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 appliquées

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 November 2009
What is (quadratic) Functional Quantization?

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

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

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

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

▷ 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]

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

using
\[ N \text{-quantizer (or } N \text{-codebook)} : \]

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

- When \( H = \mathbb{R}^d \), each \( \alpha_i \) 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.

\[ \text{Discretization by } \alpha\text{-quantization} \]

\[ X \sim \hat{X}^\alpha : \Omega \to \alpha := \{\alpha_1, \ldots, \alpha_N\}. \]

\[ \hat{X}^\alpha := \text{Proj}_\alpha(X) \]

where

\[ \text{Proj}_\alpha \text{ denotes the projection on } \alpha \text{ following the nearest neighbour rule.} \]
Fig. 1: A 2-dimensional 10-quantizer $\alpha = \{\alpha_1, \ldots, \alpha_{10}\}$ and its Voronoi diagram.
What do we know about $X - \hat{X}^\alpha$ and $\hat{X}^\alpha$?

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

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

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

$$e_N(X, H, \alpha) = \|X - \hat{X}^\alpha\|_2 = \sqrt{\mathbb{E} \left( \min_{1 \leq i \leq N} |X - \alpha_i|_H^2 \right)}.$$  

▷ **Distribution of $\hat{X}^\alpha$**: weights associated to each $\alpha_i$:

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

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

$$C_i(\alpha) := \left\{ \xi \in H : |\xi - \alpha_i|_H = \min_{1 \leq j \leq N} |\xi - \alpha_j|_H \right\}.$$
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 $v$s some paths.

(with S. Corlay)

$$X_t = \int_{-\infty}^{t} e^{-(t-s)} dW_s \quad \| \quad dX_t = -X_t dt + dW_t, \quad X_0 \sim \mathcal{N}(0; \frac{1}{2})$$
Fig. 5: A \( N = 20 \)-quantizers of Brownian bridge vs some paths.

(with S. Corlay)

\[
X_t = W_t - tW_1, \quad t \in [0, 1]
\]
non Gaussian diffusion processes? etc.

Some questions

▷ What is the connection between blue chaotic lines and pink smooth lines?

▷ How to get the pink smooth lines from the blue chaotic lines?

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

▷ 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)

$$D^X_N : H^N \rightarrow \mathbb{R}_+$$

$$\alpha = (\alpha_1, \ldots, \alpha_N) \mapsto \|X - \hat{X}^\alpha\|_2^2 = \mathbb{E}\left( \min_{1 \leq i \leq N} |X - \alpha_i|^2_H \right)$$

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

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

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

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

**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**

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

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

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

\[ e_{r,N}(X, E) := \inf \left\{ \| X - \hat{X}_\alpha \|_r, \alpha \subset E, \text{card}(\alpha) \leq N \right\} \]
▶ **Examples**: Non-Euclidean norms on $E = \mathbb{R}^d$, $E = L^p_T := L^p([0, T], dt)$, $1 \leq p < \infty$, $E = C([0, T]), \| \cdot \|_{\text{sup}}$, etc.

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

▶ Recent existence results, see Graf-Luschgy-P. (2006, *J. of Approx.*).
Stationary Quantizers

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

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

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

$$\nabla D_N^X(\alpha) = 0 \iff \hat{X}^\alpha = \mathbb{E}(X \mid \hat{X}^\alpha)$$

since

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

- **An optimal** quantizer $\alpha$ is *stationary*

(First by-product: $\mathbb{E}X = \mathbb{E}\hat{X}^\alpha$).
Numerical Integration/Conditional expectation (I) : cubature formulae

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

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

▷ If $F$ is Lipshitz continuous, then

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

in fact

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

Likewise

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

Taylor expansion yields

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

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

$$\leq [DF]_{\text{Lip}} \mathbb{E}\left| X - \hat{X}^\alpha \right|^2 $$

since

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

so that

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

Likewise

$$\left| \mathbb{E}(F(X) | \hat{X}^\alpha) - F(\hat{X}^\alpha) \right| \leq [DF]_{\text{Lip}} \mathbb{E}\left( \|X - \hat{X}^\alpha\|_2^2 | \hat{X}^\alpha \right)$$
The key for numerical applications: $F$ Lipschitz continuous

\[
\mathbb{E}(F(X) \mid Y) = \varphi_F(Y) \quad \varphi \text{ Lipschitz continuous.}
\]

