The coding complexity of diffusion processes under supremum norm distortion

Steffen Dereich*

Institut für Mathematik, MA 7-5, Fakultät II, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

Received 5 April 2006; received in revised form 3 July 2007; accepted 3 July 2007
Available online 12 July 2007

Abstract

We investigate the high resolution quantization and entropy coding problem for solutions of stochastic differential equations under supremum norm distortion. Tight asymptotic formulas are found under mild regularity assumptions. The main technical tool is a decoupling method which allows us to relate the complexity of the diffusion process to that of the Wiener process. The technique is also applicable when considering the $L^p[0, 1]$-norm distortion.

MSC: 60G35; 41A25; 94A15

Keywords: High resolution quantization; Complexity; Stochastic processes; Stochastic differential equations

1. Introduction

In this article, we study the high resolution quantization and entropy coding problem for $\mathbb{R}$-valued stochastic processes $X$ (original) that are solutions of stochastic differential equations. For $t > 0$ we let $C[0, t]$ denote the set of real-valued continuous functions defined on $[0, t]$, and let $\| \cdot \|_{[0,t]}$ denote the corresponding supremum norm that is $\|f\|_{[0,t]} = \sup_{u \in [0,t]} |f(u)|$. Mostly we will consider $\| \cdot \| = \| \cdot \|_{[0,1]}$. Moreover, we shall write $\| \cdot \|_{L^p[0,t]}$ and $\| \cdot \|_{L^p(P)}$ for the $L^p$-norm on the interval $[0, t]$ and the $L^p$-norm induced by the measure $P$, respectively.

* Tel.: +49 3031425784; fax: +49 3031421695.
E-mail address: dereich@math.tu-berlin.de.

0304-4149/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2007.07.003
The article is devoted to the analysis of the quantization error
\[ D^{(q)}(r|s) = \inf \{ \mathbb{E}[\|X - \hat{X}\|^s]^{1/s} : \hat{X} \in \mathcal{C}[0,1] \text{-valued r.v. with } |\text{range } \hat{X}| \leq e^r \}, \]
and the entropy coding error
\[ D^{(e)}(r|s) = \inf \{ \mathbb{E}[\|X - \hat{X}\|^s]^{1/s} : \hat{X} \in \mathcal{C}[0,1] \text{-valued r.v. with } H(\hat{X}) \leq r \}. \]
Both approximation quantities depend on two parameters: the rate \( r \geq 0 \) and the moment \( s > 0 \).

Here and elsewhere \( H(\hat{X}) \) denotes the entropy of \( \hat{X} \) in the natural basis, that is
\[ H(\hat{X}) = \begin{cases} \sum_{x \in \text{range}(\hat{X})} p_x \log(1/p_x) & \text{if } \hat{X} \text{ is discrete} \\ \infty & \text{otherwise,} \end{cases} \]
where the \((p_x)\) denote the probability weights of \( \hat{X} \). Former research on quantization and entropy coding comprises the construction of efficient approximation schemes, properties of optimal schemes and asymptotic formulas for the corresponding approximation quantities. The quantization problem and entropy coding problem appear naturally in information theory for instance when digitizing analog signals or for reducing the amount of information due to a given channel capacity constraint. Beside these applications good quantization schemes can be used to carry out a variance reduction for certain Monte Carlo methods or to obtain quasi-Monte Carlo algorithms (see for instance \cite{20}). In this article we investigate the asymptotic behavior of the above approximation quantities when the rate tends to infinity: the high resolution coding problem. Our analysis is intended to shed new light on the functional coding problems and to provide benchmarks for the efficiency of particular coding schemes.

An overview on quantization can be found in the monograph by Graf and Luschgy \cite{10} (see also \cite{9,11}). For a general account on information theory one might consult the books by Cover and Thomas \cite{2} and by Ihara \cite{12}.

The research on the functional quantization problem started at the beginning of the 21st century with an article by Luschgy and Pagès \cite{17} and the dissertation by Fehringer \cite{7}. As was found in \cite{4} (see also \cite{5}) for Gaussian originals on separable infinite dimensional Banach spaces, the quantization and entropy coding errors are typically weakly equivalent to the inverse of the small ball function
\[ \varphi(\varepsilon) = -\log \mathbb{P}(\|X\| \leq \varepsilon) \quad (\varepsilon > 0). \]
Thanks to the research on small ball probabilities, the weak asymptotics are known for many Gaussian processes (see for instance \cite{16}). Moreover, the above approximation numbers are related to several other approximation quantities describing the complexity of the Gaussian original or its generating operator (see for instance \cite{3}).

In the particular case where \( X \) is a Brownian motion in \( C[0,1] \), stronger results are known due to \cite{6}. In that case one has
\[ \lim_{r \to \infty} \sqrt{r} D^{(q)}(r|s) = \lim_{r \to \infty} \sqrt{r} D^{(e)}(r|s) = K \] (1)
for some constant \( K \in [\pi/\sqrt{8}, \pi] \) not depending on the moment \( s > 0 \).

