



**HAL**  
open science

# Constructive quadratic functional quantization and critical dimension

Harald Luschgy, Gilles Pagès

► **To cite this version:**

Harald Luschgy, Gilles Pagès. Constructive quadratic functional quantization and critical dimension. Electronic Journal of Probability, 2014, 19, pp.50. 10.1214/EJP.v19-3010 . hal-01316652

**HAL Id: hal-01316652**

**<https://hal.sorbonne-universite.fr/hal-01316652>**

Submitted on 17 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

## Constructive quadratic functional quantization and critical dimension

Harald Luschgy\*      Gilles Pagès†

### Abstract

We propose a constructive proof for the sharp rate of optimal quadratic functional quantization and we tackle the asymptotics of the critical dimension for quadratic functional quantization of Gaussian stochastic processes as the quantization level goes to infinity, *i.e.* the smallest dimensional truncation of an optimal quantization of the process which is “fully” quantized. We first establish a lower bound for this critical dimension based on the regular variation index of the eigenvalues of the Karhunen-Loève expansion of the process. This lower bound is consistent with the commonly shared sharp rate conjecture (and supported by extensive numerical experiments). Moreover, we show that, conversely, optimized quadratic functional quantizations based on this critical dimension rate are always asymptotically optimal (strong admissibility result).

**Keywords:** quadratic functional quantization ; Karhunen-Loève expansion ; Gaussian process ; optimal quantizer ; asymptotically optimal quantizer ; Shannon entropy.

**AMS MSC 2010:** 60G15; 60G99; 94A29.

Submitted to EJP on September 16, 2013, final version accepted on May 12, 2014.

## 1 Introduction

The aim of this paper is two-fold: first we aim at providing a *constructive proof* of the sharp rate of optimal functional quantization in the quadratic case for (a wide class of) Gaussian processes or more generally of Gaussian random vectors  $X$  taking values in a separable Hilbert space  $(H, (\cdot|\cdot)_H)$ . Secondly, we provide several results about the *critical quantization dimension* in this framework, especially a “sharp” asymptotic lower bound for the *genuine critical dimension* (sharp with respect to a conjecture supported by extensive numerical experiments carried out in [14] with the Brownian motion and the Brownian bridge) and the sharp asymptotics of the *asymptotic critical dimension*.

Before defining precisely what we mean by “constructive” and what the above two notions of critical dimension represent in an optimal functional quantization problem, let us briefly recall few basic facts on this theory. First, the  $L^r$ -mean quantization

---

\*Universität Trier, Germany. E-mail: luschgy@uni-trier.de

†UPMC-Paris 06, France. E-mail: gilles.pages@upmc.fr

error of a random variable  $X$  defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  having a finite  $r^{\text{th}}$  moment and taking values in a separable Hilbert space  $H$  is defined by

$$e_{n,r}(X) = \inf \left\{ \left\| \min_{a \in \alpha} |X - a|_H \right\|_r, \alpha \subset H, |\alpha| \leq n \right\} \tag{1.1}$$

where  $|\alpha|$  denotes the cardinality of the set  $\alpha$ ,  $r \in (0, \infty)$  and  $\|\cdot\|_r$  denotes the  $L^r(\mathbb{P})$ - (quasi-)norm. It can also be characterized as

$$e_{n,r}(X) = \min \{ \| |X - Y|_H \|_r, Y : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow H, |Y(\Omega)| \leq n \}$$

where  $|\cdot|_H$  denotes the norm on the Hilbert space  $H$ . Any random vector which achieves the above minimum (there is always at least one) is called an  $(L^r)$ -optimal  $n$ -quantization. One can show that such an optimal quantization is always of the form  $Y^* = \pi_{\alpha^*}(X)$  where  $\pi$  is a Borel projection on  $\alpha^{*,n} = Y^*(\Omega)$  following the nearest neighbour rule. The subset  $\alpha^{*,n}$  is called an *optimal Voronoi  $n$ -quantizer* (or *optimal Voronoi quantizer* at level  $n$ ). By extension any random vector of the form  $\pi_\alpha(X)$  is called a *Voronoi quantization* whereas  $\alpha$  is often called an  $n$ -quantizer (if  $|\alpha| = n$ ). The term Voronoi refers to the nearest neighbour projection. A sequence  $(\alpha_n)_{n \geq 1}$  of  $n$ -quantizers is called *asymptotically optimal* if

$$\lim_n \frac{\| \min_{a \in \alpha_n} |X - a|_H \|_r}{e_{n,r}(X)} = 1.$$

In the quadratic framework ( $r = 2$ ), we will drop the subscript  $r$  for simplicity.

When  $H$  is an infinite dimensional separable Hilbert space (typically  $H$  is function space like  $L^2([0, T], dt)$ ), one often speaks of *functional quantization*.

The first problem of interest (beyond the existence of optimal  $n$ -quantizers for every  $n \in \mathbb{N}$ ) is the rate of decay of the mean (quadratic) quantization error. This sharp rate problem in a Hilbert setting has been solved for a wide class of  $H$ -valued Gaussian random vectors/processes  $X$  in [13] (see Theorem 2.2 below). To be precise, when the nonzero eigenvalues  $(\lambda_n^X)_{n \geq 1}$  of the Karhunen-Loève ( $K$ - $L$ ) eigensystem  $(\lambda_n^X, e_n^X)_{n \geq 1}$  of an  $H$ -valued random vector  $X$  ordered in a non-increasing order (each written as many times as is its multiplicity) read  $\lambda_n^X = \varphi(n)$  where  $\varphi$  is a non-increasing regularly varying function with index  $-b$ ,  $b \geq 1$ . Thus, we know that if  $X$  is a standard Brownian motion  $W$  on the interval  $[0, T]$ , then  $\varphi(n) = \frac{T^2}{\pi^2(n-\frac{1}{2})^2}$  so that  $b = 2$  which implies

$$\lim_n \sqrt{\log n} e_n(W) = \frac{\sqrt{\pi}}{2} T.$$

Note that the same sharp rate holds true for any Gaussian process whose parameter  $b$  is equal to 2, like the Ornstein-Uhlenbeck process. For a more general statement depending on the value of  $b \geq -1$ , we refer to [13] where  $b$  is made explicit on many examples (Theorem 2.2 in the next section yields the general asymptotic result with an explicit limit, depending on  $b \geq -1$  and  $\varphi$ ). This also applies to the fractional Brownian motion  $W^H$  with Hurst constant  $H \in (0, 1)$  for which  $b = 2H + 1$  since  $\lambda_n \sim \kappa_H n^{-(1+2H)}$  where  $\kappa_H \in (0, +\infty)$  is explicit, as established *e.g.* in [13]. The sharp rate for quadratic functional quantization can be solved by this approach for other processes like the fractional Ornstein-Uhlenbeck process, Gaussian sheets, etc.

Rather unexpectedly, the proof of this general theorem which was first stated in [13] is not constructive in the sense that it never needs nor exhibits any sequence of (asymptotically) optimal grids.

The notion of critical dimension at level  $n$ , denoted  $d_n^X$  and referred to as *genuine critical dimension* in what follows, can be introduced in an intuitive way as follows

$$d_n^X = \min \left\{ \dim(\text{span}(\alpha^{*,n})), \alpha^{*,n} \text{ optimal } n\text{-quantizer} \right\}. \tag{1.2}$$

As established in [12], the optimal quantizers at level  $n$  live in finite dimensional subspaces spanned by the (first components of the) Karhunen-Loève expansion of the Gaussian process of interest. So, in our Gaussian framework, the genuine critical dimension at level  $n$  is defined equivalently as the lowest dimension of such a vector subspace which contains an optimal  $n$ -quantizer, namely

$$d_n = d_n^X = \min \{d \in \mathbb{N}^* : \exists \alpha^{*,n} \text{ optimal } n\text{-quantizer s.t. } \alpha^{*,n} \subset \text{span}\{e_k^X, 1 \leq k \leq d\}\}. \quad (1.3)$$

The “asymptotic critical dimension” corresponds to asymptotically optimal  $n$ -quantizers and will be precisely defined later on in Section 2. Like the mean quantization rate, the asymptotics of these critical dimensions are ruled by the rate of decay of the  $K$ - $L$  eigenvalues. The genuine critical dimension can be considered, at level  $n$ , as the true dimension of the above infinite dimensional optimal quantization problem (1.1) and its specification as a kind of “dimension selection”. As for the standard Brownian motion, the conjecture on the genuine critical dimension  $d_n^W$  reads

$$\lim_n (\log n)^{-1} d_n^W = 1.$$

and, more generally,

$$\lim_n (\log n)^{-1} d_n^X = \frac{2}{b}$$

if the eigenvalues of the covariance operator of  $X$  involved in its  $K$ - $L$  regularly vary with exponent  $-b$ .

