Electronic Journal of Probability

Vol. 11 (2006), Paper no. 28, pages 700–722.

Journal URL
http://www.math.washington.edu/~ejpecp/

High-resolution quantization and entropy coding for fractional Brownian motion

S. Dereich and M. Scheutzow
Institut für Mathematik, TU-Berlin
Straße des 17. Juni 136
D-10623 Berlin
{dereich,ms}@math.tu-berlin.de

Abstract

We establish the precise asymptotics of the quantization and entropy coding errors for fractional Brownian motion with respect to the supremum norm and $L^p[0,1]$-norm distortions. We show that all moments in the quantization problem lead to the same asymptotics. Using a general principle, we conclude that entropy coding and quantization coincide asymptotically. Under supremum-norm distortion, our proof uses an explicit construction of efficient codebooks based on a particular entropy constrained coding scheme.

Key words: High-resolution quantization; complexity; stochastic process; entropy; distortion rate function

AMS 2000 Subject Classification: Primary 60G35, 41A25, 94A29.

Submitted to EJP on March 22 2005, final version accepted June 1 2006.

1 Introduction

Functional quantization and entropy coding concern the identification of “good” discrete approximations to a non-discrete random signal (*original*) in a Banach space of functions. These approximations are required to satisfy a range constraint in the context of quantization and an entropy constraint in the context of entropy coding. Such discretization problems arise naturally when digitizing analog signals in order to allow storage on a computer or transmission over a channel with finite capacity.

As another application, the approximating functions of good *quantizers* may serve as evaluation points of quasi Monte Carlo methods. Moreover, some Monte Carlo methods use appropriate quantizers to carry out a variance reduction (see for instance (18) and the references therein, or (10)).

Previous research addressed, for instance, the problem of constructing good approximation schemes, the evaluation of the theoretically best approximation under an information constraint, existence of optimal quantizers and regularity properties of the paths of optimal approximations.

For Gaussian measures in Hilbert spaces optimal quantizers exist and all its approximating functions are elements of the reproducing kernel Hilbert space (16). Under mild assumptions the best achievable distortion in both problems (the quantization and entropy coding problem) coincide asymptotically and do not depend on the moment under consideration ((17), (4)). Moreover, this (optimal) approximation error is asymptotically equivalent to the *distortion rate function* which can be expressed implicitly in terms of the eigenvalues of the covariance operator.

When the underlying space is a Banach space, the approximation errors of both problems are weakly asymptotically equivalent to the inverse of the small ball function in many cases ((8), (5)). Thus asymptotic estimates for the small ball function can be translated into asymptotic estimates for the above coding problems (see for instance (15) for a summary of results on small ball probabilitites). Moreover, many approximation quantities of Gaussian measures are tightly connected to the quantization numbers (see (14), (2)). See also (12) for existence and pathwise regularity results of optimal quantizers.

The above questions are treated for Gaussian measures in Hilbert spaces by Luschgy and Pagès ((16), (17)) and by the first-named author in (4). For Gaussian originals in Banach spaces, these problems have been addressed by the authors and collaborators in (8), (9), (4), (5) and by Graf, Luschgy and Pagès in (12). For general accounts on quantization and coding theory in finite dimensional spaces, see (11) and (1) (see also (13)).

In this article, we consider the asymptotic coding problem for fractional Brownian motion under supremum and $L^p[0,1]$-norm distortion. We derive the asymptotic quality of optimal approximations. In particular, we show that efficient entropy constrained quantizers can be used to construct close to optimal high resolution quantizers when considering the supremum norm. Moreover, for all of the above norm-based distortions, all moments and both information constraints lead to the same asymptotic approximation error. In particular, quantization is asymptotically just as efficient as entropy coding. The main impetus to the present work was provided by the necessity to understand the coding complexity of Brownian motion in order to solve the quantization (resp. entropy constrained coding) problem for diffusions (see (7) and (6)).

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $H \in (0,1)$ and let $X = (X_t)_{t\geq 0}$ denote fractional Brownian motion with Hurst index H on $(\Omega, \mathcal{A}, \mathbb{P})$, i.e. $(X_t)_{t\geq 0}$ is a centered continuous Gaussian process with covariance kernel

$$K(t,u) = \frac{1}{2}[t^{2H} + u^{2H} - |t-u|^{2H}], \qquad t,u \geq 0.$$

For $a > 0$, let $\mathbb{C}[0,a]$ and $\mathbb{D}[0,a]$ denote the space of continuous real-valued functions on the interval $[0,a]$ and the space of càdlàg functions on $[0,a]$, respectively. Moreover, we let $(L^p[0,a], \|\cdot\|_{L^p[0,a]})$ denote the standard L^p-space of real-valued functions defined on $[0,a]$. Finally, $\|\cdot\|_s$, $s \in (0,\infty]$, denotes the L^s-norm induced by the probability measure $\mathbb{P}$ on the set of real-valued random variables.

Let us briefly introduce the concepts of quantization and entropy coding. For fixed $a > 0$ let $d : \mathbb{C}[0,a] \times \mathbb{D}[0,a] \to [0,\infty)$ be a measurable function For a $\mathbb{C}[0,a]$-valued r.v. Y (*original*) and *moment* $s > 0$, the aim is to minimize

$$\left\| d(Y, \pi(Y)) \right\|_s \tag{1}$$

over all measurable functions $\pi : \mathbb{C}[0,a] \to \mathbb{D}[0,a]$ with discrete image (*strategies*) that satisfy a particular information constraint parameterized by the *rate* $r \geq 0$.

Often we associate a sequence of probability weights $(p_w)_{w \in \mathrm{im}(\pi)}$ to a strategy π. Then due to Kraft's inequality, there exists a prefix-free representation for $\mathrm{im}(\pi)$ which needs less than $(-\log_2 p_w) + 1$ bits to represent $w \in \mathrm{im}(\pi)$. Thus the pair $(\pi, (p_w))$ corresponds to a coding scheme translating the original symbol x into a prefix-free representation for $\pi(x)$. The best average code length is achieved for $p_w = \mathbb{P}(\pi(Y) = w)$, which leads to an average code length of about $\mathbb{H}(\pi(Y))/\log 2$ (see for instance (1), Theorem 5.2.1).

Entropy coding (also known as *entropy constrained quantization* in the literature) concerns the minimization of (1) over all strategies π having entropy $\mathbb{H}(\pi(Y))$ at most r. Recall that the entropy of a discrete r.v. Z with probability weights $(p_w^{(Z)})$ is defined as

$$\mathbb{H}(Z) = -\sum_w p_w^{(Z)} \log p_w^{(Z)} = \mathbb{E}[-\log p_Z^{(Z)}].$$

The entropy constraint represents an *average case* complexity constraint.

In the *quantization problem*, one is considering strategies π satisfying the range constraint: $|\operatorname{range}(\pi(Y))| \leq e^r$ which is a *static* complexity constraint. The corresponding approximation quantities are the *entropy-constrained quantization error*

$$D^{(e)}(r|Y,d,s) := \inf_\pi \left\| d(Y, \pi(Y)) \right\|_s, \tag{2}$$

where the infimum is taken over all strategies π with entropy rate $r \geq 0$, and the *quantization error*

$$D^{(q)}(r|Y,d,s) := \inf_\pi \left\| d(Y, \pi(Y)) \right\|_s, \tag{3}$$

the infimum being taken over all strategies π having quantization rate $r \geq 0$. Often, all or some of the parameters Y, d, s are clear from the context, and will therefore be omitted. The *quantization information constraint* is more restrictive, so that the quantization error always dominates the entropy coding error. Moreover, the coding error increases with the moment under consideration.

Unless otherwise stated, we choose as original the fractional Wiener process $Y = X$. We are mainly concerned with two particular choices for the distortion d. First we analyse the supremum norm distortion that is $d(f,g) = \|f-g\|_{[0,1]}$. In this setting we find:

Theorem 1.1. *There exists a constant $\kappa = \kappa(H) \in (0,\infty)$ such that for all $s_1 \in (0,\infty]$ and $s_2 \in (0,\infty)$,*

$$\lim_{r\to\infty} r^H\, D^{(e)}(r|s_1) = \lim_{r\to\infty} r^H\, D^{(q)}(r|s_2) = \kappa.$$

Remark 1.2. In the above theorem, general càdlàg functions are allowed as reconstructions. Since the original process is continuous, it might seem more natural to use continuous functions as approximations. The following argument shows that confining oneself to continuous approximants does not change the corresponding quantization and entropy quantity, when $s \in [1,\infty)$. Let $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$ be an arbitrary strategy and let $\tau_n : \mathbb{D}[0,1] \to \mathbb{C}[0,1]$ denote the linear operator mapping f to its piecewise linear interpolation with supporting points $0, \frac{1}{n}, \frac{2}{n} \dots, 1$. Then

$$\begin{aligned}\big\|\|X-\tau_n\circ\pi(X)\|_{[0,1]}\big\|_s &\le \big\|\|\tau_n(X)-\tau_n\circ\pi(X)\|_{[0,1]}\big\|_s + \big\|\|X-\tau_n(X)\|_{[0,1]}\big\|_s \\ &\le \big\|\|X-\pi(X)\|_{[0,1]}\big\|_s + \big\|\|X-\tau_n(X)\|_{[0,1]}\big\|_s.\end{aligned}$$

Note that the second term vanishes when n tends to infinity and that $\tau_n\circ\pi$ satisfies the same information constraint as π. The argument can be easily modified in order to show the statement for $s \in (0,1)$.

