Convergence of multi-dimensional quantized SDE's

Gilles Pagès, Afef Sellami

Convergence of multi-dimensional quantized SDE's

Gilles PAGÈS [*] Afef SELLAMI [†]

Abstract

We quantize a multidimensional SDE (in the Stratonovich sense) by solving the related system of ODE's in which the d-dimensional Brownian motion has been replaced by the components of functional stationary quantizers. We make a connection with rough path theory to show that the solutions of the quantized solutions of the ODE converge toward the solution of the SDE. On our way to this result we provide convergence rates of optimal quantizations toward the Brownian motion for $\frac{1}{q}$-Hölder distance, $q > 2$, in $L^p(\mathbb{P})$.

Key words: Functional quantization, Stochastic Differential Equations, Stratonovich stochastic integral, stationary quantizers, rough path theory, Itô map, Hölder semi-norm, p-variation.

1 Introduction

Quantization is a way to discretize the path space of a random phenomenon: a random vector in finite dimension, a stochastic process in infinite dimension. Optimal Vector Quantization theory (finite-dimensional) random vectors finds its origin in the early 1950's in order to discretize some emitted signal (see [10]). It was further developed by specialists in Signal Processing and later in Information Theory. The infinite dimensional case started to

*LPMA, Université Paris 6, case 188, 4 pl. Jussieu, 75252 Paris Cedex 5, France gilles.pages@upmc.fr

†JP Morgan, London & LPMA, Université Paris 6, case 188, 4 pl. Jussieu, 75252 Paris Cedex 5, France afef.x.sellami@jpmorgan.com

be extensively investigated in the early 2000's by several authors (see *e.g.* [18], [5], [19], [20], [19], [4], [12], etc).

In [20], the functional quantization of a class of Brownian diffusions has been investigated from a constructive point of view. The main feature of this class of diffusions was that the diffusion coefficient was the inverse of the gradient of a diffeomorphism (both coefficients being smooth). This class contains most (non degenerate) scalar diffusions. Starting from a sequence of rate optimal quantizers, some sequences of quantizers of the Brownian diffusion are produced as solutions of (non coupled) ODE's. This approach relied on the Lamperti transform and was closely related to the Doss-Sussman representation formula of the flow of a diffusion as a functional of the Brownian motion. In many situations these quantizers are rate optimal (or almost rate optimal) *i.e.* that they quantize the diffusion at the same rate $O((\log N)^{-\frac{1}{2}})$ as the Brownian motion itself where N denotes the generic size of the quantizer. In a companion paper (see [27]), some cubature formulas based on some of these quantizers were implemented, namely those obtained from some optimal product quantizers based on the Karhunen-Loève expansion of the Brownian motion, to price some Asian options in a Heston stochastic volatility model. Rather unexpectedly in view of the theoretical rate of convergence, the numerical experiments provided quite good numerical results for some "small" sizes of quantizers. Note however that these numerical implementations included some further speeding up procedures combining the stationarity of the quantizers and the Romberg extrapolation leading to a $O((\log N)^{-\frac{3}{2}})$ rate. Although this result relies on some still pending conjectures about the asymptotics of bilinear functionals of the quantizers, it strongly pleads in favour of the construction of such stationary (rate optimal) quantizers, at least when one has in mind possible numerical applications.

Recently a sharp quantization rate (*i.e.* including an explicit constant) has been established for a class of not too degenerate 1-dimensional Brownian diffusions. However the approach is not constructive (see [4]). On the other hand, the standard rate $O((\log N)^{-\frac{1}{2}})$ has been extended in [22] to general d-dimensional Itô processes, so including d-dimensional Brownian diffusions regardless of their ellipticity properties. This latter approach, based an expansion in the Haar basis, is constructive, but the resulting quantizers are no longer stationary.

Our aim in this paper is to extend the constructive natural approach initiated in [20] to general d-dimensional diffusions in order to produce some rate optimal stationary quantizers of these processes. To this end, we will call upon some seminal results from rough path theory, namely the continuity of the Itô map, to replace the "Doss-Sussman setting". In fact we will show that if one replaces in an SDE (written in the Stratonovich sense) the Brownian motion by some elementary quantizers, the solutions of the resulting ODE's make up some rough paths which converge (in p-variation and in the Hölder metric) to the solution of the SDE. We use her the rough path theory as a tool and we do not aim at providing new insights on this theory. We can only mention that these rate optimal stationary quantizers can be seen as a new example of rough paths, somewhat less "stuck" to a true path of the underlying process.

This work is devoted to Brownian diffusions which is naturally the prominent example in view of applications, but it seems clear that this could be extended to SDE driven *e.g.* by fractional Brownian motions (however our approach requires to have an explicit form for the Karhunen-Loève basis as far as numerical implementation is concerned).

Now let us be more precise. We consider a diffusion process

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t) \circ dW_t,\ X_0 = x \in \mathbb{R}^d,\ t \in [0, T],$$

in the Stratonovich sense where $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0, T] \times \mathbb{R}^d \to \mathcal{M}(d \times d)$ are continuously differentiable with linear growth (uniformly with respect to t) and $W = (W_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P})$. (The fact that the state space and W have the same dimension is in no case a restriction since our result has nothing to do with ellipticity).

Such an SDE admits a unique strong solution denoted $X^x = (X^x_t)_{t\in[0,T]}$ (the dependency in ^{x} will be dropped from now to alleviate notations). The $\mathbb{R}^d$-valued process X is pathwise continuous and $\sup_{t\in[0,T]} |X_t| \in L^r(\mathbb{P})$, $r > 0$ (where $|\,.\,|$ denotes the canonical Euclidean norm on $\mathbb{R}^d$). In particular X is bi-measurable and can be seen as an $L^r(\mathbb{P})$-Radon random variable taking values in the Banach spaces $(L^p_{T,\mathbb{R}^d}, |\,.\,|_{L^p_T})$ where $L^p_{T,\mathbb{R}^d} = L^p_{\mathbb{R}^d}([0, T], dt)$ and $|g|_{L^p_T} = \left(\int_0^T |g(t)|^p dt\right)^{\frac{1}{p}}$ denotes the usual L^p-norm when $p \in [1, \infty)$.

For every integer $N \geq 1$, we can investigate for X the level N $(L^r(\mathbb{P}), L^p_T)$-quantization problem for this process X, namely solving the minimization of the $L^r(\mathbb{P})$-mean $L^p_{T,\mathbb{R}^d}$-quantization error

$$e_{N,r}(X,L^p) := \min\Big\{e_{N,r}(\alpha,X,L^p),\ \alpha\subset L^p_{T,\mathbb{R}^d},\ \mathrm{card}\,\alpha\le N\Big\} \tag{1.1}$$

where $e_{N,r}(\alpha,X,L^p)$ denotes the L^r-mean quantization error induced by α, namely

$$e_{N,r}(\alpha,X,L^p) := \Big(\mathbb{E}\min_{a\in\alpha}|X-a|_p^r\Big)^{\frac{1}{r}} = \Big\|\min_{a\in\alpha}|X-a|_{L^p_{T,\mathbb{R}^d}}\Big\|_{L^r(\mathbb{P})}.$$

The use of "min" in (1.1) is justified by the existence of an optimal quantizer solution to that problem as shown in [3, 13] in this infinite dimensional setting. The Voronoi diagram associated to a quantizer α is a Borel partition $(C_a(\alpha))_{a\in\alpha}$ such that

$$C_a(\alpha)\subset\Big\{x\in L^p_{T,\mathbb{R}^d}\ |\ |x-a|_{L^p_{T,\mathbb{R}^d}}\le\min_{b\in\alpha}|x-b|_{L^p_{T,\mathbb{R}^d}}\Big\}$$

and a functional quantization of X by α is defined by the nearest neighbour projection of X onto α related to the Voronoi diagram

$$\widehat{X}^\alpha := \sum_{a\in\alpha} a\mathbf{1}_{\{X\in C_a(\alpha)\}}.$$

In finite dimension (when considering $\mathbb{R}^d$-valued random vectors instead of $L^p_{T,\mathbb{R}^d}$-valued processes) the answer is provided by the so-called Zador Theorem which says (see [10]) that if $\mathbb{E}|X|^{r+\delta}<+\infty$ for some $\delta>0$ and if g denotes the absolutely continuous part of its distribution then

$$N^{\frac{1}{d}}e_{N,r}(X,\mathbb{R}^d)\to\widetilde{J}_{r,d}\|g\|^{\frac{1}{r}}_{\frac{d}{d+r}} \qquad \text{as} \qquad N\to\infty \tag{1.2}$$

where $\widetilde{J}_{r,d}$ is finite positive real constant obtained as the limit of the normalized quantization error when $X\overset{d}{=}U([0,1])$. This constant is unknown except when $d=1$ or $d=2$.

A non-asymptotic version of Zador's Theorem can be found *e.g.* in [22]: for every r, $\delta>0$ there exists a universal constant $C_{r,\delta}>0$ and an integer $N_{r,\delta}\ge$ such that, for every random vector $\Omega,\mathcal{A},\mathbb{P})\to\mathbb{R}^d$,

$$\forall\,N\ge N_{r,\delta}, \qquad e_{N,r}(X,\mathbb{R}^d)\le C_{r,\delta}\|X\|_{r+\delta}N^{-\frac{1}{d}}.$$

The asymptotic behaviour of the $L^s(P)$-quantization error of sequences of L^r-optimal quantizers of a random vector X when $s > r$ has been extensively investigated in [13] and will be one crucial tool to establish our mains results.

In infinite dimension, the case of Gaussian processes was the first to have been extensively investigated, first in the purely quadratic case ($r = p = 2$): sharp rates have been established for a wide family of Gaussian processes including the Brownian motion, the fractional Brownian motions (see [18, 19]). For these two processes sharp rates are also known for $p \in [1, \infty]$ and $r \in (0, \infty)$ (see [4]). More recently, a connection between mean regularity of $t \mapsto X_t$ (from $[0, T]$ into $L^r(\mathbb{P})$) and the quantization rate has been established (see [22]): if the above mapping is μ-Hölder for an index $\mu \in (0, 1]$, then

$$e_{N,r}(X, L^p) = O((\log N)^{-\mu}), \qquad p \in (0, r).$$

Based on this result, some universal quantization rates have been obtained for general Lévy processes with or without Brownian component some of them turning out to be optimal, once compared with the lower bound estimates derived from small deviation theory (see *e.g.* [11] or [5]). One important feature of interest of the purely quadratic case is that it is possible to construct from the Karhunen-Loève expansion of the process two families of rate optimal (stationary) quantizers, relying on

– sequences $(\alpha^{(N,prod)})_{N \geq 1}$ of optimal *product quantizers* which are rate optimal *i.e.* such that $e_{N,r}(\alpha^{(N)}, X, L^2) = O(e_{N,2}(X, L^2))$ (although not with a sharp optimal rate).

– sequences of *true* optimal quantizers (or at least some good numerical approximations) $(\alpha^{(N,*)})_{N \geq 1}$ *i.e.* such that $e_{N,r}(\alpha^{(N,*)}, X, L^2) = e_{N,2}(X, L^2)$.

We refer to Section 2.1 below for further insight on these objects (both being available on the website `www.quantize.math-fi.com`).

The main objective of this paper is the following: let $(\alpha^N)_{N \geq 1}$ denote a sequence of rate optimal stationary (see (2.8) further on) quadratic quantizers of a d'-dimensional standard Brownian motion $W = (W^1, \ldots, W^d)$. Define the sequence $x^N = (x_n^N)_{n=1,\ldots,N}$, $N \geq 1$, of solutions of the ODE's

$$x_n^N(t) = x + \int_0^t b(x_n^N(s))ds + \int_0^t \sigma(x_n^N(s))d\alpha_n^N(s), \quad n = 1, \ldots, N.$$

Then, the finitely valued-process defined by

$$\widetilde{X}^N = \sum_{n=1}^{N} x_n^N \mathbf{1}_{\{W\in C_n(\alpha^{(N)})\}}$$

converges toward the diffusion X on $[0,T]$ (at least in probability) as $N \to \infty$. This convergence will hold with respect to distance introduced in the rough path theory (see [25, 14, 6, 9, 26]) which always implies convergence with respect to the sup norm. The reason is that our result will appear as an application of (variants of the) the celebrated *Universal Limit Theorem* originally established by T. Lyons in [25]. The distances of interest in rough path theory are related to the $\frac{1}{q}$-Hölder semi-norm or the q-variation semi-norm both when $q > 2$ defined for every $x \in \mathcal{C}([0,T],\mathbb{R}^d)$ by

$$\|x\|_{q,Hol} = T^{\frac{1}{p}} \sup_{0\le s<t\le T} \frac{|x(t)-x(s)|}{|t-s|^{\frac{1}{q}}} \le +\infty,$$

and

$$\mathrm{Var}_{q,[0,T]}(x) \;:=\; \sup\Big\{\Big(\sum_{0\le \ell\le k-1} |x(t_{\ell+1})-x(t_\ell)|^q\Big)^{\frac{1}{q}} ,0\le t_0\le t_1\le\cdots\le t_k\le T, k\ge 1\Big\} \le +\infty$$

respectively. Note that

$$\|x-x(0)\|_{\sup} \le \mathrm{Var}_{p,[0,T]}(x) \le \|x\|_{p,Hol}.$$

From a technical viewpoint we aim at applying some continuity results established on the Itô map by several authors (see *e.g.* [14, 25, 6, 16]) that is the continuity of a solution x of the *ODE* (in a rough path sense)

$$dx_t = f(x_t)dy_t, \quad x_0 = x(0),$$

as a functional of y. However, the above (semi-)norms associated to a function x are not sufficient and the natural space to define such rough ODE is not the "naive" space of paths but a space of enhanced paths, which involves in the case of a multi-dimensional Brownian motion the mutual Lévy areas of its components. Convergence in this space is defined by considering appropriate $\frac{1}{q}$-Hölder and p-variation semi-norms to both the d-dimensional path and the related (pseudo-)Lévy areas (wit different values of q and p, see Section 3). Our application to quantized SDE's will make extensively use the fact that our functional quantizations of the Brownian motion W will all satisfy a stationary assumption *i.e.*

$$\widehat{W} = \mathbb{E}(W \,|\, \sigma(\widehat{W}))$$

so that we will extend the Kolmogorov criterion satisfied by W to its functional quantizers $\widehat{W}$ for free. This approach is rather straightforward and its field of application seems more general than our functional quantization purpose: thus the piecewise affine interpolations of the Brownian motion obviously satisfy such a property (see Appendix).

The paper is organized as follows. In Section 2 we provide some short background on functional quantization as well as preliminary elementary results on stochastic integration with respect to a stationary functional quantizer of a d-dimensional standard Brownian motion. In Section 3, we define a quantized approximation scheme of an SDE (in the Stratonovich sense) driven by a standard Brownian motion by its functionally quantized counterpart which turns out to be a system of (non-coupled) ODE's. To this end we recall some basic facts on rough path theory, in particular the notion of convergence we need to define on the so-called *multiplicative functionals* involved in the continuity of the Itô map which, when dealing with Brownian motion amounts, to some convergence in Hölder semi-norm of the naive path as well as, roughly speaking, the running (pseudo-)Lévy areas of its components. In Section 4 and 5, we establish successively the convergence in the Hölder distance of sequences of optimal stationary quantizations $\widehat{W}$ of the Brownian motion toward W: Section 4 is devoted to the convergence of the "regular" paths whereas Section 5 deals with the convergence of the running (pseudo-)Lévy areas (and to the global convergence of the couple). In both cases we provide some convergence rate in the $(\log N)^{-a}$, $a \in (0, \frac{1}{2})$ scale which is the natural scale for such convergences since optimal functional quantizations of the Brownian motion are known to converge at a $(\log N)^{-\frac{1}{2}}$-rate for most usual norms (like quadratic pathwise norm on $L^2([0,T], dt)$).

