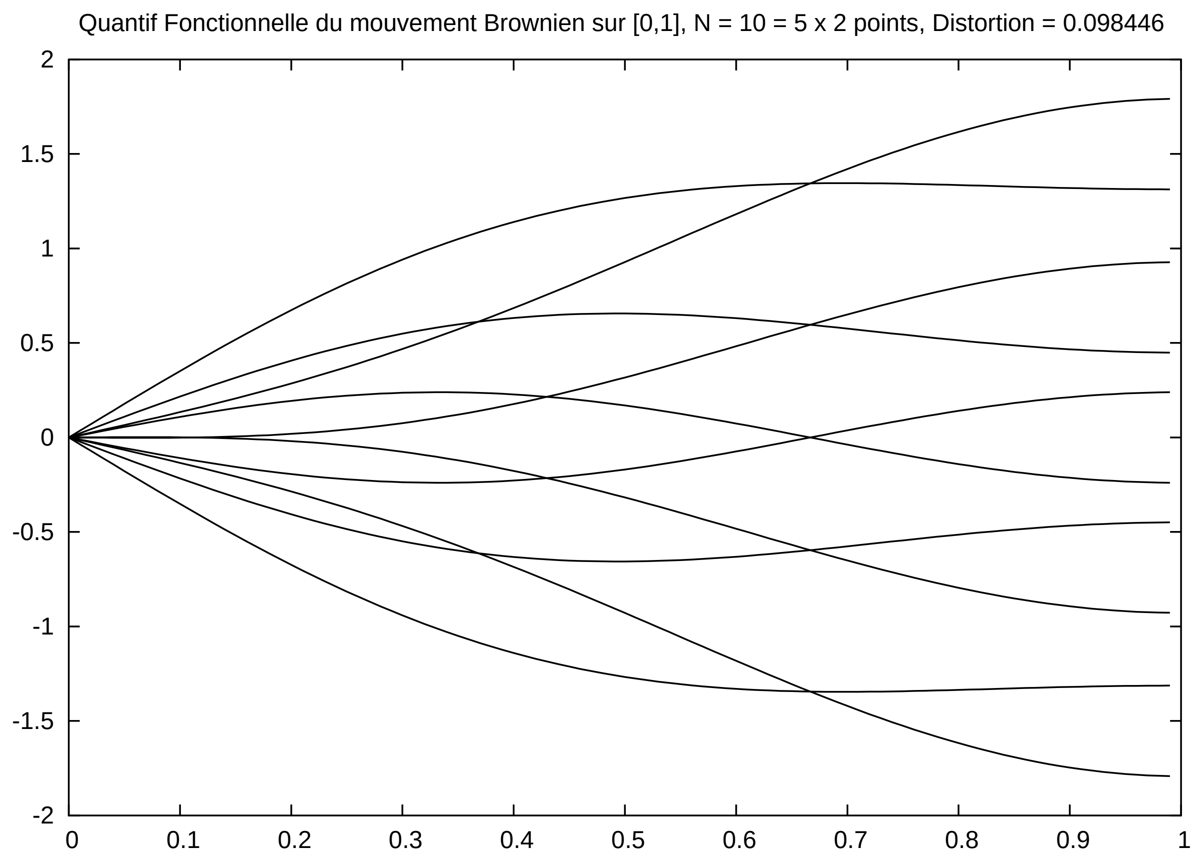


FIG. 11: The $N_{\rm rec}$-quantizer $\alpha^{(N)}$ for $N = 10$ ($N_{\rm rec} = 10$).