Let us now focus on the coding complexity of solutions \((X_t)_{t \in [0,1]}\) of stochastic differential equations. Luschgy and Pagès \cite{18} considered a class of one-dimensional diffusions with continuously differentiable diffusion coefficients. Their coding strategy is based on the Lamperti
transform which maps the original \((X_t)\) onto a process \((\tilde{X}_t)\) being a Brownian motion plus drift term. Approximating the process \(\tilde{X}\) by some close process \(\hat{X}\) and inverting the Lamperti transform for \(\hat{X}\) leads to a “good” reconstruction of the original. Under a regularity assumption on the Lamperti transform (assumption (3.8)) and the assumption that the diffusion coefficient is strictly bounded away from 0, they are able to prove that

\[
D^{(q)}(r|s) \approx \frac{1}{\sqrt{r}}, \quad r \to \infty,
\]

for any \(s \in [1, \infty)\).

In contrast to [18], we use the Doob–Meyer decomposition to decompose the diffusion \(X\) into a bounded variation term and a martingale which we represent as a time-changed Wiener process. Our approach leads to an explicit formula for the asymptotics in the quantization and entropy coding problem in terms of the average diffusion coefficient seen by the process. The techniques rely on the above representation. Since such decompositions may also be presented for multidimensional diffusions having a scalar diffusion coefficient, this setting is also covered by the present investigations. However, we shall only carry out the proof in the one-dimensional setting for the sake of notational simplicity.

Let us now fix the notation. Let \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) be a complete filtered probability space satisfying the usual hypotheses, i.e. \(\mathcal{F}_0\) contains all \(\mathbb{P}\)-null sets of \(\mathcal{F}\) and \((\mathcal{F}_t)\) is right continuous. Let \((\tilde{W}_t)_{t \geq 0}\) be a one-dimensional \((\mathcal{F}_t)\)-Wiener process. We denote by \(\sigma : \mathbb{R} \times [0, \infty) \to \mathbb{R}\) and \(b : \mathbb{R} \times [0, \infty) \to \mathbb{R}\) two measurable functions, and assume that \((X_t)_{t \geq 0}\) is an \((\mathcal{F}_t)\)-adapted semimartingale solving the integral equation

\[
X_t = x_0 + \int_0^t b(X_u, u)du + \int_0^t \sigma(X_u, u)d\tilde{W}_u \quad (t \geq 0)
\]

for some deterministic starting point \(x_0 \in \mathbb{R}\). For ease of notation, we use \(b_t := b(X_t, t)\) and \(\sigma_t := \sigma(X_t, t)\) for \(t \geq 0\). \((X_t)_{t \in [0,1]}\) represents the original process which is to be approximated by some discrete r.v. \(\hat{X}\), the reconstruction.

We need to introduce the above approximation numbers also for processes other than the original diffusion. Moreover, we shall use distortion measures over random time horizons, defined in the following way: For a \(C[0, \infty)\)-valued random vector \(Z\), a \([0, \infty)\)-valued r.v. \(\tau\), \(s > 0\), and \(r \geq 0\), let

\[
D^{(q)}(r|Z, \tau, s) = \inf_{\hat{Z}} \mathbb{E}[\|Z - \hat{Z}\|_{\tau;[0,1]}^s]^{1/s},
\]

where the infimum is taken over all discrete, \(C[0, \infty)\)-valued r.v.’s \(\hat{Z}\) with

\[
|\text{range}(\hat{Z})| \leq e^r.
\]

We call \(D^{(q)}(r|Z, \tau, s)\) the \(s\)-th moment quantization error for the rate \(r\), source \(Z\) and time \(\tau\). We use analogous notation for the entropy coding error.

From now on, we make the following technical assumption:

**Assumption C.** There exist constants \(\beta \in (0, 1]\) and \(L < \infty\) such that for \(x, x' \in \mathbb{R}\) and \(t, t' \in [0, 1]\),

\[
|b(x, t)| \leq L(|x| + 1), \quad |\sigma(0, 0)| \leq L \quad \text{and}
|\sigma(x, t) - \sigma(x', t')| \leq L(|x - x'|^{\beta} + |x - x'| + |t - t'|^{\beta})
\]

(3)
Note that one can translate the integral equation into an integral equation with starting point 0 by applying a shift along \(-x_0\). On doing so condition (C) remains valid and we can and will assume that the process starts at the origin. As a consequence of Assumption C all moments \(\mathbb{E}[^kX](s \geq 1)\) are finite, which we shall use without further mention. Additionally, we assume that the process \((\sigma_t)_{t \in [0, 1]}\) is not indistinguishable from the constant 0-function, since otherwise the problem is trivial.

Note that Assumption C ensures neither existence nor uniqueness of the solution of the stochastic differential equation (2). More information on existence and uniqueness of stochastic differential equations can be found for instance in [19]. Our main objective is to prove

**Theorem 1.1.** For each \(s > 0\) one has

\[
\lim_{r \to \infty} \sqrt{r} \, D^{(q)}(r|s) = K \| \sigma \|_{L^2[0, 1]} \| L^1(\mathbb{P})
\]

and

\[
\lim_{r \to \infty} \sqrt{r} \, D^{(c)}(r|s) = K \| \sigma \|_{L^2[0, 1]} \| L^{2/(s+2)}(\mathbb{P})
\]

where \(K\) is the real constant appearing in (1).