NOTATIONS: • For every $d \geq 1$, one denotes $\xi = (\xi^1, \dots, \xi^d)$ a row vector of $\mathbb{R}^d$. $\mathcal{M}(d \times d)$ will denote the set of square matrices with d lines.

• $|\,.\,|$ denotes the canonical Euclidean norm on $\mathbb{R}^d$.
• We denote $(\mathcal{F}_t^X)_{t\geq 0}$ the augmented natural filtration of a process $X = (X_t)_{t\geq 0}$ (so that it satisfies the usual conditions).
• For a bounded function $f : [0,T] \to \mathbb{R}^d$, $\|\, f\,\|_{\sup} := \sup_{t\in[0,T]} |f(t)|$. If f is a Borel function and $p\in[1,+\infty)$, $\|f\|_{L^p_{T,\mathbb{R}^d}} := \left(\int_0^T |f(t)|^p dt\right)^{\frac{1}{p}}$.
• For an $\mathbb{R}^d$-valued bi-measurable process X and $p \in [1,+\infty)$, we denote $\|X\|_p := \|\, |X|_{L^p_{T,\mathbb{R}^d}} \|_p = \left(\mathbb{E}\int_0^T |X_t|^p dt\right)^{1/p}$.
• We denote $t_k^n = \frac{kT}{2^n}$, $k = 0,\ldots,2^n$, the uniform mesh of the interval $[0,T]$, $T > 0$ and $I_k^n = [t_k^n, t_{k+1}^n]$, $k = 0,\ldots,2^n - 1$.
• $\lfloor x \rfloor$ denotes the lower integral part of $x\in\mathbb{R}$.
• Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be two sequences of real numbers: $a_n \sim b_n$ if $a_n = b_n + o(b_n)$ and $a_n \asymp b_n$ if $a_n = O(b_n)$ and $b_n = O(a_n)$.

2 Background and preliminary results on functional quantization

2.1 Some background on functional quantization

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

Let $(H, \langle\cdot,\cdot\rangle)$ 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 $\mathbb{P}_X$. Assume the integrability condition

$$\mathbb{E}\,|X|^2 < +\infty. \tag{2.3}$$

For $N \geq 1$, the L^2-optimal N-quantization problem for X consists in minimizing

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

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

$$e_N(X,H) := \inf\left\{(\mathbb{E}\min_{a\in\alpha} |X-a|^2)^{1/2} : \alpha\subset H,\ \mathrm{card}(\alpha)\leq N\right\}. \tag{2.4}$$

For a given N-quantizer α one defines an associated nearest neighbour projection

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

and the induced α-(Voronoi)quantization of X by setting

$$\hat{X}^\alpha := \pi_\alpha(X), \tag{2.5}$$

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

$$C_a(\alpha) \subset \{x \in H : |x-a| = \min_{b\in\alpha} |x-b|\} \tag{2.6}$$

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

$$\begin{aligned} e_n(X,H) &= \inf\left\{\left\||X-q(X)|\right\|_{L^2(\mathbb{P})},\ q : H \overset{Borel}{\to} H, \mathrm{card}(q(H)) \leq N\right\} \\ &= \inf\left\{\left\||X-Y|\right\|_{L^2(\mathbb{P})},\ Y:(\Omega,\mathcal{A}) \overset{r.v.}{\to} H, \mathrm{card}(Y(\Omega)) \leq N\right\}. \end{aligned} \tag{2.7}$$

A typical setting for functional quantization is $H = L^2_T := L^2_{\mathbb{R}}([0,1],dt)$ (equipped with $\langle f, g\rangle_2 := \int_0^T fg(t)dt$ and $|f|_{L^2_T} := \sqrt{\langle f, f\rangle_2}$). Thus any (bi-measurable, real-valued) process $X = (X_t)_{t\in[0,T]}$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that

$$\int_0^T \mathbb{E}(X_t^2)dt < +\infty$$

is a random variable $X : (\Omega, \mathcal{A}, \mathbb{P}) \to L^2_T$. But this Hilbert setting is not the only possible one for functional quantization (see *e.g.* [21], [12], [5], etc) since natural Banach spaces like $L^p_{\mathbb{R}}([0,T],dt)$ or $\mathcal{C}([0,T],\mathbb{R})$ are natural path-spaces.

In the purely Hilbert setting the existence of (at least) one *optimal N-quantizer* for every integer $N \geq 1$ is established so that the infimum in (2.4) holds as a minimum. A typical feature of this quadratic Hilbert framework is the so-called *stationarity* (or self-consistency) property satisfied by such an optimal N-quantizer $\alpha^{(N,*)}$:

$$\hat{X}^{\alpha^{(N,*)}} = \mathbb{E}(X \,|\, \hat{X}^{\alpha^{(N,*)}}). \tag{2.8}$$

This property, known as stationarity, will be used extensively throughout the paper.

This existence property holds true in any reflexive Banach space and L^1 path spaces (see [12] for details).

2.2 Constructive aspects of functional quantization of the Brownian motion

2.2.1 Karhunen-Loève basis ($d = 1$)

First we consider a scalar Brownian motion $(W_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The two main classes of rate optimal quantizers of the Brownian motion are the product optimal quantizers and the true optimal quantizers. Both are based on the Karhunen-Loève expansion of the Brownian motion given by

$$W_t = \sum_{k\geq 1} \sqrt{\lambda_k}\, \xi_k\, e_k^W(t) \tag{2.9}$$

where, for every $k \geq 1$,

$$\lambda_k = \left(\frac{T}{\pi(k-1/2)}\right)^2 \quad \text{and} \quad e_k^W(t) = \sqrt{\frac{2}{T}} \sin\left(\frac{t}{\sqrt{\lambda_k}}\right) \tag{2.10}$$

and

$$\xi_k = \frac{(W \,|\, e_k^W)_2}{\sqrt{\lambda_k}} = \sqrt{\frac{2}{T}} \int_0^T W_t \,\sin(t/\sqrt{\lambda_k}) \frac{dt}{\sqrt{\lambda_k}}.$$

The sequence $(e_k^W)_{k\geq 1}$ is an orthonormal basis of L_T^2. The system $(\lambda_k, e_k^W)_{k\geq 1}$ can be characterized as the eigensystem of the symmetric positive trace class covariance operator of $f \mapsto (t \mapsto \int_0^T (s\wedge t)\, f(s) ds) \equiv (t \mapsto \mathbb{E}(<f \,|\, W>_2 W_t)$. In particular this implies that the Gaussian sequence $(\xi_k)_{k\geq 1}$ is pairwise uncorrelated hence i.i.d., $\mathcal{N}(0;1)$-distributed. The Karhunen-Loève expansion of W plays the role of PCA of the process: it is the fastest way to exhaust the variance of W among all expansions on an orthonormal basis.

The convergence of the series in the right hand side of (2.9) holds in L^2_T for every $\omega \in \Omega$ and $\mathbb{P}(d\omega)$-*a.s.* for every $t \in [0, T]$. In fact this convergence also holds in $L^2(\mathbb{P})$ and $\mathbb{P}(d\omega)$-*a.s.* for the sup norm over $[0, T]$. The first convergence follows from Theorem 4.3(a) further on applied with $X = W$ and $\mathcal{G}_N = \sigma(\xi_1, \ldots, \xi_N)$ and the second one follows *e.g.* from [21] $\mathbb{P}(d\omega)$-*a.s.*. In particular the convergence holds in $L^2(d\mathbb{P} \otimes dt)$ or equivalently in $L^2_{L^2_T}(\mathbb{P})$. Note that this basis has already been used in the framework of rough path theory for Gaussian processes, see *e.g.* [2, 7, 8].

2.2.2 Optimal product quantization ($d \geq 1$)

$\triangleright$ *The one-dimensional case* $d = 1$. The previous expansion of the Brownian motion suggests to define a product quantization of W at level N by

$$\widehat{W}_t^{(N_1,\ldots,N_L)} := \sqrt{\frac{2}{T}} \sum_{k=1}^{L} \sqrt{\lambda_k}\, \widehat{\xi}_k^{N_k} \sin\left(\frac{t}{\sqrt{\lambda_k}}\right) \tag{2.11}$$

where $N_1, \ldots, N_L$ are non zero integers satisfying $N_1 \cdots N_L \leq N$ and $\widehat{\xi}_1^{N_1}, \ldots, \widehat{\xi}_L^{N_L}$ are optimal quadratic quantizations of $\xi_1, \ldots, \xi_L$. The resulting (squared) quadratic quantization error reads

$$\|W - \widehat{W}^{(N_1,\ldots,N_L)}\|_2^2 = \sum_{k \geq 1} \frac{\lambda_k}{N_k^2}. \tag{2.12}$$

An *optimal product N-quantization* $\widehat{W}^{N,prod}$ is obtained as a solution to the following *integral bit allocation optimization* problem for the sequence $(N_k)_{k \geq 1}$:

$$\min\left\{\|W - \widehat{W}^{(N_1,\ldots,N_L)}\|_2,\ N_1, \ldots, N_L \geq 1,\ N_1 \cdots N_L \leq N,\ L \geq 1\right\} \tag{2.13}$$

(see [18] for further details and [27] for the numerical aspects). It is established in [18] (as a special case of a more general result on Gaussian processes) that

$$\frac{1}{T}\|W - \widehat{W}^{N,prod}\|_2 \asymp (\log N)^{-\frac{1}{2}} \tag{2.14}$$

Furthermore, the critical dimension $L = L_W(N)$ satisfies $L_W(N) \sim \log N$. Numerical experiments carried out in [27] show that

$$\frac{1}{T}\|W - \widehat{W}^{N,prod}\|_2 \approx c_W (\log N)^{-\frac{1}{2}}$$

with $c_W \approx 0.5$ (at least up to $N \leq 10000$).

It is possible to get a closed form for the underlying optimal product quantizers α^N. First, note that the normal distribution on the real line being log-concave, there is exactly one stationary quadratic quantizer of full size M for every $M \geq 1$ (hence it is the optimal one). So, let $N \geq 1$ and let $(N_k)_{k\geq 1}$ denote its optimal integral bit allocation for the Brownian motion W. For every $N_k \geq 1$, we denote by $\beta^{(N_k)} := \{\beta^{(N_k)}_{i_k},\ 1 \leq i_k \leq N_k\}$ the unique optimal quantizer of the normal distribution: thus $\alpha(0) = \{0\}$ by symmetry of the normal distribution. Then, the optimal quadratic product N-quantizer $\alpha^{N,prod}$ (of "true size" $N_1 \times \cdots \times N_{L_W(N)} \leq N$) can be described using a multi-indexation as follows:

$$\alpha^{N,prod}_{(n_1,\ldots,n_k,\ldots)}(t) = \sum_{k\geq 1} \beta^{(N_k)}_{n_k} \sqrt{\lambda_k} e^W_k(t), \qquad n_k \in \{1,\ldots,N_k\},\ \ k \geq 1.$$

These sums are in fact all finite so that all the functions $\alpha^{N,prod}_{(i_1,\ldots,i_n,\ldots)}$ are $\mathcal{C}^\infty$ with finite variation on every interval of $\mathbb{R}_+$.

Explicit optimal *integral bit allocations* as well as optimal quadratic quantizations (quantizers and their weights) of the scalar normal distribution are available on the website [28]. Note for practical applications that this optimal product quantization is based on 1-dimensional quantizations of small size of the scalar normal distribution $\mathcal{N}(0;1)$. This kind of functional quantization has been applied in [27] to price Asian options in a Heston stochastic volatility model.

$\rhd$ *The d-dimensional case.* Assume now $W = (W^1,\ldots,W^d)$ is a d-dimensional Brownian motion. Its optimal product quantization at level $N \geq 1$ will be defined as the optimal product quantization at level $\lfloor N^{\frac{1}{d}} \rfloor$ of each of its d components.

$\rhd$ *Additional results on optimal vector quantization of the normal distribution on $\mathbb{R}^d$.* We will extensively make use of the *distortion mismatch* result established in [13] that we recall here only in the d-dimensional Gaussian case. Let Z be an $\mathcal{N}(0;I_d)$ random vector and let α^N be an optimal quadratic quantizer at level N of Z (hence of size N). Then

$$(i)\quad \forall\, p\in(0,2+d),\ \forall\, N\geq 1,\ \|Z-\widehat{Z}^{\alpha^N}\|_p\leq C_{Z,p}N^{-\frac{1}{d}}, \tag{2.15}$$

$$(ii)\quad \forall\, p\in[2+d,+\infty),\ \forall\,\eta\in(0,d+2),\forall\, N\geq 1,$$
$$\|Z-\widehat{Z}^{\alpha^N}\|_p\leq C_{Z,p,\eta}N^{-\frac{2+d-\eta}{dp}} \tag{2.16}$$

where $C_{Z,p}$ and $C_{Z,p,\eta}$ are two positive real constants.

2.2.3 Optimal quantization ($d = 1$)

It is established in [18] (Theorem 3.2) that the quadratic optimal quantization of the one-dimensional Brownian motion reads

$$\widehat{W}_t^{N,opt}=\sqrt{\frac{2}{T}}\sum_{k=1}^{d_W(N)}\sqrt{\lambda_k}\,(\widehat{\zeta}^N_{d_W(N)})^k\sin\left(\frac{t}{\sqrt{\lambda_k}}\right) \tag{2.17}$$

where, for every integer $d\geq 1$, $\zeta_d=\mathrm{Proj}^{\perp}_{E_d}(W)\sim\mathcal{N}(0;\mathrm{Diag}(\lambda_1,\ldots,\lambda_d))$ with $E_d:=\mathbb{R}\text{-span}\left\{\sin\left(./\sqrt{\lambda_1}\right),\ldots,\sin\left(./\sqrt{\lambda_d}\right)\right\}$ and $\widehat{\zeta}^N_d$ is an optimal quadratic quantization of ζ_d at level (or of size) N.

If one considers an optimal quadratic N-quantizer $\beta^N=\{\beta^N_n,\ n=1,\ldots,N\}\subset\mathbb{R}^{d_W(N)}$ of the distribution $\mathcal{N}\left(0;\mathrm{Diag}(\lambda_1,\ldots,\lambda_{d_W(N)})\right)$ (*a priori* not unique)

$$\alpha_n^{N,opt}(t)=\sum_{k=1}^{d_W(N)}(\beta_n^{(N)})^k\sqrt{\lambda_k}\,e_k^W(t),\qquad n=1,\ldots,N.$$

Once again this defines a $\mathcal{C}^\infty$ function with finite variation on every interval of $\mathbb{R}_+$.

A sharp rate has been obtained in [19] for the resulting optimal quantization error

$$\|W-\widehat{W}^{N,opt}\|_2\sim T c_W^{opt}(\log N)^{-\frac{1}{2}}\quad\text{ as }\quad N\to\infty \tag{2.18}$$

where $c_W^{opt}=\frac{\sqrt{2}}{\pi}\approx 0.4502$.

The true value of the critical dimension $d_W(N)$ is unknown. A conjecture supported by numerical evidences is that $d_W(N)\sim\log N$. Recently a first step to this conjecture has been established in [23] by showing that

$$\liminf_N \frac{d_W(N)}{\log(N)} \geq \frac{1}{2}.$$

Large scale computations of optimal quadratic quantizers of the Brownian motion have been carried out (up to $N = 10\,000$ and $d = 10$). They are available on the website [28].

In the d-dimensional setting, several definitions of an *optimal quantization* of the Brownian motion $W = (W^1, \ldots, W^d)$ can be given. For our purpose, it is convenient to adopt the following one:

$$\widehat{W}^{N,opt} := \left(\widehat{W^i}^{\lfloor N^{\frac{1}{d}} \rfloor, opt}\right)_{1 \leq i \leq d}.$$

Its property of interest is that this definition preserves the componentwise independence as well as a stationarity property (see below) since

$$\mathbb{E}\Big(W^i \,|\, \widehat{W}^{N,opt}\Big) = \mathbb{E}\Big(W^i \,|\, \widehat{W^i}^{\lfloor N^{\frac{1}{d}} \rfloor, opt}\Big) = \widehat{W^i}^{\lfloor N^{\frac{1}{d}} \rfloor, opt}, \quad i = 1, \ldots, d.$$

2.2.4 Wiener like integral with respect to a stationary functional quantization ($d = 1$)

Both types of quantizations defined above share an important property of quantizers: stationarity.