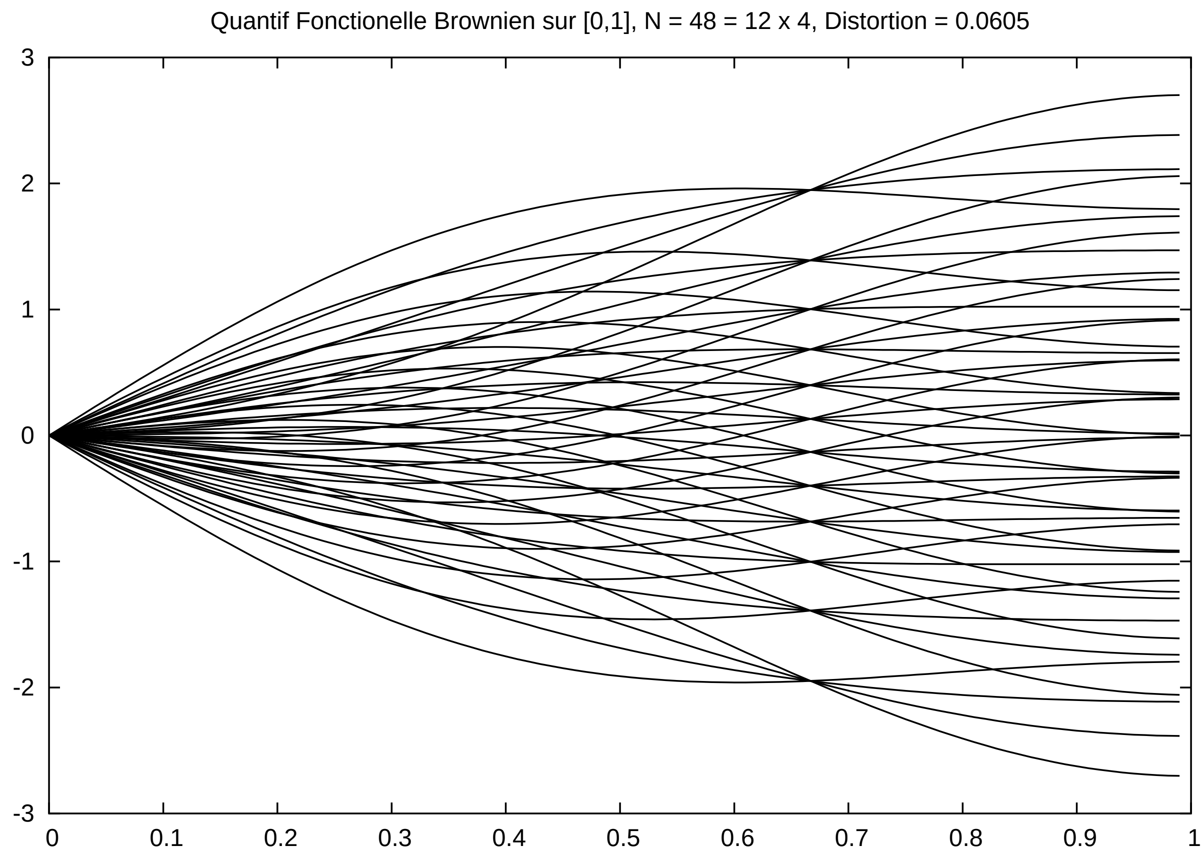


FIG. 12: The N_{rec}-quantizer $\alpha^{(N)}$ for $N = 50$ ($N_{\text{rec}} = 12 \times 4 = 48$).