Let us now describe the coding scheme. We write \((X_t)\) in its Doob–Meyer decomposition \(X_t = M_t + A_t\), where

\[
M_t = \int_0^t \sigma(X_s, s) \, d\hat{W}_s \quad \text{and} \quad A_t = \int_0^t b(X_s, s) \, ds.
\]

We shall see that the dominant term in the quantization problem is the continuous martingale \((M_t)\). As is well known, we can represent \((M_t)\) as a time-changed Wiener process. Let

\[
\varphi(t) = \int_0^t \sigma_u^2 \, du,
\]

and observe that one can ensure that \(\lim_{t \to \infty} \varphi(t) = \infty\) by changing the diffusion coefficient outside the time window \([0, 1]\) without changing \((X_t)_{t \in [0, 1]}\). Letting

\[
\varphi^{-1}(t) := \inf \left\{ u \geq 0 : \int_0^u \sigma_v^2 \, dv \geq t \right\}
\]

the process \((W_t)_{t \geq 0}\) defined as \(W_t = M_{\varphi^{-1}(t)}\) is a \((\mathcal{F}^W_t)\)-Wiener process.

Now the coding scheme for \(M\) can be decomposed into the following two steps:

1. Approximate the real time transform \(\varphi\) by some random monotone, continuous function \(\hat{\varphi} \in C[0, 1]\).
2. Approximate \((W_t)_{t \in [0, \tau]}\) \((\tau := \hat{\varphi}(1))\) by \((\hat{W}_t)_{t \in [0, \tau]}\).

Then \(\hat{M} = \hat{W}_{\hat{\varphi}(\cdot)}\) is considered as the reconstruction for \(M\), and the coding error can be controlled by

\[
\| M - \hat{M} \|_{[0, 1]} \leq \| W_{\varphi(\cdot)} - W_{\hat{\varphi}(\cdot)} \|_{[0, 1]} + \| W_{\hat{\varphi}(\cdot)} - \hat{W}_{\hat{\phi}(\cdot)} \|_{[0, 1]}
\]

\[
= \| W_{\varphi(\cdot)} - W_{\hat{\phi}(\cdot)} \|_{[0, 1]} + \| W - \hat{W} \|_{[0, \tau]}.
\]

We shall see that the first term in the above sum is asymptotically negligible, so that the asymptotics are governed by the second term. We need strong estimates for the second term, whereas weak estimates suffice for the first term.
Remark 1.2. Our analysis works equally well when $X$ is a $d$-dimensional diffusion with scalar diffusion coefficient $\sigma$. In that case $X$ can be written as a sum of a finite variation term and a time-changed $d$-dimensional Wiener process. The same techniques can be applied and thus Theorem 1.1 is also valid for $d$-dimensional diffusions with scalar diffusion coefficient with a different constant $K \in (0, \infty)$.

The article is outlined as follows. The proof of Theorem 1.1 relies on a particular representation given in Theorem 7.1. Once we have proven Theorem 7.1 we use it to conclude all assertions. However, the proof of Theorem 7.1 requires a couple of estimates. Section 2 starts with an upper bound for the quantization error based on entropy numbers of compact embeddings. This estimate enables us to control some of the asymptotically negligible terms. Next, we provide an estimate for the moments of the $\alpha$-Hölder norm of continuous martingales. The coding scheme for $\varphi$ is introduced and analyzed in Section 4. The next section is devoted to the analysis of $E[\|W(\cdot) - W(\hat{\varphi}(\cdot))\|_{s}^{1/s}]$ for “good” reconstructions $\hat{\varphi}$ of $\varphi$. Finally, results in the theory of enlargements of filtrations are used to decouple the approximate time transform $\hat{\varphi}$ and the Wiener process $W$.

Thereafter we write $f \sim g$ iff $\lim f / g = 1$, while $f \precsim g$ stands for $\limsup f / g \leq 1$. Finally, $f \approx g$ means

$$0 < \liminf \frac{f}{g} \leq \limsup \frac{f}{g} < \infty,$$

and $f \precsim g$ means

$$\limsup \frac{f}{g} < \infty.$$

Moreover, we use the Landau symbols $o$ and $O$.

2. Entropy numbers and the quantization problem

In this section we construct codebooks based on appropriate $\varepsilon$-nets and control their asymptotic efficiency. The corresponding estimate provides the main technique for controlling the complexity of the asymptotically negligible terms when coding diffusions.

Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ denote normed vector spaces such that $E$ is compactly embedded into $F$. We endow $E$ with its Borel $\sigma$-field and denote by $e_n = e_n(E, F)$ the entropy numbers of the embedding, that is

$$e_n(E, F) := \inf \left\{ \varepsilon > 0 : \exists x_1, \ldots, x_{2^n-1} \in F \text{ s.t. } B_E(0, 1) \subset \bigcup_{i=1}^{2^{n-1}} B_F(x_i, \varepsilon) \right\}.$$

Lemma 2.1. Let $\alpha > 0$, and assume that $E$ is compactly embedded into $F$ with

$$e_n(E, F) \precsim n^{-\alpha}, \quad n \to \infty. \quad (4)$$

Then for all $\tilde{s} > s > 0$, there exists a constant $c = c(s, \tilde{s}) < \infty$ such that for all $E$-valued r.v.’s $Z$ and $r \geq 0$, one has

$$D^{(q)}(r|s) \leq c E[\|Z\|_{E}^{\tilde{s}}]^{1/\tilde{s}} \frac{1}{1 + r^{\alpha}}, \quad (5)$$

where $D^{(q)}(r|s)$ denotes the $s$-th moment quantization error of $Z$ under the $\| \cdot \|_F$-norm based distortion.
Proof. Fix $\tilde{s} > s > 0$. Notice that it suffices to prove the existence of a constant $c < \infty$ such that for any $E$-valued r.v. $Z$ with $E[\|Z\|_E^{\tilde{s}}] = 1$,
\[
D^{(q)}(r|s) \leq c \frac{1}{1 + r^\alpha},
\]
since the general statement then follows by a scaling argument.