There is a connection between the (quadratic) mean quantization error  $e_n(X)$  and the critical dimension, still through the eigenvalues of the  $K$ - $L$  expansion of  $X$ . Let  $(\lambda_n^X, e_n^X)_{n \geq 1}$  be the  $K$ - $L$  eigensystem of  $X$ , where the eigenvalues  $\lambda_n$  are ordered in a non-increasing order. Then  $d_n^X$  is the *lowest dimension* satisfying

$$\begin{aligned} e_n^2(X) &= e_n^2 \left( \sum_{k=1}^{d_n^X} \lambda_k e_k^X \right) + \sum_{k \geq d_n^X+1} \lambda_k \\ &= e_n^2 \left( \bigotimes_{k=1}^{d_n^X} \mathcal{N}(0, \lambda_k^X) \right) + \sum_{k \geq d_n^X+1} \lambda_k. \end{aligned}$$

This connection, originally established in [12], is briefly recalled in Section 2.3 as a starting point of our approach. Rather unexpectedly, the proof of the sharp rate for  $e_n(X)$  in [13] does not provide asymptotic information on the critical dimension as  $n$  grows, at least in a straightforward way. In this paper, we fill these two gaps for this class of Gaussian processes: first, we propose a constructive proof of the sharp rate theorem for quantization error exhibiting asymptotically optimal grids and, secondly, we provide the first rigorous, though partial, results on the asymptotics of the critical dimension to our knowledge.

In particular we provide a first theoretical justification of the use of optimal quantizers at level  $n$  of the distribution  $\bigotimes_{k=1}^{\lfloor \frac{2}{b} \log n \rfloor} \mathcal{N}(0, \lambda_k^X)$  to functionally quantize a Gaussian process  $X$  (provided its  $K$ - $L$  system is explicit, in particular the eigenvectors). Such results are used to produce quadrature formulas to compute expectations of Lipschitz continuous or twice  $|\cdot|_{L_T^2}$ -differentiable functionals of Gaussian processes like the standard or the fractional Brownian motion, the Ornstein-Uhlenbeck process, etc, see *e.g.* [15]. It also has been used when  $X$  is diffusion solution to *SDEs* in the Stratonovich sense, see [15, 16, 4]:  $X$  is quantized by solutions of *ODEs* mimicking the *SDE* in which

the Brownian  $W$  motion is replaced by its functional quantization. Several applications to the pricing of exotic path-dependent options based on this functional discretization method have been implemented. Functional quantization also has been used to as a stratification procedure for Monte Carlo simulation (see [11, 3]).

The paper is organized as follows: in Section 2 we provide some more rigorous background on  $K$ - $L$  expansions of Gaussian random vectors and functional quantization. Then, we state our main results by exhibiting a sequence of asymptotically optimal quantization grids and provide an asymptotic lower bound for the genuine critical dimension. In Section 3 and 4 we establish some upper and lower bounds respectively for the mean quantization error. Section 5 is devoted to the proofs and constructive aspects. We conclude by few numerical illustrations which support the conjecture.

Our main tools, beyond discrete optimization techniques used for the upper bounds, are Shannon's source coding Theorem and the connection between mean quantization error and Shannon  $\varepsilon$ -entropy (or rate-distortion function, see [5]).

Notations. •  $|\cdot|$  denotes the canonical Euclidean norm on  $\mathbb{R}^d$ .

•  $\mathbb{N}^* = \{1, 2, 3, \dots\}$  denotes the set of positive integers.

•  $L_T^2 = L^2([0, T], dt)$  is equipped with its Hilbert norm  $\|f\|_{L_T^2} = \left(\int_0^T f^2(t)dt\right)^{\frac{1}{2}}$ .

• Let  $(a_n)$  and  $(b_n)$  be two sequences of real numbers.  $a_n \sim b_n$  if there exists a sequence  $(u_n)$  such that  $a_n = u_n b_n$  and  $\lim_n u_n = 1$ .

•  $o(1)$  denotes a sequence indexed by  $n \in \mathbb{N}^*$  going to 0 as  $n \rightarrow +\infty$  (which may vary from line to line).

## 2 Background on optimal functional quantization and main result

### 2.1 Karhunen-Loève expansion and main running assumption

Let  $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow H$  be a centered Gaussian random vector taking values in a separable Hilbert space  $(H, |\cdot|_H)$  satisfying

$$\dim K_X = +\infty \text{ where } K_X \text{ is the reproducing kernel Hilbert space of } X.$$

A typical example is the case of a Gaussian stochastic process  $X = (X_t)_{t \in [0, T]}$  with continuous paths. Clearly, for such a process *a.s.*  $t \mapsto X_t(\omega)$  lies in  $L_T^2$  so that  $X$  can be seen as a random vector taking values in  $(L_T^2, |\cdot|_{L_T^2})$ .

Let  $(\lambda_k^X, e_k^X)_{k \geq 1}$  be the orthonormal eigensystem of the (positive trace) covariance operator of  $X$ , also known as the Karhunen-Loève ( $K$ - $L$ ) orthonormal system of  $X$ . Since the sequence  $(\lambda_n^X)_{n \geq 1}$  has only one limiting value, 0, one may assume without loss of generality that the  $K$ - $L$  eigensystem is indexed so that the sequence of (nonzero) eigenvalues is non-increasing. To alleviate notations we will drop the dependency of the eigenvalues in  $X$  by simply noting  $\lambda_n$  instead of  $\lambda_n^X$ .

Throughout the paper, the main results are obtained under the following assumption about the  $K$ - $L$  eigenvalues:

$$(R) \equiv \begin{cases} \text{There exists } b \in [1, +\infty) \text{ and a non-increasing function } \varphi : (0, +\infty) \rightarrow (0, +\infty) \\ \text{with regular variations at infinity of index } -b \text{ (hence going to 0 at infinity)} \\ \text{such that } \lambda_k = \varphi(k), k \geq 1. \end{cases}$$

Then, the Karhunen-Loève decomposition of  $X$  reads

$$X = \sum_{k \geq 1} \sqrt{\lambda_k} \xi_k e_k^X$$

where  $\xi_k = \frac{(X, e_k^X)_H}{\sqrt{\lambda_k}}$ ,  $k \geq 1$ , defines an i.i.d. sequence of  $\mathcal{N}(0; 1)$ -distributed random variables defined on  $(\Omega, \mathcal{A}, \mathbb{P})$ .  $(\cdot, \cdot)_K$  denotes the inner product in  $H$ .) The convergence holds *a.s.* in  $H$ .

Moreover, note that the function  $\varphi$  satisfies  $\int_A^{+\infty} \varphi(y)dy < +\infty$  for large enough  $A > 0$  since  $\sum_k \lambda_k < +\infty$ .

**2.2 Optimal quadratic functional quantization**

Let  $L_H^2(\mathbb{P}) = \{Y : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow H, \text{ Borel measurable, } \mathbb{E}|Y|_H^2 < +\infty\}$ . The optimal quantization problem for  $X \in L_H^2(\mathbb{P})$  reads

$$e_n(X) = \inf \left\{ \left\| \min_{\alpha \in \mathcal{C}} |X - \alpha|_H \right\|_2, \alpha \subset H, |\alpha| \leq n \right\}. \tag{2.1}$$

Let us introduce few additional notations. For every integer  $d \geq 1$ , we denote by  $X^{(d)}$  the  $H$ -orthogonal projection of  $X$  on the vector space  $\text{span}\{e_1^X, \dots, e_d^X\}$ , namely

$$X^{(d)} = \sum_{k=1}^d \sqrt{\lambda_k} \xi_k e_k^X$$

and

$$e_n(X^{(d)}) = \inf \left\{ \left\| \min_{\alpha \in \mathcal{C}} |X^{(d)} - \alpha|_H \right\|_2, \alpha \subset \bigoplus_{1 \leq k \leq d} \mathbb{R}e_k^X, |\alpha| \leq n \right\}. \tag{2.2}$$

Note that one also has

$$e_n(X^{(d)}) = \inf \left\{ \left\| \min_{\alpha \in \mathcal{C}} |(\sqrt{\lambda_k} Z_k)_{1 \leq k \leq d} - \alpha| \right\|_2, \alpha \subset \mathbb{R}^d, |\alpha| \leq n \right\}$$

where  $Z = (Z_1, \dots, Z_d) \stackrel{d}{=} \mathcal{N}(0; I_d)$ .