Under $L^p[0,1]$-norm distortion ($p \ge 1$) that is $d(f,g) = \|f-g\|_{L^p[0,1]}$, we prove the following analog to Theorem 1.1:

Theorem 1.3. *For every $p \ge 1$ there exists a constant $\kappa_p = \kappa_p(H) \in (0,\infty)$ such that for all $s \in (0,\infty)$,*

$$\lim_{r\to\infty} r^H\, D^{(e)}(r|s) = \lim_{r\to\infty} r^H\, D^{(q)}(r|s) = \kappa_p.$$

As in Remark 1.2, a simple convolution type argument shows that allowing $L^p[0,1]$-approximations yields the same coding errors as restricting oneself to $\mathbb{C}[0,1]$-approximations.

For ease of notation, the article is restricted to the analysis of 1-dimensional processes. However, when replacing (X_t) by a process $(X_t^{(1)}, \dots, X_t^{(d)})$ consisting of d independent fractional Brownian motions, the proofs can be easily adapted, and one obtains analogous results. In particular, it is possible to prove analogs of the above theorems for a multi-dimensional Brownian motion.

Let us summarize some of the known estimates for the constant κ in the case where X is standard Brownian motion, i.e. $H = 1/2$.

- Under supremum-norm distortion, the relationship between the small ball function and the quantization problem (see (8)) shows that

$$\kappa \in \big[\frac{\pi}{\sqrt{8}}, \pi\big].$$

- Under $L^p[0,1]$-norm distortion, κ_p may again be estimated via a connection to the small ball function. Indeed, letting

$$\lambda_1 = \inf\Big\{\int_{-\infty}^{\infty} |x|^p \varphi^2(x)\,dx + \tfrac{1}{2}\int_{-\infty}^{\infty} (\varphi'(x))^2\,dx\Big\},$$

where the infimum is taken over all weakly differentiable $\varphi \in L^2(\mathbb{R})$ with unit norm, one has

$$\kappa_p \in [c, \sqrt{8}\, c]$$

for $c = 2^{1/p}\sqrt{p}\left(\frac{\lambda_1}{2+p}\right)^{(2+p)/2p}$.

In the case where $p = 2$, the constant κ_2 is known explicitly: $\kappa_2 = \frac{\sqrt{2}}{\pi}$ (see (17) and (4)).

The article is organized as follows. In Sections 2 to 5 we consider the approximation problems under the supremum norm. We start in Section 2 by introducting a coding scheme which plays an important role in the sequel. In Section 3, we use the construction of Section 2 and the self similarity of X to establish a polynomial decay for $D^{(e)}(\cdot|\infty)$. In the following section, the asymptotics of the quantization error are computed. The proof relies on a concentration property for the entropies of "good" coding schemes (Proposition 4.4). In Section 5, we use the equivalence of moments in the quantization problem to establish a lower bound for the entropy coding problem. In the last section, we treat the case where the distortion is based on the $L^p[0,1]$-norm, i.e. $d(f,g) = \|f-g\|_{L^p[0,1]}$; we introduce the distortion rate function and prove Theorem 1.3 with the help of Shannon's source coding theorem.

It is convenient to use the symbols $\sim$, $\lesssim$ and $\approx$. We write $f \sim g$ iff $\lim \frac{f}{g} = 1$, while $f \lesssim g$ stands for $\limsup \frac{f}{g} \le 1$. Finally, $f \approx g$ means

$$0 < \liminf \frac{f}{g} \le \limsup \frac{f}{g} < \infty\ .$$

2 The coding scheme

This section is devoted to the construction of strategies $\pi^{(n)} : \mathbb{C}[0,n] \to \mathbb{D}[0,n]$ which we will need later in our discussion. The construction depends on three parameters: $M \in \mathbb{N}\backslash\{1\}$, $d > 0$ and a strategy $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$.

The coding scheme is motivated as follows: Due to the self similarity of the fractional Wiener process, coding X on $[0,1]$ with accuracy εn^{-H} is as hard as coding X on the time interval $[0,n]$ with accuracy ε (see the argument at the end of the proof of Lemma 3.4). Intuitively, one may decompose the coding scheme for $(X_t)_{t\in[0,n]}$ into two steps. First store information on the values $(X_j)_{j=1,\dots,n-1}$ and then approximate the paths $X^{(j)} = (X_{j+t} - X_j)_{t\in[0,1)}$ by $\pi(X^{(j)})$ for $j = 0,\dots,n-1$. The parameter M governs the rate spent on coding the first part, and we shall see that for ε small most rate is spent for coding the second part. As in Shannon's source coding theorem, the ergodicity of $(X^{(j)})_{j\in\mathbb{N}_0}$ can be used to construct close to optimal codebooks when n is large. This will be done in the proof of Theorem 4.1 to derive an upper bound for the quantization error. Moreover, the coding scheme leads to a weak form of subadditivity which we use to prove polynomial decay of $D^{(e)}(\cdot|\infty)$ (see Theorem 3.1 and Lemma 3.4).

We define the maps $\pi^{(n)}$ by induction. Let $w \in \mathbb{C}[0,\infty)$ and set $(w_t^{(n)})_{t\in[0,1]} := (w_{t+n} - w_n)_{t\in[0,1]}$ and $\hat{w}_t := \pi(w^{(0)})(t)$ for $t \in [0,1)$. Assume that $(\hat{w}_t)_{t\in[0,n)}$ $(n \in \mathbb{N})$ has already been defined. Then we choose ξ_n to be the smallest number in $\{-d + 2kd/(M-1) : k = 0,\dots,M-1\}$ minimizing

$$|w_n - (\hat{w}_{n-} + \xi_n)|,$$

and extend the definition of $\hat{w}$ on $[n,(n+1))$ by setting

$$\hat{w}_{n+t} := \hat{w}_{n-} + \xi_n + \pi(w^{(n)})(t), \qquad t \in [0,1).$$

Note that $(\hat{w}_t)_{t\in[0,n)}$ depends only upon $(w_t)_{t\in[0,n)}$, so that the above construction induces strategies

$$\pi^{(n)} : \mathbb{C}[0,n] \to \mathbb{D}[0,n],\ w \mapsto (\bar{w}_t^{(n)})_{t\in[0,n]},$$

where $\bar{w}_t^{(n)} = \hat{w}_t$ for $t \in [0,n)$ and $\bar{w}_n^{(n)} = \hat{w}_{n-}$. Moreover, we can write

$$(\bar{w}_t^{(n)})_{t\in[0,n]} = \pi^{(n)}(w) = \varphi_n(\pi(w^{(0)}), \dots, \pi(w^{(n-1)}), \xi_1, \dots, \xi_{n-1}) \tag{4}$$

for an appropriate measurable function $\varphi_n : (\mathbb{D}[0,n])^n \times \mathbb{R}^{n-1} \to \mathbb{D}[0,n]$.

The main motivation for this construction is the following property. If one has, for some $(w_t) \in \mathbb{C}[0,\infty)$ and $n \in \mathbb{N}$,

$$\|w - \pi^{(n)}(w)\|_{[0,n]} \le \frac{M}{M-1}\, d$$

and $\|w^{(n)} - \pi(w^{(n)})\|_{[0,1]} \le d$, then

$$|w_n - (\hat{w}_{n-} + \xi_n)| \le \frac{d}{M-1},$$

whence,

$$\begin{aligned}
\|w - \hat{w}\|_{[n,n+1)} &= \|w_n + w_t^{(n)} - (\hat{w}_{n-} + \xi_n + \pi(w^{(n)})(t))\|_{[0,1)} \\
&\le |w_n - (\hat{w}_{n-} + \xi_n)| + \|w^{(n)} - \pi(w^{(n)})\|_{[0,1)} \\
&\le d/(M-1) + d = \frac{M}{M-1}\, d.
\end{aligned}$$

In particular, if $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$ satisfies

$$\big\| \|X - \pi(X)\|_{[0,1]} \big\|_\infty \le d,$$

then for any $n \in \mathbb{N}$,

$$\big\| \|X - \pi^{(n)}(X)\|_{[0,n]} \big\|_\infty \le \frac{M}{M-1}\, d. \tag{5}$$

3 Polynomial decay of $D^{(e)}(r|\infty)$

The objective of this section is to prove the following theorem.

Theorem 3.1. *There exists a constant $\kappa = \kappa(H) \in (0,\infty)$ such that*

$$\lim_{r\to\infty} r^H\, D^{(e)}(r|\infty) = \kappa. \tag{6}$$

Thereafter, $\kappa = \kappa(H)$ will always denote the finite constant defined via equation (6). In order to simplify notation, we abridge $\|\cdot\| = \|\cdot\|_{[0,1]}$.

Remark 3.2. It was found in (4) (see Theorem 3.5.2) that for finite moments $s \geq 1$ the entropy coding error is related to the asymptotic behavior of the small ball function of the Gaussian measure. In particular, for fractional Brownian motion, one obtains that

$$D^{(e)}(r|s) \approx \frac{1}{r^H}, \qquad r \to \infty.$$

In order to show that $D^{(e)}(r|\infty)$ is of the order r^{-H}, we still need to prove an appropriate upper bound. We prove a stronger statement which will be useful later on.

Lemma 3.3. *There exist strategies $\pi^{(r)} : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$, $r \geq 0$, and probability weights $(p_w^{(r)})_{w \in \mathrm{im}(\pi^{(r)})}$ such that for any $s \geq 1$,*

$$\big\| \|X - \pi^{(r)}(X)\| \big\|_\infty \leq \frac{1}{r^H} \quad \text{and} \quad \mathbb{E}[(-\log p^{(r)}_{\pi^{(r)}(X)})^s]^{1/s} \approx r. \tag{7}$$

In particular, $D^{(e)}(r|\infty) \approx r^{-H}$.