Notice that $e_n = e_n(E, F)$ is bounded by the norm $\|\text{id} : E \to F\| =: \xi$. Using assumption (4), there exists $c_1 < \infty$ with
\[
e_n \leq c_1 n^{-\alpha} \quad (n \in \mathbb{N}). \tag{6}
\]
Let $U = B_E(0, 1)$ and
\[
N(\varepsilon, A) = \min \{|C| : A \subset C + B_F(0, \varepsilon), C \subset F\} \quad (A \subset F, \varepsilon > 0).
\]
Due to (6) one has
\[
\log N(2c_1 n^{-\alpha}, vU) \leq v \log N(2e_n, U) \leq (n - 1)v \log 2
\]
so that for any $\varepsilon, v > 0$
\[
\log N(\varepsilon, vU) \leq \frac{c_2 v^{1/\alpha}}{\varepsilon^{1/\alpha}}, \tag{7}
\]
where $c_2 = (2c_1)^{1/\alpha} \log 2$. Now fix $\eta > 0$ such that $(1 + \eta)s < \tilde{s}$, let $\varepsilon > 0$ be arbitrary and consider $\varepsilon_i := \varepsilon_i(\varepsilon) := \varepsilon e^{(1 + \eta)i}$ and $v_i = e^i$ for $i \in \mathbb{N}_0$ and $v_{-1} = 0$. We use $\varepsilon_i$-nets of the sets $v_i U$ to generate an appropriate codebook. Note that $\varepsilon_i \geq \xi v_i$ if
\[
i \geq \left[ \frac{1}{\eta} \log(\xi/\varepsilon) \right] \vee 0 =: M.
\]
For each $i = 1, \ldots, M - 1$ let $C_i(\varepsilon)$ denote an optimal $\varepsilon_i$-net of $v_i U$. Since $\xi \|x\|_E \geq \|x\|_F$ for $x \in E$, the set $\{0\}$ is an optimal $\varepsilon_i$-net of $v_i U$ for $i \geq M$ and we consider as the codebook
\[
C(\varepsilon) = \{0\} \cup \bigcup_{i=0}^{M-1} C_i(\varepsilon).
\]
Then
\[
E[d_F(Z, C(\varepsilon))^s] \leq \sum_{i=0}^{\infty} E[1_{[v_{i-1}, v_i)}(\|Z\|_E) d(Z, C_i(\varepsilon))^s]
\]
\[
\leq \sum_{i=0}^{\infty} P(\|Z\|_E \geq v_{i-1}) \varepsilon_i^s
\]
\[
= \varepsilon^s + \sum_{i=1}^{\infty} P\left(\frac{\|Z\|_E^s}{v_{i-1}^s} \geq 1\right) \varepsilon_i^s
\]
\[
\leq \varepsilon^s + \sum_{i=1}^{\infty} P(\|Z\|_E^s) \sum_{i=1}^{\infty} \frac{\varepsilon_i^s}{v_{i-1}^s}
\]
\[
= \varepsilon^s \left(1 + \sum_{i=1}^{\infty} e^{s-(\tilde{s}-(1+\eta)s)i}\right).
\]
Since $\tilde{s} > (1 + \eta)s$, the series converges and we obtain
\[ \mathbb{E}[d_F(Z, C(\varepsilon))^s]^{1/s} \leq c_3 \varepsilon \]
for some constant $c_3 = c_3(s, \tilde{s}) < \infty$. It remains to compute an upper bound for the size of $C(\varepsilon)$.
If $\varepsilon \geq \xi$, then $M = 0$ and $|C(\varepsilon)| = 1$. For $\varepsilon < \xi$, Eq. (7) implies
\[
|C(\varepsilon)| \leq 1 + \sum_{i=0}^{M-1} |C_i(\varepsilon)| \leq 1 + \sum_{i=0}^{M-1} \exp \left\{ c_2 \left( \frac{v_i}{\varepsilon} \right)^{1/\alpha} \right\}
\]
\[
= 1 + \sum_{i=0}^{M-1} \exp \left\{ c_2 \frac{1}{\varepsilon^{1/\alpha}} e^{-i \eta/\alpha} \right\}
\]
\[
\leq 1 + M \exp \left\{ c_2 \frac{1}{\varepsilon^{1/\alpha}} \right\}
\]
\[
\leq 1 + \left( 1 + \frac{1}{\eta} \log(\xi/\varepsilon) \right) \exp \left\{ c_2 \frac{1}{\varepsilon^{1/\alpha}} \right\},
\]
so that there exists a constant $c_4 = c_4(\xi, s, \tilde{s}) < \infty$ for which
\[
|C(\varepsilon)| \leq \exp \left\{ c_4 \frac{1}{\varepsilon^{1/\alpha}} \right\}.
\]
For an arbitrary rate $r > 0$, one chooses $\varepsilon = (c_4/r)^\alpha$ and applies the estimates above:
\[
D_r(r|s) \leq c_3 (c_4/r)^\alpha.
\]
Note that $D_r(r|s) \leq \xi \mathbb{E}[\| Z \|^{\tilde{s}}]^{1/\tilde{s}} = \xi$ now implies the assertion. □

3. Hölder continuity of $M$

In this section, we provide estimates for the moments of the $\alpha$-Hölder norm of continuous martingales. The analysis uses a Sobolev embedding type argument based on the GRR inequality (see [8]).