Finally we set,

$$e_n^2(X, d) = e_n^2(X^{(d)}) + \sum_{k \geq d+1} \lambda_k \tag{2.3}$$

and

$$C(d) = \sup_{k \geq 1} k^2 e_k^2(\mathcal{N}(0; I_d)).$$

We know from [12] (see Proposition 2.1) that, for every  $n \in \mathbb{N}^*$ , the infimum in (2.1) holds as a minimum: there exists at least one optimal quantizer  $\alpha^{*,n}$  which turns out to have full size  $n$ . Furthermore  $\alpha^{*,n}$  lies in a finite dimensional space spanned by finitely many elements of the  $K$ - $L$  basis.

Now we are in position to come back to the *genuine critical dimension*  $d_n = d_n^X$  defined by (1.2) and characterized in (1.3) as the smallest dimension of a vector subspace of  $\text{span}\{e_n^X, n \geq 1\}$  in which an optimal  $n$ -quantizer lies. The sequence  $(d_n)_{n \geq 1}$  makes up a sequence satisfying

$$e_n^2(X) = e_n^2(X, d_n).$$

It is clear that  $d_n$  goes to infinity, otherwise one could extract a subsequence  $d_{n'}$  such that  $d_{n'} \leq \tilde{d} < +\infty$ . If so, we would have

$$e_n^2(X) \geq \sum_{k \geq \tilde{d}+1} \lambda_k$$

which contradicts the obvious fact that  $e_n^2(X)$  goes to zero as  $n$  goes to infinity (see e.g. [12]). This last claim is a consequence of the fact that, if  $(z_n)_{n \geq 1}$  is everywhere dense in  $H$ , then

$$e_n^2(X) \leq \mathbb{E} \left( \min_{1 \leq i \leq n} |X - z_i|_H^2 \right) \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

Otherwise very little is known on the sequence  $(d_n)_{n \geq 1}$ , in particular we do not know whether this sequence is monotone.

We will use the following easy lemma

**Lemma 2.1.** *Let  $n \geq 1$  be an integer. The sequence  $d \mapsto e_n^2(X, d)$  is non-increasing (and hence is constant for  $d \geq d_n$ ).*

*Proof.* Let  $d \leq d'$ . Let  $\alpha^{*,n}(d)$  be an optimal quadratic quantizer for  $X^{(d)}$  of size (at most)  $n$ . It is clear that for every  $a \in \alpha^{*,n}(d)$ ,

$$|X^{(d')} - a|_H^2 = |X^{(d)} - a|_H^2 + \sum_{k=d+1}^{d'} \lambda_k$$

since  $\alpha^{*,n}(d) \subset \text{span}\{e_1^X, \dots, e_d^X\}$ . As a consequence  $e_n^2(X^{(d')}) \leq e_n^2(X^{(d)}) + \sum_{k=d+1}^{d'} \lambda_k$  and one concludes by adding the tail  $\sum_{k \geq d'+1} \lambda_k$ .  $\square$

It still holds as a conjecture that, under Assumption (R),

$$\lim_n \frac{d_n}{\log n} = \frac{2}{b} \tag{2.4}$$

whereas the sharp rate of optimal quadratic quantization has been elucidated for long in [13] (with several extension to more general Banach settings obtained ever since *e.g.* for  $L^p([0, T], dt)$ -norms,  $1 \leq p \leq +\infty$  and in an  $L^r(\mathbb{P})$ -sense, see [6], etc).

Extensive computations carried out in [14] provide strong evidence that in fact, as concerns the standard Brownian motion and the Brownian bridge (which corresponds to  $b = 2$ ), we even have that

$$d_n \in \{ \lfloor \log n \rfloor, \lceil \log n \rceil \}.$$

These conjectures are also supported by results obtained for optimal "block quantization" with blocks of either constant or varying dimensions (see [13, 14]).

The aims of this paper can now be summed up as follows : firstly to provide a *constructive* proof of the sharp rate theorem 2.2 recalled below (with, as a result, the exhibition of natural sequences of asymptotic quantizers), secondly to provide a partial answer to the above conjecture and finally to provide a complete answer to the "asymptotic" dimension problem.

For the self-completeness of the paper we briefly recall the definitions of *regularly* and *slowly varying* functions at  $+\infty$  (see [2]). A Borel function  $f : [A, +\infty) \rightarrow \mathbb{R}$  is regularly varying with index  $a \in \mathbb{R} \setminus \{0\}$  if

$$\forall t > 0, \quad \lim_{x \rightarrow +\infty} \frac{f(tx)}{f(x)} = t^a$$

and is slowly varying if

$$\forall t > 0, \quad \lim_{x \rightarrow +\infty} \frac{f(tx)}{f(x)} = 1.$$

**Theorem 2.2.** (see [12]) *Assume (R). Let  $\psi : (0, +\infty) \rightarrow (0, +\infty)$  be defined by*

$$\psi(x) = \begin{cases} \frac{1}{x\varphi(x)} & \text{if } b > 1 \text{ so that } \psi \text{ is a regularly varying function with index } b - 1, \\ \frac{1}{\int_x^\infty \varphi(y)dy} & \text{if } b = 1 \text{ so that } \psi \text{ is a slowly varying function.} \end{cases} \tag{2.5}$$

*Then*  $\lim_n \psi(\log n) e_n^2(X) = \begin{cases} \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1} & \text{if } b > 1, \\ 1 & \text{if } b = 1. \end{cases}$

**2.3 Main result: constructive rates and critical dimension**

Now we state our two main results on the “constructive rates” and the critical dimension. While the genuine critical dimension  $d_n$  is mostly of theoretical interest, for numerical purpose the “asymptotic critical dimension, precisely defined below, is more interesting. It corresponds to the lowest sequence of dimensions  $(\delta_n)_{n \geq 1}$  which produce asymptotically optimal  $n$ -quantizers.

**Theorem 2.3** (Constructive proof of the sharp rate). *Assume (R). Let  $(\delta_n)_{n \geq 1}$  be a sequence of positive integers.*

(a) *If  $\liminf_n \frac{\delta_n}{\log n} \geq \frac{2}{b}$ , then*

$$\lim_n \psi(\log n) e_n^2(X) = \lim_n \psi(\log n) e_n^2(X, \delta_n) = \begin{cases} \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1} & \text{if } b > 1 \\ 1 & \text{if } b = 1. \end{cases} \quad (2.6)$$

*If  $b > 1$ , the converse is true and, furthermore,  $\liminf_n \psi(\delta_n) e_n^2(X^{(\delta_n)}) \geq 1$ .*

(b) *Assume  $b > 1$ . If the sequence of integers  $(\delta_n)_{n \geq 1}$  goes to infinity and produces asymptotically optimal quantizers i.e.  $\lim_n \frac{e_n^2(X, \delta_n)}{e_n^2(X)} = 1$ , then*

$$\lim_n \frac{\delta_n}{\log n} = \frac{2}{b} \iff \lim_n \psi(\log n) e_n^2(X^{(\delta_n)}) = \left(\frac{b}{2}\right)^{b-1} \iff \lim_n \psi(\delta_n) e_n^2(X^{(\delta_n)}) = 1.$$

In fact this theorem can be reformulated equivalently in terms of critical dimension, at least when  $b > 1$ . To this end we introduce the following definitions.

**Definition 2.4.** *Assume (R). Let  $\psi$  be defined as in Theorem 2.2. Let  $(\delta_n)_{n \geq 1}$  be a sequence of positive integers going to infinity.*

(a) *A sequence  $(\delta_n)_{n \geq 1}$  is an admissible sequence of dimensions for  $X$  if*

$$\lim_n \frac{e_n^2(X, \delta_n)}{e_n^2(X)} = 1.$$

(b) *If  $b > 1$ , an admissible sequence  $(\delta_n)_{n \geq 1}$  is an asymptotic critical dimension sequence or strongly admissible for  $X$  if*

$$\lim_n \psi(\delta_n) e_n^2(X^{(\delta_n)}) = 1.$$

*Admissibility* simply means that such a sequence produces asymptotic optimal quantizers in practice. For computational purpose, it is clear that the “lowest” choice  $\delta_n = \lfloor \frac{2}{b} \log n \rfloor$  seems the more appropriate. This is made more precise by the second definition in (b), even if it looks less intuitive.

As concerns *asymptotic critical dimension (strong admissibility)*, we know from (2.3) that for every fixed dimension  $d \in \mathbb{N}^*$ , there is a balance in  $e_n^2(X, d)$  between the *a priori* quantized part  $e_n^2(X^{(d)})$  and the tail of the series  $\sum_{k \geq d+1} \lambda_k$  which represents the variance of the non-quantized part (or, equivalently, trivially quantized by 0). But for a fixed  $d$  this does not prejudice of what really occurs. However, an admissible sequence  $(\delta_n)_{n \geq 1}$  being given, the smaller  $e_n^2(X^{(\delta_n)})$  is, the more quantized  $X^{(\delta_n)}$  is in practice. As illustrated by Theorem 2.3, the definition of asymptotic critical dimension (strong admissibility) suggests that in that case  $X^{(\delta_n)}$  is “fully” quantized, at least asymptotically. This is confirmed by numerical computations (see [14] and [15] for more insight on these numerical aspects).