Definition 2.1. *Let $\alpha \subset L^2_T$, $\alpha \neq \emptyset$, be a quantizer. The quantizer α is stationary for the (one-dimensional) Brownian motion W if there is a Voronoi quantization $\widehat{W} := \widehat{W}^\alpha$ induced by α such that*

$$\widehat{W} = \mathbb{E}(W \,|\, \sigma(\widehat{W})) \qquad a.s. \tag{2.19}$$

where $\mathbb{E}(\,.\,|\mathcal{G})$ denotes the functional conditional expectation given the σ-field $\mathcal{G}$ on $L^2_{L^2_T}(\mathbb{P})$ (see Appendix) and $\sigma(\widehat{W})$ is the σ-field spanned by $\widehat{W}$.

Note that if α is stationary for one Brownian motion, so it is for any Brownian motion since this stationarity property only depends on the Wiener distribution.

In the case of product quantization $\widehat{W}^{N,prod}$, this follows from the stationarity property of the optimal quadratic quantization of the marginals ξ_n

(see [18] or [27]). In the case of optimal quadratic quantization $\widehat{W}^{N,opt}$ this follows from the optimality of the quantization of $\zeta_{d_W(N)}$ itself.

We will now define a kind of *Wiener integral with respect* to such *a stationary quantization* $\widehat{W}$ of a *one-dimensional* W. So we assume that $d = 1$ until the end of this Section.

First, we must have in mind that if W is an $(\mathcal{F}_t)$-Brownian motion where the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions, one can define the Wiener stochastic integral (on $[0,T]$) of any process $\varphi \in L^2([0,T]\times\Omega, \mathcal{B}([0,T])\otimes\mathcal{F}_0, dt\otimes d\mathbb{P})$ with respect to W. The non-trivial case is when $\mathcal{F}_t^W \neq \mathcal{F}_t$, typically when $\mathcal{F}_t = \mathcal{F}_T^B \vee \mathcal{F}_t^W$, $t\in[0,T]$ where B and W are independent. One can see it as a special case of Itô stochastic integral or as an extended Wiener integral: if $(\varphi(t,\omega))_{(\omega,t)\in\Omega\times[0,T]}$ denotes an elementary process of the form

$$\varphi(t,\omega) := \sum_{k=1}^{n} \varphi_k(\omega)\mathbf{1}_{s_k<t\leq s_{k+1}},\ 0 = s_0 < s_1 < \cdots < s_{n-1} < s_n = T$$

where the random variables φ_i are $\mathcal{F}_0$-measurable (hence independent of W). Set

$$I_T(\varphi) := \sum_{k=1}^{n} \varphi_k (W_{s_{k+1}} - W_{s_k}).$$

Then, I_T is an isometry from $L^2_{L^2_T}(\mathbb{P})$ into $L^2(\mathcal{F}_T, \mathbb{P})$. Furthermore, one easily checks that

$$\mathbb{E}\left(\int_0^T \varphi(s,.)dW_s \,|\, \mathcal{F}_T^W\right) = \int_0^T \mathbb{E}\left(\varphi(s,.)\,|\,\mathcal{F}_T^W\right)dW_s$$

where $\mathcal{F}_T^W$ denotes the augmented filtration of W at time T. We follow the same lines to define the stochastic integral with respect to a stationary quantizer. Set for the same elementary process φ

$$\widehat{I}_T(\varphi) = \sum_{k=1}^{n} \xi_k(\widehat{W}_{s_{k+1}} - \widehat{W}_{s_k})$$

so that

$$\begin{aligned}\widehat{I}_T(\varphi) &= \sum_{k=1}^{n} \xi_i\, \mathbb{E}(W_{s_{k+1}} - W_{s_k} \,|\, \widehat{W})\\ &= \sum_{k=1}^{n} \mathbb{E}\big(\xi_k (W_{s_{k+1}} - W_{s_k}) \,|\, \mathcal{F}_0 \vee \sigma(\widehat{W})\big)\\ &= \mathbb{E}\left(\int_0^T \varphi(t,.)dW_t \,|\, \mathcal{F}_0 \vee \sigma(\widehat{W})\right)\end{aligned}$$

where we used that the σ-fields $\sigma(\widehat{W})$ and $\mathcal{F}_0$ are independent since $\widehat{W}$ is a Borel function of W. As a consequence,

$$\|\widehat{I}_T(\varphi)\|_2^2 \le \|I_T(\varphi)\|_2^2 = \|\,|\varphi|_{L^2_T}\|_2^2.$$

Hence, the linear transformation $\widehat{I}_T$ extends into a linear continuous mapping on the whole set $L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$. Furthermore, one checks, first on elementary processes, then on $L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$ by continuity of the (functional) conditional expectation, that

$$\mathbb{E}\left(I_T(\varphi) \,|\, \mathcal{F}_0 \vee \sigma(\widehat{W})\right) = \widehat{I}_T(\varphi).$$

We will denote from now on $\widehat{I}_T(\varphi)(\omega)$ as an integral, namely

$$\widehat{I}_T(\varphi)(\omega) := \int_0^T \varphi(t,\omega)d\widehat{W}_t(\omega).$$

Now set as usual, for every $t \in [0,T]$,

$$\int_0^t \varphi(s,\omega)d\widehat{W}_s(\omega) := \int_0^T \mathbf{1}_{[0,t]}(s)\varphi(s,\omega)d\widehat{W}_s(\omega).$$

One checks using Jensen and Doob Inequality that,

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t \varphi(s,.)d\widehat{W}_s\right|^2 &\le \mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t \varphi(s,.)dW_s\right|^2\\ &\le 4\,\mathbb{E}\int_0^T \varphi^2(s,.)\,ds. \qquad (2.20)\end{aligned}$$

Furthermore, as soon as the underlying stationary quantizer α (such that $\widehat{W} = \widehat{W}^{\alpha}$) is made up with pathwise continuous elements, for every elementary process φ, its integral process

$$\int_0^t \varphi(s,.)\, dW_s = \sum_{k=1}^{n} \xi_k (\widehat{W}_{s_{k+1}\wedge t} - \widehat{W}_{s_k \wedge t})$$

pathwise continuous as well since $\widehat{W}$ is α-valued. One classically derives, by combining this result with (2.20) and the everywhere density of elementary processes, that, for every $\varphi \in L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$, the process

$$\left(\int_0^t \varphi(s,.) d\widehat{W}_s\right)_{t\in[0,T]} \qquad \text{admits a continuous modification.}$$

This is always this modification that will be considered from now on. As a matter of fact, if φ_n denotes a sequence of elementary processes in $L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$ converging to φ, *i.e.* satisfying

$$\mathbb{E} \int_0^T (\varphi - \varphi_n)^2(s,.) ds \longrightarrow 0 \qquad \text{as } n \to \infty.$$

It follows from (2.20) that the convergence also holds in $L^2_{L^\infty_T}(\mathcal{F}_0, \mathbb{P})$. In particular, there is a subsequence that converges $\mathbb{P}$-*a.s.* for the $\|\,.\,\|_{\sup}$ which implies the existence of a continuous modification for $\int_0^t \varphi(s,\omega) d\widehat{W}_s(\omega)$.

Finally, using the characterization of functional conditional expectation (see Appendix), it follows that

$$\mathbb{E}\left(\int_0^{\cdot} \varphi(s,.) d\widehat{W}_s\ ,\, | \,\mathcal{F}_0 \vee \sigma(\widehat{W})\right) = \int_0^{\cdot} \varphi(s,.) d\widehat{W}_s. \tag{2.21}$$

Proposition 2.1. *Let W be a (real-valued) $\mathcal{F}_t$-standard Brownian motion.*

(a) For every $\varphi \in L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$

$$\int_0^t \varphi(s,.) dW_s = \sqrt{\frac{2}{T}} \sum_{k\geq 1} \xi_k \int_0^t \varphi(s,.) \cos(s/\sqrt{\lambda_k}) ds \tag{2.22}$$

where $\xi_k := (W|e^W_k)_2/\sqrt{\lambda_k}$ are independent, $\mathcal{N}(0;1)$-distributed (see (2.9) and (2.10)) and independent of φ.

(b) *Let $\widehat{W}$ be a stationary quantization of W. For every $\varphi \in L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$*

$$\int_0^t \varphi(s,.)d\widehat{W}_s = \sqrt{\frac{2}{T}} \sum_{k\geq 1} \frac{(\widehat{W}|e_k^W)_2}{\sqrt{\lambda_k}} \int_0^t \varphi(s,.)\cos(s/\sqrt{\lambda_n})ds. \tag{2.23}$$

In particular if $\widehat{W}$ is a product quantization, then

$$\frac{(\widehat{W}|e_k^W)_2}{\sqrt{\lambda_k}} = \frac{\widehat{(W|e_k^W)_2}}{\sqrt{\lambda_k}} = \widehat{\xi}_k.$$

Proof. (a) Set for every $\varphi \in L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$,

$$\begin{aligned} J_T(\varphi) &:= \sqrt{\frac{2}{T}} \sum_{k\geq 1} \xi_k \sqrt{\lambda_k} \int_0^T \varphi(s,.)d\sin(s/\sqrt{\lambda_k}) \qquad (2.24) \\ &= \sqrt{\frac{2}{T}} \sum_{k\geq 1} \xi_k \int_0^T \varphi(s,.)\cos(s/\sqrt{\lambda_k})ds. \end{aligned}$$

This defines clearly an isometry from $L^2_{L^2_T}(\mathcal{F}_0, \mathbb{P})$ into the Gaussian space spanned by $(\xi_n)_{n\geq 1}$ since

$$\mathbb{E}(J_T(\varphi)^2) = \frac{2}{T} \sum_{k\geq 1} \mathbb{E}(\xi_k^2)\,\mathbb{E}\left(\int_0^T g(s)\frac{1}{\sqrt{\lambda_k}}\cos(s/\sqrt{\lambda_k})ds\right)^2 = \mathbb{E}\int_0^T g^2(t)dt.$$

The last equality uses that the sequence $\left(\sqrt{\frac{2}{T}}\cos(\pi(k-\frac{1}{2})t/T)\right)_{k\geq 1}$ is an orthonormal basis of L^2_T. Finally, note that for every $t \in [0,T]$, $J_T(\mathbf{1}_{[0,t]}) = \sqrt{\frac{2}{T}} \sum_{k\geq 1} \sqrt{\lambda_k}\, \xi_k \sin(t/\sqrt{\lambda_k}) = W_t$. This proves that $J_T = I_T$ *i.e.* is but the (extended) Wiener integral with respect to W.

(b) This follows by taking the (functional) conditional expectation of (2.22).
$\diamondsuit$

2.2.5 Application to multi-dimensional Brownian motions ($d \geq 2$)

Now we apply the above result to a componentwise (stationary) functional quantization of a multi-dimensional standard Brownian motion.

Proposition 2.2. *Let $W =: (W^1,\dots,W^d)$ denote a d-dimensional standard Brownian motion and let $\widehat{W} := (\widehat{W}^1,\dots,\widehat{W}^d)$ be a pathwise continuous stationary quantization of W (no optimality is requested here). Then, $\mathbb{P}$-a.s., for every $i \neq j$, $i, j \in \{1,\dots,d\}$, for every $s,t \in [0,T]$, $0 \le s \le t$,*

$$\mathbb{E}\left(\int_s^t (W_u^i - W_s^i)dW_u^j \,|\, \sigma(\widehat{W})\right) = \int_s^t (\widehat{W}_s^i - \widehat{W}_s^i)d\widehat{W}_u^j.$$

Proof. All the components of $\widehat{W}$ being independent, it is clear one can replace $\sigma(\widehat{W})$ by $\sigma(\widehat{W}^i, \widehat{W}^j)$. Then, the stochastic integral $\int_0^\cdot W_s^i dW_s^j$ coincides with the (extended) Wiener integral defined with respect to the filtration $\mathcal{G}_{i,t}^j := \sigma(\mathcal{F}_T^{W^i}, \mathcal{F}_t^{W^j})$ (it is clear that W^j is a $\mathcal{G}_{i,t}^j$-standard Brownian motion still by independence). The result is then a straightforward consequence of (2.21). $\diamond$

Remark. The above result still holds if one considers an additional "0^{th}" component $W_t^0 = t$ to the Brownian motion and to its functional quantization by setting $\widehat{W}_t^0 = t$ as well.

3 Convergence of quantized SDE's: a rough path approach

3.1 From Itô to Stratonovich

An SDE

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t, \quad X_0 \in L^p_{\mathbb{R}^d}(\mathbb{P})$$

where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to \mathcal{M}(d \times q)$ are smooth enough functions (*e.g.* continuously differentiable with bounded differentials) and $W = (W_t)_{t \in [0,T]}$ is a q-dimensional Brownian motion. First note that without loss of generality one may assume that $q = d$ by increasing the dimension of W or adding some identically zero components to X (no ellipticity like assumption is needed here). This SDE can be written in the Stratonovich sense as follows

$$dX_t = f(X_t) \circ dW_t, \quad X_0 \in L^p_{\mathbb{R}^d}(\mathbb{P}), \tag{3.25}$$

where, for notational convenience $W = (W^0, W^1, \dots, W^q)$ stands for (t, W_t), $X_t = (X_t^0, X_t^1, \dots, X_t^d)$ stands for (t, X_t) and $f : [0,T] \times \mathbb{R}^d \to \mathcal{M}((d+1) \times$

$(d+1)$) (with $f^{0\cdot}(t,x) = (1,0,\dots,0)$ as 0^{th} line) is a differentiable function with bounded differentials.