Let $M = (M_t)_{t \in [0,1]}$ be an $\mathbb{R}$-valued, $(\mathcal{F}_t)$-adapted continuous martingale given as $M_t = \int_0^t \sigma_u d\tilde{W}_u$, where $(\sigma_t)$ is an $(\mathcal{F}_t)$-adapted process with
\[
\int_0^1 \sigma_u^2 du < \infty, \quad \text{a.s.}
\]
In this section we do not require that $(\sigma_t)$ be given by $\sigma_t = \sigma(X_t, t)$.

We denote by $| \cdot |_{\alpha}$ the $\alpha$-Hölder semi-norm over the interval $[0, 1]$, that is
\[
|f|_{\alpha} := \sup_{0 \leq s < t \leq 1} \frac{|f(t) - f(s)|}{|t - s|^{\alpha}}.
\]

On the basis of the GRR inequality (see [8]) we derive an upper bound for the moments of $|M|_{\alpha}$:

**Theorem 3.1.** Let $\alpha \in (0, 1/2)$ and $\kappa > 2/(1 - 2\alpha)$. Then there exists a constant $c = c(\kappa, \alpha)$ such that
\[
\mathbb{E}[|M|_{\alpha}^\kappa] \leq c \int_0^1 \mathbb{E}[|\sigma_u|^{\kappa}] du.
\]
The constant \( c \) may be chosen independently of the martingale \( M \).

**Proof.** Fix \( \alpha \in (0, 1/2) \). Let \( f : [0, 1] \to \mathbb{R} \) be a continuous function and let \( \beta, \gamma > 0 \) with \( \alpha = \gamma - 2/\beta \). We consider \( \Psi(x) = |x|^\beta \) and \( p(x) = |x|^{\gamma}, x \in \mathbb{R} \). Then the GRR lemma states that provided that

\[
B := B(f) := \int_0^1 \int_0^1 \Psi \left( \frac{f(s) - f(t)}{p(s-t)} \right) dsdt = \int_0^1 \int_0^1 \frac{|f(s) - f(t)|^\beta}{|s-t|^{\beta \gamma}} dsdt \tag{8}
\]

is finite, one has

\[
|f(s) - f(t)| \leq 8 \int_0^{[s-t]} \Psi^{-1} \left( \frac{4B}{\xi^2} \right) dp(\xi) = 8 \int_0^{[s-t]} \frac{(4B)^{1/\beta}}{\xi^{2/\beta}} dp(\xi) = 8\gamma (4B)^{1/\beta} \int_0^{[s-t]} \xi^{\gamma-1-2/\beta} d\xi = 8\frac{\gamma}{\gamma-2/\beta} (4B)^{1/\beta} |s-t|^{\gamma-2/\beta}
\]

for all \( s, t \in [0, 1] \). Consequently,

\[
|f|_\alpha \leq 4^{1/\beta} 8\frac{\gamma}{\alpha} B^{1/\beta}.
\]

Now replace \( f \) by \( M \), define \( B := B(M) \) in analogy to (8), and let \( \kappa \geq \beta \vee 2 \). In order to control the \( \kappa \)-th moment of \( |M|_\alpha \) we derive an upper bound for \( \mathbb{E}B^{\kappa/\beta} \). Due to Jensen’s inequality and the Burkholder–Davis–Gundy (BDG) inequality, there exists a constant \( c_1 = c_1(\kappa) < \infty \) such that

\[
\mathbb{E}[B^{\kappa/\beta}] \leq \mathbb{E} \left[ \int_0^1 \int_0^1 \frac{|M_u^i - M_t^i|^\kappa}{|u-t|^{\kappa \gamma}} dudt \right] \\
\leq \int_0^1 \int_0^1 \mathbb{E}[|M_u^i - M_t^i|^\kappa] \frac{c_1 \mathbb{E}[\int_u^t \sigma_v^2 dv]^{\kappa/2}}{|u-t|^{\kappa \gamma}} dudt \\
\leq 2c_1 \int_0^1 \int_0^1 1 - \delta^{\kappa \gamma/2} \frac{\mathbb{E}[\int_u^t \sigma_v^2 dv]^{\kappa/2}}{\delta^{\kappa \gamma}} dud\delta.
\]

Applying again Jensen’s inequality leads to

\[
\mathbb{E}[B^{\kappa/\beta}] \leq 2c_1 \int_0^1 \delta^{-\kappa \gamma + \kappa/2} \int_0^{1-\delta} \mathbb{E} \left[ \int_u^{u+\delta} \delta^{-1} |\sigma_v|^\kappa dv \right] dud\delta.
\]

Next, applying Fubini’s Theorem on the inner two integrals yields for any \( \delta \in (0, 1) \)

\[
\int_0^{1-\delta} \int_u^{u+\delta} \delta^{-1} \mathbb{E}[|\sigma_v|^\kappa] dvdu \leq \int_0^1 \mathbb{E}[|\sigma_u|^\kappa] du
\]

so that

\[
\mathbb{E}[B^{\kappa/\beta}] \leq 2c_1 \int_0^1 \delta^{-\kappa \gamma + \kappa/2} d\delta \cdot \int_0^1 \mathbb{E}[|\sigma_u|^\kappa] du.
\]
Thus if $-\kappa\gamma + \kappa/2 > -1$, there exists a constant $c_2 = c_2(\kappa, \gamma) < \infty$ for which
\[
\mathbb{E}[|M|^\kappa] \leq c_2 \int_0^1 \mathbb{E}[|\sigma_u|^\kappa]du.
\]
(9)

It remains to determine the values of $\kappa \in \mathbb{R}_+$ for which there exist appropriate values of $\beta$ and $\gamma$ yielding inequality (9) for a finite $c_2$: $\beta$ and $\gamma$ need to satisfy
(i) $\gamma - 2/\beta = \alpha$,
(ii) $\kappa \geq 2 \land 2$ and (iii) $-\kappa\gamma + \kappa/2 > -1$.