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

\[
\| \mathbb{E}(F(X) \mid Y) - \mathbb{E}(F(\hat{X}) \mid \hat{Y}) \|_2 \leq [F]_{\text{Lip}} \| X - \hat{X} \|_2 + [\varphi_F]_{\text{Lip}} \| Y - \hat{Y} \|_2.
\]
Vector Quantization rate \( (H = \mathbb{R}^d) \)

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

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

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

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

**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 \( \tilde{J}_{r,d} \sim \tilde{J}_{2,d} \) 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)

▷ Local inertia: $a \mapsto \mathbb{E}|X - a|^2 1_{X \in C_a(\alpha)} \approx \text{Constant}$
The 1-dimension...

▷ **Theorem** (Kiefer (82)) \( H = \mathbb{R} \). If \( P_X(d\xi) = \varphi(\xi)\,d\xi \) with \( \log \varphi \) concave, then there is exactly one stationary quantizer. Hence

\[ \forall N \geq 1, \quad \text{argmin} D_X^N = \{\alpha^{(N)}\} \]

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

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

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

**Hessian** : \( D^2(D_X^N)(\alpha) = \ldots \ldots \quad \text{only involves} \int_0^x \varphi(\xi)d\xi \text{ and } \int_0^x \xi\varphi(\xi)d\xi \)

▷ 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 optimal $N$-quantizers and companion parameters:

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

and

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

Download at our WEBSITE:

[www.quantize.maths-fi.com](http://www.quantize.maths-fi.com)

For $d = 1$ up to 10? Also available for Gaussian $\mathcal{N}(0, I_d)$ ($1 \leq N \leq 4000$). How? Stochastic optimization methods, see further on...
Optimal Functional Quantization (of the Brownian motion)

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

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

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

\[
\begin{align*}
e^W_n(t) &= \sqrt{2T} \sin \left( (n - \frac{1}{2})\pi \frac{t}{T} \right), \\
\lambda_n &= \left( \frac{T}{\pi(n - \frac{1}{2})} \right)^2, \quad n \geq 1
\end{align*}
\]

\[
W_t \overset{L^2_T}{=} \sum_{n \geq 1} (W | e^W_n) e^W_n(t) = \sum_{n \geq 1} \sqrt{\lambda_n} \xi_n e^W_n(t)
\]

$\xi_n \sim \mathcal{N}(0; 1), \quad n \geq 1, \quad \text{i.i.d.}$
\textbf{Theorem (Luschgy-P., JFA (2002) and AP (2003))} Let $\alpha^N$, $N \geq 1$, be a sequence of optimal $N$-quantizers.

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

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

$\triangleright \quad e_N(W, L^2_T) = \|W - \hat{W}^{\alpha^N}\|_2 \sim \frac{\sqrt{2}}{\pi} \frac{1}{\sqrt{\log N}}. \quad (\frac{\sqrt{2}}{\pi} = \sqrt{0.2026\ldots})$

$\triangleright \quad$ Reduction to finite dimension (Pythagore)

\[
\begin{align*}
\mathcal{O}_N \left\{ \begin{array}{l}
\|W - \hat{W}^{\alpha^N}\|_2^2 = \|Z - \hat{Z}^{\beta(N)}\|_2^2 + \sum_{k=d(N)+1}^{d(N)} \lambda_k \\
Z \sim \bigotimes_{k=1}^{d(N)} N(0, \lambda_k) \quad \& \quad \|Z - \hat{Z}^{\beta(N)}\|_2 = e_N(Z, \mathbb{R}^{d(N)})
\end{array} \right.
\end{align*}
\]