Following rough paths theory initiated by T. Lyons ([25]) and developed with many co-authors (see *e.g.* [26, 14, 16, 26, 9] for an introduction), one can also solve this equation in the sense of rough paths with finite p-variation, $p \geq 2$, since we know (*e.g.* from the former Kolmogorov criterion) that W *a.s.* does have finite $\frac{1}{q}$-Hölder norm, for any $q > 2$. Namely this means solving an equation formally reading

$$dx_t = f(x_t)d\mathbf{y}_t, \quad x_0 \in \mathbb{R}^d. \tag{3.26}$$

In this equation $\mathbf{y}$ does not represent the path (null at 0) $y_t = W_t(\omega)$, $t \in [0,T]$ itself but an *enhanced path* embedded in a larger space, also called *geometric multiplicative functional lying on* y with controlled $\frac{1}{q}$-Hölder semi-norm, namely a couple $\mathbf{y} = ((\mathbf{y}^1_{s,t})_{0\leq s\leq t\leq T}, (\mathbf{y}^2_{s,t})_{0\leq s\leq t\leq T})$ where $\mathbf{y}^1_{s,t} = y_t - y_s \in \mathbb{R}^{d+1}$, $0 \leq s \leq t \leq T$, can be identified with the path (y_t) and $(\mathbf{y}^2_{s,t})_{0\leq s\leq t\leq T}$ satisfies, $\mathbf{y}^2_{s,t} \in \mathbb{R}^{(d+1)^2}$ for every $0 \leq s \leq u \leq t \leq T$ and the following tensor multiplicative property

$$\mathbf{y}^2_{s,t} = \mathbf{y}^2_{s,u} + \mathbf{y}^2_{u,t} + \mathbf{y}^1_{s,u} \otimes \mathbf{y}^1_{u,t}.$$

Different choices for this functional are possible, leading to different solutions to the above Equation (3.26). The choice that makes coincide *a.s.* the solution of (3.25) and the pathwise solutions of (3.26) is given by

$$\mathbf{y}^1_{s,t} = W_t(\omega) - W_s(\omega),\ \mathbf{y}^2_{s,t} := \left(\int_s^t (W^i_u - W^i_s) \circ dW^j_u\right)_{i,j=0,\dots,d}(\omega) \tag{3.27}$$

so that

$$\mathbf{y}^1_{s,u} \otimes \mathbf{y}^1_{u,t} = \left(\mathbf{y}^{1,i}_{s,u}\mathbf{y}^{1,j}_{s,u}\right)_{i,j=0,\dots,d}.$$

The term $\mathbf{y}^2_{s,t}$ is but the "running" Lévy areas related to the components of the Brownian mtion W. The enhanced path of W will be denoted $\mathbf{W}$ (although we will keep the notation $\mathbf{y}$ in some proofs for notational convenience). One defines, for every $q \geq 1$, the $\frac{1}{q}$-Hölder distance by setting

$$\rho_q(\mathbf{y} - \mathbf{x}) = \|\mathbf{y}^1 - \mathbf{x}^1\|_{q,Hol} + \|\mathbf{y}^2 - \mathbf{x}^2\|_{q/2,Hol}$$

where

$$\|\mathbf{x}^2\|_{q/2,Hol} := T^{\frac{2}{q}} \sup_{0\leq s<t\leq T} \frac{|\mathbf{x}^2_{s,t}|}{|t-s|^{\frac{2}{q}}}.$$

Remark. Likewise, when $p \in [2,3)$, one defines the *p-variation distance* between two such multiplicative functionals $\mathbf{y}$, $\mathbf{z}$ is defined by

$$\delta_p(\mathbf{y},\mathbf{z}) = \mathrm{Var}_{p,[0,T]}(\mathbf{y}^1 - \mathbf{z}^1) + \mathrm{Var}_{p/2,[0,T]}(\mathbf{y}^2 - \mathbf{z}^2)$$

where

$$\mathrm{Var}_{q,[0,T]}(\mathbf{y}^2) := \sup\left\{\left(\sum_{\ell=0}^{k-1} |\mathbf{y}^2_{t_\ell,t_{\ell+1}}|^q\right)^{\frac{1}{q}}, 0 \le t_0 \le t_1 \le \cdots \le t_k \ \le T, k \ge 1\right\}.$$

The distance ρ_q has been introduced in [24] although rough path theory was originally developed for the distance δ_p in p-variation. Recently several authors came back to Hölder distances ρ_q (see *e.g.* [16, 6, 9]).

The following so-called universal limit theorem theorem (including variants) describes the continuity of the so-called Itô map $\mathbf{y} \mapsto x$ with respect to both δ_p and ρ_p-distances and will be the key for our main result. It was the starting point of rough path theory initiated by T. Lyons. Several statements (or improvements) can be found *e.g.* in [25, 14, 15, 26, 9]. We state here some versions coming from [14] and [16].

Theorem 3.1. *Let* $\alpha \in (0,1]$.
(*a*) *(See [16]) Let* $f : [0,T] \times \mathbb{R}^d \to \mathcal{M}((d+1) \times (d+1)$, *twice differentiable with a bounded first differential and an* α*-Hölder second differential. Suppose the multiplicative functional* $\mathbf{y}$ *satisfies* $\|\mathbf{y}^1 - \mathbf{x}^1\|_{q,Hol} + \|\mathbf{y}^2 - \mathbf{x}^2\|_{q/2,Hol} < +\infty$ *for* $q \in (2, 2+\alpha)$. *Then Equation (3.26) has a unique solution starting at* x_0.

When $\mathbf{y} = \mathbf{W}(\omega)$ *(i.e. given by (3.27)), the first component* $\mathbf{x}^1 = x$ *of the solution solution* $\mathbf{x} = (\mathbf{x}^1, \mathbf{x}^2)$ *a.s. coincides with* $(X_t(\omega))_{t\in[0,T]}$, *solution to the SDE in the Stratonovich sense.*

Furthermore, the Itô map $\mathbf{y} \mapsto \mathbf{x}$ *is continuous for the Hölder* ρ_q *distance (and locally Lipschitz in sense described in [16]).*
(*b*) *(See [9, 17]) If* $f \in \mathcal{C}^2\big([0,T] \times \mathbb{R}^d, \mathcal{M}((d+1) \times (d+1)\big)$ *is such that* $f.\nabla f$ *is bounded with an* α*-Hölder differential, then the conclusions of claim* (*a*)*still hold.*

3.2 Quantization of the SDE and main result

Let $(\alpha^N)_{N\geq 1}$ denote a sequence of quantizers of the Brownian motion. Each α^N is made up of N functions (or elementary quantizer) $\alpha_n^N : [0,T] \to \mathbb{R}^d$, $n = 1, \ldots, N$. For convenience a component "0" will be added accordingly to each elementary quantizer α_n^N by setting $\alpha_n^{N,0}(t) = t$ (which exactly quantizes the function $W_t^0 = t$). We assume that every elementary quantizer α_n^N is a continuous function with finite variation over $[0,T]$. The resulting Voronoi quantizer $\widehat{W} = \widehat{W}^{\alpha^N}$ of W reads

$$\widehat{W}_t = \sum_{n=1}^{N} \alpha_n^N(t)\mathbf{1}_{\{W\in C_n(\alpha^N)\}}, \quad t \in [0,T].$$

Our aim is to approximate the diffusion process $(x_t)_{t\in[0,T]}$ solution to the SDE (3.25) by the solution $\widetilde{X}^N$ of the equation

$$d\widetilde{X}_t^N = f\big(\widetilde{X}_t^N\big)d\widehat{W}_t,\ \widetilde{X}_0^N = x_0.$$

as $N \to \infty$. In fact, a less formal expression is available for the process $\widetilde{X}^N$, namely

$$\widetilde{X}^N = \sum_{n=1}^{N} \widetilde{x}_n^N \mathbf{1}_{\{W\in C_n(\alpha^N)\}}$$

where each x_n^N is solution to the ODE

$$d\widetilde{x}_n^N(t) = f(\widetilde{x}_n^N(t))\, d\,\alpha_n^N(t), \quad \widetilde{x}_n^N(0) = x_0, \quad n = 1, \ldots, N. \tag{3.28}$$

Note that X^N is a non-Voronoi quantization of (x_t) (at level N). The starting natural idea was to hope that X^N converges to (x_t) owing to the convergence of $\widehat{W}^N$ toward W... in an appropriate sense. Since we will use the above Theorem 3.1, we need to prove the convergence of the geometric functional $\widehat{\mathbf{W}}^N$ related to $\widehat{W}$ toward that of W. The quantity $\widehat{\mathbf{W}}^N$ is formally defined by mimicking the definition of $\mathbf{W}$, namely, for every $(s,t) \in [0,T]$, $0 \leq s < t \leq T$,

$$\widehat{\mathbf{W}}^{1,N}(\omega) := \widehat{W}_t(\omega) - \widehat{W}_s(\omega),\ \widehat{\mathbf{W}}_{s,t}^{2,N}(\omega) := \left(\int_s^t (\widehat{W}_u^i - \widehat{W}_s^i)d\widehat{W}_u^j\right)_{i,j=0,\ldots,d}(\omega)$$

still with the convention $\widehat{W}_t^{0,N} = t$. The integral must be understood in the usual Stieltjes sense.

Theorem 3.2. *Let $(\widehat{W}^N)_{N\geq 1}$ be a sequence of stationary quadratic functional quantizers of the Brownian motion converging to W in $L^2_{L^2_T}(\mathbb{P})$.*

Let f be like in claims (a) or (b) in Theorem 3.1. Consider for every $N \geq 1$, the solutions of the quantized ODE

$$d\widetilde{X}^N_t = f(\widetilde{X}^N_t)\, d\widehat{W}^N_t, \qquad N \geq 1.$$

as defined by (3.28). Let $\mathbf{X}$ and $\widetilde{\mathbf{X}}^N$ denote the enhanced paths of X, solution to (3.25), and $\widetilde{X}^N$ respectively. Then, for every $q \in (2, 2+\alpha)$,

$$\rho_q(\widetilde{\mathbf{X}}^N, \mathbf{X}) \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

Furthermore if $r > \frac{2}{3}$ then

$$\rho_q(\widetilde{\mathbf{X}}^{\lfloor e^{N^r} \rfloor}, \mathbf{X}) \stackrel{a.s.}{\longrightarrow} 0.$$

In view of what precedes this result is, as announced, a straightforward corollary of the continuity of the Itô map established Theorem 3.1, once the convergence $\rho_q(\widehat{\mathbf{W}}^N, \mathbf{W})$ in probability is established for any $q \in (2, 3)$. A slightly more derailed proof is proposed at the end of Section 5.

In fact we will prove a much precise statement concerning the Brownian motion since we will establish for every $q > 2$ the convergence in every $L^p(\mathbb{P})$, $0 < p < \infty$, of $\rho_q(\widehat{\mathbf{W}}^N, \mathbf{W})$ with an explicit $L^p(\mathbb{P})$-rate of convergence in the scale $(\log N)^{-\theta}$, $\theta \in (0, 1)$.

These rates can be transferred to the convergence of the quantized SDE, conditionally to some events on which the Itô map is itself Lipschitz continuous for the distances ρ_q. Several results of local Lipschitz continuity have been established recently, especially in [6], [9], [16], [17], although not completely satisfactory from a practical point of view. So we decided not to reproduce (and take advantage of) them here.

The proof is divided into two steps: the convergence for the Hölder semi-norm) of the regular path component is established in Section 4 (in which more general processes are considered) and the convergence of approximate Lévy areas in Section 5 (entirely devoted to the Brownian case for the sake of simplicity).

Remarks. • There is a small abuse of notation in the above Theorem since $\widetilde{X}^N$ is not a Voronoi quantizer of X: this quantization of X is defined on the Voronoi partition (for the $L^2_{T,\mathbb{R}^d}$-norm) induced by the quantization of the Brownian motion W.

• The same results holds for the Brownian bridge, the Ornstein-Uhlenbeck process and more generally for continuous Gaussian semi-martingales that satisfy the Kolmogorov criterion.

4 Convergence of the paths of processes in Hölder semi-norm

4.1 A general setting including stationary functional quantization

In this section we investigate the connections between the celebrated Kolmogorov criterion and the tightness of some classes of sequences of processes for the topology of $\frac{1}{q}$-Hölder convergence. In fact this connection is somehow the first step of the rough path theory, but we will look at it in a slightly different way. Whatsoever this naive pathwise convergence is not sufficient to get the continuity of the Itô map in a Brownian framework and we will also have to deal for our purpose with the multiplicative functional (see Section 5).

But at this stage we aim at showing that when a sequence $(Y^N)_{N\geq 1}$ satisfies some "stationarity property" with respect to a process Y, several properties of Y can be transferred to the Y^N. Indeed, the same phenomenon will occur for the multiplicative function (see the next section).

If Y satisfies the Kolmogorov criterion and $(\mathcal{G}_N)_{N\geq 1}$ denotes a sequence of sub-σ-fields of $\mathcal{A}$, then a sequence of processes defined by

$$Y^N := \mathbb{E}(Y \,|\, \mathcal{G}^N), \quad N \geq 1,$$

where the conditional expectation is considered in the functional sense (see Appendix) is (C-tight and) tight for a whole family of topologies induced by convergence in $\frac{1}{q}$-Hölder sense.

Definition 4.2. *Let $p \geq 1$, $\theta > 0$. A process $Y = (Y_t)_{t\in[0,T]}$ satisfies the Kolmogorov criterion $(K_{p,\theta})$ if there is a real constant $C_T^{Kol} > 0$ such that*

$$\forall\, s,\, t \in [0,T], \qquad \mathbb{E}|Y_t - Y_s|^p \leq C_T^{Kol}\, |t-s|^{1+\theta} \qquad \textit{and} \qquad Y_0 \in L^p(\mathbb{P}).$$

Theorem 4.3. *Let $Y := (Y_t)_{t\in[0,T]}$ be a pathwise continuous process defined on $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the Kolmogorov criterion $(K_{p,\theta})$. Let $(\mathcal{G}_N)_{N\geq 1}$ be a sequence of sub-σ-fields of $\mathcal{A}$. For every $N \geq 1$ set*

$$Y^N := \mathbb{E}(Y \,|\, \mathcal{G}_N).$$

For every $N \geq 1$, Y^N has a pathwise continuous version satisfying

$$\forall\, t \in [0, T], \qquad Y_t^N = \mathbb{E}(Y_t \,|\, \mathcal{G}_{N}) \quad a.s.$$

Furthermore, if one of the following conditions is satisfied
(a) $\mathcal{G}_N \subset \mathcal{G}_{N+1}$,
(b) There exists an everywhere dense subset $D \subset [0, T]$ such that

$$\forall\, t \in [0, T], \quad Y_t^N \xrightarrow{\mathbb{P}} Y_t.$$

(c) $|Y^N - Y|_{L^r_T} \xrightarrow{\mathbb{P}} 0$ for some $r \geq 1$,
then

$$\forall\, q > \frac{1}{\theta}, \qquad \forall\, p \in [1, q\theta), \qquad \|Y - Y^N\|_{\sup} + \|Y - Y^N\|_{q,Hol} \xrightarrow{L^p} 0.$$

The proof of the theorem is a variant of the proof of the Kolmogorov criterion for functional tightness of processes. It consists in a string of several lemmas. For the following classical lemma, we refer to [14] (where it is stated and proved for semi-norms in p-variation).

Lemma 4.1. *Let $x, y \in \mathcal{C}([0, T], \mathbb{R}^d)$ and let $q \geq 1$. Then*
(a) $\|x - x(0)\|_{\sup} \leq \|x\|_{q,Hol}$.
(b) $\|x + y\|_{q,Hol} \leq \|x\|_{p,Hol} + \|y\|_{q,Hol}$ if $q \geq 1$,
(c) For every $q > q' \geq 1$, $\|x\|^q_{q,Hol} \leq (2\|x\|_{\sup})^{q-q'} \|x\|^{q'}_{q',Hol}$.
(d) Claims (a)-(b)-(c) remain true with the p-variation semi-norm $\mathrm{Var}_{q,[0,T]}$ instead of the $\frac{1}{q}$-Hölder semi-norm.

Lemma 4.2. *Let $p \in [1, \infty)$. If Y satisfies the Kolmogorov criterion $(K_{p,\theta})$ then, for every $N \geq 1$, the process Y^N defined by $Y_t^N = \mathbb{E}(Y_t \,|\, \mathcal{G}_{N})$ has a continuous modification which is $\frac{\theta'}{p}$-Hölder continuous for every $\theta' \in (0, \theta)$*

(i.e. $\|Y^N\|_{\frac{p}{\theta'},Hol} < +\infty$ a.s.). Furthermore, the sequence $(Y^N)_{N\geq 1}$ is C-tight and for every $\theta' \in (0,\theta)$, there exists a random variable $Z_{\theta'} \in L^p_{\mathbb{R}}(\mathbb{P})$ such that

$$\mathbb{P}(d\omega)\text{-}a.s.\ \|Y(\omega)\|_{\frac{p}{\theta'},Hol} \leq Z_{\theta'} \tag{4.29}$$

and

$$\forall\, N \geq 1,\ \|Y^N(\omega)\|_{\frac{p}{\theta'},Hol} \leq \mathbb{E}(Z_{\theta'}\,|\,\mathcal{G}_N)(\omega). \tag{4.30}$$

In particular, the sequence of Hölder semi-norms $(\|Y^N\|_{\frac{p}{\theta'},Hol})_{N\geq 1}$ is L^p-uniformly integrable.

Remark. As a by-product of the proof we also get that

$$\mathbb{E}(Z^p_{\theta'}) \leq C_{p,T,\theta,\theta'} C^{Kol}_T$$

where $C_{T,p,\theta,\theta'}$ is a finite real constant that only depends upon p, T, θ and θ' (and not on Y or the σ-fields $\mathcal{G}_{N}$).