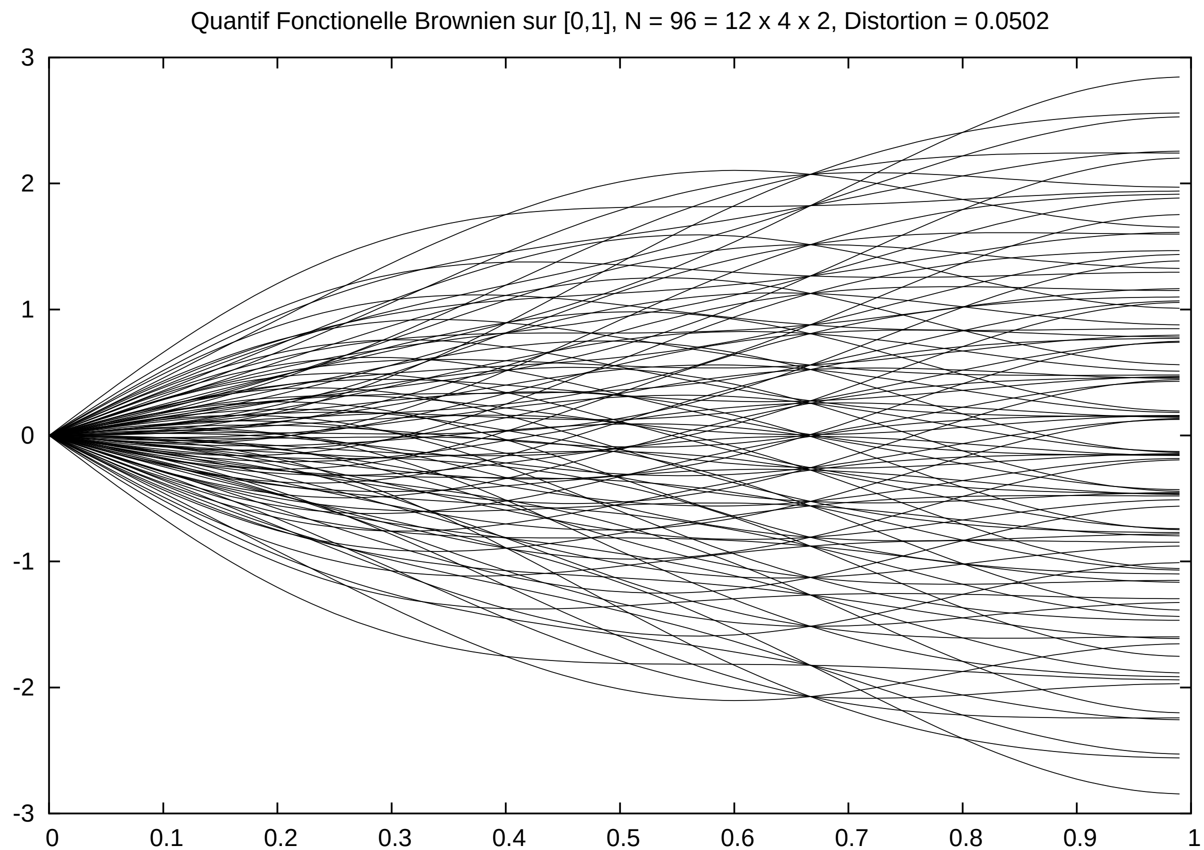


FIG. 13: The $N_{\rm rec}$-quantizer $\alpha^{(N)}$ for $N = 100$ ($N_{\rm rec} = 12 \times 4 \times 2 = 96$).

A cherry on the cake : stationarity again

The quantization-product in the K-L basis provides a stationary quantizer (although sub-optimal).

$$\begin{aligned}
\widehat{W} &= \sum_{n\geq 1} \sqrt{\lambda_n}\, \xi_n^{(N_n)} e_n(t) \\
\text{so that} \qquad \sigma(\widehat{W}) &= \sigma(\widehat{\xi}_k^{(N_k)},\ k \geq 1). \\
\text{and} \qquad \mathbb{E}(W \,|\, \widehat{W}) &= \mathbb{E}(W \,|\, \sigma(\widehat{\xi}_k^{(N_k)},\ k \geq 1)) \\
\mathbb{E}(W \,|\, \widehat{W}) &= \sum_{n\geq 1} \sqrt{\lambda_n}\, \mathbb{E}\left(\xi_n \,|\, \sigma(\widehat{\xi}_k^{(N_k)},\ k \geq 1)\right) e_n \\
&\stackrel{i.i.d.}{=} \sum_{n\geq 1} \sqrt{\lambda_n}\, \mathbb{E}\left(\xi_n \,|\, \widehat{\xi}_n^{(N_n)}\right) e_n \\
&= \sum_{n\geq 1} \sqrt{\lambda_n}\, \widehat{\xi}_n^{(N_n)} e_n = \widehat{W}.
\end{aligned}$$

Comparison with optimal quadratic functional quantization

– (Numerical) Optimal Quantization (in average over $1 \le N \le 10.000$)

$$e_{_N}(W, L^2_{_T})^2 \approx \frac{0.2195}{\log N}$$

– Optimal Product quantization :

$$\min\left\{\| \, |W - \widehat{W}|_{L^2_{_T}} \|^2_{_2},\ 1 \le N_1 \cdots N_m \le N,\ m \ge 1\right\} \approx \frac{0.25}{\log N}$$

– Optimal quantization significantly more accurate on numerical experiments but more demanding (keeping large files off-line).

– Both methods are included in the option pricer Premia soft released by INRIA.

Rate optimal FQ of "Doss-Sussman" diffusions $(\supset d = 1)$

▷ Diffusion process : $dX_t = b(t, X_t)dt + \vartheta(t, X_t)dW_t$

b, ϑ Lipschitz continuous, $\vartheta(t, .) = (\nabla S_t(.))^{-1}$ bounded,etc.

▷ α^N, $N \geq 1$, sequence of stationary rate optimal N-quantizers of W.

▷ $d\, x_i^{(N)}(t) = \left(b(t, x_i^{(N)}(t)) - \frac{1}{2}\vartheta\vartheta'(t, x_i^{(N)}(t)) \right) dt + \vartheta(t, x_i^{(N)}(t))d\,\alpha_i^N(t)$.

▷ Theorem (Luschgy-P., SPA (2006)) $(x^{(N)})_{N\geq 1}$ is rate optimal *i.e.*

$$\|\, |X - \widetilde{X}^{x^{(N)}}|_{L^2_T} \|_2 = O\left(\frac{1}{(\log N)^{\frac{1}{2}}}\right) \qquad (\asymp \text{ if } \vartheta \geq \varepsilon_0 > 0)$$

where
$$\widetilde{X}_t^{x^{(N)}} = \sum_{k=1}^{N} x_i^{(N)}(t)\mathbf{1}_{\{\widehat{W}^{\alpha^N} = \alpha_i^N\}}$$

is a (computable) *non-Voronoi* quantizer.

▷ Sharp rate $c(\log N)^{-\frac{1}{2}}$ (Dereich, SPA, 2008), non constructive.