When choosing $\gamma \in (1/2, \infty)$, conditions (i)–(iii) are equivalent to
\[
\kappa \geq \frac{2}{\gamma - \alpha} \land 2 \land \kappa < \frac{1}{\gamma - 1/2},
\]
and elementary analysis implies the existence of an estimate like (9) for each
\[
\kappa > \frac{2}{1 - 2\alpha}.
\]

Remark 3.2. The condition $\kappa > 2/(1 - 2\alpha)$ is necessary for the validity of the lemma. If the condition is not satisfied, a counterexample is obtained as follows: fix $\varepsilon \in (0, 1]$ and let $\sigma_t := 1_{[0, \varepsilon]}(t)\varepsilon^{-1/\kappa}$ ($t \in [0, 1]$); then the right hand side of the inequality is equal to $c$, whereas $\mathbb{E}[|M|^\kappa]$ tends to infinity when letting $\varepsilon$ to zero.

4. Coding scheme for $\varphi$

We assume again the setting of Section 1. In this section we introduce the coding schemes used for the random time transform $\varphi$. The construction depends on a parameter $\alpha \in (0, \beta/2)$.

For $n \in \mathbb{N}$ denote by $\hat{\varphi}^{(n)} = (\hat{\varphi}^{(n)}_t)_{t \in [0, 1]}$ a random increasing and continuous function that is linear on each interval $[i/n, i+1/n]$ ($i = 0, \ldots, n-1$) and satisfies
\[
\hat{\varphi}^{(n)}(i/n) = \arg\min_{y \in \mathbb{I}(n)} |\varphi(i/n) - y| \quad (i = 0, \ldots, n),
\]
where $\mathbb{I}(n)$ is defined as
\[
\mathbb{I}(n) = \left\{ j \frac{1}{n^{1+\alpha}} : j \in \mathbb{N}_0, j \leq n^{2(1+\alpha)} \right\}.
\]

Proposition 4.1. For any $s > 0$ there exists a constant $C = C(s) < \infty$ such that
\[
\log |\text{range} (\hat{\varphi}^{(n)})| \leq Cn(1 + \log n)
\]
and
\[
\mathbb{E}[\|\varphi(\cdot) - \hat{\varphi}^{(n)}(\cdot)\|_{[0,1]}^s]^{1/s} \leq C \frac{1}{n^{1+\alpha}}.
\]

The proof of the proposition relies on the following regularity result for $(\sigma_t^2)$ which itself is essentially a consequence of Assumption C, Theorem 3.1 and the finiteness of all moments $\mathbb{E}[\|X\|_{[0,1]}^s]$ ($s > 1$).
Lemma 4.2. One has for every $s > 0$
\[
\mathbb{E}[|\sigma_s^2|^s] < \infty.
\]
In particular, $\mathbb{E}[\varphi(1)^s]$ is finite for all $s > 0$.

Proof. It suffices to consider $s \geq 1$. By Theorem 3.1, it is true that for every $\alpha' \in (0, 1/2)$
\[
\mathbb{E}[|M_{\alpha'}^2|^s] < \infty.
\]
Moreover, by Assumption C
\[
\mathbb{E}[|A_{\alpha'}^{2s}|^{1/2s}] \leq \mathbb{E}[|A_{\alpha'}^{2s}|^{1/2s}] \leq \mathbb{E}[\|b_s\|^{2s}]^{1/2s} \leq L + L \mathbb{E}[\|X\|^{2s}]^{1/2s} < \infty
\]
and hence
\[
\mathbb{E}[|X_{\alpha'}^{2s}|^{1/2s}] \leq \mathbb{E}[|M_{\alpha'}^2|^s] + \mathbb{E}[|A_{\alpha'}^{2s}|^{1/2s}] < \infty.
\]
(10)

Since in general $|\sigma^n_s| \leq 2\|\varphi\|\|\sigma\|_{\alpha}$, one also has
\[
\mathbb{E}[|\sigma^n_s|^s] \leq 2^s \mathbb{E}[\|\sigma\|^s|\sigma\|^s_{\alpha}],
\]
and due to the Cauchy–Schwarz inequality it suffices to establish the finiteness of $\mathbb{E}[\|\sigma\|^{2s}]$ and $\mathbb{E}[|\sigma_s|^s]$. First note that by Assumption C
\[
\mathbb{E}[\|\sigma\|^{2s}] \leq 3L + 2L \mathbb{E}[\|X\|^{2s}]^{1/2s} < \infty.
\]
On the other hand, elementary analysis implies that
\[
|\sigma_s| \leq L \left(|X_s^\beta/\beta| + |X_s| + 2\right),
\]
so that (10) and the inequality $\alpha/\beta < 1/2$ imply that $\mathbb{E}[|\sigma_s|^s]$ is finite. □

