# Functional quantization and metric entropy for Riemann-Liouville processes

HARALD LUSCHGY* AND GILLES PAGÈS †

January 24, 2005

**Abstract**

We derive a high-resolution formula for the $L2$-quantization errors of Riemann-Liouville processes and the sharp Kolmogorov entropy asymptotics for related Sobolev balls. We describe a quantization procedure which leads to asymptotically optimal functional quantizers. Regular variation of the eigenvalues of the covariance operator plays a crucial role.

*Keywords: Functional quantization, metric entropy, Gaussian process, Riemann-Liouville process, optimal quantizer.*

*MSC: 60G15, 60E99, 41A46.*

## 1 Introduction

Functional quantization of stochastic processes can be seen as a discretization of the path-space of a process and the approximation (coding) of a process by finitely many deterministic functions from its path-space. In a Hilbert space setting this reads as follows.

Let $(H, <\cdot,\cdot>)$ be a separable Hilbert space with norm $\|\cdot\|$ and let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to H$ be a random vector taking its values in $H$ with distribution $\mathbb{P}_X$. For $n \in \mathbb{N}$, the $L2$-quantization problem for $X$ of level $n$ (or of nat-level $\log n$) consists in minimizing

$$\left(\mathbb{E} \min_{a\in\alpha} \|X-a\|^2\right)^{1/2} = \| \min_{a\in\alpha} \|X-a\| \|_{L2(\mathbb{P})}$$

over all subsets $\alpha \subset H$ with $\mathrm{card}(\alpha) \le n$. Such a set $\alpha$ is called $n$-codebook or $n$-quantizer. The minimal $n$th quantization error of $X$ is then defined by

$$e_n(X) := \inf\left\{(\mathbb{E} \min_{a\in\alpha} \|X-a\|^2)^{1/2} : \alpha \subset H,\ \mathrm{card}(\alpha) \le n\right\}. \tag{1.1}$$

Under the integrability condition

$$\mathbb{E}\,\|X\|^2 < \infty \tag{1.2}$$

the quantity $e_n(X)$ is finite.

*Universität Trier, FB IV-Mathematik, D-54286 Trier, Germany. E-mail: luschgy@uni-trier.de

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

Dagstuhl Seminar Proceedings 04401
Algorithms and Complexity for Continuous Problems
http://drops.dagstuhl.de/opus/volltexte/2005/144

For a given $n$-quantizer $\alpha$ one defines an associated closest neighbour projection

$$\pi_\alpha := \sum_{a\in\alpha} a\mathbf{1}_{C_a(\alpha)}$$

and the induced $\alpha$-quantization (Voronoi quantization) of $X$ by

$$\hat{X}^\alpha := \pi_\alpha(X), \tag{1.3}$$

where $\{C_a(\alpha) : a \in \alpha\}$ is a Voronoi partition induced by $\alpha$, that is a Borel partition of $H$ satisfying

$$C_a(\alpha) \subset V_a(\alpha) := \{x \in H : \|x-a\| = \min_{b\in\alpha} \|x-b\|\} \tag{1.4}$$

for every $a \in \alpha$. Then one easily checks that, for any random vector $X' : \Omega \to \alpha \subset H$,

$$\mathbb{E}\,\|X - X'\|^2 \geq \mathbb{E}\,\|X - \hat{X}^\alpha\|^2 = \mathbb{E}\,\min_{a\in\alpha} \|X-a\|^2$$

so that finally

$$\begin{aligned} e_n(X) &= \inf\Big\{(\mathbb{E}\,\|X-\hat{X}\|^2)^{1/2} : \hat{X} = f(X), f : H \to H \text{ Borel measurable,} \\ &\qquad \operatorname{card}(f(H)) \leq n\} \\ &= \inf\Big\{(\mathbb{E}\,\|X-\hat{X}\|^2)^{1/2} : \hat{X} : \Omega \to H \text{ random vector, } \operatorname{card}(\hat{X}(\Omega)) \leq n\Big\}. \end{aligned} \tag{1.5}$$

Observe that the Voronoi cells $V_a(\alpha), a \in \alpha$ are closed and convex (where convexity is a characteristic feature of the underlying Hilbert structure). Note further that there are infinitely many $\alpha$-quantizations of $X$ which all produce the same quantization error and $\hat{X}^\alpha$ is $\mathbb{P}$-a.s. uniquely defined if $\mathbb{P}_X$ vanishes on hyperplanes.

A typical setting for functional quantization is $H = L2([0,1], dt)$ but is obviously not restricted to the Hilbert space setting. Functional quantization is the natural extension to stochastic processes of the so-called optimal vector quantization of random vectors in $H = \mathbb{R}^d$ which has been extensively investigated since the late 1940's in Signal processing and Information Theory (see [4], [8]). For the mathematical aspects of vector quantization in $\mathbb{R}^d$, one may consult [5], for algorithmic aspects see [15] and "non-classical" applications can be found in [14], [16]. For a first promising application of functional quantization to the pricing of financial derivatives through numerical integration on path-spaces see [17].

We address the issue of high-resolution quantization which concerns the performance of $n$-quantizers and the behaviour of $e_n(X)$ as $n \to \infty$. The asymptotics of $e_n(X)$ for $\mathbb{R}^d$-valued random vectors has been completely elucidated for non-singular distributions $\mathbb{P}_X$ by the Zador Theorem (see [5]) and for a class of self-similar (singular) distributions by [6]. In infinite dimensions no such global results hold, even for Gaussian processes.