General Multi-dimensional diffusions

(Joint work with A. Sellami)

Diffusion in the Stratanovich sense :

$$dX_t = b(t, X_t)\, dt + \vartheta(t, X_t) \circ dW_t \qquad X_0 = x \in \mathbb{R}^d$$

▷ $W = (W^1, \ldots, W^d)$ is a d-dimensional B.M.

$$\min_{|\alpha| \le N} \|W - \widehat{W}^{\alpha}\|_2 \sim C_d \frac{1}{\sqrt{\log N}} \quad \text{as} \quad N \to \infty.$$

▷ $\frac{1}{p}$-Hölder norm : $\mathbf{x}_{s,t} = (x^1_s, x^2_{s,t})$, $s \le t$.

$$\|\mathbf{x}\|_{q,Hol} = \sup_{s,t \in [0,T]} \frac{|x^1(t) - x^1(s)|}{|t-s|^{\frac{1}{q}}} + \sup_{s,t \in [0,T]} \frac{|x^2(s,t)|}{|t-s|^{\frac{2}{q}}}$$

Thus $\mathbf{W} = (W_t, \int_s^t (W_u - W_s) dW_u)$

$\triangleright$ THEOREM (P.-Sellami, (2006), (2009)) (a) Let $\alpha^N = (\alpha_1^N, \cdots, \alpha_N^N)$ be a sequence of optimal (stationary) *N-product* quantizers of W. Then

$$\forall\, p > 2,\ \forall\, q > \frac{p}{p-2}, \quad \|\|\mathbf{W} - \widehat{\mathbf{W}}\|_{q,Hol}\|_{L^p(\mathbb{P})} = O\left(\frac{1}{\sqrt{\log N}}\right).$$

(b) Assume b and ϑ are $\mathcal{C}^{2+\alpha}([0,T]\times\mathbb{R}^d)$, $\alpha > 0$.

$$ODE \qquad \equiv \qquad dx_i^{(N)}(t) = b(t, x_i^{(N)}(t))dt + \vartheta(t, x_i^{(N)}(t))d\,\alpha_i^N(t),\ i = 1,\ldots,N.$$

Set
$$\widetilde{X}_t := \sum_{i=1}^{N} x_i^{(N)}(t)\mathbf{1}_{\{W\in C_i(\alpha^N)\}}$$

$$\forall\, p > 2,\ \forall\, q > \frac{p}{p-2}, \quad \left\|\|\widetilde{X}_t - X\right\|_{Hol,q}\|_{L^p(\mathbb{P})} = O\left(\frac{1}{\sqrt{\log N}}\right)$$

(topology of $\frac{1}{q}$-Holder-convergence).

$\triangleright$ The keys : connection with rough paths theory, Kolmogorov criterion, (pseudo-)stationarity.

Typical functionals

– Fonctionals $|\,.\,|_{L^2_T}$-continuous at every $\omega \in \mathcal{C}([0,T])$?

$$F(\omega) := \int_0^T f(t,\omega(t))dt$$

where$f$ is *locally Lipschitz continuous*, namely

$$|f(t,u) - f(t,v)| \le C_f |u-v|(1 + g(t,u) + g(t,v)).$$

EXAMPLE : The Asian payoff in B-S model

$$F(\omega) = \exp(-rT)\left(\frac{1}{T}\int_0^T \exp(\sigma\omega(t) + (r-\sigma^2/2)t)dt - K\right)_+ .$$

Numerical Integration (II) : log-Romberg extrapolation

▷ $F : L^2_T \longrightarrow \mathbb{R}$, 3 times $|\,.\,|_{L^2_T}$-differentiable with bounded differentials.

▷ $\widehat{W}^{(N)}$, $N \geq 1$, stationary rate-optimal quantizations

▷ Higher order Taylor expansion yields

$$\begin{aligned} F(W) &= F(\widehat{W}^{(N)}) + DF(\widehat{W}^{(N)}).(W - \widehat{W}^{(N)}) \\ &\quad + \frac{1}{2}D^2F(\widehat{W}^{(N)}).(W - \widehat{W}^{(N)})^{\otimes 2} + \frac{1}{6}D^3F(\widehat{W}^{(N)}).(W - \widehat{W}^{(N)})^{\otimes 3}. \end{aligned}$$

$$\boxed{\mathbb{E}F(W) = \mathbb{E}F(\widehat{W}^{(N)}) + \frac{1}{2}\mathbb{E}\left(D^2F(\widehat{W}^{(N)}).(W - \widehat{W}^{(N)})^{\otimes 2}\right) + o\left((\log N)^{-\frac{3}{2}+\varepsilon}\right).}$$

$$\boxed{\textsc{Conjecture :} \qquad \mathbb{E}\left(D^2F(\widehat{W}^{(N)}).(W-\widehat{W}^{(N)})^{\otimes 2}\right) \sim \frac{c}{\log N}, \qquad N \to \infty}$$