Then

$$\hat{W}^{\alpha^N} = \sum_{k=1}^{d(N)} (\hat{Z}^{\beta(N)})_k e^W_k.$$
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 $(\lambda^X_n, e^X_n)_{n \geq 1}$. Let $\alpha^N, N \geq 1$, be a sequence of quadratic optimal $N$-quantizers for $X$. If

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

$\triangleright \alpha^N = (\alpha^N_1, \cdots, \alpha^N_N) \subset \text{span}\{e^X_1, \ldots, e^X_{d^X(N)}\}$ with

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

$\triangleright e_N(X, L^2_{[0,1]}) = \|X - \hat{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 $\lambda^X_n \begin{pmatrix} \leq \\ \geq \end{pmatrix} \varphi(n), \quad \varphi \text{ regularly varying, index } -b \leq -1.$
APPLICATIONS TO CLASSICAL (CENTERED) GAUSSIAN PROCESSES

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

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

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

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

- 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

  $$e_{N,r}(W^H, L_T^p) \sim \frac{c_p}{(\log N)^H}.$$
Optimal quadratic Functional Quantization (of $W$): numerical aspects ($T = 1$)

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

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

▷ A way out:

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

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

so that

\[
(\alpha^N)(t + 1) = (\alpha^N)(t) - \frac{c}{t+1} \nabla D_N Z((\alpha^N)(k), \zeta_{t+1}), \quad (\alpha^N)(0) \subset \mathbb{R}^d
\]

\[
= \text{nearest neighbor search} + \text{Dilatation}_{\zeta_{t+1}, 1 - \frac{c}{t+1}} \text{(winner)}
\]
— “Lloyd I procedure”: randomized fixed point procedure based on the stationarity equality:

\[
\hat{Z}(\alpha^N)(t+1) = \mathbb{E}(Z \mid \hat{Z}(\alpha^N)(t)), \quad (\alpha^N)(0) \subset \mathbb{R}^d.
\]

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

\[
x_i^{(t+1)} := \mathbb{E}(X \mid X^{\alpha(t)} \in C_i(\Gamma(\ell))), \quad i = 1, \ldots, N
\]

\[
= \lim_{M \to \infty} \frac{\sum_{m=1}^{M} X_m 1\{X_m \in C_i(\alpha(t))\}}{|\{1 \leq m \leq M, X_m \in C_i(\alpha(t))\}|}
\]

based on repeated nearest neighbour searches.

Then \(\alpha(t + 1) = \{x_i(t + 1)\}, \ i = 1, \ldots, N\}, \text{ etc.}
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 : $K$-$d$-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

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

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

▷ Quantization by (infinite) product-quantizers

\[
\hat{W}^{(N)}_t \overset{\text{def}}{=} \sum_{n \geq 1} \sqrt{\lambda_n} \hat{\xi}^{(N_n)} e^n_W(t) = \sum_{n=1}^m \sqrt{\lambda_n} \hat{\xi}^{(N_n)} e^n_W(t)
\]

where \(\prod_{n=1}^m N_n \leq N\) and \(\hat{\xi}^{(N_n)} = \text{Proj}_{\beta(N_n)}(\xi_n)\) optimal \(N_n\)-quantization of \(\xi_n\)
Alternative expression: multi-index

\[
\hat{W}^{(N)}_t = \sum_{1 \leq i_1 \leq N_1, \ldots, 1 \leq i_m \leq N_m} 1 \{ \xi^{(Nn)}_n = \beta^{(Nn)}_{i_n}, n=1,\ldots,m \} \sum_{n=1}^{m} \sqrt{\lambda_n} \beta^{(Nn)}_{i_n} e_n(t)
\]

Elementary Quantizer \( \alpha^{(N)}_i \):

\[
\alpha^{(N)}_i(t) := \sum_{n=1}^{m} \sqrt{\lambda_n} \beta^{(Nn)}_{i_n} e_n(t)
\]

Voronoi cell of \( \alpha^{(N)}_i \):

\[
C_i^{(\alpha^{(N)})} = \prod_{n=1}^{m} [\beta^{(Nn)}_{i_n - \frac{1}{2}}, \beta^{(Nn)}_{i_n + \frac{1}{2}}]
\]
Quantization rate by product quantizers