The proof of the lemma is based on an asymptotic estimate for the mass concentration in randomly centered small balls, to be found in (9). Let $\tilde{X}_1$ denote a fractional Brownian motion that is independent of X with $\mathcal{L}(X) = \mathcal{L}(\tilde{X}_1)$. Then, for any $s \in [1, \infty)$, one has

$$\mathbb{E}[(-\log \mathbb{P}(\|X - \tilde{X}_1\| \leq \varepsilon | X))^s]^{1/s} \approx -\log \mathbb{P}(\|X\| \leq \varepsilon) \approx \varepsilon^{-1/H} \tag{8}$$

as $\varepsilon \downarrow 0$ (see (9), Theorem 4.2 and Corollary 4.4).

Proof. For a given $\mathbb{D}[0,1]$-valued sequence $(\tilde{w}_n)_{n \in \mathbb{N} \cup \{\infty\}}$, we consider the following coding strategy $\pi^{(r)}(\cdot|(\tilde{w}_n))$: let

$$T^{(r)}(w) := T^{(r)}(w|(\tilde{w}_n)) := \inf\{n \in \mathbb{N} : \|w - \tilde{w}_n\| \leq 1/r^H\},$$

with the convention that the infimum of the empty set is ∞, and set

$$\pi^{(r)}(w) := \pi^{(r)}(w|(\tilde{w}_n)) := \tilde{w}_{T^{(r)}(w)}.$$

Moreover, let $(p_n)_{n \in \mathbb{N}}$ denote the sequence of probability weights defined as

$$p_n = \frac{6}{\pi^2} \frac{1}{n^2}, \qquad n \in \mathbb{N},$$

and set $p_\infty := 0$.

Now we let $(\tilde{X}_n)_{n \in \mathbb{N} \cup \{\infty\}}$ denote independent FBM's that are also independent of X, and analyze the random coding strategies $\pi^{(r)}(\cdot) := \pi^{(r)}(\cdot|(\tilde{X}_n))$. With $T^{(r)} := T^{(r)}(X|(\tilde{X}_n))$ we obtain

$$\hat{X}^{(r)} := \pi^{(r)}(X) = \tilde{X}_{T^{(r)}},$$

and

$$\mathbb{E}[(-\log p_{T^{(r)}})^s]^{1/s} \leq 2\mathbb{E}[(\log T^{(r)})^s]^{1/s} + \log \frac{\pi^2}{6}. \tag{9}$$

Given X, the random time $T^{(r)}$ is geometrically distributed with parameter $\mathbb{P}(\|X - \tilde{X}_1\| \leq 1/r^H|X)$, and due to Lemma A.2 there exists a universal constant $c_1 = c_1(s) < \infty$ for which

$$\mathbb{E}[(\log T^{(r)})^s|X]^{1/s} \leq c_1\,[1 + \log \mathbb{E}[T^{(r)}|X]] = c_1\,[1 + \log 1/\mathbb{P}(\|X - \tilde{X}_1\| \leq 1/r^H|X)].$$

Consequently,

$$\begin{aligned}\mathbb{E}[(\log T^{(r)})^s]^{1/s} &= \mathbb{E}\big[\mathbb{E}[(\log T^{(r)})^s|X]\big]^{1/s} \\ &\leq c_1\,\mathbb{E}[(1 + \log 1/\mathbb{P}(\|X - \tilde{X}_1\| \leq 1/r^H|X))^s]^{1/s} \\ &\leq c_1\,(1 + \mathbb{E}[(-\log \mathbb{P}(\|X - \tilde{X}_1\| \leq 1/r^H|X))^s]^{1/s}).\end{aligned} \tag{10}$$

Due to (8), one has

$$\mathbb{E}[(-\log \mathbb{P}(\|X - \tilde{X}_1\| \leq 1/r^H|X))^s]^{1/s} \approx r,$$

so that (9) and (10) imply that $\mathbb{E}[(-\log p_{T^{(r)}})^s]^{1/s} \lesssim c_2 r$ for some appropriate constant $c_2 < \infty$. In particular, for any $r \geq 0$, we can find a $\mathbb{C}[0,1]$-valued sequence $(\tilde{w}^{(r)})_{n\in\mathbb{N}}$ of pairwise different elements such that

$$\mathbb{E}[(-\log p_{T^{(r)}(X|(\tilde{w}_n^{(r)}))})^s]^{1/s} \leq \mathbb{E}[(-\log p_{T^{(r)}})^s]^{1/s} \lesssim c_2\,r.$$

Now the strategies $\pi^{(r)}(\cdot|(\tilde{w}_n^{(r)}))$ with associated probability weights $p^{(r)}_{\tilde{w}_n^{(r)}} := p_n$ $(n \in \mathbb{N})$ satisfy (7). Moreover, $D^{(e)}(r|\infty) \approx r^{-H}$ follows since

$$\mathbb{H}(\pi^{(r)}(X|(\tilde{w}_n^{(r)}))) \leq \mathbb{E}\big[-\log p^{(r)}_{\pi^{(r)}(X|(\tilde{w}_n^{(r)}))}\big].$$

□

Next we use the coding scheme of Section 2 to prove

Lemma 3.4. *Let $n \in \mathbb{N}$, $r \geq 0$ and $\Delta r \geq 1$. Then*

$$D^{(e)}(n(r + \Delta r)|\infty) \leq n^{-H}\,\frac{e^{\Delta r}}{e^{\Delta r} - 2}\,D^{(e)}(r|\infty). \tag{11}$$

Proof. Fix $\varepsilon > 0$ and let $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$ be a strategy satisfying

$$\big\|\|X - \pi(X)\|_{[0,1]}\big\|_\infty \leq (1 + \varepsilon)D^{(e)}(r|\infty) =: d$$

and

$$\mathbb{H}(\pi(X)) \leq r.$$

Choose $M := \lfloor e^{\Delta r} \rfloor$ and let $\pi^{(n)}$ be as in Section 2. Note that $\Delta r \geq 1$ guarantees that $M \geq e^{\Delta r} - 1 \geq e^{\Delta r}/2$, so that

$$\big\|\|X - \pi^{(n)}(X)\|_{[0,n]}\big\|_\infty \leq \frac{M}{M-1}\,d \leq \frac{e^{\Delta r}}{e^{\Delta r} - 2}(1 + \varepsilon)D^{(e)}(r|\infty).$$

We let $(X_t^{(i)})_{t\in[0,1]} = (X_{i+t} - X_i)_{t\in[0,1]}$ for $i = 1, \ldots, n$, and $(\xi_i)_{i=1,\ldots,n-1}$ be as in Section 2 for $w = X$. Observe that, due to the representation (4),

$$\begin{aligned}\mathbb{H}(\pi^{(n)}(X)) &\le \mathbb{H}(\pi(X^{(0)}), \ldots, \pi(X^{(n-1)}), \xi_1, \ldots, \xi_{n-1})\\ &\le \mathbb{H}(\pi(X^{(0)})) + \cdots + \mathbb{H}(\pi(X^{(n-1)})) + \mathbb{H}(\xi_1, \ldots, \xi_{n-1})\\ &\le nr + \log|\operatorname{range}(\xi_1, \ldots, \xi_{n-1})| \le nr + n\log M\\ &\le n(r + \Delta r).\end{aligned} \tag{12}$$

Now let

$$\alpha_n : \mathbb{D}[0,1] \to \mathbb{D}[0,n],\ f \mapsto \alpha_n(f)(s) = n^H f(s/n)$$

and consider the strategy

$$\tilde\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1],\ f \mapsto \alpha_n^{-1} \circ \pi^{(n)} \circ \alpha_n(f).$$

Since $\alpha_n(X)$ is again a fractional Brownian motion on $[0,n]$, it follows that, a.s.

$$\|X - \tilde\pi(X)\|_{[0,1]} = n^{-H}\,\|\alpha_n(X) - \pi^{(n)}(\alpha_n(X))\|_{[0,n]} \le (1+\varepsilon)\,n^{-H}\frac{e^{\Delta r}}{e^{\Delta r} - 2} D^{(e)}(r|\infty).$$

Moreover,

$$\mathbb{H}(\tilde\pi(X)) = \mathbb{H}(\alpha_n^{-1} \circ \pi^{(n)}(\alpha_n(X))) = \mathbb{H}(\pi^{(n)}(X)) \le r.$$

Since $\varepsilon > 0$ is arbitrary, the proof is complete. □

Proof of Theorem 3.1. For $r \ge 0$, $\Delta r \ge 1$ and $n \in \mathbb{N}$, Lemma 3.4 yields

$$D^{(e)}(n(r+\Delta r)|\infty) \le \frac{1}{n^H}\,\frac{e^{\Delta r}}{e^{\Delta r} - 2}\,D^{(e)}(r|\infty).$$

Now set $\kappa := \liminf_{r\to\infty} r^H\, D^{(e)}(r|\infty)$ which lies in $(0,\infty)$ due to Lemma 3.3. Let $\varepsilon \in (0, 1/2)$ be arbitrary, and choose $r_0, \Delta r \ge 1$ such that

$$\begin{cases} r_0^H\, D^{(e)}(r_0|\infty) \le (1+\varepsilon)\kappa,\\ \Delta r \le \varepsilon r_0 \quad \text{and}\\ e^{-\Delta r} \le \varepsilon.\end{cases}$$

Then

$$\begin{aligned}D^{(e)}((1+\varepsilon)nr_0|\infty) &\le \frac{1}{n^H}\,\frac{1}{1-2\varepsilon} D^{(e)}(r_0|\infty)\\ &\le \frac{1}{\big((1+\varepsilon)nr_0\big)^H}\,\frac{1}{1-2\varepsilon}(1+\varepsilon)^{1+H}\,\kappa\end{aligned}$$

and we obtain that

$$\limsup_{n\to\infty}\big((1+\varepsilon)nr_0\big)^H\, D^{(e)}((1+\varepsilon)nr_0|\infty) \le \frac{(1+\varepsilon)^{1+H}}{1-2\varepsilon}\,\kappa.$$