Set

$$M \ll N \quad (e.g.\ M \approx N/4)$$

and $\forall\, \varepsilon > 0$

$$\boxed{\mathbb{E}(F(W)) = \frac{\log N \times \mathbb{E}(F(\widehat{W}^{(N)})) - \log M \times \mathbb{E}(F(\widehat{W}^{(M)}))}{\log N - \log M} + o\left((\log N)^{-\frac{3}{2}+\varepsilon}\right),}$$

Variant (mainly for *product quantizations*, B.Wilbertz (Trier, 2005)) :

$$\text{Replace} \qquad \log N \qquad \text{by} \qquad 1/\|W-\widehat{W}^{(N)}\|_2^2.$$

Application : Asian option in a Heston stochastic volatility model

▷ The dynamics : Let ϑ, k, a s.t. $\vartheta^2/(4ak) < 1$.

$$\begin{aligned} dS_t &= S_t(r\,dt + \sqrt{v_t})dW_t^1, \qquad S_0 = s_0 > 0, \quad \text{(risky asset)} \\ dv_t &= k(a - v_t)dt + \vartheta\sqrt{v_t}dW_t^2,\ v_0 > 0 \quad \text{with } <W^1, W^2>_t = \rho\, t,\ \rho \in [-1, 1]. \end{aligned}$$

▷ The payoff and the premium :

$$\mathrm{AsCall}^{Hest} = e^{-rT}\mathbb{E}\left(\left(\frac{1}{T}\int_0^T S_s ds - K\right)_+\right).$$

▷ The procedure : • Projection of W^1 on W^2

$$S_t = s_0 \exp\left((r - \frac{1}{2}\bar{v}_t)t + \rho\int_0^t \sqrt{v_s}dW_s^2\right)\exp\left(\sqrt{1-\rho^2}\int_0^t \sqrt{v_s}d\widetilde{W}_s^1\right)$$

- Chaining rule for conditional expectations

$$\mathrm{AsCall}^{Hest}(s_0, K) = e^{-rT}\mathbb{E}\left(\mathbb{E}\left(\left(\frac{1}{T}\int_0^T S_s ds - K\right)_+ |\sigma(W_t^2, 0 \le t \le T)\right)\right)$$

- State process $= (\widetilde{W}_t^1, v_t)$.
- Solving the quantization ODE's for (v_t) (by a Runge-Kuta scheme)

$$dy_i(t) = \left(k(a - y_i(t) - \frac{\vartheta^2}{4k})\right) dt + \vartheta\sqrt{y_i(t)}d\alpha_i^N(t),\ i = 1, \dots, N.$$

Set the (non-Voronoi rate optimal) N-quantization of (v_t, S_t) by

$$\widetilde{v}_t^{n,N} = \sum_i y_i^{n,N}(t)\mathbf{1}_{C_i(\alpha^N)}(W^2).$$

$$\text{and} \qquad \widetilde{S}_t^{n,N} = \sum_{1\le i,j\le N} s_{i,j}^{n,N}(t)\mathbf{1}_{\alpha_i^N}(\widetilde{W}^1)\mathbf{1}_{\alpha_j^N}(W^2).$$

with

$$\begin{aligned} s_{i,j}^{n,N}(t) &= s_0 \exp\left(t\left((r - \frac{\rho a k}{\vartheta}) + \bar{y}_j^{n,N}(t)(\frac{\rho k}{\vartheta} - \frac{1}{2})\right) + \frac{\rho}{\vartheta}(y_j^{n,N}(t) - v_0)\right) \\ &\quad \times \exp\left(\sqrt{1-\rho^2}\int_0^t \sqrt{y_j^{n,N}} d\alpha_i^N\right). \end{aligned}$$

- Computation of *crude* quantized premium for N and M.
- Space Romberg log-extrapolation $\widehat{\text{RCrAsCall}}^{Hest}(s_0, K)$.
- K-linear interpolation $\widehat{\text{IRAsCall}}^{Hest}(s_0, K)$ based on the (Asian) forward moneyness Ke^{-rT} and the Asian Call-Put parity formula

$$\text{AsianCall}^{Hest}(s_0, K) - \text{AsianPut}(s_0, K) = s_0\frac{1-e^{-rT}}{rT} - Ke^{-rT}.$$