Proof. First it follows form the Kolmogorov criterion that for every $N \geq 1$, Y^N admits a continuous modification which is $\frac{\theta'}{p}$-Hölder for every $\theta' \in (0,\theta)$. Moreover the sequence $(Y^N)_{N\geq 1}$ is C-tight since every Y^N satisfies the same Kolmogorov criterion $(K_{p,\theta})$ and $Y_0^N = \mathbb{E}(Y_0|\mathcal{G}_{N})$ is tight on $\mathbb{R}$ (see [1], [29] p.26). Now, let $s,\, t \in [0,T]$, let $m,\, n \geq 1$ be two fixed integers. First note that

$$\sup_{s,t\in[0,T],\, t\leq s\leq t+\frac{T}{2^n}} |Y_t - Y_s| \quad\leq\quad 2\sum_{m\geq 0} \max_{0\leq k\leq 2^{n+m}-1} |Y_{t^{n+m}_{k+1}} - Y_{t^{n+m}_k}| \tag{4.31}$$

and

$$\max_{0\leq k\leq 2^{n+m}-1} |Y_{t^{n+m}_{k+1}} - Y_{t^{n+m}_k}|^p \leq \sum_{k=0}^{2^{n+m}-1} |Y_{t^{n+m}_{k+1}} - Y_{t^{n+m}_k}|^p.$$

For every $\theta' \in (0,\theta)$, set

$$Z_{\theta'} := \frac{2}{T}\left(\sum_{n\geq 0} 2^{n\frac{\theta'}{p}} \sup_{s,t\in[0,T],\, t\leq s\leq t+\frac{T}{2^n}} |Y_t - Y_s|\right). \tag{4.32}$$

Taking the L^p-norm in (4.31) yields

$$\begin{aligned}\|Z_{\theta'}\|_p &\le \left(\frac{2}{T}\right)^{\frac{\theta'}{p}} \sum_{n\ge 0} 2^{n\frac{\theta'}{p}} \| \sup_{s,t\in[0,T],\, t\le s\le t+\frac{T}{2^n}} |Y_t - Y_s| \|_p \\ &\le 2\left(\frac{2}{T}\right)^{\frac{\theta'}{p}} \sum_{n\ge 0} 2^{n\frac{\theta'}{p}} \sum_{\ell\ge 0} \| \max_{0\le k\le 2^{n+m}-1} |Y_{t^{n+m}_{k+1}} - Y_{t^{n+m}_k}| \|_p.\end{aligned}$$

On the other hand, owing to the the Kolmogorov criterion $(K_{p,\theta})$,

$$\begin{aligned}\mathbb{E} \max_{0\le k\le 2^{n+m}-1} |Y_{t^{n+m}_{k+1}} - Y_{t^{n+m}_k}|^p &\le \sum_{k=0}^{2^{n+m}-1} \mathbb{E}|Y_{t^{n+m}_{k+1}} - Y_{t^{n+m}_k}|^p \\ &\le 2^{n+m} C_T^{Kol} 2^{-(n+m)(1+\theta)} T^{-(1+\theta)} \\ &= C_T^{Kol} T^{-(1+\theta)} 2^{-(n+m)\theta}.\end{aligned}$$

Hence

$$\mathbb{E}\, Z^p_{\theta'} \le C_T^{Kol} C_{p,T,\theta,\theta'} \left(\sum_{n\ge 0}\sum_{m\ge 0} 2^{n\frac{\theta'-\theta}{p}} 2^{-m\frac{\theta}{p}}\right)^p < +\infty$$

where the finite real constant $C_{p,T,\theta,\theta'}$ only depends on p, T, θ and θ'. On the other hand, for every $\delta \in [0,T]$, there exists a integer $n_\delta \ge 1$ such that $2^{-(1+n_\delta)} \le \delta/T \le 2^{-n_\delta}$. Hence,

$$\delta^{-\theta'} \sup_{s,t\in[0,T],\, t\le s\le t+\delta} |Y_t - Y_s|^p \le 2^{(1+n_\delta)\theta'} T^{-\theta'} \times \sup_{s,t\in[0,T],\, t\le s\le t+\frac{T}{2^n}} |Y_t - Y_s|^p \le Z^p_{\theta'}.$$

Consequently, for every s, $t \in [0,T]$, and every $\omega \in \Omega$,

$$|Y_t(\omega) - Y_s(\omega)| \le Z_{\theta'}(\omega)|t-s|^{\frac{\theta'}{p}}$$

i.e.

$$\|Y(\omega)\|_{\frac{p}{\theta'},Hol} \le Z_{\theta'}(\omega).$$

Finally, it follows from Jensen's Inequality that for every s, $t \in \mathbb{Q}\cap[0,T]$,

$$\mathbb{P}(d\omega)\text{-}a.s. \qquad |Y^N_t(\omega) - Y^N_s(\omega)| \le \mathbb{E}(Z_{\theta'}\,|\,\mathcal{G}_N)(\omega)|t-s|^{\theta'}.$$

In particular this means that, for every $p \geq 1$ and every $\theta' \in (0, \theta)$,

$$\mathbb{P}(d\omega)\text{-}a.s. \qquad \|Y^N(\omega)\|_{\frac{p}{\theta'},Hol} \leq \mathbb{E}(Z_{\theta'} \,|\, \mathcal{G}_N)(\omega) < +\infty$$

and satisfies the L^p-uniform integrability assumption. $\diamond$

Proof of Theorem 4.3. The sequence $(Y^N)_{N\geq 1}$ being C-tight on $(\mathcal{C}([0,T], \mathbb{R}^d), \|\,.\,\|_{\sup})$, so is the case of the the sequence $(Y^N, Y)_{N\geq 1}$ on $(\mathcal{C}([0,T], \mathbb{R}^{2d}), \|\,.\,\|_{\sup})$ since the product topology coincides with the uniform topology. Let $\mathbb{Q} = w\text{-}\lim_N \mathbb{P}_{(Y^{N'},Y)}$ denote a weak functional limiting value of $(Y^N, Y)_{N\geq 1}$. If $\Xi = (\Xi^1, \Xi^2)$ denotes the canonical process on $(\mathcal{C}([0,T], \mathbb{R}^{2d}), \|\,.\,\|_{\sup})$, it is clear that $\mathbb{Q}_{\Xi^2} = \mathbb{P}_Y$.

$\triangleright$ *Convergence of the sup-norm.* Assume that (c) holds: the functional $y \mapsto |y^1(t) - y^2(t)|_{L^r_T}$ is continuous on $(\mathcal{C}([0,T], \mathbb{R}^{2d}), \|\,.\,\|_{\sup})$, consequently, $|\Xi^1 - \Xi^2|_{L^r_T} = 0$ $\mathbb{Q}$-$a.s.$ $i.e.$ $\mathbb{Q} = \mathbb{P}_{(Y,Y)}$ so that $(Y^N, Y) \stackrel{\mathcal{L}(\|\,.\,\|_{\sup})}{\longrightarrow} (Y, Y)$ as $N \to \infty$ which simply means that $\|Y^N - Y\|_{\sup} \stackrel{\mathbb{P}}{\longrightarrow} 0$. On the other hand,it follows from Lemma 4.2 that, for every $N \geq 1$,

$$\|Y^N - Y\|^p_{\sup} \leq C_{p,T}\left(\mathbb{E}(Z^p_{\theta'} \,|\, \mathcal{G}_N) + Z^p_{\theta'}\right)\ a.s.$$

(for a given fixed $\theta' \in (0, \theta)$) which implies that $(\|Y^N - Y\|^p_{\sup})_{N\geq 1}$ is uniformly integrable. Finally,

$$\mathbb{E}\,\|Y^N - Y\|^p_{\sup} \longrightarrow 0 \qquad \text{as} \qquad N \to \infty.$$

Assume that (b) holds: it follows that, for every $t_1, \ldots, t_k \in D$, one has $(Y^N_{t_1}, \ldots, Y^N_{t_k}) \stackrel{\mathbb{P}}{\longrightarrow} (Y_{t_1}, \ldots, Y_{t_k})$, which in turn implies that the convergence $(Y^N_{t_1}, \ldots, Y^N_{t_k}, Y_{t_1}, \ldots, Y_{t_k}) \stackrel{\mathcal{L}}{\longrightarrow} (Y_{t_1}, \ldots, Y_{t_k}, Y_{t_1}, \ldots, Y_{t_k})$. This means that $\mathbb{Q}$ and $\mathbb{P}_{(Y,Y)}$ have the same finite dimensional marginals $i.e.$ $\mathbb{Q} = \mathbb{P}_{(Y,Y)}$. One concludes like in (c).

If (a) holds, for every $t \in [0,T]$, $Y^N_t \to Y_t$ $\mathbb{P}$-$a.s.$, so that (b) is satisfied.

$\triangleright$ *Convergence of the Hölder semi-norm.* Let $q \geq 1$. As concerns the convergence of the $\frac{1}{q}$-Hölder semi-norm, one proceeds as follows. Let $q' \in (\frac{p}{\theta}, q)$ and set $\theta' := \frac{p}{q'} \in (0, \theta)$. It follows from Lemma 4.1(b)-(c) that

$$\|Y - Y^N\|_{q,Hol} \;\;\leq\;\; 2^{1-\frac{q'}{q}} \|Y - Y^N\|_{\sup}^{1-\frac{q'}{q}} \times \left(\|Y\|_{q',Hol} + \|Y^N\|_{q',Hol}\right)^{\frac{q'}{q}}.$$

Now let $Z := Z_{\theta'}$ be defined by (4.32). Then,

$$\|Y\|_{q',Hol} + \|Y^N\|_{q',Hol} \le Z + \big(\mathbb{E}(Z\,|\,\mathcal{G}_{_N})\big).$$

Hence, the sequence $(\|Y\|_{q',Hol} + \|Y^N\|_{q',Hol})_{N\ge1}$, is tight since it is L^p-bounded. On the other hand, $\|Y-Y^N\|_{\sup} \stackrel{L^p}{\longrightarrow} 0$ so that $\|Y-Y^N\|_{q,Hol} \stackrel{\mathbb{P}}{\longrightarrow} 0$ as $N\to\infty$.

Now let $\widetilde\theta = \frac{p}{q} \in (0,\theta)$. The same argument as above shows that $\|Y - Y^N\|_{q,Hol} \le \widetilde Z + \mathbb{E}(\widetilde Z\,|\,\mathcal{G}_{_N})$ where $\widetilde Z = Z_{\widetilde\theta}$ is still given by (4.32). As a consequence, $(\|Y-Y^N\|^p_{q,Hol})_{N\ge1}$ is uniformly integrable since, for every $N\ge1$, Jensen's Inequality implies

$$\|Y-Y^N\|^p_{q,Hol} \le 2^{p-1}\big(\widetilde Z^p + \mathbb{E}(\widetilde Z^p\,|\,\mathcal{G}_{_N})\big)$$

which finally implies that $\|Y-Y^N\|_{q,Hol} \stackrel{L^p}{\longrightarrow} 0.$ $\diamondsuit$

4.2 Application to stationary quantizations of Brownian motion: convergence and rates

Theorem 4.4. *(a) Let $(\widehat W^N)_{N\ge1}$ be a sequence of stationary quadratic functional quantizers of a standard d-dimensional Brownian motion W defined by (2.11) or (2.17) converging to W in a (purely) quadratic sense, namely $\|\,|W-\widehat W^N|_{L^2_T}\|_2 \to 0$ as $N\to\infty$. Then, for every $q>2$,*

$$\forall\, p\in(0,\infty),\qquad \|W-\widehat W^N\|_{q,Hol} \stackrel{L^p}{\longrightarrow} 0 \quad as \quad N\to\infty.$$

(b) Let $q>2$. If, for every $N\ge1$, $\widehat W^N$ is an optimal product quantization *at level N. Then, for every $p\in(0,\infty)$,*

$$\Big\|\|W-\widehat W^N\|_{q,Hol}\Big\|_p = o\Big((\log N)^{-\frac32\min\left(\frac15(1-\frac2q),\frac1p\right)+\alpha}\Big),\quad \forall\,\alpha>0.$$

The proof of this Theorem is a consequence of the above Theorem 4.3. So we need to get accurate estimates for the increments of the processes $W-\widehat W^N$. This is the aim of the following lemma.

Lemma 4.3. *Let $p\in[2,+\infty)$. Let $\widehat W^N$, $N\ge1$, denote a sequence of optimal* product *quadratic quantizers. For every $\rho\in(0,\frac12)$ and every $\varepsilon\in(0,3)$, for every $s,\,t\in[0,T]$, $s\le t$,*

$$\Big\|(W_t-W_s)-(\widehat W^N_t-\widehat W^N_s)\Big\|_p \le C_{\rho,p,T,d,\varepsilon}|t-s|^\rho(\log N)^{-(\frac12-\rho)\wedge(\frac{3-\varepsilon}{2p})}. \tag{4.33}$$

In particular, if $p \in (2,3)$*, then*

$$\Big\|(W_t - W_s) - (\widehat{W}_t^N - \widehat{W}_s^N)\Big\|_p \le C_{\rho,p,T,d}|t-s|^\rho (\log N)^{-(\frac{1}{2}-\rho)}. \tag{4.34}$$

Proof. We may assume without loss of generality that we deal with a one-dimensional Brownian motion W, quantized at level $N' = \lfloor N^{\frac{1}{d}} \rfloor$ since everything is done component by component. Set for every $k \ge 1$, $\widetilde{\xi}_k := \xi_k - \widehat{\xi}_k^{N_k}$ where $N_1, \dots, N_k, \dots$ denotes the optimal bit allocation of an optimal product quadratic quantization at level N'. Keep in mind that for every $k > L_W(N')$, $N_k = 1$ and that of course $N_1 \cdots N_{L_W(N')} \le N'$. The random vectors $(\widetilde{\xi}_k)_{k \ge 1}$ are independent and centered.