It is convenient to use the symbols $\sim$ and $\stackrel{<}{\sim}$, where $a_n \sim b_n$ means $a_n/b_n \to 1$ and $a_n \stackrel{<}{\sim} b_n$ means $\limsup_{n\to\infty} a_n/b_n \leq 1$. A measurable function $\varphi : (s,\infty) \to (0,\infty)\,(s \geq 0)$ is said to be regularly varying at infinity with index $b \in \mathbb{R}$ if, for every $c > 0$,

$$\lim_{x\to\infty} \frac{\varphi(cx)}{\varphi(x)} = c^b.$$

Now let $X$ be centered Gaussian. Denote by $K_X \subset H$ the reproducing kernel Hilbert space (Cameron-Martin space) associated to the covariance operator

$$C_X : H \to H,\ C_X y := \mathbb{E}\,(<y, X> X) \tag{1.6}$$

of $X$. Let $\lambda_1 \geq \lambda_2 \geq \ldots > 0$ be the ordered nonzero eigenvalues of $C_X$ and let $\{u_j : j \geq 1\}$ be the corresponding orthonormal basis of $\operatorname{supp}(\mathbb{P}_X)$ consisting of eigenvectors (Karhunen-Loève basis). If $d := \dim K_X < \infty$, then $e_n(X) = e_n\left(\bigotimes_{j=1}^{d} N(0,\lambda_j)\right)$, the minimal $n$th $L2$-quantization error of $\bigotimes_{j=1}^{d} N(0,\lambda_j)$ with respect to the $l_2$-norm on $\mathbb{R}^d$, and thus we can read off the asymptotic behaviour of $e_n(X)$ from the high-resolution formula

$$e_n(\bigotimes_{j=1}^{d} N(0,\lambda_j)) \sim q(d)\sqrt{2\pi}\left(\Pi_{j=1}^{d}\lambda_j\right)^{1/2d}\left(\frac{d+2}{d}\right)^{(d+2)/4} n^{-1/d} \quad \text{as} \quad n \to \infty \tag{1.7}$$

where $q(d) \in (0,\infty)$ is a constant depending only on the dimension $d$ (see [5]). Except in dimension $d = 1$ and $d = 2$, the true value of $q(d)$ is unknown. However, one knows (see [5]) that

$$q(d) \sim \left(\frac{d}{2\pi e}\right)^{1/2} \quad \text{as} \quad d \to \infty. \tag{1.8}$$

Assume $\dim K_X = \infty$. Under regular behaviour of the eigenvalues the sharp asymptotics of $e_n(X)$ can be derived analogously to (1.7). In view of (1.8) it is reasonable to expect that the limiting constants can be evaluated. The recent high-resolution formula is as follows.

**Theorem 1** *([11]) Let $X$ be a centered Gaussian. Assume $\lambda_j \sim \varphi(j)$ as $j \to \infty$, where $\varphi : (s,\infty) \to (0,\infty)$ is a decreasing, regularly varying function at infinity of index $-b < -1$ for some $s \geq 0$. Set, for every $x > s$,*

$$\psi(x) := \frac{1}{x\varphi(x)}.$$

*Then*

$$e_n(X) \sim \left(\left(\frac{b}{2}\right)^{b-1}\frac{b}{b-1}\right)^{1/2} \psi(\log n)^{-1/2} \quad \text{as} \quad n \to \infty.$$

A high-resolution formula in case $b = 1$ is also available (see [11]). Note that the restriction $-b \leq -1$ on the index of $\varphi$ is natural since $\sum_{j=1}^{\infty} \lambda_j < \infty$. The minimal $L^r$-quantization errors of $X$, $0 < r < \infty$, are strongly equivalent to the $L2$-errors $e_n(X)$ (see [2]) and thus exhibit the same high-resolution behaviour.

A related quantization problem is the Kolmogorov metric entropy problem for the closed unit ball

$$U_X := \left\{x \in K_X : \|x\|_{K_X} \leq 1\} = \{x \in \operatorname{supp}(\mathbb{P}_X) : \sum_{j\geq 1} \frac{<x,u_j>^2}{\lambda_j} \leq 1\right\} \tag{1.9}$$

of $K_X$ (Strassen ball). Note that $U_X$ is a compact subset of $H$. For $n \in \mathbb{N}$, the metric entropy problem for $U_X$ consists in minimizing

$$\max_{x\in U_X} \min_{a\in\alpha} \|x - a\| = \|\min_{a\in\alpha} \|X' - a\|\|_{L^\infty(\mathbb{P})}$$

over all subsets $\alpha \subset H$ with $\operatorname{card}(\alpha) \leq n$, where $X'$ is any $H$-valued random vector with $\operatorname{supp}(\mathbb{P}_{X'}) = U_X$. The $n$th entropy number is then defined by

$$e_n(U_X) := \inf\left\{\max_{x\in U_X}\min_{a\in\alpha}\|x-a\| : \alpha \subset H,\ \operatorname{card}(\alpha) \leq n\right\}. \tag{1.10}$$

If $d := \dim K_X < \infty$, then $e_n(U_X) = e_n(\mathcal{E}_d)$, the $n$th entropy number of the ellipsoid

$$\mathcal{E}_d := \left\{ x \in \mathbb{R}^d : \sum_{j=1}^{d} \frac{x2_j}{\lambda_j} \leq 1 \right\}$$

with respect to the $l_2$-norm on $\mathbb{R}^d$. Thus we can read off the asymptotic behaviour of $e_n(U_X)$ from the formula

$$e_n(\mathcal{E}) \sim p(d)(\Pi_{j=1}^d \lambda_j)^{1/2}(\text{vol } (B_d(0,1)))^{1/d} n^{-1/d} \quad \text{as} \quad n \to \infty \tag{1.11}$$

where the constant $p(d) \in (0, \infty)$ is unknown for $d \geq 3$ and $p(d) \sim q(d), d \to \infty$ (see [9], [5]).

If $\dim K_X = \infty$, the recent solution of the Kolmogorov metric entropy problem for $U_X$ is as follows.

**Theorem 2** *([12]) Assume the situation of Theorem 1. Then*

$$e_n(U_X) \sim \left(\frac{b}{2}\right)^{b/2} \varphi(\log n)^{1/2} \quad as \quad n \to \infty.$$

This formula is still valid for $b = 1$ and, ignoring the probabilistic interpretation, also for $b \geq 0$ $(00 := 1)$ provided $\lambda_j \to 0$ as $j \to \infty$. (see [7], [12]). A different approach via the inverse of $e_n(U_X)$, the Kolmogorov $\varepsilon$-entropy, is due to Donoho [3]. (However, his result does not provide the correct constant.)