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

\[
\min \left\{ \| W - \hat{W} \|_{L^2_T}^2, 1 \leq N_1 \cdots N_m \leq N, \ m \geq 1 \right\} \leq \frac{c_W}{(\log N)^{1/2}}
\]

▷ **Proof**: \[\| W - \hat{W} \|_{L^2_T}^2 = \sum_{n \geq 1} \lambda_n \| \hat{\xi}(N_n) - \xi_n \|_2^2\]

\[\leq C \left( \sum_{n=1}^{m} \frac{1}{n^2 N_n^2} + \sum_{n \geq m+1} \lambda_n \right)\]

with \(\prod_n N_n \leq N\). Set

\(m = [\log N], \ N_k = \left[ \frac{(m!N)^{1/m}}{k} \right], k = 1, \ldots, m.\)

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 $\mathbb{P}(\hat{\xi}^{(N_n)}_n = \beta^{(N_n)}_{i_n}, n = 1, \ldots, m(N))$ are explicit too . . .

since the normalized coordinates $\xi_n$ are independent so that

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

The distribution of a $K$-$L$ product quantization $\hat{W}$ is known.
Numerical aspects: optimal “integer bit allocation” i.e. solving

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

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

<table>
<thead>
<tr>
<th>(N)</th>
<th>(N_{\text{rec}})</th>
<th>Quant. Error</th>
<th>Opti. Alloc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.7071</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0.3138</td>
<td>5-2</td>
</tr>
<tr>
<td>100</td>
<td>96</td>
<td>0.2264</td>
<td>12-4-2</td>
</tr>
<tr>
<td>1,000</td>
<td>966</td>
<td>0.1881</td>
<td>23-7-3-2</td>
</tr>
<tr>
<td>10,000</td>
<td>9,984</td>
<td>0.1626</td>
<td>26-8-4-3-2-2</td>
</tr>
<tr>
<td>100,000</td>
<td>97,920</td>
<td>0.1461</td>
<td>34 – 10 – 6 – 4 – 3 – 2 – 2</td>
</tr>
</tbody>
</table>
Brownian product quantizations

Fig. 11: The $N_{\text{rec}}$-quantizer $\alpha^{(N)}$ for $N = 10$ ($N_{\text{rec}} = 10$).
Fig. 12: The $N_{\text{rec}}$-quantizer $\alpha^{(N)}$ for $N = 50$ ($N_{\text{rec}} = 12 \times 4 = 48$).
Fig. 13: The $N_{\text{rec}}$-quantizer $\alpha^{(N)}$ for $N = 100$ ($N_{\text{rec}} = 12 \times 4 \times 2 = 96$).
A cherry on the cake: stationarity again

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

\[
\hat{W} = \sum_{n \geq 1} \sqrt{\lambda_n} \xi_n^{(N_n)} e_n(t)
\]

so that

\[
\sigma(\hat{W}) = \sigma(\hat{\xi}_k^{(N_k)}, k \geq 1).
\]

and

\[
\mathbb{E}(W | \hat{W}) = \mathbb{E}(W | \sigma(\hat{\xi}_k^{(N_k)}, k \geq 1))
\]

\[
\mathbb{E}(W | \hat{W}) = \sum_{n \geq 1} \sqrt{\lambda_n} \mathbb{E} \left( \xi_n | \sigma(\hat{\xi}_k^{(N_k)}, k \geq 1) \right) e_n
\]

\( i.i.d. \)

\[
= \sum_{n \geq 1} \sqrt{\lambda_n} \mathbb{E} \left( \xi_n | \hat{\xi}_n^{(N_n)} \right) e_n
\]

\[
= \sum_{n \geq 1} \sqrt{\lambda_n} \hat{\xi}_n^{(N_n)} e_n = \hat{W}.
\]
Comparison with optimal quadratic functional quantization

- (Numerical) **Optimal Quantization** (in average over $1 \leq N \leq 10.000$)

$$e_N(W, L_T^2)^2 \approx \frac{0.2195}{\log N}$$

- Optimal **Product quantization**:

$$\min \left\{ \| W - \hat{W} \|_{L_T^2}^2, 1 \leq N_1 \cdots N_m \leq N, \ m \geq 1 \right\} \approx \frac{0.25}{\log N}$$

- 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 ($d = 1$)

- **Diffusion process:** $dX_t = b(t, X_t)dt + \vartheta(t, X_t)dW_t$
  - $b, \vartheta$ Lipschitz continuous, $\vartheta(t, .) = (\nabla S_t(.))^{-1}$ bounded, etc.