Let now $r \geq (1+\varepsilon)r_0$ and introduce $\bar{r} = \bar{r}(r) = \min\{(1+\varepsilon)nr_0 : n \in \mathbb{N}, r \leq (1+\varepsilon)nr_0\}$ as well as $\underline{r} = \underline{r}(r) = \max\{(1+\varepsilon)nr_0 : n \in \mathbb{N}, (1+\varepsilon)nr_0 \leq r\}$. Using the monotonicity of $D^{(e)}(r|\infty)$, we conclude that

$$\begin{aligned}\limsup_{r\to\infty} r^H\, D^{(e)}(r|\infty) &\leq \limsup_{r\to\infty} \bar{r}^H\, D^{(e)}(\underline{r}|\infty)\\ &\leq \limsup_{r\to\infty} (\underline{r} + (1+\varepsilon)r_0)^H\, D^{(e)}(\underline{r}|\infty)\\ &\leq \frac{(1+\varepsilon)^{1+H}}{1-2\varepsilon}\,\kappa.\end{aligned}$$

Noticing that $\varepsilon > 0$ is arbitrary finishes the proof. $\square$

4 The quantization problem

Theorem 4.1. *One has for any $s \in (0,\infty)$,*

$$D^{(q)}(r|s) \sim \kappa \frac{1}{r^H}, \qquad r \to \infty.$$

Recall that a strategy π and probability weights (p_w) on the image of π intuitively correspond to a coding scheme which maps an original symbol x onto a prefix-free representation for $\pi(x)$ with codelength of about $-\log_2 p_{\pi(x)}$. The proof of Theorem 4.1 relies on Proposition 4.4. There we show that for good coding schemes $-\log_2 p_{\pi(X)}$ is strongly concentrated around some typical value when r is large. In order to prove the proposition we combine Lemma 3.3 with the following lemma.

Lemma 4.2. *There exist strategies $(\pi^{(r)})_{r\geq 0}$ and probability weights $(p_w^{(r)})$ such that*

$$\big\| \|X - \pi^{(r)}(X)\| \big\|_\infty \leq \kappa \frac{1}{r^H} \quad \textit{and} \quad -\log p^{(r)}_{\pi^{(r)}(X)} \lesssim r, \quad \textit{in probability.}$$

Proof. Let $\varepsilon > 0$ and choose $r_0 \geq 2$ such that

$$\Big(\frac{r_0+1}{r_0-1}\Big)^{1/H} \leq 1 + \frac{\varepsilon}{2}$$

By Theorem 3.1,

$$D^{(e)}((1+\varepsilon/2)r|\infty) \lesssim \kappa \frac{r_0-1}{r_0+1}\,\frac{1}{r^H}$$

In particular, there exists $r_1 \geq r_0 \vee \frac{2}{\varepsilon}\log(r_0+1)$ and a map $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$ such that

$$\big\| \|X - \pi(X)\|_{[0,1]} \big\|_\infty \leq \kappa \frac{r_0-1}{r_0}\,\frac{1}{r_1^H} =: d \quad \text{and} \quad \mathbb{H}(\pi(X)) \leq (1+\varepsilon/2)r_1.$$

For $n \in \mathbb{N}$, let $\pi^{(n)}$ and φ_n be as in Section 2 for $M = \lceil r_0 \rceil$, d and π. Then by (5)

$$\big\| \|X - \pi^{(n)}(X)\|_{[0,n]} \big\|_\infty \leq \kappa \frac{(r_0-1)M}{r_0(M-1)}\,\frac{1}{r_1^H} \leq \kappa\,\frac{1}{r_1^H}. \tag{13}$$

For $\hat{w}^{(0)}, \dots, \hat{w}^{(n-1)} \in \mathrm{im}(\pi)$ and $k_1, \dots, k_{n-1} \in \{-d + \frac{2kd}{M-1} : k = 0, \dots, M-1\}$, let

$$p^{(n)}_{\varphi_n(\hat{w}^{(0)}, \dots, \hat{w}^{(n-1)}, k_1, \dots, k_{n-1})} = \frac{1}{M^{n-1}} \prod_{i=0}^{n-1} \mathbb{P}(\pi(X) = \hat{w}^{(i)}).$$

The $(p^{(n)}_w)$ define probability weights on the image of φ_n. Moreover,

$$-\log p^{(n)}_{(\hat{X}_t)_{t\in[0,n]}} = (n-1)\log M - \sum_{i=0}^{n-1} \log p_{\pi(X^{(i)})}$$

and the ergodic theorem implies

$$\lim_{n\to\infty} -\frac{1}{n} \log p^{(n)}_{(\hat{X}_t)_{t\in[0,n]}} = \log M + \mathbb{H}(\pi(X)), \qquad \text{a.s.}$$

Note that $\log M + \mathbb{H}(\pi(X)) \le (1+\varepsilon) r_1$.

Just as in the proof of Lemma 3.4, we use the self similarity of X to translate the strategy $\pi^{(n)}$ into a strategy for encoding $(X_t)_{t\in[0,1]}$. For $n \in \mathbb{N}$, let

$$\alpha_n : \mathbb{D}[0,1] \to \mathbb{D}[0,n],\ f \mapsto (\alpha_n f)(t) = n^H f(t/n)$$

and consider $\tilde{p}^{(n)}_w := p^{(n)}_{\alpha_n(w)}$ and $\tilde{\pi}^{(n)}(w) := \alpha_n^{-1} \circ \pi^{(n)} \circ \alpha_n(w)$. Then

$$-\log \tilde{p}^{(n)}_{\tilde{\pi}^{(n)}(X)} = -\log p^{(n)}_{\pi^{(n)}(\alpha_n(X))} \lesssim (1+\varepsilon) n r_1, \qquad \text{in probability}$$

and by (13)

$$\begin{aligned}
\big\| \|X - \tilde{\pi}^{(n)}(X)\|_{[0,1]} \big\|_\infty &= \big\| \|\alpha_n^{-1}(\alpha_n(X) - \pi^{(n)}(\alpha_n(X)))\|_{[0,1]} \big\|_\infty \\
&= \frac{1}{n^H} \big\| \|\alpha_n(X) - \pi^{(n)}(\alpha_n(X))\|_{[0,n]} \big\|_\infty \\
&= \frac{1}{n^H} \big\| \|X - \pi^{(n)}(X)\|_{[0,n]} \big\|_\infty \le \kappa \frac{1}{(n r_1)^H}.
\end{aligned}$$

By choosing $\bar{\pi}^{(r)} = \tilde{\pi}^{(n)}$ and $(\bar{p}^{(r)}) = (\tilde{p}^{(n)})$ for $r \in ((n-1)r_1, n r_1]$, one obtains a coding scheme satisfying

$$\big\| \|X - \bar{\pi}^{(r)}(X)\| \big\|_\infty \le \kappa \frac{1}{r^H}$$

and

$$-\log \bar{p}^{(r)}_{\bar{\pi}^{(r)}(X)} \lesssim (1+\varepsilon) r, \quad \text{in probability},$$

so that the assertion follows by a diagonalization argument. □

Remark 4.3. In the above proof, we have constructed a high resolution coding scheme based on a strategy $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$, using the identity $\tilde{\pi}^{(n)} = \alpha_n^{-1} \circ \pi^{(n)} \circ \alpha_n$. This coding scheme leads to a coding error which is at most

$$\frac{M}{M-1} \big\| \|X - \pi(X)\|_{[0,1]} \big\|_\infty n^{-H}. \tag{14}$$

Moreover, the ergodic theorem implies that, for large n, $\tilde\pi^{(n)}(X)$ lies with probability almost one in the typical set $\{w \in \mathbb{D}[0,1] : -\log \tilde p_w^{(n)} \le n(\mathbb{H}(\pi(X)) + \log M + \varepsilon)\}$, where $\varepsilon > 0$ is arbitrarily small. This set is of size $\exp\{n(\mathbb{H}(\pi(X)) + \log M + \varepsilon)\}$, and will serve as a close to optimal high resolution codebook. It remains to control the case where $\tilde\pi^{(n)}(X)$ is not in the typical set. We will do this in the proof of Theorem 4.1 at the end of this section (see (19)).

Proposition 4.4. *For $s \ge 1$ there exist strategies $(\pi^{(r)})_{r\ge 0}$ and probability weights $(p_w^{(r)})$ such that*

$$\big\| \|X - \pi^{(r)}(X))\| \big\|_\infty \le \kappa \frac{1}{r^H} \quad \textit{and} \quad \lim_{r\to\infty} \frac{\mathbb{E}[(-\log p^{(r)}_{\pi^{(r)}(X)})^s]^{1/s}}{r} = 1. \tag{15}$$

In addition, for any $\varepsilon > 0$ one has

$$\lim_{r\to\infty} \sup_{\pi,(p_w)} \mathbb{P}\Big(-\log p_{\pi(X)} \le (1-\varepsilon)r, \|X - \pi(X)\| \le \kappa \frac{1}{r^H}\Big) = 0, \tag{16}$$

where the supremum is taken over all strategies $\pi : \mathbb{C}[0,1] \to \mathbb{D}[0,1]$ and over all sequences of probability weights (p_w).