From Theorems 1 and 2 we conclude that functional quantization and metric entropy are related by

$$e_n(X) \sim \left(\frac{2 \log n}{b-1}\right)^{1/2} e_n(U_X) \quad \text{as} \quad n \to \infty. \tag{1.12}$$

The paper is organized as follows. In Section 2 we investigate Riemann-Liouville processes in $H = L2([0,1], dt)$. For $\rho \in (0, \infty)$, the Riemann-Liouville process $X^\rho = (X_t^\rho)_{t \in [0,1]}$ on $[0,1]$ is defined by

$$X_t^\rho := \int_0^t (t-s)^{\rho - \frac{1}{2}} dW_s \tag{1.13}$$

where $W$ is a standard Brownian motion. We derive a high-resolution formula for $X^\rho$ and correspondingly, the precise entropy asymptotics for fractional Sobolev balls. As a consequence we obtain a new result for fractionally integrated Brownian motions. In Section 3 we describe a quantization procedure which furnishes asymptotically optimal quantizers in the situation of Theorem 1. Here the Karhunen-Loève expansion plays a crucial rôle. In Section 4 we discuss a dimension conjecture.

## 2 Riemann-Liouville processes

Let $X^\rho = (X_t^\rho)_{t \in [0,1]}$ be the Riemann-Liouville process of index $\rho \in (0, \infty)$ as defined in (1.13). Its covariance function is given by

$$\mathbb{E}\, X_s^\rho X_t^\rho = \int_0^{s \wedge t} (t-r)^{\rho - \frac{1}{2}} (s-r)^{\rho - \frac{1}{2}} dr. \tag{2.1}$$

Using $\rho \wedge \frac{1}{2}$-Hölder continuity of the application $t \mapsto X_t^\rho$ from [0,1] into $L2(\mathbb{P})$ and the Kolmorogov criterion one checks that $X^\rho$ has a pathwise continuous modification so that we may assume without loss of generality that $X^\rho$ is pathwise continuous. In particular, $X^\rho$ can be seen as a centered Gaussian random vector with values in

$$H = L2([0,1], dt).$$

The following high-resolution formula relies on a theorem by Vu and Gorenflo [18] on singular values of Riemann-Liouville integral operators

$$R_\beta\, g(t) = \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} g(s) ds, \qquad \beta \in (0,\infty). \tag{2.2}$$

**Theorem 3** *For every $\rho \in (0,\infty)$,*

$$e_n(X^\rho) \sim \pi^{-(\rho+\frac{1}{2})}(\rho+1/2)^\rho (\frac{2\rho+1}{2\rho})^{1/2}\Gamma(\rho+1/2)(\log n)^{-\rho} \quad \textit{as} \quad n \to \infty.$$

**Proof.** For $\beta > 1/2$, the Riemann-Liouville fractional integral operator $R_\beta$ is a bounded operator from $L2([0,1], dt)$ into $L2([0,1], dt)$. The covariance operator

$$C_\rho : L2([0,1], dt) \to L2([0,1], dt)$$

of $X^\rho$ is given by the Fredholm transformation

$$C_\rho g(t) = \int 1_0 g(s) E X_s^\rho X_t^\rho ds.$$

Using (2.1), one checks that $C_\rho$ admits a factorization

$$C_\rho = S_\rho S_\rho^*,$$

where

$$S_\rho = \Gamma(\rho+1/2) R_{\rho+\frac{1}{2}}.$$

Consequently, it follows from Theorem 1 in [18] that the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots > 0$ of $C_\rho$ satisfy

$$\lambda_j \sim \Gamma(\rho+1/2)2(\pi j)^{-(2\rho+1)} \quad \text{as} \quad j \to \infty. \tag{2.3}$$

Now the assertion follows from Theorem 1 (with $\varphi(x) = \Gamma(\rho+1/2)2\pi^{-b}x^{-b}$ and $b = 2\rho+1$). □

An immediate consequence for fractionally integrated Brownian motions on $[0,1]$ defined by

$$Y_t^\beta := \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} W_s ds \tag{2.4}$$

for $\beta \in (0,\infty)$ is as follows.

**Corollary 1** *For every $\beta \in (0,\infty)$,*

$$e_n(Y^\beta) \sim \pi^{-(\beta+1)}(\beta+1)^{\beta+\frac{1}{2}} (\frac{2\beta+2}{2\beta+1})^{1/2} (\log n)^{-(\beta+\frac{1}{2})} \quad \textit{as} \quad n \to \infty.$$

**Proof.** For $\rho > 1/2$, the Ito formula yields

$$X_t^\rho = (\rho - \frac{1}{2}) \int_0^t (t-s)^{\rho - \frac{3}{2}} W_s ds.$$

Consequently,

$$Y_t^\beta = \frac{1}{\beta \Gamma(\beta)} \beta \int_0^t (t-s)^{\beta + \frac{1}{2} - \frac{3}{2}} W_s ds = \frac{1}{\Gamma(1+\beta)} X_t^{\beta + \frac{1}{2}}.$$

The assertion follows from Theorem 3. □

**Remark.** The preceding corollary provides new high-resolution formulas for $e_n(Y^\beta)$ in the cases $\beta \in (0, \infty) \setminus \mathbb{N}$.

One further consequence is a precise relationship between the quantization errors of Riemann-Liouville processes and fractional Brownian motions. The fractional Brownian motion with Hurst exponent $\rho \in (0,1]$ is a centered pathwise continuous Gaussian process $Z^\rho = (Z_t^\rho)_{t \in [0,1]}$ having the covariance function

$$\mathbb{E}\, Z_s^\rho Z_t^\rho = \frac{1}{2}(s^{2\rho} + t^{2\rho} - \mid s - t \mid^{2\rho}). \tag{2.5}$$

**Corollary 2** *For every $\rho \in (0,1)$,*

$$e_n(X^\rho) \sim \frac{\Gamma(\rho + 1/2)}{(\Gamma(2\rho+1)\sin(\pi\rho))^{1/2}} e_n(Z^\rho) \quad as \quad n \to \infty.$$

**Proof.** By [11], we have

$$e_n(Z^\rho) \sim \pi^{-(\rho+\frac{1}{2})} (\rho + 1/2)^\rho \left(\frac{2\rho+1}{2\rho}\right)^{1/2} (\Gamma(2\rho+1)\sin(\pi\rho))^{1/2} (\log n)^{-\rho}, n \to \infty.$$