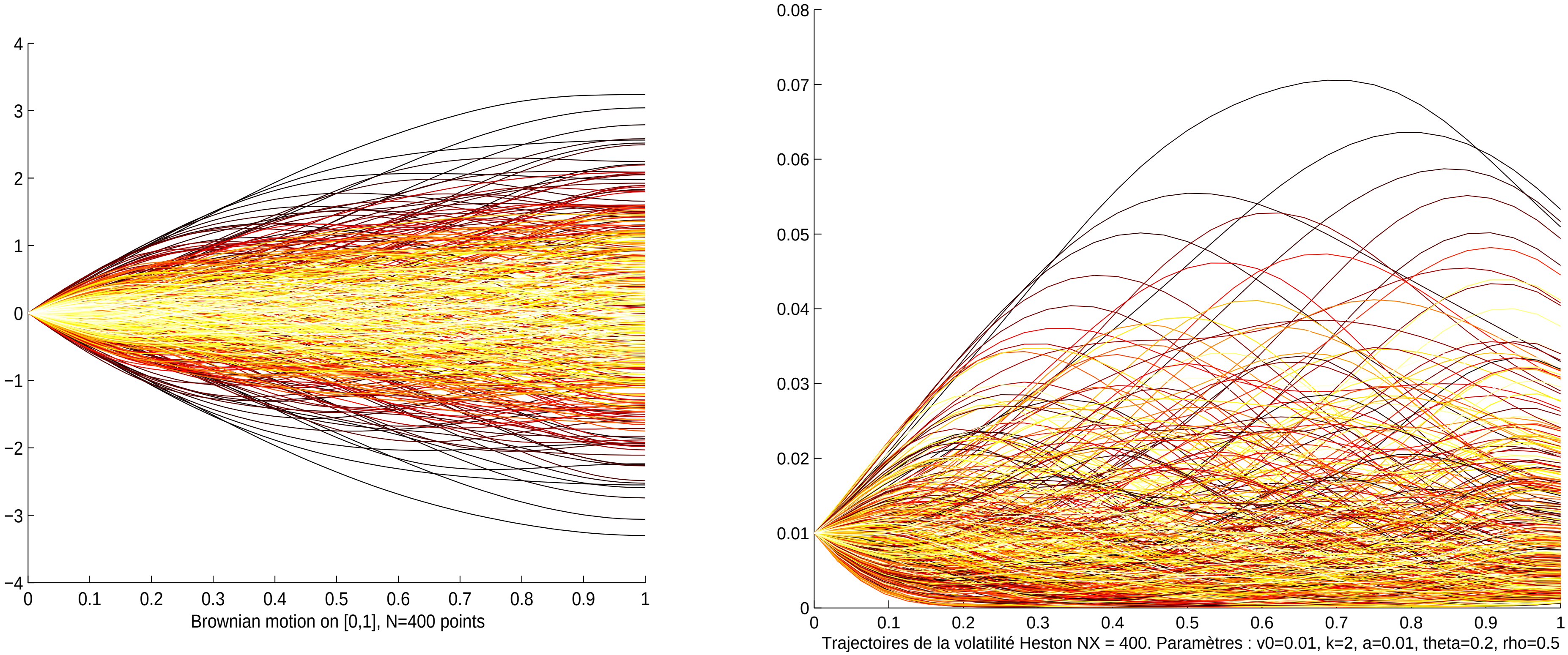


FIG. 14: Optimized Quantizer of the Heston volatility process $N = 400$

▷ Parameters of the Heston model :

$$\boxed{\quad s_0 = 100,\ k = 2,\ a = 0.01,\ \rho = 0.5,\ v_0 = 10\%,\ \vartheta = 20\%.}$$

▷ Parameters of the option portfolio :

$$T = 1,\ K = 99, \cdots, 111 \qquad\qquad (13 \text{ strikes}).$$

▷ Reference price : computed by a 10^8 trial Monte Carlo simulation (including a time Romberg extrapolation with $2n = 256$).

▷ Parameters of the quantization cubature formulae :

$$\Delta t = 1/32, \qquad\qquad (N, M) = (400, 100),\ (1000, 100) \ \text{ or } \ (3200, 400)$$