Proof. Fix $s > 1$ and let for each $R \ge 0$, $\pi_1^{(r)}$ be a strategy and $(p_w^{(r,1)})$ be a sequence of probability weights as in Lemma 4.2. Moreover, let $\pi_2^{(r)}$ and $(p_w^{(r,2)})$ be as in Lemma 3.3 for the moment $2s$. We consider the maps $\kappa_1^{(r)}(w) := -\log p^{(r,1)}_{\pi_1^{(r)}(w)}$ and $\kappa_2^{(r)}(w) := -\log p^{(r,2)}_{\pi_2^{(r)}(w)}$, and set

$$\pi^{(r)}(w) := \begin{cases} \pi_1^{(r)}(w) & \text{if } \kappa_1^{(r)}(w) \le (1+\delta)r, \\ \pi_2^{(r)}(w) & \text{otherwise,} \end{cases}$$

for some fixed $\delta > 0$. Then one obtains, for $p_w^{(r)} = \frac{1}{2}(p_w^{(r,1)} + p_w^{(r,2)})$ and $\mathcal{T}_r := \{w \in \mathbb{C}[0,1] : \kappa_1^{(r)}(w) \le (1+\delta)r\}$,

$$\begin{aligned} \mathbb{E}[(-\log 2p^{(r)}_{\pi^{(r)}(X)})^s]^{1/s} &\le \mathbb{E}[1_{\mathcal{T}_r}(X)\kappa_1^{(r)}(X)^s]^{1/s} + \mathbb{E}[1_{\mathcal{T}_r^c}(X)\kappa_2^{(r)}(X)^s]^{1/s} \\ &\le (1+\delta)r + \mathbb{P}(X \in \mathcal{T}_r^c)^{1/2s}\, \mathbb{E}[\kappa_2^{(r)}(X)^{2s}]^{1/2s}. \end{aligned}$$

The definitions of $\pi_1^{(r)}$ and $\pi_2^{(r)}$ imply that $\lim_{r\to\infty} \mathbb{P}(X \in \mathcal{T}_r^c) = 0$ and $\mathbb{E}[\kappa_2^{(r)}(X)^{2s}]^{1/2s} \approx r$. Consequently,

$$\mathbb{E}[(-\log p^{(r)}_{\pi^{(r)}(X)})^s]^{1/s} \lesssim (1+\delta)r.$$

Since $\delta > 0$ can be chosen arbitrarily small, a diagonalization procedure leads to strategies $\tilde\pi^{(r)}$ and probability weights $(\tilde p_w^{(r)})$ with

$$\big\| \|X - \tilde\pi^{(r)}(X)\|_{[0,1]} \big\|_\infty \le \kappa \frac{1}{r^H} \ \text{ and } \ \mathbb{E}[(-\log \tilde p_{\tilde\pi^{(r)}(X)})^s]^{1/s} \lesssim r.$$

Now the first assertion follows from (16).

It remains to show that for arbitrary strategies $\bar\pi^{(r)}$, $r \ge 0$, and probability weights $(\bar p_w^{(r)})$:

$$\lim_{r\to\infty} \mathbb{P}\Big(-\log \bar p^{(r)}_{\bar\pi^{(r)}(X)} \le (1-\varepsilon)r, \|X - \bar\pi^{(r)}(X)\| \le \kappa \frac{1}{r^H}\Big) = 0. \tag{17}$$

Without loss of generality, we can assume that

$$\big\| \|X - \bar{\pi}^{(r)}(X)\|_{[0,1]} \big\|_\infty \le \kappa \frac{1}{r^H}. \tag{18}$$

Otherwise we modify the map $\bar{\pi}^{(r)}$ for all $w \in \mathbb{C}[0,1]$ with $\|w - \bar{\pi}^{(r)}(w)\| > \kappa\, r^{-H}$ in such a way that (18) be valid. Hereby the probability in (17) increases and it suffices to prove the statement for the modified strategy. Let us consider

$$\pi^{(r)}(w) = \begin{cases} \bar{\pi}^{(r)}(w) & \text{if } \bar{p}^{(r)}_{\bar{\pi}^{(r)}(w)} \ge \tilde{p}^{(r)}_{\tilde{\pi}^{(r)}(w)} \\ \tilde{\pi}^{(r)}(w) & \text{else.} \end{cases}$$

Then the probability weights $p^{(r)} := \frac{1}{2}(\bar{p}^{(r)} + \tilde{p}^{(r)})$ satisfy

$$\mathbb{E}[(-\log 2p^{(r)}_{\pi(X)})^s]^{1/s} \le \mathbb{E}[(-\log \tilde{p}^{(r)}_{\tilde{\pi}(X)})^s]^{1/s} \lesssim r.$$

Recall that

$$\big\| \|X - \pi^{(r)}(X)\|_{[0,1]} \big\|_\infty \le \kappa \frac{1}{r^H},$$

hence by Theorem 3.1, one has $\mathbb{E}[-\log p^{(r)}_{\pi^{(r)}(X)}] \ge \mathbb{H}(\pi^{(r)}(X)) \gtrsim r$. Now the equivalence of moments (see Lemma A.1) implies that

$$-\log p^{(r)}_{\pi^{(r)}(X)} \sim r, \quad \text{in probability,}$$

and

$$-\log \bar{p}^{(r)}_{\bar{\pi}^{(r)}(X)} \ge -\log 2p^{(r)}_{\pi^{(r)}(X)} \gtrsim r, \quad \text{in probability,}$$

which gives (17). $\square$

Proof of Theorem 4.1. We start by proving the lower bound. Fix $s > 0$, let $\mathcal{C}_r$, $r \ge 0$, denote arbitrary codebooks of size e^r, and let $\pi^{(r)} : \mathbb{C}[0,1] \to \mathcal{C}_r$ denote arbitrary strategies. Moreover, let $(p^{(r)}_w)$ be the sequence of probability weights defined as $p^{(r)}_w = 1/|\mathcal{C}_r|$, $w \in \mathcal{C}_r$. Then $-\log p^{(r)}_{\pi^{(r)}(X)} \le r$ a.s., and the above lemma implies that for any $\varepsilon \in (0,1)$,

$$\lim_{r\to\infty} \mathbb{P}\Big(\|X - \pi^{(r)}(X)\| \le \kappa \frac{(1-\varepsilon)^H}{r^H} \Big) = 0.$$

Therefore,

$$\mathbb{E}[\|X - \pi^{(r)}(X)\|^s]^{1/s} \ge \kappa \frac{(1-\varepsilon)^H}{r^H}\, \mathbb{P}\Big(\|X - \pi^{(r)}(X)\| \ge \kappa \frac{(1-\varepsilon)^H}{r^H} \Big)^{1/s} \sim \kappa \frac{(1-\varepsilon)^H}{r^H},$$

which proves the lower bound.

It remains to show that $D^{(q)}(r|s) \lesssim \kappa / r^H$. By Lemma 4.2, there exist strategies $\pi^{(r)}$ and probability weights $(p^{(r)}_w)$ such that

$$\big\| \|X - \pi^{(r)}(X)\| \big\|_\infty \le \kappa \frac{1}{r^H} \quad \text{and} \quad -\log p_{\pi^{(r)}(X)} \lesssim r, \quad \text{in probability.}$$

Furthermore, due to Theorem 4.1 in (8), there exist codebooks $\bar{\mathcal{C}}_r$ of size e^r with

$$\mathbb{E}[\min_{\hat{w}\in\bar{\mathcal{C}}_r} \|X-\hat{w}\|^{2s}]^{1/2s} \approx \frac{1}{r^H}.$$

We consider the codebook $\mathcal{C}_r := \bar{\mathcal{C}}_r \cup \{\hat{w} : -\log p_{\hat{w}}^{(r)} \le (1+\varepsilon/2)r\}$. Clearly, $\mathcal{C}_r$ contains at most $e^r + e^{(1+\varepsilon/2)r}$ elements. Moreover,

$$\begin{aligned}\mathbb{E}[\min_{\hat{w}\in\mathcal{C}_r} \|X-\hat{w}\|^s]^{1/s} &\le \mathbb{E}[1_{\mathcal{C}_r}(\pi^{(r)}(X))\,(\kappa\frac{1}{r^H})^s]^{1/s} \\ &\quad + \mathbb{E}[1_{\mathcal{C}_r^c}(\pi^{(r)}(X)) \min_{\hat{w}\in\bar{\mathcal{C}}_r} \|X-\hat{w}\|^s]^{1/s} \\ &\le \kappa\frac{1}{r^H} + \mathbb{P}(\pi^{(r)}(X)\notin\mathcal{C}_r)^{1/2s}\,\mathbb{E}[\min_{\hat{w}\in\bar{\mathcal{C}}_r}\|X-\hat{w}\|^{2s}]^{1/2s}.\end{aligned} \tag{19}$$

Since $\lim_{r\to\infty}\mathbb{P}(\pi^{(r)}(X)\notin\mathcal{C}_r) = 0$ and the succeeding expectation is of order $\mathcal{O}(1/r^H)$, the second summand is of order $o(1/r^H)$. Therefore, for $r \ge 2/\varepsilon$

$$D^{(q)}((1+\varepsilon)r|s) \le \mathbb{E}[\min_{\hat{w}\in\mathcal{C}_r}\|X-\hat{w}\|^s]^{1/s} \lesssim \kappa\frac{1}{r^H}.$$

By switching from r to $\tilde{r} = (1+\varepsilon)r$, we obtain

$$D^{(q)}(\tilde{r}|s) \lesssim \kappa\,(1+\varepsilon)^H\,\frac{1}{\tilde{r}^H}.$$

Since $\varepsilon > 0$ was arbitrary, the proof is complete. □

5 Implications of the equivalence of moments

In this section we complement Theorem 4.1 by

Theorem 5.1. *For arbitrary $s \in (0,\infty]$, one has*

$$D^{(e)}(r|s) \sim \kappa\frac{1}{r^H}.$$

The proof of this theorem is based on the following general principle: *if the asymptotic quantization error coincides for two different moments $s_1 < s_2$, then all moments $s \le s_2$ lead to the same asymptotic quantization error and the entropy coding problem coincides with the quantization problem for all moments $s \le s_2$.*

Let us prove this relationship in a general setting. E and $\hat{E}$ denoting arbitrary measurable spaces and $d : E\times\hat{E}\to[0,\infty)$ a measurable function, the quantization error for a general E-valued r.v. X under the distortion d is defined as

$$D^{(q)}(r|s) = \inf_{\mathcal{C}\subset E}\mathbb{E}[\min_{\hat{x}\in\mathcal{C}} d(X,\hat{x})^s]^{1/s},$$

where the infimum is taken over all codebooks $\mathcal{C}\subset\hat{E}$ with $|\mathcal{C}| \le e^r$. In order to simplify notations, we abridge

$$d(x,A) = \inf_{y\in A} d(x,y), \qquad x\in E,\ A\subset\hat{E}.$$

Analogously, we denote the entropy coding error by

$$D^{(e)}(r|s) = \inf_{\hat{X}} \mathbb{E}[d(X,\hat{X})^s]^{1/s},$$

where the infimum is taken over all discrete $\hat{E}$-valued r.v. $\hat{X}$ with $\mathbb{H}(\hat{X}) \le r$.
Then Theorem 5.1 is a consequence of Theorem 4.1 and the following theorem.