Combining this formula with Theorem 3 yields the assertion □

Observe that strong equivalence $e_n(X^\rho) \sim e_n(Z^\rho)$ as $n \to \infty$ is true for exactly two values of $\rho \in (0,1)$, namely for $\rho = 1/2$ where even $e_n(X^{1/2}) = e_n(Z^{1/2}) = e_n(W)$ and, a bit mysterious, for $\rho = 0.81557\ldots$

Now consider the Strassen ball $U_\rho$ of $X^\rho$. Since the covariance operator $C_\rho$ satisfies $C_\rho = \Gamma(\rho + \frac{1}{2}) R_{\rho+\frac{1}{2}} (\Gamma(\rho+\frac{1}{2}) R_{\rho+\frac{1}{2}})^*$, one gets

$$\begin{aligned} U_\rho &= \Gamma(\rho + 1/2) R_{\rho + \frac{1}{2}} (B_{L2}(0,1)) \\ &= \left\{ R_{\rho+1/2} g : g \in L2([0,1], dt), \int 1_0 g(t)^2 dt \le \Gamma(\rho+1/2)2 \right\}, \end{aligned} \tag{2.6}$$

a fractional Sobolev ball. Theorem 2 and (2.3) yield the solution of the entropy problem for fractional Sobolev balls.

**Theorem 4** *For every $\rho \in (0, \infty)$,*

$$\begin{aligned} e_n(U_\rho) &\sim \left(\frac{\rho + \frac{1}{2}}{\pi}\right)^{\rho+\frac{1}{2}} \Gamma(\rho + 1/2)(\log n)^{-(\rho+\frac{1}{2})} \\ &\sim \left(\frac{\rho}{\log n}\right)^{1/2} e_n(X^\rho) \quad as \quad n \to \infty. \end{aligned}$$

## 3 Asymptotically optimal functional quantizers

Let $X$ be a $H$-valued random vector satisfying (1.2). For every $n \in \mathbb{N}$, $L2$-optimal $n$-quantizers $\alpha \subset H$ exist, that is

$$(\mathbb{E} \min_{a \in \alpha} \|X - a\|^2)^{1/2} = e_n(X).$$

If card $(\mathrm{supp}(\mathbb{P}_X)) \geq n$, optimal $n$-quantizers $\alpha$ satisfy $\mathrm{card}(\alpha) = n$, $\mathbb{P}(X \in C_a(\alpha)) > 0$ and the stationarity condition

$$a = \mathbb{E}\,(X \mid \{X \in C_a(\alpha)\}),\ a \in \alpha$$

or what is the same

$$\hat{X}^{\alpha} = \mathbb{E}\,(X \mid \hat{X}^{\alpha}) \tag{3.1}$$

for every Voronoi partition $\{C_a(\alpha) : a \in \alpha\}$ (see [10]). In particular, $\mathbb{E}\,\hat{X}^{\alpha} = \mathbb{E}\,X$.

Now let $X$ be centered Gaussian with $\dim K_X = \infty$. The Karhunen-Loève basis $\{u_j : j \geq 1\}$ consisting of normalized eigenvectors of $C_X$ is optimal for the quantization of Gaussian random vectors (see [10]). So we start with the Karhunen-Loève expansion

$$X \overset{H}{=} \sum_{j=1}^{\infty} \lambda_j^{1/2} Z_j u_j,$$

where $Z_j =< X, u_j > /\lambda_j^{1/2}, j \geq 1$ are i.i.d. $N(0,1)$-distributed random variables. The design of an asymptotically optimal quantization of $X$ is based on optimal quantizing blocks of coefficients of variable ($n$-dependent) block length. Let $n \in \mathbb{N}$ and fix temporarily $m, l, n_1, \ldots, n_m \in \mathbb{N}$ with $\Pi_{j=1}^m n_j \leq n$, where $m$ denotes the number of blocks, $l$ the block length and $n_j$ the size of the quantizer for the $j$th block

$$Z^{(j)} := (Z_{(j-1)l+1}, \ldots, Z_{jl}), \quad j \in \{1, \ldots, m\}.$$

Let $\alpha_j \subset \mathbb{R}^l$ be an $L2$-optimal $n_j$-quantizer for $Z^{(j)}$ and let $\widehat{Z^{(j)}} = \widehat{Z^{(j)}}^{\alpha_j}$ be a $\alpha_j$-quantization of $Z^{(j)}$. Then, define a quantized version of $X$ by

$$\hat{X}^n := \sum_{j=1}^{m} \sum_{k=1}^{l} \lambda_{(j-1)l+k}^{1/2} (\widehat{Z^{(j)}})_k u_{(j-1)l+k}. \tag{3.2}$$

It is clear that $\mathrm{card}(\hat{X}^n(\Omega)) \leq n$. Using (3.1) for $Z^{(j)}$, one gets $\mathbb{E}\,\hat{X}^n = 0$. If

$$\widehat{Z^{(j)}} = \sum_{b \in \alpha_j} b \mathbf{1}_{C_b(\alpha_j)}(Z^{(j)}),$$

then

$$\hat{X}^n = \sum_{a \in \times_{j=1}^m \alpha_j} (\sum_{j=1}^{m} \sum_{k=1}^{l} \lambda_{(j-1)l+k}^{1/2} a_k^{(j)} u_{(j-1)l+k}) \Pi_{j=1}^m \mathbf{1}_{C_{a^{(j)}}(\alpha_j)}(Z^{(j)})$$

where $a = (a^{(1)}, \ldots, a^{(m)}) \in \times_{j=1}^m \alpha_j$. Observe that in general, $\hat{X}^n$ is not a Voronoi quantization of $X$ since it is based on the (less complicated) Voronoi partitions for $Z^{(j)}, j \leq m$. ($\hat{X}^n$ is a Voronoi quantization if $l = 1$ or if $\lambda_{(j-1)l+1} = \ldots = \lambda_{jl}$ for every $j$.) Using again (3.1) for $Z^{(j)}$ and the independence structure, one checks that $\hat{X}^n$ satisfies a kind of stationarity equation:

$$\mathbb{E}\,(X \mid \hat{X}^n) = \hat{X}^n.$$

**Lemma 1** *Let $n \geq 1$. For every $l \geq 1$ and every $m \geq 1$*

$$\mathbb{E}\,\|X - \hat{X}^n\|^2 \leq \sum_{j=1}^{m} \lambda_{(j-1)l+1} e_{n_j}(N(0, I_l))2 + \sum_{j \geq ml+1} \lambda_j. \tag{3.3}$$