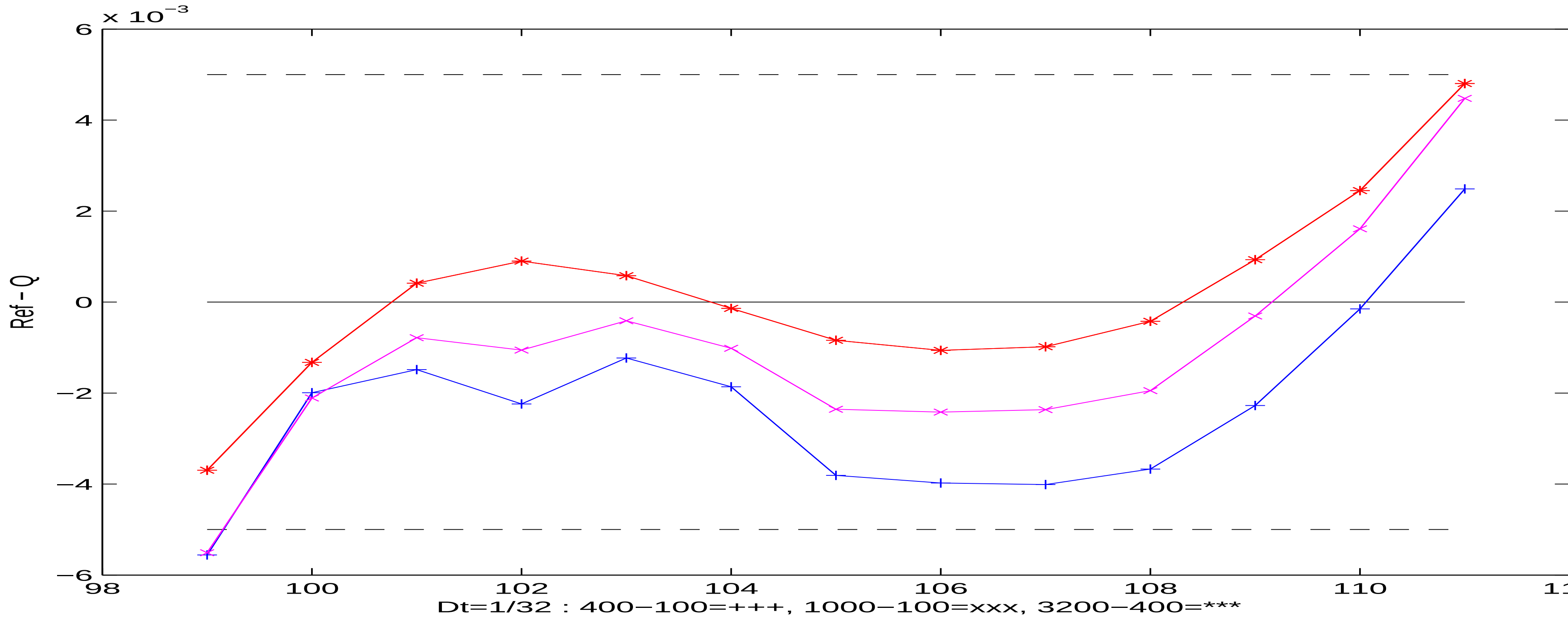


FIG. 15: K-Interpolated-log-Romberg extrapolated- FQ price : The error with $(N, M) = (400, 100)$, $(N, M) = (1000, 100)$, $(N, M) = (3200, 400)$

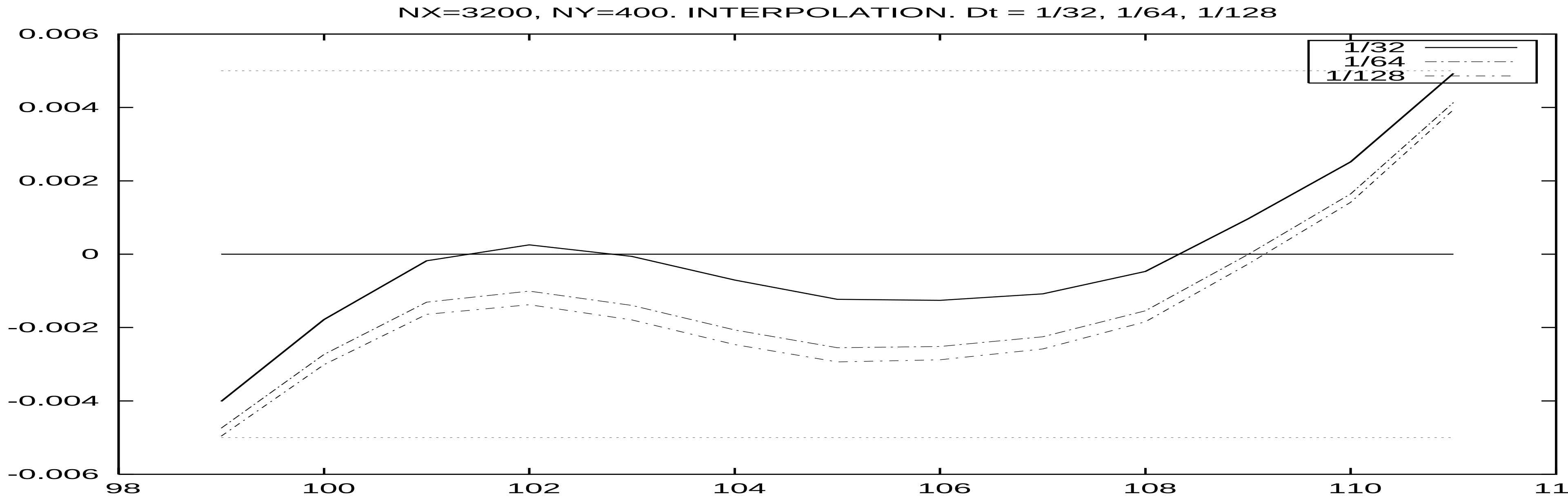


FIG. 16: K-Interpolated-log-Romberg extrapolated- FQ price : Convergence as $\Delta t \to 0$ with $(N, M) = (3200, 400)$

▷ Functional Quantization can compute a whole vector (more than 10) option premia for the Asian option in the Heston model.