Theorem 5.2. *Assume that $f : [0,\infty) \to \mathbb{R}_+$ is a decreasing, convex function satisfying*

$$\limsup_{r\to\infty} \frac{-r\frac{\partial^+}{\partial r} f(r)}{f(r)} < \infty, \tag{20}$$

and suppose that, for some $0 < s_1 < s_2$,

$$D^{(q)}(r + \log 2|s_1) \sim D^{(q)}(r|s_2) \gtrsim f(r).$$

Then for any $s > 0$,

$$D^{(e)}(r|s) \gtrsim f(r).$$

We need two technical lemmas.

Lemma 5.3. *Let $0 < s_1 < s_2$ and $f : [0,\infty) \to \mathbb{R}_+$. If*

$$D^{(q)}(r + \log 2|s_1) \sim D^{(q)}(r|s_2) \sim f(r),$$

then for any $\varepsilon > 0$,

$$\lim_{r\to\infty} \sup_{\substack{\mathcal{C}\subset E:\\ |\mathcal{C}|\le e^r}} \mathbb{P}(d(X,\mathcal{C}) \le (1-\varepsilon)f(r)) = 0.$$

Proof. For $r \ge 0$, let $\mathcal{C}_r^*$ denote codebooks of size e^r with

$$\mathbb{E}[d(X,\mathcal{C}_r^*)^{s_2}]^{1/s_2} \sim f(r). \tag{21}$$

Now let $\mathcal{C}_r$ denote arbitrary codebooks of size e^r, and consider the codebooks $\bar{\mathcal{C}}_r := \mathcal{C}_r^* \cup \mathcal{C}_r$. Using (21) and the inequality $s_1 \le s_2$, it follows that

$$f(r) \gtrsim \mathbb{E}[d(X,\bar{\mathcal{C}}_r)^{s_2}]^{1/s_2} \ge \mathbb{E}[d(X,\bar{\mathcal{C}}_r)^{s_1}]^{1/s_1} \ge D^{(q)}(r + \log 2|s_1) \sim f(r).$$

Thus the s_1-th and the s_2-th moment coincide asymptotically and it follows by Lemma A.1 that

$$d(X,\bar{\mathcal{C}}_r) \sim f(r), \quad \text{in probability},$$

so that in particular,

$$d(X,\mathcal{C}_r) \gtrsim f(r), \quad \text{in probability}.$$

□

Lemma 5.4. *Assume that* $f:[0,\infty)\to\mathbb{R}_+$ *is a decreasing, convex function satisfying (20) and*

$$\lim_{r\to\infty}\sup_{\substack{\mathcal{C}\subset\hat{E}:\\ |\mathcal{C}|\le e^r}}\mathbb{P}(d(X,\mathcal{C})\le f(r))=0.$$

Then for any $s>0$,

$$D^{(e)}(r|s)\gtrsim f(r).$$

Proof. The result is a consequence of the technical Lemma A.3. Consider the family $\mathcal{F}$ consisting of all random vectors

$$(A,B)=(d(X,\hat{X})^s,-\log p_{\hat{X}}),$$

where $\hat{X}$ is an arbitrary discrete $\hat{E}$-valued r.v. and (p_w) is an arbitrary sequence of probability weights on the range of $\hat{X}$. Let $\tilde{f}(r)=f(r)^s$, $r\ge 0$. Then for any choice of $\hat{X}$ and (p_w) and an arbitrary $r\ge 0$, the set $\mathcal{C}:=\{w\in\hat{E}:-\log p_w\le r\}$ contains at most e^r elements. Consequently,

$$\mathbb{P}(d(X,\hat{X})^s\le\tilde{f}(r),-\log p_{\hat{X}}\le r)=\mathbb{P}(d(X,\hat{X})\le f(r),\hat{X}\in\mathcal{C})\le\mathbb{P}(d(X,\mathcal{C})\le f(r)).$$

By assumption the right hand side converges to 0 as $r\to\infty$, independently of the choice of $\hat{X}$ and (p_w). Since $\tilde{f}$ satisfies condition (27), Lemma A.3 implies that

$$D^{(e)}(r|s)=\inf_{\hat{X}:\mathbb{H}(\hat{X})\le r}\mathbb{E}[d(X,\hat{X})^s]^{1/s}=\inf_{A\in\mathcal{F}_r}\mathbb{E}[A]^{1/s}\gtrsim\tilde{f}(r)^{1/s}=f(r),$$

where $\mathcal{F}_r=\{A:(A,B)\in\mathcal{F},\ \mathbb{E}B\le r\}$. $\square$

Theorem 5.2 is now an immediate consequence of Lemma 5.3 and Lemma 5.4.

6 Coding with respect to the $L^p[0,1]$-norm distortion

In this section, we consider the coding problem for the fractional Brownian motion X under $L^p[0,1]$-norm distortion for some fixed $p\in[1,\infty)$. In order to treat this approximation problem, we need to introduce Shannon's distortion rate function. It is defined as

$$D(r|s)=\inf\big\|\|X-\hat{X}\|_{L^p[0,1]}\big\|_s,$$

where the infimum is taken over all $\mathbb{D}[0,1]$-valued r.v.'s $\hat{X}$ satisfying the mutual information constraint $I(X;\hat{X})\le r$. Here and elsewhere I denotes the *Shannon mutual information*, defined as

$$I(X;\hat{X})=\begin{cases}\int\log\frac{d\mathbb{P}_{X,\hat{X}}}{d\mathbb{P}_X\otimes\mathbb{P}_{\hat{X}}}\,d\mathbb{P}_{X,\hat{X}} & \text{if } \mathbb{P}_{X,\hat{X}}\ll\mathbb{P}_X\otimes\mathbb{P}_{\hat{X}}\\ \infty & \text{else.}\end{cases}$$

The objective of this section is to prove

Theorem 6.1. *The following limit exists*

$$\kappa_p=\kappa_p(H)=\lim_{r\to\infty}r^H\,D(r|p)\in(0,\infty),\tag{22}$$

and for any $s>0$, *one has*

$$D^{(q)}(r|s)\sim D^{(e)}(r|s)\sim\kappa_p\frac{1}{r^H}.\tag{23}$$

We will first prove that statement (23) is valid for

$$\kappa_p := \liminf_{r\to\infty} r^H\, D(r|p).$$

Since $D(r|p)$ is dominated by $D^{(q)}(r|p)$, the existence of the limit in (22) then follows immediately. Due to Theorem 1.2 in (5), the distortion rate function $D(\cdot|p)$ has the same weak asymptotics as $D^{(q)}(\cdot|p)$. In particular, $D(r|p) \approx r^{-H}$ and κ_p lies in $(0,\infty)$.

We proceed as follows: decomposing X into the two processes

$$X^{(1)} = (X_t - X_{\lfloor t\rfloor})_{t\ge 0} \quad \text{and} \quad X^{(2)} = (X_{\lfloor t\rfloor})_{t\ge 0},$$

we consider the coding problem for $X^{(1)}$ and $X^{(2)}$ in $L^p[0,n]$ ($n\in\mathbb{N}$ being large). We control the coding complexity of the first term via Shannon's source coding theorem (SCT) and use a limit argument in order to show that the coding complexity of $X^{(2)}$ is asymptotically negligible. We recall the SCT in a form which is appropriate for our discussion; for $n\in\mathbb{N}$, let

$$d_p(f,g) = \Big(\int_0^1 |f(t)-g(t)|^p\,dt\Big)^{1/p}$$

and

$$d_{n,p}(f,g) = \Big(\int_0^n |f(t)-g(t)|^p\,\frac{dt}{n}\Big)^{1/p}.$$

Then $d_{n,p}(f,g)^p$, $n\in\mathbb{N}$, is a single letter distortion measure, when interpreting the function $f|_{[0,n)}$ as the concatenation of the "letters" $f^{(0)},\dots,f^{(n-1)}$, where $f^{(i)} = (f(i+t))_{t\in[0,1)}$. Analogously, the process $X^{(1)}$ corresponds to the letters $X^{(1,i)} := (X_{i+t})_{t\in[0,1)}$, $i\in\mathbb{N}_0$. Since $(X^{(1,i)})_{i\in\mathbb{N}_0}$ is an ergodic stationary $\mathbb{C}[0,1)$-valued process, the SCT implies that for fixed $r\ge 0$ and $\varepsilon>0$ there exist codebooks $\mathcal{C}_n\subset\mathbb{D}[0,n]$, $n\in\mathbb{N}$, with at most $\exp\{(1+\varepsilon)nr\}$ elements such that

$$\lim_{n\to\infty}\mathbb{P}(d_{n,p}(X^{(1)},\mathcal{C}_n)^p \le (1+\varepsilon)D(r|p)^p) = 1. \tag{24}$$

The statement is an immediate consequence of the asymptotic equipartition property as stated in (3) (Theorem 1) (see also (1) and (3)).