*Furthermore, (3.3) stands as an equality if $l = 1$ (or $\lambda_{(j-1)l+1} = \ldots = \lambda_{jl}$ for every $j$, $l \geq 1$).*

**Proof.** The claim follows from the orthonormality of the basis $\{u_j : j \geq 1\}$. We have

$$\begin{aligned}
\mathbb{E}\,\|X - \hat{X}^n\|^2 &= \sum_{j=1}^{m}\sum_{k=1}^{l} \lambda_{(j-1)l+k}\mathbb{E} \mid Z_k^{(j)} - (\widehat{Z^{(j)}})_k \mid 2 + \sum_{j \geq ml+1} \lambda_j \\
&\leq \sum_{j=1}^{m} \lambda_{(j-1)l+1} \sum_{k=1}^{l} \mathbb{E} \mid Z_k^{(j)} - \widehat{Z^{(j)}})_k \mid 2 + \sum_{j \geq ml+1} \lambda_j \\
&= \sum_{j=1}^{m} \lambda_{(j-1)l+1} e_{n_j}(Z^{(j)})2 + \sum_{j \geq ml+1} \lambda_j.
\end{aligned}$$

□

Set

$$C(l) := \sup_{k \geq 1} k^{2/l} e_k(N(0, I_l))2. \tag{3.4}$$

By (1.7), $C(l) < \infty$. For every $l \in \mathbb{N}$,

$$e_{n_j}(N(0, I_l)2 \leq n_j^{-2/l} C(l) \tag{3.5}$$

Then one may replace the optimization problem which consists, for fixed $n$, in minimizing the right hand side of Lemma 1 by the following optimal allocation problem:

$$\min\{C(l) \sum_{j=1}^{m} \lambda_{(j-1)l+1} n_j^{-2/l} + \sum_{j \geq ml+1} \lambda_j : m, l, n_1, \ldots, n_m \in \mathbb{N}, \Pi_{j=1}^m n_j \leq n\}. \tag{3.6}$$

Set

$$m = m(n,l) := \max\{k \geq 1 : n^{1/k} \lambda_{(k-1)l+1}^{l/2} (\Pi_{j=1}^k \lambda_{(j-1)l+1})^{-l/2k} \geq 1\}, \tag{3.7}$$

$$n_j = n_j(n,l) := [n^{1/m} \lambda_{(j-1)l+1}^{l/2} (\Pi_{i=1}^m \lambda_{(i-1)l+1})^{-l/2m}],\ j \in \{1, \ldots, m\}, \tag{3.8}$$

where $[x]$ denotes the integer part of $x \in \mathbb{R}$ and

$$l = l_n := [(\max\{1, \log n\})^{\vartheta}],\ \vartheta \in (0,1). \tag{3.9}$$

In the following theorem it is demonstrated that this choice is at least asymptotically optimal provided the eigenvalues are regularly varying.

**Theorem 5** *Assume the situation of Theorem 1. Consider $\hat{X}^n$ with tuning parameters defined in (3.7)-(3.9). Then $\hat{X}^n$ is asymptotically $n$-optimal, i.e.*

$$(\mathbb{E}\,\|X - \hat{X}^n\|^2)^{1/2} \sim e_n(X) \ \text{ as } \ n \to \infty.$$

Note that no block quantizer with fixed block length is asymptotically optimal (see [11]). As mentioned above, $\hat{X}^n$ is not a Voronoi quantization of $X$. If $\alpha_n := \hat{X}^n(\Omega)$, then the Voronoi quantization $\hat{X}^{\alpha_n}$ is clearly also asymptotically $n$-optimal.

The key property for the proof is the following $l$-asymptotics of the constants $C(l)$ defined in (3.4). It is interesting to consider also the smaller constants

$$Q(l) := \lim_{k\to\infty} k^{2/l} e_k(N(0,I_l))2 \tag{3.10}$$

(see (1.7)).

**Proposition 1** *The sequences $(C(l))_{l\geq 1}$ and $(Q(l))_{l\geq 1}$ satisfy*

$$\lim_{l\to\infty} \frac{C(l)}{l} = \lim_{l\to\infty} \frac{Q(l)}{l} = \inf_{l\geq 1} \frac{C(l)}{l} = \inf_{l\geq 1} \frac{Q(l)}{l} = 1.$$

**Proof.** From [11] it is known that

$$\liminf_{l\to\infty} \frac{C(l)}{l} = 1. \tag{3.11}$$

Furthermore, it follows immediately from (1.7) and (1.8) that

$$\lim_{l\to\infty} \frac{Q(l)}{l} = 1. \tag{3.12}$$

(The proof of the existence of $\lim_{l\to\infty} C(l)/l$ we owe to S. Dereich.) For $l_0, l \in \mathbb{N}$ with $l \geq l_0$, write

$$l = n\, l_0 + m \text{ with } n \in \mathbb{N}, m \in \{0, \ldots, l_0 - 1\}.$$

Since for every $k \in \mathbb{N}$,

$$[k^{l_0/l}]^n \; [k^{1/l}]^m \leq k,$$

one obtains by a block-quantizer design consisting of $n$ blocks of length $l_0$ and $m$ blocks of length 1 for quantizing $N(0, I_l)$,

$$e_k(N(0,I_l))2 \leq n e_{[k^{l_0/l}]}(N(0,I_{l_0}))2 + m e_{[k^{1/l}]}(N(0,1))2. \tag{3.13}$$

This implies

$$\begin{aligned} C(l) &\leq nC(l_0) \sup_{k\geq 1} \frac{k^{2/l}}{[k^{l_0/l}]^{2/l_0}} + mC(1) \sup_{k\geq 1} \frac{k^{2/l}}{[k^{1/l}]^2} \\ &\leq 4^{1/l_0} nC(l_0) + 4mC(1). \end{aligned}$$