The theorem below and the corollary that follows reformulate the above results in term of critical dimensions problem, at least for  $b > 1$ .

**Theorem 2.5.** Assume (R). Let  $\psi$  be defined as in Theorem 2.2.

(a) If  $b > 1$ ,

$$(\delta_n)_{n \geq 1} \text{ is admissible} \iff \liminf_n \frac{\delta_n}{\log n} \geq \frac{2}{b}$$

and

$$(\delta_n)_{n \geq 1} \text{ is strongly admissible} \iff \lim_n \frac{\delta_n}{\log n} = \frac{2}{b}.$$

(b) If  $b = 1$ ,

$$(\delta_n)_{n \geq 1} \text{ is admissible} \implies \liminf_n \frac{\psi(\delta_n)}{\psi(\log n)} \geq 1$$

and

$$\liminf_n \frac{\delta_n}{\log n} > 0 \implies (\delta_n)_{n \geq 1} \text{ is admissible.}$$

As for the genuine critical dimension, we thus obtain the following lower bounds.

**Corollary 2.6.** Assume (R). (a) If  $b > 1$ ,

$$\liminf_n \frac{d_n}{\log n} \geq \frac{2}{b}.$$

(b) If  $b = 1$ ,

$$\liminf_n \frac{\psi(d_n)}{\psi(\log n)} \geq 1.$$

Following the definition of asymptotic critical dimension, it seems intuitive to guess that the sequence  $(d_n)_{n \geq 1}$  of genuine critical dimension is an asymptotic critical dimension sequence. In fact such a claim is just a reformulation of the conjecture (2.4). In more mathematical terms, it means that the conjecture is true if and only if

$$\lim_{n \rightarrow +\infty} \psi(d_n) e_n^2(X^{(d_n)}) = 1.$$

### 3 Upper bound (proofs)

Since we are trying to provide fairly a new *constructive* proof of Theorem 2.2 *i.e.* the sharp rate for quadratic functional quantization, we emphasize that we will not use any of its claims. In particular, we are not in position at this stage to claim that  $\lim_n \psi(\log n) e_n^2(X)$  does exist. This is the reason why the claims in the proposition below involve  $\limsup_n \psi(\log n) e_n^2(X)$  which always exists (the same will be true with Proposition 4.3 in the next section).

**Proposition 3.1.** Assume (R). Let  $(\delta_n)_{n \geq 1}$  be a sequence of integers going to infinity.

(a) If  $b > 1$  and  $\liminf_n \frac{\delta_n}{\log n} \geq \frac{2}{b}$ , then

$$\limsup_n \psi(\log n) e_n^2(X) \leq \limsup_n \psi(\log n) e_n^2(X, \delta_n) \leq \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1}. \quad (3.1)$$

Furthermore, if  $\lim_n \frac{\delta_n}{\log n} = \frac{2}{b}$ , then

$$\limsup_n \psi(\log n) e_n^2(X^{(\delta_n)}) \leq \left(\frac{b}{2}\right)^{b-1}.$$

(b) If  $b = 1$  and  $\liminf_n \frac{\delta_n}{\log n} = \kappa \in (0, +\infty)$ , then

$$\limsup_n \psi(\log n) e_n^2(X) \leq \limsup_n \psi(\log n) e_n^2(X, \delta_n) \leq 1. \tag{3.2}$$

First we need two lemmas devoted two block quantization and their critical dimension which are the key of the proof. For every integer  $d \geq 1$ , we define set  $\lambda_k^{(d)} = \lambda_{(k-1)d+1}$ .

**Lemma 3.2** (Block quantization). *Let  $d, d_0 \in \mathbb{N}^*$ ,  $d > d_0$ . Then, for every  $k \in \mathbb{N}^*$ ,  $k \leq \frac{d}{d_0}$ , we have*

$$e_n^2(X^{(d)}) \leq C(d_0) \min \left\{ \sum_{\ell=1}^k \lambda_{\ell}^{(d_0)} n_{\ell}^{-\frac{2}{d_0}}, n_1, \dots, n_k \in \mathbb{N}^*, \prod_{\ell=1}^k n_{\ell} \leq n \right\} + \sum_{i=kd_0+1}^d \lambda_i$$

*Proof.* We introduce the (sub-optimal)  $d_0$ -block product quantizer defined as follows

$$\tilde{X}^{(d, d_0, k)} = \sum_{\ell=1}^k \sum_{i=1}^{d_0} \sqrt{\lambda_{(\ell-1)d_0+i}} \left( \text{Proj}_{\alpha^{(\ell)}} \left( (\xi_j)_{(\ell-1)d_0+1 \leq j \leq \ell d_0} \right) \right)_i e_{(\ell-1)d_0+i}^X$$

where  $\alpha^{(\ell)} \subset \mathbb{R}^{d_0}$  is an optimal quadratic quantizer of size (or at level)  $n_{\ell}$  of  $\mathcal{N}(0; I_{d_0})$  and  $\text{Proj}_{\alpha^{(\ell)}} : \mathbb{R}^{d_0} \rightarrow \alpha^{(\ell)}$  is a Borel nearest neighbour projection on  $\alpha^{(\ell)}$ .

Elementary computations based on the Pythagoras theorem (see Lemma 4.2 in [13]) show that

$$\begin{aligned} \left\| X - \tilde{X}^{(d, d_0, k)} \right\|_2^2 &= \sum_{\ell=1}^k \sum_{i=1}^{d_0} \lambda_{(\ell-1)d_0+i} \mathbb{E} \left| \left( \text{Proj}_{\alpha^{(\ell)}} \left( (\xi_j)_{(\ell-1)d_0+1 \leq j \leq \ell d_0} \right) \right)_i - \xi_{(\ell-1)d_0+i} \right|^2 \\ &\quad + \sum_{i=kd_0+1}^d \lambda_i, \\ &\leq \sum_{\ell=1}^k \lambda_{(\ell-1)d_0+1} \mathbb{E} \left| \left( \text{Proj}_{\alpha^{(\ell)}} \left( (\xi_j)_{(\ell-1)d_0+1 \leq j \leq \ell d_0} \right) \right)_i - \xi_{(\ell-1)d_0+i} \right|^2 \\ &\quad + \sum_{i=kd_0+1}^d \lambda_i, \\ &\leq \sum_{\ell=1}^k \lambda_{(\ell-1)d_0+1} e_{n_{\ell}}^2 \left( \mathcal{N}(0; I_{d_0}) \right) + \sum_{i=kd_0+1}^d \lambda_i. \end{aligned}$$

The definition of  $C(d_0)$  completes the proof. □

This optimal integer bit allocation has a formal almost optimal solution given by

$$n_{\ell} = \lfloor x_{\ell} \rfloor \quad \text{with} \quad x_{\ell} = \left( \lambda_{\ell}^{(d_0)} \right)^{\frac{d_0}{2}} \left( \prod_{j=1}^k \lambda_j^{(d_0)} \right)^{-\frac{d_0}{2k}} n^{\frac{1}{k}}, \quad \ell = 1, \dots, k,$$

as suggested by considering the problem on  $(\mathbb{R}_+)^k$  instead of  $(\mathbb{N}^*)^k$ . This solution is admissible as soon as all the  $n_{\ell}$ s are nonzero or equivalently since they are non-increasing in  $\ell$  as soon as  $n_k \geq 1$ .

**Lemma 3.3** (Critical dimension for block quantization). *Let  $d, d_0$  and  $k$  be like in Lemma 3.2.*

$$A(n, d_0) = \left\{ k \geq 1 : \left( \lambda_k^{(d_0)} \right)^{\frac{d_0}{2}} \left( \prod_{j=1}^k \lambda_j^{(d_0)} \right)^{-\frac{d_0}{2k}} n^{\frac{1}{k}} \geq 1 \right\}.$$

(a)  $A(n, d_0) = \{1, \dots, k_n(d_0)\}$ .

(b) Assume (R). Then  $k_n(d_0) \sim \frac{2}{bd_0} \log n$  as  $n \rightarrow +\infty$ .

(c) For every integer  $k \leq \min(k_n(d_0), \frac{d}{d_0})$ ,

$$e_n^2(X^{(d)}) \leq 4^{\frac{1}{d_0}} C(d_0) k \lambda_k^{(d_0)} + \sum_{i=kd_0+1}^d \lambda_i.$$

*Proof.* (a) This follows from the fact that the sequence

$$a_k = a_k^{(d_0)} := \frac{d_0}{2} \left( \sum_{\ell=1}^k \log \lambda_\ell^{(d_0)} - k \log \lambda_k^{(d_0)} \right) \tag{3.3}$$

is non-decreasing since  $A(n, d_0) = \{k : a_k \leq \log n\}$ .