It follows from the K-L expansion of W and its product quantization that

$$(W_t - W_s) - (\widehat{W}_t^{N'} - \widehat{W}_s^{N'}) = \sum_{k \ge 1} \lambda_k \widetilde{\xi}_k \big(e_k^W(t) - e_k^W(s)\big).$$

Then, it follows from the B.D.G. Inequality for discrete time martingales that

$$\begin{aligned}
\Big\|(W_t - W_s) - (\widehat{W}_t^{N'} - \widehat{W}_s^{N'})\Big\|_p &\le C_{p,T} \left\| \sum_{k \ge 1} \lambda_k \widetilde{\xi}_k (e_k^W(t) - e_k^W(s))^2 \right\|_{\frac{p}{2}}^{\frac{1}{2}} \\
&\le C_{p,T} \left(\sum_{k \ge 1} \lambda_k^{1-\rho} \|\widetilde{\xi}_k\|_p^2 \right)^{\frac{1}{2}} |t-s|^\rho
\end{aligned}$$

since, for every $k \ge 1$,

$$(e_k^W(t) - e_k^W(s))^2 = \frac{8}{T} \sin^2\left(\frac{t-s}{\sqrt{\lambda_k}}\right) \cos^2\left(\frac{t-s}{\sqrt{\lambda_k}}\right) \le \frac{8}{T} |t-s|^{2\rho} \lambda_k^{-\rho}.$$

The random variables $\widehat{\xi}_k^{N_k}$ being an optimal quadratic quantization of the one-dimensional normal distribution for every $k \in \{1, \dots, L_W(N')\}$, it follows from (2.16) that, there exists for every $\varepsilon \in (0,3)$, a constant $\kappa_{p,\varepsilon}$ such that

$$\forall\, m \ge 1, \qquad \|\widetilde{\xi}_k\|_p = \|\xi - \widehat{\xi}_k^{N_k}\|_p \le \kappa_{p,\varepsilon} \frac{1}{N_k^{1 \wedge \frac{3-\varepsilon}{p}}}$$

where $\widehat{\xi}^m$ denotes the (unique) optimal quadratic quantization at level m of a normally distributed scalar random variable ξ. As a consequence,

$$\left\|(W_t - W_s) - (\widehat{W}_t^{N'} - \widehat{W}_s^{N'})\right\|_p \le C_{p,T,\varepsilon}|t-s|^\rho \left(\sum_{k\ge 1} \lambda_k^{1-\rho} \frac{1}{N_k^{2(1\wedge\frac{3-\varepsilon}{p})}}\right)^{\frac{1}{2}}.$$

$\triangleright$ Temporarily assume that $p \in [2,3)$. One may choose ε so that $1\wedge\frac{3-\varepsilon}{p} = 1$. Now, keeping in mind that $L' := L_{_W}(N') \sim \log N'$ and $\lambda_k \le c\,k^{-2}$ for a real constant $c > 0$, one gets

$$\begin{aligned}\sum_k \lambda_n^{1-\rho}\frac{1}{N_k^2} &\le \lambda_{L'}^\rho \sum_{k=1}^{L'} \frac{\lambda_k}{N_k^2} + \sum_{k>L'} \lambda_k^{1-\rho}\\ &\le C_\rho \left((\log N')^{2\rho} \sum_{k=1}^{L'} \frac{\lambda_k}{N_k^2} + (\log N)^{2\rho-1}\right).\end{aligned}$$

Now, following *e.g.* [18], we know that the optimal bit allocation yields

$$\sum_{k=1}^{L'} \frac{\lambda_k}{N_k^2} \le \frac{C}{T}(\log N')^{-1}$$

so that, finally

$$\left\|(W_t - W_s) - (\widehat{W}_t^{N'} - \widehat{W}_s^{N'})\right\|_p \le C_{\rho,p,T}|t-s|^\rho (\log N')^{\rho-\frac{1}{2}}.$$

$\triangleright$ Assume now that $p \in [3,+\infty)$ and $\varepsilon \in (0,3)$. Set $\tilde{p} = \frac{p}{3-\varepsilon} > 1$ and $\tilde{q}$ its conjugate exponent. Then, Hölder Inequality implies

$$\sum_{k=1}^{L'} \frac{\lambda_k^{1-\rho}}{N_k^{\frac{2}{\tilde{p}}}} \le \left(\sum_{k=1}^{L'} \frac{\lambda_k}{N_k^2}\right)^{\frac{1}{\tilde{p}}} \left(\sum_{k=1}^{L'} \lambda_k^{1-\frac{\rho p}{p-3+\varepsilon}}\right)^{\frac{1}{\tilde{q}}}.$$

We inspect now three possibles cases for ρ.

• If $0 < \rho < \frac{1}{2}(1 - \frac{3-\varepsilon}{p})$, then $1 - \frac{\rho p}{p-3+\varepsilon} > \frac{1}{2}$ so that $\sum_{k\ge 1} \lambda_k^{1-\frac{\rho p}{p-3+\varepsilon}} < +\infty$, which in turn implies that

$$\sum_{k=1}^{L'} \frac{\lambda_k^{1-\rho}}{N_k^{\frac{2}{\tilde{p}}}} \le C_{\rho,p,T}\Big(\log N'\Big)^{-\frac{3-\varepsilon}{p}}.$$

Furthermore $1-\frac{\rho}{2} > \frac{3-\varepsilon}{p}$.

• If $\frac{1}{2}(1-\frac{3-\varepsilon}{p}) < \rho < \frac{1}{2}$, then, $1-\frac{\rho}{2} < \frac{3-\varepsilon}{p}$ and $1-\frac{\rho p}{p-3+\varepsilon} = \frac{1}{2}$ so that $\sum_{k\ge 1}\lambda_k^{1-\frac{\rho p}{p-3+\varepsilon}} < +\infty$

$$\begin{aligned}\sum_{k=1}^{L'}\frac{\lambda_k^{1-\rho}}{N_k^{\frac{2}{p}}} &\le C_{\rho,p,T}\Big(\log N'\Big)^{-\frac{3-\varepsilon}{p}}\times\Big(L_W(N')^{\frac{2\rho p}{p-3+\varepsilon}-1}\Big)^{1-\frac{3-\varepsilon}{p}}\\ &= C_{\rho,p,T}\Big(\log N'\Big)^{2\rho-1}.\end{aligned}$$

• If $\frac{1}{2}(1-\frac{3-\varepsilon}{p}) = \rho < \frac{1}{2}$, then $1-\frac{\rho}{2} = \frac{3-\varepsilon}{p}$ and $1-\frac{\rho p}{p-3+\varepsilon} = \frac{1}{2}$ so that $\sum_{k=1}^{L'}\lambda_k^{1-\frac{\rho p}{p-3+\varepsilon}} \le C_{\rho,p,T}\log\log N'$ (keep in mind $L' = L_W(N') \sim \log N'$). Hence, for every $\varepsilon' \in (0,\varepsilon)$,

$$\sum_{k=1}^{L'}\frac{\lambda_k^{1-\rho}}{N_k^{\frac{2}{p}}} = o\Big((\log N')^{-\frac{3-\varepsilon'}{p}}\Big).$$

As conclusion, we get that

$$\begin{aligned}\Big(\sum_k \lambda_k^{1-\rho}\|\widetilde{\xi}_k\|_p^2\Big)^{\frac{1}{2}} &\le \Big(\sum_k \lambda_k^{1-\rho}\frac{1}{N_k^{2(1\wedge\frac{3-\varepsilon}{p})}}\Big)^{\frac{1}{2}}\\ &= O\Big((\log N')^{-(\frac{1}{2}-\rho)\wedge(\frac{3-\varepsilon}{2p})}\Big) \qquad (4.35)\end{aligned}$$

which completes the proof since $\log(1+N') > \frac{1}{d}\log N$ (which implies $\log N' > \frac{1}{d}\log(N/2)$). $\diamond$

Proof of Theorem 4.4. (a) Owing to the monotonicity of the L^p-norms, it is enough to show that, the announced convergence holds for every $q > 2$ and every $p > \frac{2q}{q-2}$ or equivalently for every $p > 2$ and every $q > \frac{2p}{p-2}$. This statement follows for the $\frac{1}{q}$-Hölder (semi-)norm follows from Theorem 4.3(c). Indeed W satisfies the Kolmogorov $K_{p,\theta}$ with $\theta = p/2-1$. On the other hand, it follows from [13] that, for any sequence of (Voronoi) quantizations $\widehat{W}^N$ at level N converging in $L^2_{L^2_T}(\mathbb{P})$ toward W, this convergence also holds in the $a.s.$ sense. So Criterion(c) is fulfilled.

(b) Let $q > 2$. The process $W - \widehat{W}^N$ satisfies $K_{p,\rho p-1}$ for every $\rho \in (\frac{1}{p}, \frac{1}{2})$ with "Kolmogorov constants"

$$C^{Kol}_{T,p} = C_{p,T,\rho,d,\varepsilon}(\log N)^{-p[(\frac{1}{2}-\rho)\wedge(\frac{3-\varepsilon}{2p})]},\ \varepsilon \in (0,3).$$

We wish to apply Lemma 4.2 (and the remark that follows).

$\triangleright$ Assume $0 < p < \frac{5q}{q-2}$. Then there exists $\eta > 0$ such that $p < p' = \frac{5q}{q-2+\eta}$. Set $\theta' = \frac{p'}{q}$. One cheks that $\frac{1}{p'} + \frac{1}{q} < \frac{1}{2}$ so that there exists $\eta' > 0$ such that $\rho = \frac{1}{p'} + \frac{1}{q} + \eta' < \frac{1}{2}$. Elementary computations show that $\frac{1}{2} - \rho < \frac{3}{2p}$. Let $\varepsilon \in (0,3)$ such that $\frac{1}{2} - \rho < \frac{3}{2p} - \varepsilon$. Consequently, Lemma 4.2 (and the remark that follows) imply that

$$\Big\| \|W - \widehat{W}^N\|_{q,Hol} \Big\|_{p'} \le C_{q,\eta,\eta',T,\varepsilon}(\log N)^{-(\frac{1}{2}-\rho)}$$

and for any small enough $\alpha > 0$, one my specify η, η' and ε so that $\frac{1}{2} - \rho = \frac{3}{10}(1 - \frac{2}{q}) - \alpha$. Finally this bounds holds true for $p \in (0, p')$ since the the L^p-norm is non-decreasing.

$\triangleright$ Now, if $p \ge \frac{5q}{q-2}$, one checks that $\frac{3}{2p} \ge \frac{1}{2} - \left(\frac{1}{p} + \frac{1}{q}\right)$. It becomes impossible to specify $\rho \in (0, \frac{1}{2})$ so that $\theta' = \frac{p}{q} < \theta = \rho p - 1$ and $1 - \rho > \frac{3}{2p}$. So the same specifications as above lead to

$$\Big\| \|W - \widehat{W}^N\|_{q,Hol} \Big\|_{p'} \le C_{q,\eta,\eta',\varepsilon,T}(\log N)^{-\frac{3-\varepsilon}{2p}}$$

which yields the announced result. $\diamondsuit$

5 Convergence of stationary quantizations of the Brownian motion for the ρ_q-Hölder distance.

In view of what will be needed to apply this theorem to the Brownian motion and its functional quantizations, we need to prove a counterpart of Lemmas 4.2 and 4.3 for $\mathbf{W}^2_{s,t}$. However, for the sake of simplicity, by contrast with the previous section, we will only deal with the case of the Brownian motion and its stationary quantizations.

The main result of this section is the following Theorem.

Theorem 5.5. *Let $q > 2$.*

(a) *Let* $(\widehat{W}^N)_{N\geq 1}$ *be a sequence of stationary quadratic functional quantizers of a standard d-dimensional Brownian motion* W *defined by (2.11) or (2.17) converging to* W *in a (purely) quadratic sense, namely* $\|\,|W-\widehat{W}^N|_{L^2_T}\|_2 \to 0$ *as* $N\to\infty$. *Then,*

$$\forall\, q>2,\quad \forall\, p>0,\qquad \Big\|\rho_q(\mathbf{W},\widehat{\mathbf{W}}^N)\Big\|_p \longrightarrow 0 \quad \text{as } N\to\infty.$$

(b) *Let* $q>2$. *Assume that, for every* $N\geq 1$, $\widehat{W}^N$ *is an optimal* product quantization *at level* N *of* W. *Then, for every* $q>2$ *and every* $p>0$,

$$\Big\|\|\mathbf{W}^2-\widehat{\mathbf{W}}^{2,N}\|_{\frac{q}{2},Hol}\Big\|_p \;=\; o\Big((\log N)^{-\frac{3}{2}\min\left(\frac{2}{7}(1-\frac{2}{q}),\frac{1}{p}\right)+\alpha}\Big),\ \forall\,\alpha>0,$$

so that, finally,

$$\Big\|\rho_q(\mathbf{W},\widehat{\mathbf{W}}^N)\Big\|_p = o\Big((\log N)^{-\frac{3}{2}\min\left(\frac{1}{7}(1-\frac{2}{q}),\frac{1}{p}\right)+\alpha}\Big),\ \forall\,\alpha>0.$$

(c) *If* $r>\frac{2}{3}$, *then*

$$\rho_q(\mathbf{W},\widehat{\mathbf{W}}^{\lfloor e^{N^r}\rfloor}) = o\Big(N^{-(\frac{3}{2}r-1)\frac{q-2}{7q}+\alpha}\Big)\ \forall\,\alpha>0,\ \mathbb{P}\text{-}a.s.$$

Note that the result of interest for our purpose (convergence on multi-dimensional stochastic integrals) corresponds to $q\in(2,3)$. The proposition below appears as the counterpart of Lemma 4.2 on the way to the proof.

Proposition 5.3. *Let* $p>2$.
(a) *Let* $\mathbf{W}^2_{s,t}$ *be defined by (3.27). For every* $\tilde\theta'\in(0,p-1)$, *there exists a random variable* $Z^{(2)}_{\tilde\theta'}\in L^p$ *such that*

$$\mathbb{P}\text{-}a.s.\quad \forall\, s,\,t\in[0,T],\qquad |\mathbf{W}^2_{s,t}|\leq Z^{(2)}_{\tilde\theta'}|t-s|^{\frac{\tilde\theta'}{p}}.$$

(b) *Let*

$$\widehat{\mathbf{W}}^{2,N}_{s,t}(\omega) = \left(\int_s^t(\widehat{W}^i_u-\widehat{W}^i_s)d\widehat{W}^j_u\right)_{i,j=0,\ldots,d}(\omega),\qquad s,t\in[0,T],\ s\leq t,$$

where $\widehat{W}=\widehat{W}^N$ *is a stationary quantization of* W *(the integration holds in the Stieltjes sense). Then, for every* $p>2$ *and every* $\tilde\theta'\in(0,p-1)$,

$$\mathbb{P}\text{-}a.s.\quad \forall\, s,\,t\in[0,T],\qquad |\widehat{\mathbf{W}}^{2,N}_{s,t}|\leq \mathbb{E}(Z^{(2)}_{p,\tilde\theta'}\,|\,\mathcal{G}_N)|t-s|^{\frac{\tilde\theta'}{p}}.$$

(c) *Let* $\widetilde{\mathbf{W}}^{2,N}_{s,t} = \mathbf{W}^2_{s,t} - \widehat{\mathbf{W}}^{2,N}_{s,t}$ *where* $\widehat{W} = \widehat{W}^N$ *is now an* optimal quadratic product quantization *of* W *at level* N*. Then, if* $p > \frac{1}{\rho}$*, for every* $\tilde{\theta}' \in (0, p(\rho + \frac{1}{2}) - 2)$*, for every* $\varepsilon \in (0,3)$ *and every* $\delta > 0$*, there exists a real constant* $C_{\rho,p,T,d,\varepsilon,\delta} > 0$ *such that*

$$\left\| \sup_{s,t\in[0,T]} \frac{|\widetilde{\mathbf{W}}^{2,N}_{s,t}|}{|t-s|^{\frac{\tilde{\theta}'}{p}}} \right\|_p \leq C_{\rho,p,T,d,\varepsilon,\delta}\big(\log N\big)^{-(\frac{1}{2}-\rho)\wedge\frac{3-\varepsilon}{2(p+\delta)}}.$$

Proof. (a) The random variable $Z^{(2)}_{\tilde{\theta}'}$ of interest is defined by

$$Z^{(2)}_{\tilde{\theta}'} := \frac{2}{T}\sum_{n\geq 0} 2^{n\frac{\tilde{\theta}'}{p}} \sup_{s\leq t\leq s+\frac{T}{2^n}} |\mathbf{W}^2_{s,t}|.$$

Let $s,\ t \in [0,T],\ s \leq t \leq s + \frac{T}{2^n}$. We know from the multiplicative tensor property that, for every $u \in [s,t]$,

$$\mathbf{W}^2_{s,t} = \mathbf{W}^2_{s,u} + \mathbf{W}^2_{u,t} + W_{s,u} \otimes W_{u,t}$$

and that, for every $i,\ j \in \{0,\ldots,d\}$,

$$|W^i_{s,u} \otimes W^j_{u,t}| \leq \frac{1}{2}(|W^i_{s,u}|^2 + |W^j_{u,t}|^2).$$

To evaluate $\sup_{t\in[s,s+\frac{T}{2^n}]}|\mathbf{W}^{2,}_{s,t}|$, we may restrict to dyadic numbers owing to the continuity in (s,t) of $\mathbf{W}^2_{s,t}$. As a consequence, we have, still following the classical scheme of Kolmogorov criterion

$$\begin{aligned}\sup_{t\in[s,s+\frac{T}{2^n}]} |\mathbf{W}^2_{s,t}| \quad &\leq \quad 2\sum_{m\geq 0} \max_{0\leq k\leq 2^{n+m}-1} |\mathbf{W}^2_{t^{n+m}_k, t^{n+m}_{k+1}}| \\ &\quad + \max_{0\leq k\leq 2^{n+m}-1} |W_{t^{n+m}_k, t^{n+m}_{k+1}}|^2.\end{aligned}$$

Now

$$\mathbb{E} \max_{0\leq k\leq 2^{n+m}-1} |\mathbf{W}^2_{t^{n+m}_k, t^{n+m}_{k+1}}|^p \leq \sum_{\ell=0}^{2^{m+n}-1} \mathbb{E}\,|\mathbf{W}^2_{t^{n+m}_\ell, t^{n+m}_{\ell+1}}|^p$$

and

$$\mathbb{E} \max_{0\leq k\leq 2^{n+m}-1} |W_{t^{n+m}_k, t^{n+m}_{k+1}}|^p \leq \sum_{\ell=0}^{2^{m+n}-1} \mathbb{E}\,|W_{t^{n+m}_\ell, t^{n+m}_{\ell+1}}|^p$$

where the norms $|\,.\,|$ are the canonical Euclidean norms on the spaces $\mathcal{M}((d+1),(d+1))$ and $\mathbb{R}^{d+1}$ respectively.