Consequently, using $n/l \leq 1/l_0$,

$$\frac{C(l)}{l} \leq \frac{4^{1/l_0} C(l_0)}{l_0} + \frac{4mC(1)}{l}$$

and hence

$$\limsup_{l\to\infty} \frac{C(l)}{l} \leq \frac{4^{1/l_0} C(l_0)}{l_0}.$$

This yields

$$\limsup_{l\to\infty} \frac{C(l)}{l} \leq \liminf_{l_0\to\infty} \frac{C(l_0)}{l_0} = 1. \tag{3.14}$$

It follows from (3.13) that

$$Q(l) \leq nQ(l_0) + mQ(1).$$

Consequently

$$\frac{Q(l)}{l} \le \frac{Q(l_0)}{l_0} + \frac{mQ(1)}{l}$$

and therefore

$$1 = \lim_{l\to\infty} \frac{Q(l)}{l} \le \frac{Q(l_0)}{l_0}.$$

This implies

$$\inf_{l_0\ge 1} \frac{Q(l_0)}{l_0} = 1. \tag{3.15}$$

Since $Q(l) \le C(l)$, the proof is complete. □

The $n$-asymptotics of the number $m(n,l_n)l_n$ of quantized coefficients in the Karhunen-Loève expansion in the quantization $\hat{X}^n$ is as follows.

**Lemma 2** *([12], Lemma 4.8) Assume the situation of Theorem 1. Let $m(n,l_n)$ be defined by (3.7) and (3.9). Then*

$$m(n,l_n)l_n \sim \frac{2\log n}{b} \quad \textit{as} \quad n\to\infty.$$

**Proof of Theorem 5.** For every $n \in \mathbb{N}$,

$$\begin{aligned}
\sum_{j=1}^{m} \lambda_{(j-1)l+1} n_j^{-2/l} \quad &\le \quad \sum_{j=1}^{m} \lambda_{(j-1)l+1} (n_j+1)^{-2/l} (\frac{n_j+1}{n_j})^{2/l} \\
&\le \quad 4^{1/l} m n^{-2/ml} (\Pi_{j=1}^{m} \lambda_{(j-1)l+1})^{1/m} \\
&\le \quad 4^{1/l} m \lambda_{(m-1)l+1}.
\end{aligned}$$

Therefore, by Lemma 1 and (3.5),

$$\mathbb{E}\,\|X-\hat{X}^n\|^2 \le 4^{1/l}\frac{C(l)}{l} ml\lambda_{(m-1)l+1} + \sum_{j\ge ml+1} \lambda_j$$

for every $n \in \mathbb{N}$. By Lemma 2, we have

$$ml = m(n,l_n)l_n \sim \frac{2\log n}{b} \quad \text{as} \quad n\to\infty.$$

Consequently, using regular variation at infinity with index $-b < -1$ of the function $\varphi$,

$$ml\lambda_{(m-1)l+1} \sim ml\lambda_{ml} \sim \left(\frac{2}{b}\right)^{1-b} \psi(\log n)^{-1}$$

and

$$\sum_{j\ge ml+1} \lambda_j \sim \frac{ml\varphi(ml)}{b-1} \sim \frac{1}{b-1}\left(\frac{2}{b}\right)^{1-b} \psi(\log n)^{-1} \quad \text{as} \quad n\to\infty,$$

where, like in Theorem 1, $\psi(x) = 1/x\varphi(x)$. Since by Proposition 1,

$$\lim_{n\to\infty} \frac{4^{1/l_n} C(l_n)}{l_n} = 1,$$

one concludes

$$\mathbb{E}\,\|X-\hat{X}^n\|^2 \lesssim \left(\frac{2}{b}\right)^{1-b} \frac{b}{b-1}\,\psi(\log n)^{-1} \quad \text{as} \quad n\to\infty.$$

The assertion follows from Theorem 1. □

Numerical and computational aspects: As soon as the Karhunen-Loève basis $(u_j)_{j\geq 1}$ of a Gaussian process $X$ is explicit, it is possible to compute the asymptotically optimal functional quantization (3.2) which solves the minimization problem (3.6) as well as its distribution and induced quantization error (at least for a given $\vartheta \in (0,1)$). This is possible since some optimal (or at least locally optimal) vector quantizations of the $N(0,I_d)$-distribution has been already computed and kept off line. Let us be more specific.

In 1-dimension, the normal distribution $N(0,1)$ has only one stationary $n$-quantizer – hence optimal – since its probability density is log-concave (for this result due to Kiefer, see *e.g.* [5]). Deterministic methods to compute these optimal quantizers are based on the stationary equation (3.1). They are very easy to implement, converge very fast with a very high accuracy. The Newton-Raphson algorithm is a possible choice (see [15] for details). Closed forms for the lowest quadratic quantization error $\mathbb{E}\|Z-\widehat{Z}\|^2$ and for the distribution of the optimal $n$-quantization $\widehat{Z}^\alpha$ as a function of the optimal $n$-quantizer $\alpha$ are also available in [15]. These three quantities have been tabulated up to very high values of $n$. A file can be downloaded at the URL `www.proba.jussieu.fr/pageperso/pages.html`.

In higher dimension, one still relies on the stationary equation (3.1) which reads:

$$\mathbb{E}\left(\mathbf{1}_{C_a(\alpha)}(Z)(a-Z)\right)=0, \qquad a\in\alpha.$$

One must keep in mind that the left hand side of the above equation is but the gradient of the (squared) quantization error $\mathbb{E}\|Z-\widehat{Z}^\alpha\|^2$ viewed as a function of the quantizer $\alpha$ (assumed to be of full size $n$). A stochastic gradient descent based on this integral representation can be implemented easily since the normal distribution $N(0,I_d)$ can be simulated on a computer from (pseudo-)random numbers (*e.g.* by the Box-Muller method). This algorithm is known as the Competitive Learning Vector Quantization (or $CLVQ$) algorithm. It has been extensively investigated both from a theoretical (see *e.g.* [14], [1]) and numerical (see *e.g.* [15] as concerns normally distributed vectors) viewpoints. The algorithm reads as follows: let $(\zeta(t))_{t\geq 1}$ be an i.i.d. sequence of $N(0,I_d)$-distributed random vectors, let $(\gamma_t)_{t\geq 1}$ be a decreasing sequence of positive *gain* parameters satisfying $\sum_t \gamma_t = +\infty$ and $\sum_{t\geq 1}\gamma 2_t < +\infty$ and let $\alpha(0)\in(\mathbb{R}^d)^n$ denote a starting $n$-quantizer. Then, at time $t\in\mathbb{N}$, one updates the running $n$-quantizer $\alpha(t-1) := (\alpha_1(t-1),\ldots,\alpha_n(t-1))$ as follows

$$\begin{array}{llcl}
\text{Competitive phase: select} & i(t) & \in & \operatorname{argmin}\{i:\ \|\alpha_i(t-1)-\zeta(t)\| = \min_j \|\alpha_j(t-1)-\zeta(t)\|\}\\
\text{Learning phase:} & \alpha_{i(t)}(t-1) & = & (1-\gamma_t)\alpha_{i(t)}(t-1)+\gamma_t\,\zeta(t)\\
 & \alpha(t)_j & = & \alpha_{j-1}(t-1), \qquad j\neq i(t).
\end{array}$$