Proof of Proposition 4.1. Again it suffices to consider $s \geq 1$. First note that
\[
\log |\text{range } (\hat{\varphi}^{(n)})| \leq n \log |\Pi(n)| \leq n \log (2n^{2(\alpha+1)})
\]
which immediately implies the first estimate. Now let $\varphi^{(n, \text{int})}$ denote the interpolation of $\varphi$ with supporting points $\{0, 1/n, \ldots, 1\}$, that is
\[
\varphi^{(n, \text{int})}(t) = \varphi(i/n) + n(\varphi((i+1)/n) - \varphi(i/n))(t - i/n)
\]
for $i \in \{0, \ldots, n - 1\}$ and $t \in [i/n, (i+1)/n]$. Then one has
\[
\mathbb{E}[\|\varphi(-) - \varphi^{(n, \text{int})}(-)\|^s_{[0,1]}]|^{1/s} \leq \frac{1}{n^{1+\alpha}} \mathbb{E}[|\sigma^n_{1/\alpha}|^s]|^{1/s}.
\]
Indeed, for $t \in [i/n, (i+1)/n]$ and $\tilde{\sigma}^2_t = n \int_{i/n}^{(i+1)/n} \sigma^2_u \, du$ one has
\[
|\varphi(t) - \varphi^{(n, \text{int})}(t)| \leq \int_{i/n}^t |\sigma^2_u - \tilde{\sigma}^2_t| \, du \leq \frac{1}{n^{1+\alpha}} |\sigma^2|_{\alpha}.
\]
Next, we apply the triangle inequality and conclude that
\[
\mathbb{E}[\|\varphi(-) - \hat{\varphi}^{(n)}(-)\|^s_{[0,1]}]|^{1/s} \leq \mathbb{E}[\|\varphi(-) - \varphi^{(n, \text{int})}(-)\|^s_{[0,1]}]|^{1/s} + \mathbb{E}[|\varphi^{(n, \text{int})}(-) - \hat{\varphi}^{(n)}(-)\|^s_{[0,1]}]|^{1/s} \leq \frac{1}{n^{1+\alpha}} \mathbb{E}[|\sigma^n_{1/\alpha}|^s]|^{1/s} + \frac{1}{n^{1+\alpha}} + \mathbb{E}[1_{\{\varphi(1) > n^{1+\alpha}\}} |\varphi(1)^s]|^{1/s}.\]
Due to the Cauchy–Schwarz inequality and Jensen’s inequality one has
\[
\mathbb{E}[\mathbb{1}_{\{\varphi(1) > n^{1+\alpha}\}} \varphi(1)^3]^{1/3} \leq \mathbb{P}(\varphi(1) > n^{1+\alpha})^{1/2s} \mathbb{E}[\varphi(1)^{2s}]^{1/2s} \\
\leq \mathbb{E}[\varphi(1)^{2s}]^{1/3} \frac{1}{n^{1+\alpha}},
\]
and the second assertion follows. □

5. An estimate for \( \mathbb{E}[\|W_{\varphi(\cdot)} - \hat{W}_{\varphi(\cdot)}\|_{[0,1]}^p] \)

In the previous section, we introduced and analyzed an approximation for the time-change \( \phi \).
It remains to study the behavior of \( \mathbb{E}[\|W_{\varphi(\cdot)} - \hat{W}_{\varphi(\cdot)}\|_{[0,1]}^p] \) for “good” reconstructions \( \hat{\varphi} \) of \( \varphi \).
The following analysis relies heavily on concentration properties of Gaussian measures.

Lemma 5.1. Let \( T, \rho, \varepsilon > 0 \) with \( \varepsilon \geq \sqrt{2\rho} \). We have
\[
\mathbb{P}\left( \sup_{\{u,t\in[0,T]: |u-t|\leq\rho\}} |W_t - W_u| \leq 3\varepsilon \right) \geq \left( 1 - 2e^{-\frac{\varepsilon^2}{2\rho}} \right)^n,
\]
where \( n := \lceil T/\rho \rceil \).