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

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

- **Theorem (Luschgy-P., SPA (2006))** $(x^{(N)})_{N \geq 1}$ is rate optimal i.e.
  \[
  \| |X - \tilde{X}^{x(N)}|_{L^2_T}^2 \| = O \left( \frac{1}{(\log N)^{1/2}} \right) \quad (\asymp \text{ if } \vartheta \geq \varepsilon_0 > 0)
  \]

  where
  \[
  \tilde{X}^{x(N)}_t = \sum_{k=1}^{N} x_i^{(N)}(t) 1_{\{W^{\alpha^N} = \alpha^N_i\}}
  \]

  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:

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

\[ W = (W^1, \ldots, W^d) \text{ is a } d\text{-dimensional B.M.} \]

\[ \min_{|\alpha| \leq N} \| W - \hat{W}^\alpha \|_2 \sim C_d \frac{1}{\sqrt{\log N}} \quad \text{as} \quad N \to \infty. \]

\[ \frac{1}{p}\text{-Hölder norm} \colon \mathbf{x}_{s,t} = (x^1_s, x^2_{s,t}), s \leq t. \]

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

Thus \[ W = (W_t, \int_s^t (W_u - W_s) \, dW_u) \]
Theorem (P.-Sellami, (2006), (2009) (a) Let $\alpha^N = (\alpha^N_1, \ldots, \alpha^N_N)$ be a sequence of optimal (stationary) $N$-product quantizers of $W$. Then

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

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

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

Set

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

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

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

The keys: connection with rough paths theory, Kolmogorov criterion, (pseudo-)stationarity.
Typical functionals

– Functionals $\text{F}_{L^2_T}$-continuous at every $\omega \in C([0, T])$?

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

where $f$ is locally Lipschitz continuous, namely

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

**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

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

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

$\triangleright$ Higher order Taylor expansion yields

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

$$\mathbb{E}F(W) = \mathbb{E}F(\widehat{W}^{(N)}) + \frac{1}{2} \mathbb{E} \left( D^2 F(\widehat{W}^{(N)}).(W - \widehat{W}^{(N)}) \otimes^2 \right) + o \left( (\log N)^{-\frac{3}{2} + \varepsilon} \right).$$
Conjecture: \( \mathbb{E} \left( D^2 F(\hat{W}^{(N)}). (W - \hat{W}^{(N)}) \otimes 2 \right) \sim \frac{c}{\log N}, \quad N \to \infty \)

Set

\[ M \ll N \quad (e.g. \, M \approx N/4) \]

and \( \forall \varepsilon > 0 \)

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

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

Replace \( \log N \) by \( 1/\|W - \hat{W}^{(N)}\|_2^2 \).
Application : Asian option in a Heston stochastic volatility model

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

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

▷ The payoff and the premium :

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

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

\[
S_t = s_0 \exp\left((r - \frac{1}{2}\Bar{v}_t)t + \rho \int_0^t \sqrt{v_s}dW^2_s\right) \exp\left(\sqrt{1 - \rho^2} \int_0^t \sqrt{v_s}d\Bar{W}^1_s\right)
\]
• Chaining rule for conditional expectations

\[ A_{\text{Call}}^{Hest}(s_0, K) = e^{-rT} \mathbb{E}\left( \mathbb{E}\left( \left( \frac{1}{T} \int_0^T S_s ds - K \right) \mid \sigma(W_t^2, 0 \leq t \leq T) \right) \right) \]

• State process = \((\tilde{W}_t^1, v_t)\).