(b) First note that  $\lambda_j^{(d_0)} = \varphi((j-1)d_0 + 1)$  and that  $\varphi((\cdot - 1)d_0 + 1)$  is regularly varying still with index  $b$  if  $\varphi$  is. As a consequence standard arguments on regularly varying functions show that

$$\frac{2a_k}{d_0} \sim bk \text{ or equivalently } a_k \sim \frac{bd_0 k}{2} \text{ as } k \rightarrow +\infty$$

which in turn implies that  $k_n(d_0) \sim \frac{2}{bd_0} \log n$  as  $n \rightarrow +\infty$ .

(c) It is straightforward that

$$\begin{aligned} \sum_{\ell=1}^k \lambda_\ell^{(d_0)} n_\ell^{-\frac{2}{d_0}} &\leq 2^{\frac{2}{d_0}} \sum_{\ell=1}^k \lambda_\ell^{(d_0)} (n_\ell + 1)^{-\frac{2}{d_0}} \\ &\leq 4^{\frac{1}{d_0}} \sum_{\ell=1}^k \lambda_\ell^{(d_0)} x_\ell^{-\frac{2}{d_0}} \\ &\leq 4^{\frac{1}{d_0}} k \lambda_k^{(d_0)} x_k^{-\frac{2}{d_0}} \\ &\leq 4^{\frac{1}{d_0}} k \lambda_k^{(d_0)} \end{aligned}$$

since  $\lambda_\ell^{(d_0)} x_\ell^{-\frac{2}{d_0}} = \left( \prod_{j=1}^k \lambda_j^{(d_0)} \right)^{-\frac{d_0}{2k}} n^{\frac{1}{k}}$  does not depend on  $\ell$  and  $x_k \geq 1$ . □

Here we come to the proof of Proposition 3.1.

*Proof.* (a) First assume that  $\lim_n \frac{\delta_n}{\log n} = \frac{2}{b}$ . Let  $d_0 \in \mathbb{N}^*$  be a (temporarily) fixed integer. Set  $k_n = k_n(d_0) \wedge \left\lfloor \frac{\delta_n}{d_0} \right\rfloor$  for  $n$  large enough to have  $\delta_n \geq d_0$ . It follows from Lemma 3.3(b)

that  $k_n \sim \frac{2 \log n}{b d_0}$  and  $k_n \leq k_n(d_0)$  and  $k_n \leq \frac{\delta_n}{d_0}$  so that by Lemma 3.3(c) we get as soon as  $n \geq n_{d_0}$ ,

$$\begin{aligned} e_n^2(X^{(\delta_n)}) &\leq 4^{\frac{1}{d_0}} C(d_0) k_n \lambda_{k_n}^{(d_0)} + \sum_{i=k_n d_0+1}^{\delta_n} \lambda_i \\ &\leq 4^{\frac{1}{d_0}} C(d_0) k_n \lambda_{k_n}^{(d_0)} + (\delta_n - d_0 k_n) \lambda_{k_n+1}^{(d_0)}. \end{aligned} \tag{3.4}$$

Now, mimicking arguments in [13] involving regularly varying functions, namely  $\varphi$ , we get

$$d_0 k_n \lambda_{k_n}^{(d_0)} = d_0 k_n \varphi(k_n d_0 + 1) \sim \frac{2}{b} \log n \left(\frac{2}{b}\right)^{-b} \varphi(\log n) = \left(\frac{2}{b}\right)^{1-b} \frac{1}{\psi(\log n)} \quad \text{as } n \rightarrow +\infty.$$

Moreover

$$(\delta_n - d_0 k_n) \lambda_{k_n+1}^{(d_0)} = \left(\frac{\delta_n}{k_n d_0} - 1\right) k_n d_0 \lambda_{k_n+1}^{(d_0)} = o\left(\frac{1}{\psi(\log n)}\right)$$

since  $\delta_n \sim k_n d_0 \sim \frac{2}{b} \log n$ .

Consequently, by letting  $d_0$  go to  $+\infty$ , we get

$$\limsup_n \psi(\log n) e_n^2(X^{(\delta_n)}) \leq \left(\frac{2}{b}\right)^{1-b} \limsup_{d_0} \frac{C(d_0)}{d_0}.$$

One concludes by using (see [13]) that, owing to the converse of Shannon’s source coding theorem,

$$\lim_{d \rightarrow +\infty} \frac{C(d)}{d} = 1.$$

On the other hand

$$\sum_{i \geq k_n d_0+1} \lambda_i^{(d_0)} \sim \frac{k_n d_0 \varphi(k_n d_0)}{b-1} \sim \frac{1}{(b-1) \psi\left(\frac{2}{b} \log n\right)} \sim \left(\frac{2}{b}\right)^{1-b} \frac{1}{(b-1) \psi(\log n)}$$

which yields the announced result by sub-additivity of  $\limsup_n$ .

If  $\liminf_n \frac{\delta_n}{\log n} \geq \frac{2}{b}$ , then set  $\delta'_n = \delta_n \wedge \left\lfloor \frac{2 \log n}{b} \right\rfloor$ ,  $n \geq 1$ . Then  $\lim_n \frac{\delta'_n}{\log n} = \frac{2}{b}$  whereas by Lemma 3.2  $e_n^2(X, \delta_n) \leq e_n^2(X, \delta'_n)$  which implies

$$\limsup_n \psi(\log n) e_n^2(X, \delta_n) \leq \limsup_n \psi(\log n) e_n^2(X, \delta'_n) \leq \left(\frac{2}{b}\right)^{1-b} \frac{b}{b-1}.$$

(b) Assume first that  $\lim_n \frac{\delta_n}{\log n} = \kappa \in (0, +\infty)$ . Owing to Lemma 3.2, we may assume as above that  $k_n$  defined like in (a) satisfies  $k_n \sim \kappa' \log n$  where  $\kappa' = \kappa \wedge \left(\frac{2}{d_0}\right)$ . As  $b = 1$ ,  $\psi(x) = \frac{1}{\int_x^{+\infty} \varphi(y) dy}$ . It follows from Proposition 1.5.9b in [2] (applied with  $\ell(y) = y\varphi(y)$ ) that  $\psi$  is a slowly varying function satisfying  $x\varphi(x) = o(1/\psi(x))$ . Hence, we derive that

$$\sum_{i \geq \delta_n+1} \lambda_i^{(d_0)} \sim \frac{1}{\psi(\delta_n)} \sim \frac{1}{\psi(\kappa \log n)} \sim \frac{1}{\psi(\log n)}.$$

On the other hand,

$$d_0 k_n \lambda_{k_n}^{(d_0)} = d_0 k_n \varphi(d_0 k_n + 1) = o\left(\frac{1}{\psi(d_0 k_n + 1)}\right) = o\left(\frac{1}{\psi(\log n)}\right)$$

since  $\psi$  is slowly varying, and

$$\sum_{i=d_0 k_n+1}^{\delta_n} \lambda_i \leq \left(\frac{\delta_n}{d_0 k_n} - 1\right) d_0 k_n \lambda_{k_n}^{(d_0)} = o\left(\frac{1}{\psi(\log n)}\right)$$

since  $\frac{\delta_n}{d_0 k_n} - 1$  has a finite limit  $\frac{\kappa}{d_0 \kappa'} - 1$ . As a consequence

$$\limsup_n \psi(\log n) e_n^2(X, \delta_n) \leq 1.$$

The extension to the general case  $\liminf_n \frac{\delta_n}{\log n} = \kappa \in (0, +\infty)$  is straightforward up to the extraction of a subsequence.  $\square$

**Remark.** Note that when  $b = 1$ , we do not need to let  $d_0$  go to infinity. Since this rate is optimal (in view of Theorem 2.2), this means in particular that *scalar product* quantization (*i.e.* block quantization with blocks of size  $d_0 = 1$ ) is asymptotically optimal.