Some further details concerning the numerical implementation of this procedure can be found in [15], especially some heuristics concerning the initialization and the specification of the gain parameter sequence usually choosen of the form $\gamma_t = \frac{A}{B+t}$. It converges toward some local minima of the quantization error at a $\sqrt{\gamma_t}$-rate. Some $d$-dimensional grids ($d=2$ up to 10) can be downloaded at the above URL for many values of $n$ in the range 2 up to 2 000. These quantizations were carried out to solve numerically multi-dimensional stopping time problems (pricing of American options on baskets, see [16] and the references therein).

The 1-dimensional optimal quantization of the $N(0,1)$-distribution has already been used to produce some optimal *scalar* product functional quantization - *i.e.* based on blocks of fixed length 1- in [17] with some promising applications to the pricing of path-dependent European options in stochastic volatility models (this work is also based on results about diffusion processes from [13]).

To be competitive with other methods (Monte Carlo, pde's) one needs to have good performances for not too large values of $n$. Within this range of values, it is more efficient to perform directly a numerical optimisation of (3.3) (or (3.6)) with $l = 1$ rather than using the theoretical asymptotically optimal parameters (3.7) and (3.8).

As far as numerical implementation of functional quantization with $n$-varying block length is concerned, some first numerical experiments carried out by Benedikt Wilbertz [19] for Brownian motion suggest that it slightly improves the scalar approach for high values of $n$, say $n \leq 106$, simply using up to 3-dimensional $n_j$-quantizers with some $n_j$ not greater than 100. A similar improvement can be obtained for lower values of $n$ (say $n \geq 20\,000$) by using product quantizers made of blocks with mixed lengths (1, 2 or 3).

EXAMPLES: The basic example (among Riemann-Liouville processes) is $X^{1/2} = W$ and $H = L2([0,1], dt)$, where

$$\lambda_j = (\pi(j-1/2))^{-2},\ u_j(t) = \sqrt{2}\sin\left(t/\sqrt{\lambda_j}\right),\ j \geq 1. \tag{3.16}$$

Since for $\delta, \rho \in (0, \infty)$,

$$X^{\delta+\rho} = \frac{\Gamma(\delta+\rho+\frac{1}{2})}{\Gamma(\rho+\frac{1}{2})} R_\delta(X^\rho),$$

one gets expansions of $X^{\delta+\rho}$ from Karhunen-Loève expansions of $X^\rho$. In particular,

$$X^{\delta+\frac{1}{2}} = \Gamma(\delta+1)\sum_{j=1}^{\infty}\sqrt{\lambda_j} Z_j R_\delta(u_j).$$

However, the functions $R_\delta(u_j), j \geq 1$, are not orthogonal in $H$ so that the nonzero correlation between the components of $(Z^{(j)} - \widehat{Z^{(j)}})$ prevents the previous estimates for $\mathbb{E}\|X - \widehat{X}^n\|^2$ given in Lemma 1 from working in this setting in the general case.

However, when $l = 1$ (scalar product quantizers made up with blocks of fixed length $l = 1$), one checks that these estimates still stand as equalities since orthogonality can now be substituted by the independence of $Z_j - \hat{Z}_j$ and stationarity property (3.1) of the quantizations $\hat{Z}_j, j \geq 1$. It is often good enough for applications to use scalar product quantizers (see [10], [17]). If, for instance $\delta = 1$, then

$$X := X^{3/2} = \sum_{j=1}^{\infty}\sqrt{\lambda_j} Z_j R_1(u_j),$$

where

$$R_1(u_j)(t) = \sqrt{2\lambda_j}(1 - \cos(t/\sqrt{\lambda_j})).$$

Note that $\|R_1(u_j)\|^2 = \lambda_j(3 - 4(-1)^{j-1}\sqrt{\lambda_j}),\ j \geq 1$. Set

$$\hat{X}^n = \sum_{j=1}^{m}\sqrt{\lambda_j}\hat{Z}_j R_1(u_j).$$

The quantization $\widehat{X}^n$ is non Voronoi (it is related to the Voronoi tessellation of $W$) and satisfies

$$\mathbb{E}\|X - \widehat{X}^n\|^2 = \sum_{j=1}^{m} \lambda 2_j(3 - 4(-1)^{j-1}\sqrt{\lambda_j}) e_{n_j}(N(0,1))2 + \sum_{j \geq m+1} \lambda 2_j(3 - 4(-1)^{j-1}\sqrt{\lambda_j}). \tag{3.17}$$

It is possible to optimize the (scalar product) quantization error using this expression instead of (3.6). As concerns asymptotics, if the parameters are tuned following (3.7)-(3.9) with $l = 1$ and $\lambda_j$ replaced by

$$\nu_j := \lambda 2_j(3 + 4\sqrt{\lambda_j}) \sim 3\pi^{-4} j^{-4} \quad \text{as} \quad n \to \infty,$$

and using Theorem 3 gives

$$(\mathbb{E}\, \|X - \hat{X}^n\|^2)^{1/2} \lesssim \left(\frac{3(12C(1) + 1)}{4}\right)^{1/2} e_n(X) \text{ as } n \to \infty. \tag{3.18}$$

Numerical experiments seem to confirm that $C(1) = Q(1)$. Since $Q(1) = \pi\sqrt{3}/2$ (see [5], p. 124), the above upper bound is then

$$\left(\frac{3(6\pi\sqrt{3} + 1)}{4}\right)^{1/2} = 5.02357\ldots$$

## 4 Dimension

Let $X$ be a $H$-valued random vector satisfying (1.2). For $n \in \mathbb{N}$, let $\mathcal{C}_n(X)$ be the (nonempty) set of all $L2$-optimal $n$-quantizers. Introduce the integral number

$$d_n(X) := \min \{\dim \text{span } (\alpha)) : \alpha \in \mathcal{C}_n(X)\}. \tag{4.1}$$

It represents the dimension at level $n$ of the functional quantization problem for $X$. Here span$(\alpha)$ denotes the linear subspace of $H$ spanned by $\alpha$. In view of Section 3, a reasonable conjecture for Gaussian random vectors is $d_n(X) \sim 2 \log n/b$ in regular cases, where $-b$ is the regularity index. We have at least the following lower estimate in the Gaussian case.