It is clear that, for every $i\neq j$, $i,\,j\geq 1$ and every $t\geq s$,

$$\begin{aligned}\|\mathbf{W}^{2,ij}_{s,t}\|_p &= \left\|\int_s^t (W^i_u-W^i_s)dW^j_u\right\|_p\\ &\leq \left\|\int_s^t (W^i_u-W^i_s)dW^j_u\right\|_p\\ &\leq C^{BDG}_p\left\|\int_s^t (W^i_u-W^i_s)^2du\right\|^{\frac12}_{\frac p2}\\ &\leq C_p|t'-t|\end{aligned}$$

whereas

$$\|\,|W_{t'}-W_t|^2\|_p=|t'-t|\|\,|W_1|\|_p=C_{p,d}|t'-t|.$$

Noting that $W^0_t=t$ and, if $i=j$, $1\leq i\leq d$, $\mathbf{W}^{2,ii}_{s,t}=\frac12(W^i_t-W^i_s)^2$ shows that the above upper-bound still holds for $i=j$ and i or $j=0$. Consequently, we also have

$$\|\mathbf{W}^{2,ij}_{s,t}\|_p\leq C_{p,d}|t'-t|.$$

Consequently

$$\mathbb{E}\max_{0\leq k\leq 2^{n+m}-1}|\mathbf{W}^2_{t^{n+m}_k,t^{n+m}_{k+1}}|^p\leq C_{p,d}\sum_{k=0}^{2^{n+m}-1}\left(\frac{T}{2^{n+m}}\right)^p=C_{p,d,T}2^{(n+m)(1-p)}$$

so that

$$\|Z^{(2)}_{\tilde\theta'}\|_p\leq C_{p,d,T}\sum_{n\geq0}2^{n\frac{\tilde\theta'}{p}}\sum_{m\geq0}2^{(n+m)(\frac1p-1)}=C_{p,d,T}\sum_{n\geq0}2^{n(\frac{\tilde\theta'}{p}-1)}<+\infty$$

since $\tilde\theta'<p-1$.

On the other hand, one has obviously

$$\sup_{s,t\in[0,T],s\neq t}\frac{|\mathbf{W}^2_{s,t}|}{|t-s|^{\frac{\tilde\theta'}{p}}}\leq Z^{(2)}_{\tilde\theta'}<+\infty\qquad a.s.$$

Lemma 4.2(a) applied to W (which satisfies $(K_{p,\frac p2-1})$) yields for every $\theta'\in(0,\frac p2-1)$ the existence of $Z^{(1)}\in L^p(\mathbb{P})$ such that

$$\sup_{s,t\in[0,T],s\leq t}\frac{|\mathbf{W}^1_{s,t}|}{|t-s|^{\frac{\theta'}{p}}}\leq Z^{(1)}_{\theta'}\qquad a.s.$$

As a consequence, combining these two results shows that, for every $q > \frac{2p}{p-2}$,

$$\rho_q(\mathbf{W}, 0) < Z = Z^{(1)}_{\theta'} + Z^{(2)}_{\tilde{\theta}'} \in L^p(\mathbb{P})$$

where $Z^{(1)}$ is related to $\theta' = \frac{p}{q} \in (0, \frac{p}{2} - 1)$ and $Z^{(2)}$ is related to $\tilde{\theta}' = \frac{2p}{q} \in (0, p-2)$.

(b) If $i \neq j$, $0 \leq i, j \leq d$, it follows from Proposition 2.2 that $\widehat{\mathbf{W}}^{2,ij,N}_{s,t} = \mathbb{E}(\widehat{\mathbf{W}}^{2,ij,N}_{s,t} \,|\, \mathcal{G}_N)$ where $\mathcal{G}_N = \sigma(\widehat{W})$ and $\widehat{W}^N = (\widehat{W}^{i,N})_{1 \leq i \leq d}$ is an optimal product quantization at level N (which means that for each component W^i, $\widehat{W}^{i,N}$ is an optimal product quantization at level $N' = \lfloor N^{\frac{1}{d}} \rfloor$).

When $i = j \geq 1$, $|\widehat{\mathbf{W}}^{2,ii,N}_{s,t}| \leq \frac{1}{2}\mathbb{E}\left((W^i_t - W^i_s)^2 \,|\, \mathcal{G}_N\right)$. One derives that

$$\frac{|\widehat{\mathbf{W}}^{2,ii,N}_{s,t}|}{|t-s|^{\frac{\tilde{\theta}'}{p}}} \leq \mathbb{E}\left(\frac{|\mathbf{W}^{2,ii}_{s,t}|}{|t-s|^{\frac{\tilde{\theta}'}{p}}} \,|\, \mathcal{G}_N\right) \leq \mathbb{E}\left((Z^{(2)}_{\tilde{\theta}'})^{\frac{\tilde{\theta}'}{p}} \,|\, \mathcal{G}_N\right).$$

When $i = j = 0$, $\widehat{\mathbf{W}}^{2,ii,N} = \mathbf{W}^{2,ii} = \frac{1}{2}(t-s)^2$.

(c) In this claim, the random variable $Z^{(2),N}_{\tilde{\theta}'}$ of interest is defined by

$$\widetilde{Z}^{(2),N}_{\theta'} = \frac{2}{T} \sum_{n \geq 0} 2^{n\frac{\tilde{\theta}'}{p}} \sup_{s \leq t \leq s + \frac{T}{2^n}} |\widetilde{\mathbf{W}}^{2,N}_{s,t}|$$

and we aim at showing that it lies in $L^p(\mathbb{P})$ with a control on its L^p-norm as a function of N. One first derives for $\widetilde{\mathbf{W}}^{2,N}_{s,t}$ the straightforward identity when $s \leq u \leq t$

$$\widetilde{\mathbf{W}}^{2,N}_{s,t} = \widetilde{\mathbf{W}}^{2,N}_{s,u} + \widetilde{\mathbf{W}}^{2,N}_{u,t} + \widetilde{W}^N_{s,u,t}$$

where

$$\begin{aligned} \widetilde{W}^N_{s,u,t} &= W_{s,u} \otimes W_{u,t} - \widehat{W}^N_{s,u} \otimes \widehat{W}^N_{u,t} \\ &= (W_{s,u} - \widehat{W}^N_{s,u}) \otimes W_{u,t} + \widehat{W}^N_{s,u} \otimes (W_{u,t} - \widehat{W}^N_{u,t}) \qquad (5.36) \end{aligned}$$

with $W_{r,s} := W_r - W_s$ if $r \geq s$, etc. One derives from (5.36) that

$$\begin{aligned} |\widetilde{\mathbf{W}}^{2,N}_{s,t}| \leq\ & 2 \sum_{m \geq 0} \max_{0 \leq k \leq 2^{n+m}-1} |\widetilde{\mathbf{W}}^{2,N}_{t^{n+m}_k, t^{n+m}_{k+1}}| \qquad (5.37) \\ & + 2 \sum_{m,m' \geq 0} \max_{\substack{0 \leq k \leq 2^{n+m}-1 \\ 0 \leq k' \leq 2^{n+m'}-1}} |W^{2,N}_{t^{n+m}_k, t^{n+m}_{k+1}} - \widehat{W}^{2,N}_{t^{n+m}_k, t^{n+m}_{k+1}}||W_{t^{n+m}_{k'}, t^{n+m}_{k'+1}}| \qquad (5.38) \\ & + 2 \sum_{m,m' \geq 0} \max_{\substack{0 \leq k \leq 2^{n+m}-1 \\ 0 \leq k' \leq 2^{n+m'}-1}} |W^{2,N}_{t^{n+m}_k, t^{n+m}_{k+1}} - \widehat{W}^{2,N}_{t^{n+m}_k, t^{n+m}_{k+1}}||\widehat{W}_{t^{n+m}_{k'}, t^{n+m}_{k'+1}}|. \qquad (5.39) \end{aligned}$$

where we used that $|u \otimes v| \le |u||v|$.

We will first deal with deal with the first term in (5.37). We note that

$$\mathbb{E} \max_{0\le k\le 2^{n+m}-1} |\widetilde{\mathbf{W}}^{2,N}_{t_k^{n+m},t_{k+1}^{n+m}}|^p \le \sum_{0\le k\le 2^{n+m}-1} \mathbb{E}|\widetilde{\mathbf{W}}^{2,N}_{t_k^{n+m},t_{k+1}^{n+m}}|^p.$$

Let s, $t\in[0,T]$, $s \le t$ and i, $j\in\{1,\dots,d\}$, $i \ne j$. One checks that the following decomposition holds

$$\widetilde{\mathbf{W}}^{2,ij,N}_{s,t} = \underbrace{\int_s^t W^i_{s,u} d(W^j_u - \widehat{W}^{j,N}_u)}_{(A)} + \underbrace{\int_s^t \widehat{W}^{j,N}_{u,t} d(W^i_u - \widehat{W}^{i,N}_u)}_{(B)}.$$

Let us focus on (A). First not that, owing to Proposition 2.1 applied with $\mathcal{F}_t = \sigma(W^i_u,\ u\in[0,T],\ W^j_s,\ s\le t)$,

$$(A) = \sum_{n\ge1} \widetilde{\xi}^j_n \int_s^t W^i_{s,u} \cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big) du.$$

Using that W^i and W^j are independent, one derives that (A) is the terminal value of a martingale with respect to the filtration $\sigma(\xi^j_k,\ k\le n,\ W^i_u, 0\le u\le T)$, $n\ge 1$ so that combining B.D.G. and Minkowski inequalities yields, with the notations of Lemma 4.3,

$$\begin{aligned}
\mathbb{E}(|(A)|^p) &\le C_p^{BDG}\mathbb{E}\left(\sum_{n\ge1}(\widetilde{\xi}^j_n)^2\left(\int_s^t W^i_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du\right)^2\right)^{\frac{p}{2}}\\
&\le C_p^{BDG}\left(\sum_{n\ge1}\|\widetilde{\xi}^j_n\|_p^2\left\|\int_s^t W^i_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du\right\|_p^2\right)^{\frac{p}{2}}
\end{aligned}$$

where $\tilde{\xi}_n = \xi_n - \widehat{\xi}_n^{N_n}$ and $N_1,\dots,N_n,\dots$ denote the optimal bit allocation of an optimal quadratic product quantization at level N' (keep in mind that $N_k = 1$, $k > L_B(N')$ and $N_1\cdots N_{L_B(N')} \le N'$ (B scalar Brownian motion). Now an elementary integration by parts yields

$$\int_s^t W^i_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du = \sqrt{\lambda_n}\int_s^t\Big(\sin\Big(\frac{t}{\sqrt{\lambda_n}}\Big) - \sin\Big(\frac{u}{\sqrt{\lambda_n}}\Big)\Big)dW^i_u$$

so that, for every $\rho\in(0,\frac{1}{2})$, one checks that, owing to the BDG Inequality,

$$\left\|\int_s^t W^i_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du\right\|_p \le C_p^{BDG}C_{p,\rho}\lambda_n^{\frac{1-\rho}{2}}|t-s|^{\frac{1}{2}+\rho}.$$

Finally, for every $\varepsilon\in(0,3)$,

$$\|(A)\|_p \le C_{p,T,\rho,\varepsilon}\left(\sum_{n\ge1}\lambda_n^{1-\rho}\|\tilde\xi_n\|_p^2\right)^{\frac12}|t-s|^{\frac12+\rho}.$$

One shows likewise the same inequality for (B) once noted that

$$\int_s^t \widehat{W}^{i,N}_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du = \mathbb{E}\Big(\int_s^t W^i_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du\,|\,\mathcal{F}_T^{\widehat{W}^{i,N}}\Big)$$

which implies

$$\Big\|\int_s^t \widehat{W}^{i,N}_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du\Big\|_p \le \Big\|\int_s^t W^i_{s,u}\cos\Big(\frac{u}{\sqrt{\lambda_n}}\Big)du\Big\|_p.$$

Consequently, for every $\varepsilon\in(0,3)$,

$$\begin{aligned}\|\widetilde{\mathbf{W}}^{2,ij,N}_{s,t}\|_p &\le C_{p,\rho,T}\left(\sum_{n\ge1}\lambda_n^{1-\rho}\|\tilde\xi_n\|_p^2\right)^{\frac12}|t-s|^{\frac12+\rho}\\ &\le C_{p,T,\rho,d,\varepsilon}\big(\log N\big)^{-(\frac12-\rho)\wedge\frac{3-\varepsilon}{2p}}|t-s|^{\frac12+\rho}. \qquad (5.40)\end{aligned}$$

If $i=j\ge1$, then

$$\widetilde{\mathbf{W}}^{2,ii,N}_{s,t} = \frac12\Big((W^i_{s,t})^{-}(\widehat{W}^{i,N}_{s,t})^2\Big)$$

so that, using again Hölder Inequality,

$$\|\widetilde{\mathbf{W}}^{2,ii,N}_{s,t}\|_p = \frac12\|W^i_{s,t}-\widehat{W}^{i,N}_{s,t}\|_{p+\delta)}\|W^i_{s,t}-\widehat{W}^{i,N}_{s,t}\|_{p(1+\frac{p}{\delta})}$$

and one gets the same bounds as in the case $i\ne j$.