First we prove a lemma which will later be used to control the coding complexity of $X^{(2)}$.

Lemma 6.2. *Let $(Z_i)_{i\in\mathbb{N}}$ be an ergodic stationary sequence of real-valued r.v.'s and let $S_n = \sum_{i=1}^n Z_i$, $n\in\mathbb{N}_0$. For $\varepsilon>0$ there exist codebooks $\mathcal{C}_n\subset 2\varepsilon\mathbb{Z}^n$ of size $\exp\{2n\mathbb{E}[\log(|Z_1|/2\varepsilon+2)]+2nc\}$ satisfying*

$$\lim_{n\to\infty}\mathbb{P}\big(\min_{\hat{s}\in\mathcal{C}_n}\|S_1^n-\hat{s}\|_{l_\infty^n})\le\varepsilon\big) = 1,$$

where S_1^n denotes $(S_i)_{i=1,\dots,n}$, c is a universal constant and $\|\cdot\|_{l_\infty^n}$ denotes the maximum norm on $\mathbb{R}^n$.

Proof. Let $c>0$ be such that $(p_n)_{n\in\mathbb{Z}}$ defined through

$$p_n = e^{-c}\,\frac{1}{(|n|+1)^2}$$

is a sequence of probability weights, and let

$$\mathcal{C}_n = \big\{\hat{s}_1^n \in 2\varepsilon\mathbb{Z}^n : -\frac{1}{n}\log p^{(n)}_{\hat{s}_1^n} \le 2\mathbb{E}[\log(|Z_1|/2\varepsilon+2)] + 2c\big\},$$

where

$$p^{(n)}_{\hat{s}_1^n} = p_{\hat{s}_1}\prod_{i=2}^n p_{(\hat{s}_i-\hat{s}_{i-1})/2\varepsilon}, \qquad \hat{s}_1^n \in 2\varepsilon\mathbb{Z}^n.$$

Since $(p^{(n)}_{\hat{s}_1^n})$ defines a sequence of probability weights on $2\varepsilon\mathbb{Z}^n$ the set $\mathcal{C}_n$ satisfies the required size constraint. Let $\hat{S}_1^n$ denote a best approximation for S_1^n in the set $2\varepsilon\mathbb{Z}$. Then always $\|S_1^n - \hat{S}_1^n\|_{l_\infty^n} \le \varepsilon$. Note that

$$-\log p^{(n)}_{\hat{S}_1^n} \le 2\sum_{i=1}^n \log(|Z_i|/2\varepsilon+2) + nc$$

so that the ergodic theorem implies that $\lim_{n\to\infty}\mathbb{P}(\hat{S}_1^n \in \mathcal{C}_n) = 1$ which implies the assertion. □

We now use the SCT combined with the previous lemma to construct codebooks that guarantee almost optimal reconstructions with a high probability.

Lemma 6.3. *For any $\varepsilon > 0$ there exist codebooks $\mathcal{C}_r$, $r \ge 0$, of size e^r such that*

$$\lim_{r\to\infty}\mathbb{P}(d_p(X,\mathcal{C}_r) \le (1+\varepsilon)\kappa_p r^{-H}) = 1.$$

Proof. Let $\varepsilon > 0$ be arbitrary and c be as in Lemma 6.2. We fix $r_0 \ge \big(\frac{4\varepsilon\kappa_p}{\mathbb{E}|X_1|}\big)^{1/H}$ such that

$$\varepsilon\kappa_p r^{-H} \ge e^{-\varepsilon r + c + \log\mathbb{E}|X_1|} \quad\text{and}\quad D(r_0|p) \le (1+\varepsilon)\kappa_p r_0^{-H}. \tag{25}$$

We decompose X into the two processes

$$X_t^{(1)} = X_t - X_{\lfloor t\rfloor} \quad\text{and}\quad X_t^{(2)} = X_{\lfloor t\rfloor}.$$

Due to the SCT (24), there exist codebooks $\mathcal{C}_n^{(1)} \subset \mathbb{D}[0,n]$ of size $\exp\{(1+\varepsilon)nr_0\}$ satisfying

$$\lim_{n\to\infty}\mathbb{P}(d_{n,p}(X^{(1)},\mathcal{C}_n^{(1)})^p \le (1+2\varepsilon)^p\kappa_p^p r_0^{-pH}) = 1.$$

We apply Lemma 6.2 for $\varepsilon' := \varepsilon\kappa_p r_0^{-H}$. Note that

$$\mathbb{E}\log\Big(\frac{|X_1|}{2\varepsilon'}+2\Big) + c \le \log\Big(\frac{\mathbb{E}|X_1|}{2\varepsilon'}+2\Big) + c$$

Since $r_0^H \ge \frac{4\varepsilon\kappa_p}{\mathbb{E}|X_1|}$, it follows that $\frac{\mathbb{E}|X_1|}{2\varepsilon'} = \frac{r_0^H\,\mathbb{E}|X_1|}{2\varepsilon\kappa_p} \ge 2$, so that

$$\begin{aligned}\mathbb{E}\log\Big(\frac{|X_1|}{2\varepsilon'}+2\Big) + c &\le \log\Big(\frac{\mathbb{E}|X_1|}{\varepsilon'}\Big) + c\\ &= -\log(\varepsilon\kappa_p r_0^{-H}) + c + \log\mathbb{E}|X_1| \le \varepsilon r,\end{aligned}$$

due to (25). Hence, there exist codebooks $\mathcal{C}_n^{(2)} \subset \mathbb{D}[0,n]$ of size $\exp\{\varepsilon n r_0\}$ with

$$\lim_{n\to\infty} \mathbb{P}\Big(d_{n,p}(X^{(2)}, \mathcal{C}_n^{(2)}) \le \varepsilon \kappa_p \frac{1}{r_0^H}\Big) = 1.$$

Let now $\tilde{\mathcal{C}}_n := \mathcal{C}_n^{(1)} + \mathcal{C}_n^{(2)}$ denote the Minkowski sum of the sets $\mathcal{C}_n^{(1)}$ and $\mathcal{C}_n^{(2)}$. Then $|\tilde{\mathcal{C}}_n| \le \exp\{(1+2\varepsilon)nr_0\}$, and one has

$$\mathbb{P}(d_{n,p}(X, \tilde{\mathcal{C}}_n) \le (1+3\varepsilon)\kappa_p r_0^{-H}) \ge \mathbb{P}(d_{n,p}(X^{(1)}, \mathcal{C}_n^{(1)}) \le (1+2\varepsilon)\kappa_p r_0^{-H} \text{ and } d_{n,p}(X^{(2)}, \mathcal{C}_n^{(2)}) \le \varepsilon\kappa_p r_0^{-H}) \to 1.$$

Consider the isometric isomorphism

$$\beta_n : L^p[0,1] \to (L^p[0,n], d_{n,p}),\ f \mapsto f(nt),$$

and the codebooks $\mathcal{C}_n \subset \mathbb{D}[0,1]$ given by

$$\mathcal{C}_n = \{n^{-H}\beta_n^{-1}(\hat{w}) : \hat{w} \in \tilde{\mathcal{C}}_n\}$$

Then $\tilde{X}^{(n)} = n^{-H}\beta_n^{-1}(X)$ is a fractional Brownian motion and one has

$$d_p(\tilde{X}^{(n)}, \mathcal{C}_n) = d_{n,p}(\beta_n(\tilde{X}^{(n)}), \beta_n(\mathcal{C}_n)) = n^{-H} d_{n,p}(X, \tilde{\mathcal{C}}_n).$$

Hence, the codebooks $\mathcal{C}_n$ are of size $\exp\{(1+2\varepsilon)nr_0\}$ and satisfy

$$\mathbb{P}(d_p(X, \mathcal{C}_n) \le (1+3\varepsilon)\kappa_p (nr_0)^{-H})) = \mathbb{P}(d_{n,p}(X, \tilde{\mathcal{C}}_n) \le (1+3\varepsilon)\kappa_p r_0^{-H}) \to 0$$

as $n \to \infty$. Now the general statement follows by an interpolation argument similar to that used at the end of the proof of Theorem 3.1. □

Proof of Theorem 6.1. Let $s \ge 1$ be arbitrary, let $\mathcal{C}_r^{(1)}$ be as in the above lemma for some fixed $\varepsilon > 0$. Moreover, we let $\mathcal{C}_r^{(2)}$ denote codebooks of size e^r with

$$\mathbb{E}[d_p(X, \mathcal{C}_r^{(2)})^{2s}]^{1/(2s)} \approx \frac{1}{r^H}.$$

Then the codebooks $\mathcal{C}_r := \mathcal{C}_r^{(1)} \cup \mathcal{C}_r^{(2)}$ contain at most $2e^r$ elements and satisfy, in analogy to the proof of Theorem 4.1 (see (19)),

$$\mathbb{E}[d_p(X, \mathcal{C}_r)^s]^{1/s} \lesssim (1+\varepsilon)\kappa_p \frac{1}{r^H}, \qquad r \to \infty.$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$D^{(q)}(r|s) \lesssim \kappa_p \frac{1}{r^H}.$$