• Solving the quantization ODE's for \((v_t)\) (by a Runge-Kutta scheme)

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

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

\[ \tilde{v}_t^{n,N} = \sum_i y_i^{n,N}(t) 1_{C_i(\alpha^N)}(W^2). \]
and 
\[ \widetilde{S}_{t}^{n,N} = \sum_{1 \leq i,j \leq N} s_{i,j}^{n,N}(t) 1_{\alpha_{i}^{N}}(\widetilde{W}^{1}) 1_{\alpha_{j}^{N}}(W^{2}). \]

with
\[ s_{i,j}^{n,N}(t) = s_{0} \exp \left( t \left( \left( r - \frac{\rho a_{k}}{\vartheta} \right) + \bar{y}_{j}^{n,N}(t) \left( \frac{\rho k}{\vartheta} - \frac{1}{2} \right) \right) + \frac{\rho}{\vartheta} (y_{j}^{n,N}(t) - v_{0}) \right) \times \exp \left( \sqrt{1 - \rho^{2}} \int_{0}^{t} \sqrt{y_{j}^{n,N}} \, d\alpha_{i}^{N} \right). \]

- Computation of *crude* quantized premium for $N$ and $M$.
- Space Romberg log-extrapolation $\hat{RCrAsCall}^{Hest}(s_{0}, K)$.
- $K$-linear interpolation $\hat{IRAsCall}^{Hest}(s_{0}, K)$ based on the (Asian) forward moneyness $K e^{-rT}$ and the Asian Call-Put parity formula

\[ \text{AsianCall}^{Hest}(s_{0}, K) - \text{AsianPut}(s_{0}, K) = s_{0} \frac{1 - e^{-rT}}{rT} - K e^{-rT}. \]
Fig. 14: Optimized Quantizer of the Heston volatility process $N = 400$
Parameters of the Heston model:

\[ s_0 = 100, \ k = 2, \ a = 0.01, \ \rho = 0.5, \ \nu_0 = 10\%, \ \vartheta = 20\%. \]

Parameters of the option portfolio:

\[ T = 1, \ K = 99, \cdots, 111 \quad \text{(13 strikes)}. \]

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:

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

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

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

then

\[ \forall 0 < p \leq r, \quad e_{N,r}(X, L^p_T) = O((\log N)^{-a}). \]

\[ \text{Ingredients: Haar basis (instead of K-L basis...), non asymptotic Zador Theorem (Pierce Lemma) and product functional quantization.} \]

\[ \text{Examples: • d-dim Itô processes (includes d-dim diffusions with sublinear coefficients) } a = 1/2; \]

• General Lévy process \( X \) with Lévy measure \( \nu \) (with Brownian component) \( a = 1/2; \)
• General Lévy process $X$ with Lévy measure $\nu$ (without Brownian component) with square integrable big jumps. Then

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

where

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

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

$$\forall 0 < p \leq r < \alpha, \quad e_{N,r}(X, L^p_T) \approx O((\log N)^{-\alpha})$$
A guided Monte Carlo method: hybrid “Q+MC”

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

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

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

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

Drawback: nearest neighbour search [complexity $= O(\log N)$] at each step...
Quantization based universal stratified sampling (with J. Printems (2008) and S. Corlay (2009))

- Let $\alpha$ be a product $N$-quantizer with structural dimension $d(N) = \log N$.
- The idea starts from the ability to simulate

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

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):

\[ p_i = \mathbb{P}(\hat{W} = \alpha_i) \quad \text{and} \quad \sigma_i^2 = \text{Var}(W | \hat{W} = \alpha_i) \]

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

\[ (W^m_{t_1}, \ldots, W^m_{t_n}) \sim \mathcal{L}(W_{t_1}, \ldots, W_{t_n} | \hat{W} = \alpha_i), \; 1 \leq m \leq M_i, \; i.i.d. \]

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:

\[ \frac{\|X - \hat{X}^\alpha\|_2^2}{\|X - \mathbb{E}X\|_2^2} \]

like for control variate... but no nearest neighbour search.