#### 4 Lower bound

We will rely on the famous notion in Information Theory, the Shannon  $\varepsilon$ -entropy (or rate-distortion function) of  $P$  (see [17]). Let  $P$  be a probability measure on  $H$ . For  $\varepsilon > 0$ , it is defined by

$$R(\varepsilon) = R_P(\varepsilon) = \inf \left\{ \mathcal{H}(Q|P \otimes Q_2) : Q \text{ probability measure on } H \times H \right. \\ \left. \text{with first marginal } Q_1 = P \text{ and } \int_{H \times H} \|x - y\|^2 dQ(x, y) \leq \varepsilon^2 \right\}$$

where  $\mathcal{H}(Q|P \otimes Q_2)$  classically denotes the relative entropy (mutual information)

$$\mathcal{H}(Q|P \otimes Q_2) = \begin{cases} \int_H \log \left( \frac{dQ}{dP \otimes Q_2} \right) dQ & \text{if } Q \text{ is absolutely continuous with respect to } P \otimes Q_2, \\ +\infty & \text{otherwise.} \end{cases}$$

The simple converse part of Shannon’s source coding theorem (see [1] Theorem 3.2.2, [7], p.163) says that the minimal number  $N(\varepsilon)$  of codewords needed in a codebook  $\alpha$  such that  $\mathbb{E} \min_{a \in \alpha} \|X - a\|^2 \leq \varepsilon^2$  satisfies  $\log N(\varepsilon) \geq R(\varepsilon)$  so that, in particular

$$R(e_n(X)) \leq \log n.$$

We rely here on the closed form for Shannon’s entropy of Gaussian vectors known as Kolmogorov-Ihara’s formula (see [10, 8]) that we will apply to the probability measure

$P = \mathcal{L}(X^{(d)})$  and the space  $H = \bigoplus_{k=1}^d \mathbb{R} e_k^X$  (or equivalently to the  $d$ -dimensional normal

distribution  $P = \mathcal{N}(0; \text{Diag}(\lambda_1, \dots, \lambda_d))$  on the canonical space  $H = \mathbb{R}^d$ ). Of course, the eigenvalues are still supposed to be ordered in a non-increasing way.

**Theorem 4.1** (Kolmogorov-Ihara, see [10, 8]). *Let  $d \geq 1$  and let  $P = \mathcal{N}(0; \text{Diag}(\lambda_1, \dots, \lambda_d))$  where  $\lambda_1 \geq \dots \geq \lambda_d > 0$ . For every  $\varepsilon > 0$  such that  $\varepsilon^2 \in (0, \lambda_1 + \dots + \lambda_d)$ ,*

$$R(\varepsilon) = \frac{1}{2} \sum_{k=1}^{r(\varepsilon)} \log \left( \frac{\lambda_k}{\theta(\varepsilon)} \right) = \log \left( \prod_{k=1}^{r(\varepsilon)} \frac{\lambda_k}{\theta(\varepsilon)} \right)^{\frac{1}{2}}$$

where  $r(\varepsilon) = \max\{k \in \{1, \dots, d\} : \bar{\lambda}_k^d > \varepsilon^2\}$ , with  $\bar{\lambda}_k^d = k\lambda_k + \lambda_{k+1} + \dots + \lambda_d$ ,  $k = 1, \dots, d$  and  $\bar{\lambda}_k^d = 0$ ,  $k \geq d + 1$ , and  $\theta(\varepsilon)$  is the unique solution to the equation

$$\varepsilon^2 = r(\varepsilon)\theta(\varepsilon) + \sum_{k=r(\varepsilon)+1}^d \lambda_k.$$

Note that the definition of  $r(\varepsilon)$  is consistent since  $(\bar{\lambda}_k^d)_{1 \leq k \leq d}$  is non-increasing; furthermore by construction  $\theta(\varepsilon) \in [\lambda_{r(\varepsilon)+1}, \lambda_{r(\varepsilon)}]$ .

By the definition of optimal quantization at level  $n$ , we have, as recalled above (see also [13]),

$$\forall n \geq 1, \quad R(e_n(X^{(d)})) \leq \log n.$$

**Lemma 4.2.** Let  $d, n \in \mathbb{N}^*$ . Then  $e_n^2(X^{(d)}) \geq \min \left( n^{-\frac{2}{d}} d \left( \prod_{k=1}^d \lambda_k \right)^{\frac{1}{d}}, d\lambda_d \right)$ .

*Proof.* If  $e_n^2(X^{(d)}) =: \varepsilon^2 < \bar{\lambda}_d^d$ , then  $r(\varepsilon) = d$  and  $\theta(\varepsilon) = \frac{\varepsilon^2}{d}$  so that

$$R(\varepsilon) = \log \left( \prod_{k=1}^d \lambda_k \right)^{\frac{1}{2}} - \frac{d}{2} \log \left( \frac{\varepsilon^2}{d} \right) \leq \log n$$

if and only if

$$e_n^2(X^{(d)}) = \varepsilon^2 \geq n^{-\frac{2}{d}} d \left( \prod_{k=1}^d \lambda_k \right)^{\frac{1}{d}}.$$

□

**Proposition 4.3** (Lower bound). Assume (R). Let  $\delta_n$  be a sequence of dimensions going to infinity.

(a) If  $b > 1$  and  $\kappa = \limsup_n \frac{\delta_n}{\log n} \in [0, +\infty]$  then, with standard conventions,

$$\liminf_n \psi(\log n) e_n^2(X, \delta_n) \geq \kappa^{1-b} \left( \frac{1}{b-1} + e^{-2(\frac{1}{\kappa} - \frac{b}{2})_+} \right).$$

Furthermore, if  $\limsup_n \psi(\log n) e_n^2(X, \delta_n) = \limsup_n \psi(\log n) e_n^2(X)$ , then

$$\liminf_n \frac{\delta_n}{\log n} \geq \frac{2}{b}.$$

(b) If  $b = 1$ , then

$$\liminf_n \psi(\log n) e_n^2(X) \geq \liminf_n \frac{\psi(\log n)}{\psi(\delta_n)}.$$

Furthermore, if  $\limsup_n \psi(\log n) e_n^2(X, \delta_n) = \limsup_n \psi(\log n) e_n^2(X)$ , then

$$\liminf_n \frac{\psi(\delta_n)}{\psi(\log n)} \geq 1.$$

*Proof.* (a) Having in mind that  $\psi(x) = 1/(x\varphi(x))$ , it follows from Lemma 4.2 that

$$\begin{aligned} \psi(\log n) e_n^2(X^{(\delta_n)}) &\geq \psi(\log n) \min \left( n^{-\frac{2}{\delta_n}} \delta_n \left( \prod_{k=1}^{\delta_n} \lambda_k \right)^{\frac{1}{\delta_n}}, \frac{1}{\psi(\delta_n)} \right) \\ &= \frac{\psi(\log n)}{\psi(\delta_n)} \min \left( e^{-2\frac{\log n}{\delta_n}} \frac{1}{\varphi(\delta_n)} e^{\frac{1}{\delta_n} \sum_{1 \leq k \leq \delta_n} \log \varphi(k)}, 1 \right). \end{aligned}$$

The function  $\varphi$  being regularly varying with index  $-b$ ,  $b > 1$ , one checks (see [2])

$$\frac{1}{m} \sum_{k=1}^m \log \varphi(k) = b + \log \varphi(m) + o(1) \quad \text{as } m \rightarrow +\infty$$

so that,

$$\psi(\log n) e_n^2(X^{(\delta_n)}) \geq \frac{\psi(\log n)}{\psi(\delta_n)} \min \left( e^{-2\frac{\log n}{\delta_n} + b + o(1)}, 1 \right)$$

which in turn implies that

$$\psi(\log n) e_n^2(X, \delta_n) \geq \frac{\psi(\log n)}{\psi(\delta_n)} \left( \min \left( e^{-2\frac{\log n}{\delta_n} + b + o(1)}, 1 \right) + \frac{1 + o(1)}{b - 1} \right).$$

At this stage we introduce the function  $g_b$  defined on  $[0, +\infty]$  (with the usual conventions) by

$$g_b(u) := \left( \min \left( e^{-2\left(\frac{1}{u} - \frac{b}{2}\right)}, 1 \right) + \frac{1}{b - 1} \right) u^{1-b} = u^{1-b} \left( \frac{1}{b - 1} + e^{-2\left(\frac{1}{u} - \frac{b}{2}\right)+} \right).$$

The function  $g_b$  is decreasing on  $[0, +\infty]$  with

$$g_b\left(\frac{2}{b}\right) = \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1}.$$

Let  $(n')$  be a subsequence such that  $\frac{\delta_{n'}}{\log n'} \rightarrow u \in [0, +\infty]$ . Using that  $\psi$  is regularly varying with index  $b - 1$  we derive that

$$\liminf_n \psi(\log n') e_n^2(X, \delta_{n'}) \geq g_b(u)$$

so that finally

$$\liminf_n \psi(\log n) e_n^2(X, \delta_n) \geq \sup_{u \leq \kappa} g_b(u) = g_b(\kappa) \text{ where } \kappa = \limsup_n \frac{\delta_n}{\log n}.$$

Assume now that  $\limsup_n \psi(\log n) e_n^2(X, \delta_n) = \limsup_n \psi(\log n) e_n^2(X)$ . Let

$$c := \liminf_n \frac{\delta_n}{\log n} \in [0, +\infty]$$

and let  $(\delta_{n'})_{n \geq 1}$  be a subsequence such that  $\frac{\delta_{n'}}{\log n'} \rightarrow c$ . Let  $(\tilde{\delta}_n)_{n \geq 1}$  be a sequence going to infinity and satisfying  $\tilde{\delta}_{n'} = \delta_{n'}$  and  $\limsup_n \frac{\tilde{\delta}_n}{\log n} = c$ . Then one gets

$$\limsup_n \psi(\log n) e_n^2(X, \delta_n) \geq \liminf_n \psi(\log n') e_n^2(X, \delta_{n'}) \geq \liminf_n \psi(\log n) e_n^2(X, \tilde{\delta}_n) \geq g_b(c).$$

If  $c = 0$ ,  $g_b(0) = +\infty$  and we would have that  $\psi(\log n) e_n^2(X, \delta_n) \rightarrow +\infty$  which is in contradiction with claim (a) in Proposition 3.1.