For $s \ge p$ the quantization error is greater than the distortion rate function $D(r|p)$, so that the former inequality extends to

$$\lim_{r\to\infty} r^H\, D^{(q)}(r|s) = \kappa_p.$$

In particular, we obtain the asymptotic equivalence of all moments s_1, s_2 greater or equal to p. Next, an application of Theorem 5.2 with $d(f,g) = d_p(f,g)^s$ implies that for any $s > 0$,

$$D^{(e)}(r|s) \gtrsim \kappa_p \frac{1}{r^H},$$

which establishes the assertion. □

Appendix

Lemma A.1. *For $r \geq 0$, let A_r denote $[0,\infty)$-valued r.v.'s. If one has, for $0 < s_1 < s_2$ and some function $f : [0,\infty) \to \mathbb{R}_+$,*

$$\mathbb{E}[A_r^{s_1}]^{1/s_1} \sim \mathbb{E}[A_r^{s_2}]^{1/s_2} \sim f(r), \tag{26}$$

then

$$A_r \sim f(r), \text{ in probability.}$$

Proof. Consider

$$\tilde{A}_r := A_r^{s_1}/\mathbb{E}[A_r^{s_1}],$$

and $\tilde{s}_2 = s_2/s_1$. Then (26) implies that

$$\mathbb{E}[\tilde{A}_r^{\tilde{s}_2}]^{1/\tilde{s}_2} \sim \mathbb{E}[\tilde{A}_r] = 1$$

so that by a classical result $\lim_{r\to\infty} \tilde{A}_r = 1$ in $L^{\tilde{s}_2}(\mathbb{P})$ or, equivalently, $\lim_{r\to\infty} A_r/f(r) = 1$ in $L^{s_2}(\mathbb{P})$. This immediately implies the result. □

Lemma A.2. *Let $s \geq 1$. There exists a constant $c = c(s) < \infty$ such that for all $[1,\infty)$-valued r.v.'s Z one has*

$$\mathbb{E}[(\log Z)^s]^{1/s} \leq c\,[1 + \log \mathbb{E}[Z]].$$

Proof. Using elementary analysis, there exists a positive constant $c_1 = c_1(s) < \infty$ such that $\psi(x) := (\log x)^s + c_1 \log x$, $x \in [1,\infty)$, is concave. For any $[1,\infty)$-valued r.v. Z, Jensen's inequality then yields

$$\begin{aligned}\mathbb{E}[(\log Z)^s]^{1/s} &\leq \mathbb{E}[\psi(Z)]^{1/s} \leq \psi(\mathbb{E}[Z])^{1/s}\\ &\leq \log \mathbb{E}[Z] + c_1^{1/s}(\log \mathbb{E}[Z])^{1/s} \leq c\,[1 + \log \mathbb{E}[Z]],\end{aligned}$$

where $c = c(s) < \infty$ is an appropriate universal constant. □

Lemma A.3. *Let $f : [0,\infty) \to \mathbb{R}_+$ be a decreasing, convex function satisfying $\lim_{r\to\infty} f(r) = 0$ and*

$$\limsup_{r\to\infty} \frac{-r\,\frac{\partial^+}{\partial r} f(r)}{f(r)} < \infty, \tag{27}$$

and $\mathcal{F}$ be a family of $[0,\infty]^2$-valued random variables for which

$$\lim_{r\to\infty} \sup_{(A,B)\in\mathcal{F}} \mathbb{P}(A \leq f(r), B \leq r) = 0. \tag{28}$$

Then the sets of random variables $\mathcal{F}_r$ defined for $r \geq 0$ through

$$\mathcal{F}_r := \{A : (A,B) \in \mathcal{F},\ \mathbb{E}B \leq r\}$$

satisfy

$$\inf_{A\in\mathcal{F}_r} \mathbb{E}A \gtrsim f(r)$$

as $r \to \infty$.

Proof. Fix $R > 0$, positive integers I and N, and define $\lambda := -\frac{\partial^+}{\partial r} f(R)$,

$$r_i := \frac{i+N}{N} R, \qquad i = -N, -N+1, \ldots.$$

For $(A, B) \in \mathcal{F}$ with $\mathbb{E}B \leq r$, we define

$$\mathcal{T}_{A,B} := \{\nexists i \in \{-N+1, \ldots, I\} \text{ such that } A \leq f(r_i) \text{ and } B \leq r_i\}.$$

Then we have

$$\begin{aligned}
\mathbb{E}\big[A + \lambda B\big] &\geq \sum_{i=-N}^{I-1} \mathbb{E}\big[1_{\mathcal{T}_{A,B}} 1_{[r_i, r_{i+1})}(B)(A + \lambda r_i)\big] \\
&\geq \sum_{i=-N}^{I-1} \mathbb{E}\big[1_{\mathcal{T}_{A,B}} 1_{[r_i, r_{i+1})}(B)(f(r_{i+1}) + \lambda r_i)\big] \\
&= \sum_{i=-N}^{I-1} \mathbb{E}\Big[1_{\mathcal{T}_{A,B}} 1_{[r_i, r_{i+1})}(B)(f(r_{i+1}) + \lambda r_{i+1} - \lambda \frac{R}{N})\Big] \\
&\geq \sum_{i=-N}^{I-1} \mathbb{E}\Big[1_{\mathcal{T}_{A,B}} 1_{[r_i, r_{i+1})}(B)(f(R) + \lambda R - \lambda \frac{R}{N})\Big],
\end{aligned}$$

where the last inequality follows from the fact that

$$f(R) + \lambda R = \inf_{r \geq 0} \left[f(r) + \lambda r\right]$$

by the definition of λ and the convexity of f. Now, fix $\varepsilon > 0$ and pick $N \geq 1/\varepsilon$, $I \geq 2N/\varepsilon$ and R_0 so large that

$$\mathbb{P}(\mathcal{T}_{A,B}) \geq 1 - \frac{\varepsilon}{2} \text{ for all } R \geq R_0 \text{ and all } (A, B) \in \mathcal{F}_R.$$

Using Chebychev's inequality, we then obtain for $R \geq R_0$,

$$\begin{aligned}
\mathbb{E}[A + \lambda B] &\geq (1 - \varepsilon)(f(R) + \lambda R) \left(1 - \mathbb{P}\left(\mathcal{T}^c\right) - \mathbb{P}\left(B \geq R \frac{I}{N}\right)\right) \\
&\geq (1 - \varepsilon)(f(R) + \lambda R) \left(1 - \frac{\varepsilon}{2} - \frac{\varepsilon}{2}\right).
\end{aligned}$$

Hence,

$$\lambda R + \mathbb{E}A \geq (1 - \varepsilon)^2 \left(f(R) + \lambda R\right)$$

and therefore

$$\mathbb{E}A \geq (1 - \varepsilon)^2 f(R) + \lambda R \left((1 - \varepsilon)^2 - 1\right).$$

Using the definition of λ and (27), as well as the fact that $\varepsilon > 0$ is arbitrary, the conclusion follows. $\square$

References

[1] T. M. Cover and J. A. Thomas. *Elements of information theory.* Wiley Series in Telecommunications. New York: John Wiley & Sons, Inc., 1991.

[2] J. Creutzig. Approximation of Gaussian random vectors in Banach spaces. Ph.D. Dissertation, Friedrich-Schiller-Universität Jena, 2002.

[3] A. Dembo and I. Kontoyiannis. Source coding, large deviations, and approximate pattern matching. *IEEE Trans. Inform. Theory*, 48(6):1590–1615, 2002. Special issue on Shannon theory: perspective, trends, and applications.

[4] S. Dereich. High resolution coding of stochastic processes and small ball probabilities. Ph.D. Dissertation, TU Berlin,
URL: http://edocs.tu-berlin.de/diss/2003/dereich_steffen.htm, 2003.

[5] S. Dereich. Asymptotic behavior of the distortion-rate function for Gaussian processes in Banach spaces. *Bull. Sci. Math.*, 129(10):791–803, 2005.

[6] S. Dereich. The coding complexity of diffusion processes under $L^p[0,1]$-norm distortion. Preprint, 2006.

[7] S. Dereich. The coding complexity of diffusion processes under supremum norm distortion. Preprint, 2006.

[8] S. Dereich, F. Fehringer, A. Matoussi, and M. Scheutzow. On the link between small ball probabilities and the quantization problem for Gaussian measures on Banach spaces. *J. Theoret. Probab.*, 16(1):249–265, 2003.

[9] S. Dereich and A. Lifshits. Probabilities of randomly centered small balls and quantization in Banach spaces. *Ann. Probab.*, 33(4):1397–1421, 2005.

[10] S. Dereich, T. Müller-Gronbach, and K. Ritter. Infinite-dimensional quadrature and quantization. Preprint, arXiv:math.PR/0601240, 2006.

[11] S. Graf and H. Luschgy. *Foundations of quantization for probability distributions.* Lecture Notes in Mathematics 1730, Berlin: Springer, 2000.

[12] S. Graf, H. Luschgy, and G. Pagès. Optimal quantizers for Radon random vectors in a Banach space. Preprint, 2005.

[13] R. M. Gray and D. L. Neuhoff. Quantization. *IEEE Trans. Inf. Theory*, 44(6):2325–2383, 1998.

[14] J. Kuelbs and W. V. Li. Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.*, 116(1):133–157, 1993.

[15] W. V. Li and Q.-M. Shao. Gaussian processes: inequalities, small ball probabilities and applications. In *Stochastic processes: theory and methods*, volume 19 of *Handbook of Statist.*, pages 533–597. North-Holland, Amsterdam, 2001.

[16] H. Luschgy and G. Pagès. Functional quantization of Gaussian processes. *J. Funct. Anal.*, 196(2):486–531, 2002.

[17] H. Luschgy and G. Pagès. Sharp asymptotics of the functional quantization problem for Gaussian processes. *Ann. Probab.*, 32(2):1574–1599, 2004.

[18] G. Pagès and J. Printems. Functional quantization for pricing derivatives. Preprint, Université de Paris VI, LPMA no. 930, 2004.