**Proposition 2** *Assume the situation of Theorem 1. Then*

$$d_n(X) \gtrsim \frac{1}{b^{1/(b-1)}} \frac{2 \log n}{b} \quad \text{as } n \to \infty.$$

**Proof.** For every $n \in \mathbb{N}$, we have

$$d_n(X) = \min \left\{ k \geq 0 : e_n(\bigotimes_{j=1}^{k} N(0, \lambda_j))2 + \sum_{j \geq k+1} \lambda_j \leq e_n(X)2 \right\} \tag{4.2}$$

(see [10]). Define

$$c_n := \min \left\{ k \geq 0 : \sum_{j \geq k+1} \lambda_j \leq e_n(X)2 \right\}.$$

Clearly, $c_n$ increases to infinity as $n \to \infty$ and by (4.2), $c_n \leq d_n(X)$ for every $n \in \mathbb{N}$. Using Theorem 1 and the fact that $\psi$ is regularly varying at infinity with index $b - 1$, we obtain

$$((b-1)\psi(c_n))^{-1} \sim \sum_{j \geq c_n+1} \lambda_j \sim e2_n(X) \sim \left(\frac{2}{b}\right)^{1-b} \frac{b}{b-1} \psi(\log n)^{-1}$$

and thus

$$\psi(c_n) \sim \left(\frac{2}{b}\right)^{1-b} \frac{1}{b} \psi(\log n) \sim \psi\left(\frac{1}{b^{1/(b-1)}} \frac{2 \log n}{b}\right) \quad \text{as } n \to \infty.$$

Consequently,

$$c_n \sim \frac{1}{b^{1/(b-1)}} \frac{2 \log n}{b} \quad \text{as} \quad n \to \infty.$$

This yields the assertion. □

For Riemann-Liouville processes one concludes

$$d_n(X^\rho) \gtrsim (2\rho+1)^{-1/2\rho} \frac{2 \log n}{2\rho + 1}$$

(see (2.3)).

For the metric entropy problem one may introduce the numbers $d_n(U_X)$ analogously. Then, in the situation of Theorem 1 it is known that $d_n(U_X) \gtrsim 2 \log n/b$ (see [12]). It remains an open question whether $d_n(X) \sim d_n(U_X) \sim 2 \log n/b$.

## References


[1] Bouton C., Pagès G., About the multi-dimensional Competitive Learning Vector Quantization Algorithm with constant gain, *Ann. Applied Probab.*, **7** (3) (1997), 670-710.

[2] Dereich, S., *High resolution coding of stochastic processes and small ball probabilities*, PhD thesis, TU Berlin, 2003.

[3] Donoho, D.L., *Counting bits with Kolmogorov and Shannon*, Technical Report 38, Stanford University, 2000.

[4] Gersho, A., Gray, R.M., *Vector Quantization and Signal Compression*. Kluwer, Boston, 1992.

[5] Graf, S., Luschgy, H., *Foundations of Quantization for Probability Distributions*, Lecture Notes in Mathematics, Vol. 1730, Springer, Berlin, 2000.

[6] Graf, S., Luschgy, H., The point density measure in the quantization of self-similar probabilities, *Math. Proc. Cambridge Phil. Soc.*, (to appear).

[7] Graf, S., Luschgy, H., Sharp asymptotics of the metric entropy for ellipsoids, *J. Complexity*, **20** (2004), 876-882.

[8] Gray, R.M., Neuhoff, D.L., Quantization, *IEEE Trans. Inform.*, **44**, (1998) 2325-2383.

[9] Kolmogorov, A.N., Tikhomirov, V.M., $\varepsilon$-entropy and $\varepsilon$-capacity of sets in functional spaces, *Amer. Math. Soc. Transl. Ser.*, **2**(17) (1961), 272-364.

[10] Luschgy, H., Pagès, G., Functional quantization of stochastic processes, *J. Funct. Anal.*, **196** (2002), 486-531.

[11] Luschgy, H., Pagès, G., Sharp asymptotics of the functional quantization problem for Gaussian processes, *Ann. Probab.*, **32** (2004), 1574-1599.

[12] Luschgy, H., Pagès, G., Sharp asymptotics of the Kolmorogov entropy for Gaussian measures, *J. Funct. Anal.*, **212** (2004), 89-120.

[13] Luschgy, H., Pagès, G., Functional quantization of 1-dimensional Brownian diffusions, Preprint. LPMA Université Paris 6, 853, 2003.

[14] Pagès, G., A space vector quantization method for numerical integration, *J. Computational and Applied Math.*, **89** (1997), 1-38.

[15] Pagès, G., Printems, J., Optimal quadratic quantization for numerics: the Gaussian case, *Monte Carlo Methods and Applications*, **9** (2003), 135-166.

[16] Pagès, G., Pham, H., Printems, J., Optimal quantization methods and applications to numerical methods and applications in finance, *Handbook of Computational and Numerical Methods in Finance* (ed. S. Rachev), 253-298, Birkhäuser, 2004.

[17] Pagès, G., Printems, J., Functional quadratic quantization for pricing derivatives, Preprint LMPA 930, Université Paris 6, 2004.

[18] Vu Kim Tuan, Gorenflo, R., Asymptotics of singular values of Volterra integral operators, *Numer. Funct. Anal. and Optimiz.*, **17** (1996), 453-461.

[19] Wilbertz, B., Work in progress, 2004.