Within 1 cent accuracy in less than 1 second

(implementation in C on 2.5 GHz processor).

Functional Quantization of non Gaussian processes

$\triangleright$ THEOREM (Luschgy-P. 2006, *AAP*) Let $X = (X_t)_{t\in[0,T]}$. If

$$X_0 \in L^r(\mathbb{P}), \quad \|X_t - X_s\|_{L^r(\mathbb{P})} \le C_X|t-s|^a, \qquad 0 < a \le 1$$

then $$\boxed{\forall\, 0 < p \le r, \qquad e_{N,r}(X, L^p_T) = O((\log N)^{-a}).}$$

$\triangleright$ Ingredients : Haar basis (instead of K-L basis...), non asymptotic Zador Theorem (Pierce Lemma) and product functional quantization.

$\triangleright$ Examples : • d-dim Itô processes (includes d-dim diffusions with sublinear coefficients) $a = 1/2$;

• General Lévy process X with Lévy measure ν (with Brownian component) $a = 1/2$;

• General Lévy process X with Lévy measure ν (without Brownian component) with square integrable big jumps. Then

$$a = 1/\beta^*(X)$$

where

$$\beta^*(X) := \inf\{\theta \ : \ \int |x|^\theta \nu(dx) < \infty\} \in (0,2) \qquad \text{(Blumenthal-Getoor index of } X\text{)}.$$

• *Exact rates for a wide class of subordinated Lévy processes (to the Brownian motion)* includes α-stable symmetric Lévy processes for which

$$\forall\, 0 < p \le r < \alpha, \qquad e_{N,r}(X, L^p_T) \approx O((\log N)^{-\alpha})$$

A guided Monte Carlo method : hybrid "Q+MC"

$\triangleright$ Quantization as a control variate, (P.-Printems, *MCMA*, 2005). Let X_k, $k \geq 1$, i.i.d. $X_1 \sim X$.

$\widehat{X}_k$ (optimal) N-quantization of X_k and F a Lipschitz continuous functional.

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

$$\begin{aligned}
\mathrm{Var}\left(\frac{1}{M}\sum_{k=1}^{M} F(X_k) - F(\widehat{X}_k^{\alpha})\right) &= \frac{\|F(X) - F(\widehat{X}^{\alpha})\|_2^2 - (\mathbb{E}\,F(X) - \mathbb{E}\,F(\widehat{X}^{\alpha}))^2}{M} \\
&\leq \frac{\|F(X) - F(\widehat{X}^{\alpha})\|_2^2}{M} \\
&\leq [F]_{\mathrm{Lip}} \frac{\|X - \widehat{X}^{\alpha}\|_2^2}{M}
\end{aligned}$$

Drawback : nearest neighbour search [complexity$= O(\log N)$] at each step...

$\triangleright$ Quantization based universal stratified sampling (with J.Printems (2008) and S. Corlay (2009))

– Let α be a product N-quantizer with structural dimension $d(N) = \log N$.

– The idea starts from the ability to simulate

$$\mathcal{L}(W_{t_1}, \ldots, W_{t_n} \,|\, W \in C_i(\alpha)) = \mathcal{L}(W_{t_1}, \ldots, W_{t_n} \,|\, \widehat{W} = \alpha_i)$$

from the Karhunen-Loève expansion of W :

$$W_t = \sum_{n \geq 1} \frac{1}{\pi(n - \frac{1}{2})} \xi_n e_n W(t)$$

with complexity $O(n \times d(N))$.

– Weight and intra-class variances are tabulated (up to Pythagorus Theorem) :

$$p_i = \mathbb{P}(\widehat{W} = \alpha_i) \quad \text{and} \quad \sigma_i^2 = \mathrm{Var}(W \,|\, \widehat{W} = \alpha_i)$$

so that

$$\mathbb{E} f(W_{t_1}, \ldots, W_{t_n}) = \sum_{i=1}^{N} \frac{1}{M_i} \sum_{m=1}^{M_i} f(\tilde{W}^m_{t_1}, \ldots, \tilde{W}^m_{t_n})$$

where

$$(W^m_{t_1}, \ldots, W^m_{t_n}) \sim \mathcal{L}(W_{t_1}, \ldots, W_{t_n} \,|\, \widehat{W} = \alpha_i),\ 1 \le m \le M_i,\ i.i.d.$$

and

$$M_i = M \times \frac{p_i \sigma_i}{\sum_j p_j \sigma_j},\ i = 1, \ldots, N$$

is the best “min-max” Monte Carlo estimator in the family of Lipschitz functional among all possible stratifications.

– Variance reduction factor :

$$\frac{\|X - \widehat{X}^{\alpha}\|_2^2}{\|X - \mathbb{E}\, X\|_2^2}$$

like for control variate…but no nearest neighbour search.