If  $c \in (0, +\infty]$ , the upper bound obtained in Proposition 3.1 implies  $g_b(c) \leq g_b\left(\frac{2}{b}\right)$  which in turn implies  $c \geq \frac{2}{b}$ .

(b) Using again standard results from [2] about regularly varying functions with index  $-1$ , we get

$$\sum_{i \geq \delta_n + 1} \lambda_i \sim \frac{1}{\psi(\delta_n)} \quad \text{still with} \quad \psi(x) = \frac{1}{\int_x^{+\infty} \varphi(y) dy}.$$

Hence

$$\psi(\log n)e_n^2(X, \delta_n) \geq \psi(\log n) \sum_{i=\delta_n+1} \lambda_i \sim \frac{\psi(\log n)}{\psi(\delta_n)}.$$

Using the same trick (based on the sequence  $(\tilde{\delta}_n)_{n \geq 1}$ ) as in the former case), we derive similarly that, if  $\limsup_n \psi(\log n)e_n^2(X, \delta_n) = \limsup_n \psi(\log n)e_n^2(X)$ , then

$$\limsup_n \frac{\psi(\log n)}{\psi(\delta_n)} \leq 1$$

which is the announced result. □

## 5 Synthesis

### 5.1 Proof of Theorem 2.3 (sharp rate and constructive aspects)

First we provide a proof of Theorem 2.2 based on the upper and lower bounds established in former sections and the following lemma (already established in [13] but reproduced here for the reader's convenience). Furthermore, it has to be noticed that it provides an easily tractable (and asymptotically optimal) lower bound for the quadratic quantization error, keeping in mind that the sequence  $(k_n(1))_{n \geq 1}$  is defined in Lemma 3.3.

**Lemma 5.1.** *For every  $n \in \mathbb{N}^*$ ,*

$$e_n^2(X) \geq k_n(1)\lambda_{k_n(1)+1} + \sum_{k \geq k_n(1)+1} \lambda_k.$$

*Proof.* It follows from Kolmogorov-Ihara's formula that for every  $\varepsilon^2 \in (0, \lambda_1 + \dots + \lambda_d)$ ,  $R(\varepsilon) > a_{r(\varepsilon)}^{(1)}$  since  $\theta(\varepsilon) < \lambda_{r(\varepsilon)}$  (see Equation (3.3) for a definition of  $a_k^{(1)}$ ). As a consequence,  $a_{r(e_n(X^{(d)}))}^{(1)} \leq \log n$ . Consequently, it follows from Lemma 3.3(a) that  $r(e_n(X^{(d)})) \leq k_n(1)$  which in turn implies that, for every  $d \in \mathbb{N}^*$ ,  $\bar{\lambda}_{k_n(1)+1}^d \leq e_n^2(X^{(d)})$ . Noting that  $e_n^2(X) \geq e_n^2(X^{(d)})$  and letting  $d$  go to infinity, we get, for every  $n \in \mathbb{N}^*$ ,

$$e_n^2(X) \geq (k_n(1) + 1)\lambda_{k_n(1)+1} + \sum_{k \geq k_n(1)+2} \lambda_k.$$

□

Now, we come to the proof of Theorem 2.3.

*Proof.* Step 1 (Sharp rate): Case  $b > 1$ . We know from Proposition 3.1 that

$$\limsup_n \psi(\log n)e_n^2(X) \leq \limsup_n \psi(\log n)e_n^2(X, \delta_n) \leq \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1}.$$

On the other hand, combining the fact that  $k_n(1) \sim \frac{2}{b} \log n$  and arguments based on regularly varying functions already used in Proposition 3.1 yield that

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

and

$$\sum_{k \geq k_n(1)+1} \lambda_k \sim \frac{1}{(b-1)\psi(k_n(1))} \sim \left(\frac{b}{2}\right)^{b-1} \frac{1}{(b-1)\psi(\log n)}$$



so that  $\liminf_n \psi(\log n)e_n^2(X) \geq \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1}$  which gives the sharp rate of  $e_n(X)$ .

Case  $b = 1$ . One concludes likewise since  $k_n(1)\lambda_{k_n(1)+1} \sim \frac{1}{\psi(\log n)}$  and  $\sum_{k \geq k_n(1)+1} \lambda_k = o(1/\psi(\log n))$ .

Step 2 (Constructive aspects): When  $b > 1$ , Claim (2.3) in (a) follows from (3.1) combined with the above sharp rate for  $e_n(X)$ . The converse claim follows from the second claim in Proposition 4.3(a). The last claim is as follows. First note that

$$\psi(\delta_n)e_n^2(X^{(\delta_n)}) = \frac{\psi(\delta_n)}{\psi(\log n)}\psi(\log n)e_n^2(X) - \psi(\delta_n) \sum_{k \geq \delta_n+1} \varphi(k). \tag{5.1}$$

Then,  $\varphi$  and  $\psi$  being regularly varying with index  $b$  and  $b - 1$  respectively, we derive by standard arguments on regularly varying functions that  $\psi(\delta_n) \sum_{k \geq \delta_n+1} \varphi(k) \sim \frac{1}{b-1}$  and  $\liminf_n \frac{\psi(\delta_n)}{\psi(\log n)} \geq \left(\frac{2}{b}\right)^{b-1}$ . Combining this with the sharp rate for  $\psi(\log n)e_n(X)$  yields

$$\liminf_n \psi(\delta_n)e_n^2(X^{(\delta_n)}) \geq \left(\frac{2}{b}\right)^{b-1} \left(\frac{b}{2}\right)^{b-1} \frac{b}{b-1} - \frac{1}{b-1} = 1.$$

When  $b = 1$ , Claim (2.3) in (a) follows from (3.1) combined with the above sharp rate for  $e_n(X)$ . By Proposition 3.1(b),  $\limsup_n \psi(\log n)e_n^2(X) \leq \psi(\log n)e_n^2(X, \delta_n) \leq 1$ . Combining these two inequalities yields the announced result.

(b) Assume  $b > 1$ . If  $e_n(X, \delta_n) \sim e_n(X)$ , then (a) implies

$$\liminf_n \frac{\delta_n}{\log n} \geq \frac{2}{n} \quad \text{and} \quad \liminf_n \psi(\delta_n)e_n^2(X^{(\delta_n)}) \geq 1.$$

Then the three equivalences follow from (5.1), once noted again that

$$\psi(\delta_n) \sum_{k \geq \delta_n+1} \varphi(k) \sim \frac{1}{b-1}$$

since  $\delta_n \rightarrow +\infty$ . □

### 5.2 Proof of Theorem 2.5

*Proof.* (a) When  $b > 1$ , the direct claim on admissibility is a consequence of Proposition 4.3. The converse claim follows from Proposition 3.1(a) and Theorem 2.3.

As for strong admissibility, the direct claim is as follows: from the definition of strong admissibility, we get  $e_n(X, \delta_n)^2 \sim e_n(X)^2$  (by admissibility) and

$$e_n(X, \delta_n)^2 \sim \frac{1}{\psi(\delta_n)} + \frac{1}{b-1} \frac{1}{\psi(\delta_n)} = \frac{b}{b-1} \frac{1}{\psi(\delta_n)}$$

so that  $\frac{b}{b-1} \frac{1}{\psi(\delta_n)} \sim e_n(X)^2$ . Then comparing with the sharp rate from Theorem 2.3, we get

$$\frac{1}{\psi(\delta_n)} \sim \left(\frac{b}{2}\right)^{b-1} \frac{1}{\psi(\log n)}$$

which finally implies, having in mind that  $\psi$  is regularly varying with index  $b - 1$ , that

$$\delta_n \sim \frac{b}{2} \log n.$$

The converse claim is a consequence of Proposition 3.1(a) and Theorem 2.3. Claim (b) follows the same lines and details are left to the reader. □

**5.3 Back to the conjecture(s) ( $b > 1$ )**

As concerns the conjecture  $\lim_n \frac{d_n}{\log n} = \frac{2}{b}$  on the sharp asymptotics of the critical dimension  $d_n$ , strictly speaking, we only went half way by proving that

$$\liminf_n \frac{d_n}{\log n} \geq \frac{2}{b}.$$

The reverse inequality seems out of reach with the existing technology developed so far for functional quantization. However the strong admissibility result in Theorem 2.5(a) can be seen as an answer *in the asymptotic sense* since it shows that if  $\lim_n \frac{\delta_n}{\log n} = \frac{2}{b}$ , then the resulting quadratic quantization error is asymptotically optimal and (asymptotically almost) all dimensions are used (strong admissibility).