If i or $j=0$, one gets similar bounds: we leave the details to the reader. Finally, one gets that, for every $i,j\in\{0,\dots,d\}$,

$$\|\widetilde{\mathbf{W}}^{2,N}_{s,t}\|_p \le C_{p,\rho,T,d,\varepsilon,\delta}\big(\log N\big)^{-(\frac12-\rho)\wedge\frac{3-\varepsilon}{2(p+\delta)}}|t-s|^{\frac12+\rho}.$$

By standard computations similar to those detailed in Lemma 4.2, we get

$$\sum_{m\ge0}\|\max_{0\le k\le 2^{n+m}-1}|\widetilde{\mathbf{W}}^{2,N}_{t_k^{n+m},t_{k+1}^{n+m}}|\|_p \le C_{p,\rho,T,d,\varepsilon,\delta}\big(\log N\big)^{-(\frac12-\rho)\wedge\frac{3-\varepsilon}{2(p+\delta)}}2^{-n(\frac12+\rho)}.$$

Let us pass now to the two other sums. We will focus on (5.38) since both behave and can be treated similarly.

$$\max_{\substack{0\le k\le 2^{n+m}-1\\ 0\le k'\le 2^{n+m'}-1}} |W^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}} - \widehat{W}^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}}|^p |W_{t^{n+m}_{k'},t^{n+m}_{k'+1}}|^p$$
$$\le \sum_{\substack{0\le k\le 2^{n+m}-1\\ 0\le k'\le 2^{n+m'}-1}} |W^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}} - \widehat{W}^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}}|^p |W_{t^{n+m}_{k'},t^{n+m}_{k'+1}}|^p.$$

Now for every $s,u,t\in[0,T]$, $s\le u\le t$, it follows from Hölder Inequality that

$$\begin{aligned}\|\,|W_{s,u}-\widehat{W}^N_{s,u}|\,|W_{u,t}|\|_p &\le \|W_{s,u}-\widehat{W}^N_{s,u}\|_{p+\delta}\|W_{u,t}\|_{p(1+p/\delta)}\\ &\le C_{p,\delta}\|W_{s,u}-\widehat{W}^N_{s,u}\|_{p+\delta}|t-u|^{\frac12}.\end{aligned}$$

Using Inequality (4.33) from Lemma 4.3, we get for every $p>2$, every $\rho\in(0,\frac12)$, every $\varepsilon\in(0,3)$, and every s, $t\in[0,T]$, $s\le t$,

$$\left\|W^i_{s,t}-\widehat{W}^{i,N}_{s,t}\right\|_p \le C_{\rho,p,T,d,\varepsilon}|t-s|^\rho(\log N)^{-(\frac12-\rho)\wedge(\frac{3-\varepsilon}{2p})}.$$

Now,

$$\mathbb{E}\max_{\substack{0\le k\le 2^{n+m}-1\\ 0\le k'\le 2^{n+m'}-1}} |W^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}} - \widehat{W}^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}}|^p |W_{t^{n+m}_{k'},t^{n+m}_{k'+1}}|^p$$
$$\le \left(C_{\rho,p,T,d,\delta,\varepsilon}(\log N)^{-(\frac12-\rho)\wedge(\frac{3-\varepsilon}{2(p+\delta)})}\right)^p 2^{(n+m)(1-\rho p)}2^{(n+m')(1-\frac p2)}$$

and we use that $\rho>\frac1p$ and $p>2$ to show that

$$\sum_{m,m'\ge0}\max_{\substack{0\le k\le 2^{n+m}-1\\ 0\le k'\le 2^{n+m'}-1}} \||W^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}} - \widehat{W}^{2,N}_{t^{n+m}_k,t^{n+m}_{k+1}}|^p |W_{t^{n+m}_{k'},t^{n+m}_{k'+1}}|\|_p$$
$$\le C_{\rho,p,T,d,\delta,\varepsilon}(\log N)^{-(\frac12-\rho)\wedge\frac{3-\varepsilon}{2(p+\delta)}}2^{n(\frac2p-(\frac12+\rho))}.$$

Finally, we get

$$\mathbb{E}\Big(\widetilde{Z}^{(2),N}_{\tilde\theta'}\Big)^p \le C_{\rho,p,T,d,\varepsilon,\delta}\big(\log N\big)^{-p(\frac12-\rho)\wedge\frac{3-\varepsilon}{2(p+\delta)}}$$

as soon as $\tilde\theta'\in(0,\tilde\theta)$ with $\tilde\theta=p(\rho+\frac12)-2$. Now, it follows by standard arguments that

$$\sup_{s,t\in[0,T]}|\widetilde{\mathbf{W}}^{2,N}_{s,t}|\le \widetilde{Z}^{(2),N}_{\tilde\theta'}|t-s|^{\frac{\tilde\theta'}{p}}$$

so that, finally

$$\left\|\sup_{s,t\in[0,T]}\frac{|\widetilde{\mathbf{W}}^{2,N}_{s,t}|}{|t-s|^{\frac{\tilde\theta'}{p}}}\right\|_p \le C_{\rho,p,T,d,\varepsilon,\delta}\big(\log N\big)^{-(\frac12-\rho)\wedge\frac{3-\varepsilon}{2(p+\delta)}}. \quad \diamond$$

Now, we are in position to prove the main result of this section.

Proof of Theorem 5.5. (a) Given Theorem 4.4, this amounts to proving that $\|\mathbf{W}^2-\widehat{\mathbf{W}}^{2,N}\|_{\frac{q}{2},Hol}$ converges to 0 in every $L^p(\mathbb{P})$. This easily follows from Proposition 5.3(a)-(b).

(b) We inspect successively four cases to maximize $\min(1-\rho,\frac{3}{2p})$ in ρ when it is possible.

$\rhd$ $q\in(2,4)$ and $p<\frac{7q}{2(q-2)}$. Let p' be defined by $\frac{1}{p'}=\frac{2(q-2)}{7q}+\frac{\alpha}{2}$ with $\alpha>0$ small enough so that $p'>p$ and $\frac{1}{p'}+\frac{1}{q}<\frac12$. Then set $\rho'=\frac{2}{q}+\frac{2}{p'}-\frac12+\frac{\alpha}{2}$ (note that $\rho'>\frac{1}{p'}$). One checks that $\frac12-\rho'=1-2(\frac{1}{p'}+\frac1q)=\frac37(1-\frac2q)-\alpha\in(0,\frac{3-\varepsilon}{2(p'+\delta)}\wedge\frac12)$ at least for any small enough α, $\delta=\delta(\alpha,q)>0$ and $\varepsilon=\varepsilon(\alpha,q)>0$. Now, Proposition 5.3$(c)$ applied with $\tilde\theta'=\frac{2p'}{q}<p'(\rho'+\frac12)-2$ yields the announced asymptotic rate for $\Big\|\|\mathbf{W}^2-\widehat{\mathbf{W}}^{2,N}\|_{\frac{q}{2},Hol}\Big\|_p$, $p<p'$, since $L^p(\mathbb{P})$-norms are non-decreasing in p.

$\rhd$ $q\in(2,4)$ and $p\ge\frac{7q}{2(q-2)}$. One sets the same specifications as above for ρ but with $p'=p$. Then $1/2-\rho>\frac{3}{2p}$ and choose $\varepsilon=\varepsilon(q,\alpha)>0$ and $\delta=\delta(q,\alpha)>0$ small enough so that $\frac{3-\varepsilon}{2(p+\delta)}\le\frac{3}{2p}+\alpha$.

$\rhd$ $q\in[4,20/3)$. Then $\frac{7q}{2(q-2)}<\frac{2q}{q-4}$ and one checks that the cases $p\in(2,\frac{7q}{2(q-2)})$ and $p\in[\frac{7q}{2(q-2)},\frac{2q}{q-4})$ can be solved as above. If $p\ge\frac{2q}{q-4}$ (hence ≥ 5), no optimization in ρ is possible *i.e.* any admissible ρ satisfies $\frac12-\rho>\frac{3}{2p}$.

$\rhd$ $q\ge 20/3$ *i.e.* $\frac{7q}{2(q-2)}>\frac{2q}{q-4}$. If $p<\frac{2q}{q-4}$, set p' such that $\frac{1}{p'}=\frac{q-4}{2q}+\alpha'/2$, $\alpha'>0$ small enough and $\rho'=\frac2q+\frac{2}{p'}-\frac12+\frac{\alpha}{2}$. Doing as above yields $\min(1-\rho,\frac{3}{2p})=\frac2q+\alpha$ for an arbitrary small $\alpha>0$. Note that this quantity is greater than $\frac37(1-\frac2q)+\alpha$ (so in that case our exponent is not optimal). If $p\ge\frac{2q}{q-4}$, we proceed to no optimization in ρ.

(c) This is a consequence of Borel-Cantelli's Lemma by considering $p>\frac{7q}{q-2}$. $\diamond$

Now we conclude by proving Theorem 3.2.

Proof of Theorem 3.2. First we check using Proposition 5.3 that $\rho_q(\widehat{\mathbf{W}}^N, 0)$ and $\rho_q(\mathbf{W}, 0)$ are *a.s.* finite since they are integrable. Now we may apply Theorem 3.1 which yields the announced result. $\diamond$

ACKNOWLEDGEMENT: We thank A. Lejay for helpful discussions and comments about several versions of this work and F. Delarue and S. Menozzi for initiating our first "meeting" with rough path theory.

References

[1] BILLINGSLEY P. (1968). *Convergence of Probability Measure*, Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 253p.

[2] COUTIN L., VICTOIR N. (2009). Enhanced Gaussian Processes and Applications, *ESAIM Probab. Stat.*, **13**: 247–269.

[3] CUESTA-ALBERTOS J.A., MATRÁN C. (1988). The strong law of large numbers for k-means and best possible nets of Banach valued random variables, *Probab. Theory Rel. Fields*, **78**:523–534.

[4] DEREICH S. (2008). The coding complexity of diffusion processes under $L^p[0,1]$-norm distortion, pre-print, *Stoch. proc. and their Appl.*, **118**(6):938–951.

[5] DEREICH S., FEHRINGER F., MATOUSSI A., SCHEUTZOW M. (2003). On the link between small ball probabilities and the quantization problem for Gaussian measures on Banach spaces, *J. Theoretical Probab.*, **16**:249–265.

[6] FRIZ P. (2005). Continuity of the Itô-map for Hölder rough paths with applications to the support theorem in Hölder norm. *Probability and partial differential equations in modern applied mathematics*, 117-135, IMA Vol. Math. Appl., **140**, Springer, New York.

[7] FRIZ P., VICTOIR N. (2007). Differential Equations Driven by Gaussian Signals I (2007), available at arxiv:0707.0313.

[8] FRIZ P., VICTOIR N. (2007). Differential Equations Driven by Gaussian Signals II (2007), available at arxiv:0711.0668.

[9] FRIZ P., VICTOIR N. (2010). *Multidimensional Stochastic Differential Equations as Rough Paths: Theory and Applications*, Cambridge Studies in Advanced Mathematics, 670p.

[10] GRAF S., LUSCHGY H. (2000). *Foundations of Quantization for Probability Distributions*. Lect. Notes in Math. 1730, Springer, Berlin.

[11] Graf S., Luschgy H., Pagès G. (2003). Functional quantization and small ball probabilities for Gaussian processes, *J. Theoret. Probab*, **16**(4):1047–1062.

[12] Graf S., Luschgy H., Pagès G. (2007). Optimal quantizers for Radon random vectors in a Banach space, *J. of Approximation Theory*, **144**:27–53.

[13] Graf S., Luschgy H., Pagès G. (2008). Distortion mismatch in the quantization of probability measures, *ESAIM P&S*, **12**, 127-153.

[14] Lejay A. (2003). An introduction to rough paths, *Séminaire de Probabilités YXXVII*, *Lecture Notes in Mathematics 1832*:1–59.

[15] Lejay A. (2009). Yet another introduction to rough paths, *Séminaire de Probabilités XLII*, *Lecture Notes in Mathematics 1979*, 1–101.

[16] Lejay A. (2009). On rough differential equations, *Electronic Journal of Probability*, **14**(12):341–364.

[17] Lejay A. (2010). Global solutions to rough differential equations with unbounded vector fields, pre-pub. Inria 00451193, 2010.

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

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

[20] Luschgy H., Pagès G. (2006). Functional quantization of a class of Brownian diffusions: a constructive approach, *Stoch. Proc. and their Appl.*, **116**:310-336.

[21] Luschgy H., Pagès G. (2007). High resolution product quantization for Gaussian processes under sup-norm distortion, *Bernoulli*, **13**(3):653–671.

[22] Luschgy H., Pagès G. (2008). Functional quantization and mean pathwise regularity with an application to Lévy processes, *Annals of Applied Probability*, **18**(2):427–469.

[23] Luschgy H., Pagès G., Wilbertz B. (2008). Asymptotically optimal quantization schemes for Gaussian processes, *ESAIM P&S*, **12**:127–153.

[24] Lyons T. (1995). Interpretation and Solutions of ODE's Driven by Rough signals, *Proc. Symposia Pure*, 1583.

[25] Lyons T. (1998). Differential Equations driven by rough signals, *Rev. Mat. Iberoamericana*, **14**(2): 215–310.

[26] Lyons T., Caruana M.J., Lévy T. (2007). *Differential equations driven by rough paths.* Lect. Notes in Math. 1908. Notes from T. Lyons's course at École d'été de Saint-Flour (2004).

[27] Pagès G., Printems J. (2005). Functional quantization for numerics with an application to option pricing, *Monte Carlo Methods & Applications*, **11**(4): 407–446.

[28] Pagès G., Printems J. (2005). Website devoted to vector and functional optimal quantization: `www.quantize.maths-fi.com`.

[29] Revuz D., Yor M. (1999). *Continuous martingales and Brownian motion.* Third edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **293**, Springer-Verlag, Berlin, 602 p.

Appendix: Functional conditional expectation

Let $(Y_t)_{t\in[0,T]}$ be a bi-measurable process defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that

$$\int_0^T \mathbb{E}(Y_t^2)dt < +\infty.$$

One can consider Y as a random variable $Y : (\Omega, \mathcal{A}, \mathbb{P}) \to L^2_T := L^2([0,T], dt)$ and more precisely as an element of the Hilbert space

$$L^2_{L^2_T}(\Omega, \mathcal{A}, \mathbb{P}) := \left\{ Y : (\Omega, \mathcal{A}, \mathbb{P}) \to L^2_T,\ \mathbb{E}\,|Y|^2_{L^2_T} < +\infty \right\}$$

where $|f|^2_{L^2_T} = \int_0^T f^2(t)dt$. For the sake of simplicity, one denotes $\|Y\|_2 := \sqrt{\mathbb{E}\,|Y|^2_{L^2_T}}$. If $\mathcal{B}$ denotes a sub-σ-field of $\mathcal{A}$ (containing all $\mathbb{P}$-negligible sets of $\mathcal{A}$) then $L^2_{L^2_T}(\Omega, \mathcal{B}, \mathbb{P})$ is a closed sub-space of $L^2_{L^2_T}(\Omega, \mathcal{A}, \mathbb{P})$ and one can define the *functional conditional expectation* of Y by

$$\mathbb{E}(Y\,|\,\mathcal{B}) := \mathrm{Proj}^{\perp}_{L^2_{L^2_T}(\Omega,\mathcal{B},\mathbb{P})}(Y).$$

Functional conditional expectation can be extended to bi-measurable processes Y such that $\|Y\|_1 := \mathbb{E}\,|Y|_{L^1_T} < +\infty$ following the approach used for $\mathbb{R}^d$-valued random vectors. Then, $\mathbb{E}(Y\,|\,\mathcal{B})$ is characterized by: for every $\mathcal{B}([0,T]) \otimes \mathcal{B}$-bi-measurable process $Z = (Z_t)_{t\in[0,T]}$, bounded by 1,

$$\mathbb{E}\int_0^T Z_t\, Y_t\, dt = \mathbb{E}\int_0^T Z_t\, \mathbb{E}(Y\,|\,\mathcal{B})_t\, dt.$$

In particular, owing to the Fubini theorem, this implies that as soon as the process $(\mathbb{E}(Y_t\,|\,\mathcal{B}))_{t\in[0,T]}$ has a $\mathcal{B}([0,T])\otimes\mathcal{B}$ bi-measurable version, the functional conditional expectation could be defined by setting

$$\mathbb{E}(Y\,|\,\mathcal{B})_t(\omega) = \mathbb{E}(Y_t\,|\,\mathcal{B})(\omega), \qquad (\omega, t) \in \Omega \times [0,T].$$

EXAMPLES: (*a*) Let $\mathcal{B} := \sigma(\mathcal{N}_{\mathcal{A}},\, B_i,\; i \in I)$ where $(B_i)_{i\in I}$ is a finite measurable partition of Ω such that $\mathbb{P}(B_i) > 0$, $i \in I$.

(*b*) Let $Y := (W_t)_{t\in[0,T]}$ a standard Brownian motion in $\mathbb{R}^d$ and let $\mathcal{B} := \sigma(W_{t_1}, \ldots, W_{t_n})$ where $0 = t_0 < t_1 < \ldots < t_n = T$. Then

$$\forall\, t \in [t_k, t_{k+1}), \qquad \mathbb{E}(W \,|\, \mathcal{B})_t = W_{t_k} + \frac{t - t_k}{t_{k+1} - t_k}(W_{t_{k+1}} - W_{t_k}).$$