This result is helpful from a numerical point of view since it shows that for the Brownian motion, the Brownian bridge or the Ornstein-Uhlenbeck process (and any Gaussian process for which  $b = 2$ , see below), considering a truncation at  $\delta_n = \lfloor \log n \rfloor$  or  $\delta_n = \lceil \log n \rceil$  is (at least) asymptotically optimal whatever the future of the sharper conjecture

$$d_n \in \{ \lfloor \log n \rfloor, \lceil \log n \rceil \}$$

could be.

For various other examples of families of processes satisfying Assumptions (R) (including multi-parameters processes like the Brownian sheet, we refer to [13]).

▷ *Numerical experiments on the Brownian motion.* We know that the  $K$ - $L$  eigensystem of the standard Brownian motion  $W = (W_t)_{t \in [0, T]}$  over  $[0, T]$  is given by

$$\lambda_k^W = \left( \frac{T}{\pi(k - \frac{1}{2})} \right)^2, \quad e_n^W(t) := \sqrt{\frac{2}{T}} \sin\left(\frac{t}{\sqrt{\lambda_k}}\right), \quad k \geq 1,$$

so that  $b = 2$ . Then, Theorem 2.2 yields

$$\lim_n \log(n) e_n^2(W) = \frac{2T^2}{\pi^2} \approx 0.2026 \times T^2.$$

Figure 1 depicts the graph of the  $n \mapsto \log(n) e_n^2(W)$  (with  $T = 1$ ). One can see that it looks as a piecewise affine function with breaks in the slope. Note that the exponential function  $e^x$  satisfies

$$e^3 \approx 20.09 \approx 20 \quad e^4 \approx 54.59 \approx 55 \quad e^5 \approx 148.41 \approx 148.$$

These values graphically fit with the monotony slope breaks.

The graph in Figure 1 suggests, at this (low) range of the computation, that the limiting value for  $n \mapsto \log(n) e_n^2(W)$  is higher ( $\approx 0.22$ ) than the theoretical one ( $\approx 0.2026$ ). This impression is misleading since further computations not reproduced here show that the sequence  $n \mapsto \log(n) e_n^2(W)$  starts to be slowly decreasing beyond  $n \geq 1000$ . The value 0.22 seems to be a local maximum. For further details on these (highly time consuming) computations we refer to [14].

The quantization grids, computed during these numerical experiments by stochastic optimization methods (randomized Lloyd’s procedure, Competitive Learning Vector Quantization algorithm) for  $n = 1$  up to  $10^4$  for the standard Brownian motion (when  $T = 1$ ), can be downloaded from the website

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

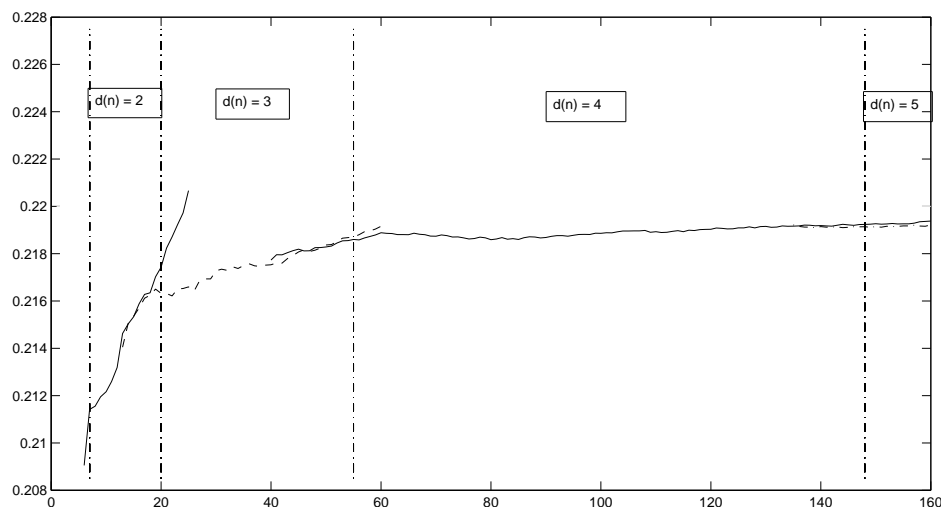


Figure 1:  $n \mapsto \log(n)e_n^2(W)$  and (conjectured) theoretical areas for  $d_n^W = d(n) = \lfloor \log n \rfloor$ .

## References

- [1] Berger, T.: Distortion Theory. *Prentice-Hall*, Englewood Cliffs, New Jersey, 1971. xiii+311 pp. MR-0408988
- [2] Bingham, N.H., Goldie, C.M. and Teugels, J.L.: Regular Variation. *Cambridge University Press*, Cambridge, 1987. xx+491 pp. MR-0898871
- [3] Corlay S.: Some aspects of optimal quantization and applications to finance. Thèse de l'Université Pierre et Marie Curie. 2011 tel-00626445
- [4] Corlay S., Lebowitz J. and Lévy-Véhel J.: Multifractional Stochastic volatility models. *Mathematical Finance*, **20**(2), (2014), 364–402.
- [5] Dembo, A. and Zeitouni, O.: Large deviations techniques and applications. (Corrected reprint of the second (1998) edition). *Stochastic Modeling and Applied Probability*, **38**. Springer-Verlag, Berlin, 2010. xvi+396p. MR-2571413
- [6] Dereich, S.: Asymptotic behavior of the distortion-rate function for Gaussian processes in Banach spaces. *Bull. Sci. Math.*, **129**(10), (2005), 791–803. MR-2178943
- [7] Graf, S. and Luschgy, H.: Foundations of Quantization for Probability Distributions. LNM 1730. *Springer*, Berlin, 2000. x+228 pp. MR-1764176
- [8] Ihara, S.: On  $\varepsilon$ -entropy of equivalent Gaussian processes. *Nagoya Math. J.*, **37**, (1970), 121–130. MR-0258107
- [9] Janson, S.: Gaussian Hilbert spaces. *Cambridge University Press*, Cambridge, 1997. x+340 pp. MR-1474726
- [10] Kolmogorov, A. N.: On the Shannon entropy of information transmission in the case of continuous signals. *IRE Trans. Inform. Theory*, **2**, (1956), 102–108.
- [11] Lejay, A. and Reutenauer, V.: A variance reduction technique using a quantized Brownian motion as a control variate. *J. Comput. Finance*, **16**(2), (2012), 61–84.
- [12] Luschgy, H. and Pagès, G.: Functional quantization of stochastic processes. *J. Funct. Anal.*, **196**(2), (2002), 486–531. MR-1943099
- [13] Luschgy, H. and Pagès, G.: Sharp asymptotics of the functional quantization problem for Gaussian processes. *Ann. Probab.*, **32**(2), (2004), 1574–1599. MR-2060310
- [14] Luschgy, H., Pagès, G. and Wilbertz, B.: Asymptotically optimal quantization schemes for Gaussian processes, *ESAIM P&S*, **14**, (2010), 93–116. MR-2654549
- [15] Pagès, G. and Printems, J.: Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods & Applications J.*, **11**(4), (2005), 407–446. MR-2186817

- [16] Pagès, G. and Sellami A.: Convergence of multi-dimensional quantized SDEs. Séminaire de Probabilités XLIII. LNM 2006 (C. Donati-Martin, A. Lejay, A. Rouault eds.). *Springer*, Berlin, 2010, 269–308. MR-2790377
- [17] Shannon, C.E. and Weaver, W.: The Mathematical Theory of Communication. *University of Illinois Press*, Urbana, 1949, 1949. vi+117 pp. MR-0032134

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

### Economical model of EJP-ECP

- Low cost, based on free software (OJS<sup>1</sup>)
- Non profit, sponsored by IMS<sup>2</sup>, BS<sup>3</sup>, PKP<sup>4</sup>
- Purely electronic and secure (LOCKSS<sup>5</sup>)

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>OJS: Open Journal Systems <http://pkp.sfu.ca/ojs/>

<sup>2</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>3</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>4</sup>PK: Public Knowledge Project <http://pkp.sfu.ca/>

<sup>5</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>