Proof. Let \( T, \rho, \varepsilon \) and \( n \) be as in the lemma. Set \( t_i = i\rho, i = 0, \ldots, n-1, \) and \( t_n = T \). Then
\[
\sup_{\{u,t\in[0,T]: |u-t|\leq\rho\}} |W_t - W_u| \leq 3 \max_{i=0,\ldots,n-1} \sup_{u\in[t_i,t_{i+1}]} |W_u - W_{t_i}|.
\]
Note that for \( M_i = \sup_{u\in[t_i,t_{i+1}]} |W_u - W_{t_i}| (i = 0, \ldots, n-1) \) one has
\[
\mathbb{P}\left( \sup_{\{u,t\in[0,T]: |u-t|\leq\rho\}} |W_t - W_u| \leq 3\varepsilon \right) \geq \mathbb{P}( \max_{i=0,\ldots,n-1} M_i \leq \varepsilon ).
\]
Since the r.v.'s \( M_0, \ldots, M_{n-1} \) are independent it follows that
\[
\mathbb{P}( \max_{i=0,\ldots,n-1} M_i \leq \varepsilon ) = \prod_{i=0}^{n-1} \mathbb{P}(M_i \leq \varepsilon ),
\]
so that
\[
\mathbb{P}(M_i > \varepsilon ) \leq 2\mathbb{P}( \sup_{u\in[t_i,t_{i+1}]} (W_u - W_{t_i}) > \varepsilon )
\]
\[
= 4 \mathbb{P}( (W_{t_{i+1}} - W_{t_i}) > \varepsilon )
\]
\[
= 4 \mathbb{P}\left( \frac{1}{\sqrt{t_{i+1} - t_i}} (W_{t_{i+1}} - W_{t_i}) > \frac{\varepsilon}{\sqrt{t_{i+1} - t_i}} \right)
\]
\[
= 4 \psi\left( \frac{\varepsilon}{\sqrt{t_{i+1} - t_i}} \right) \leq 2e^{-\frac{\varepsilon^2}{2(t_{i+1} - t_i)}},
\]
where \( \psi(t) := (2\pi)^{-1/2} \int_t^\infty \exp(-x^2/2)dx \ (t \in \mathbb{R}) \). By assumption, the last term is less than 1 and
\[
\mathbb{P}( \max_{i=0,\ldots,n-1} M_i \leq \varepsilon ) \geq \prod_{i=1}^{n-1} \left( 1 - 2e^{-\frac{\varepsilon^2}{2(t_{i+1} - t_i)}} \right) \geq \left( 1 - 2e^{-\frac{\varepsilon^2}{2\rho}} \right)^n. \]
Notice that $C_0[0, T]$ equipped with the norm
$$
\| f \|_{\rho, T} := \sup_{u, t \in [0, T], |u - t| \leq \rho} |f(t) - f(u)|
$$
is a separable Banach space, say $C_{\rho, T}$. Thus, we can interpret $(W_t)_{t \in [0, T]}$ as a centered Gaussian random vector in this space. Let $m_{\rho, T} \in \mathbb{R}_+$ denote the $7/8$-quantile of $\| W \|_{\rho, T}$. Using elementary analysis together with Lemma 5.1, we obtain:

**Lemma 5.2.** There exists a constant $c < \infty$ such that for all $T, \rho > 0$ one has
$$
m_{\rho, T} \leq c \sqrt{\rho \left( 1 + \log \left( \frac{T}{\rho} \right) \right)}.
$$

**Lemma 5.3.** For any $A \in \mathcal{F}$ and $s > 0$, one has
$$
\mathbb{E}[1_A \| W \|_{\rho, T}^s]^{1/s} \leq 2 \sqrt{2} m_{\rho, T} \left( 2 \int_{\sqrt{2} \Psi^{-1}(\mathbb{P}(A))}^\infty \frac{(x + 1)^s}{\sqrt{2\pi}} e^{-x^2/2} dx \right)^{1/s},
$$
where $\Psi(t) := \int_t^\infty \sqrt{2\pi}^{-1} \exp(-x^2/2) dx (t \in \mathbb{R})$.

**Proof.** Due to [15] (p. 99) (see also [14], pp. 202, 210) one has
$$
\sigma := \sup_{f \in C_{\rho, T}^*, \| f \|_{C_{\rho, T}^*} \leq 1} \mathbb{E}[f^2(W)]^{1/2} \leq 2 \sqrt{2} m_{\rho, T},
$$
where $C_{\rho, T}^*$ is the topological dual of $C_{\rho, T}$. As a consequence of the isoperimetric inequality, one obtains
$$
\mathbb{P}(\| W \|_{\rho, T} \geq m_{\rho, T} + t \sigma) \leq \Psi(t) \quad (t \geq 0).
$$
Thus we can find a standard normal random variable $N$ such that
$$
\| W \|_{\rho, T} \leq 2 \sqrt{2} m_{\rho, T}[1 + N^+],
$$
where $N^+ = N \vee 0$. Consequently,
$$
\mathbb{E}[1_A \| W \|_{\rho, T}^s]^{1/s} \leq 2 \sqrt{2} m_{\rho, T} \mathbb{E}[1_A (N^+ + 1)^s]^{1/s}. \quad \square
$$

**Lemma 5.4.** For $s > 0$ there exists a constant $c = c(s)$ such that for all $A \in \mathcal{F}, T, \rho > 0$ one has
$$
\mathbb{E}[1_A \| W \|_{\rho, T}^s] \leq c \left( \rho \left( 1 + \log \left( \frac{T}{\rho} \right) \right) (1 + \log(1/q)) \right)^{s/2} q,
$$
where $q := \mathbb{P}(A)$.

**Proof.** By elementary analysis one obtains
$$
\int_x^\infty (y + 1)^s e^{-u^2/2} du \sim x^s \Psi(x), \quad x \to \infty,
$$
and thus
$$
\int_{\Psi^{-1}(\varepsilon)}^\infty (u + 1)^s e^{-u^2/2} du \sim \varepsilon \Psi^{-1}(\varepsilon)^s \sim \varepsilon \sqrt{2\log(1/\varepsilon)^s}, \quad \varepsilon \downarrow 0.
$$
Consequently, there exists a constant $c_1 = c_1(s) < \infty$ such that for all $\varepsilon \in (0, 1]$
\[
\int_{0\vee \psi^{-1}(\varepsilon)}^{\infty} (u + 1)^s e^{-u^2/2} du \leq c_1 \varepsilon \sqrt{1 + \log(1/\varepsilon)^s}.
\]

Applying Lemmas 5.2 and 5.3 yields
\[
\mathbb{E}[1_A \|W\|_{p, T}^s] \leq c_2 \left( \rho T \left( 1 + \log \frac{T}{\rho} \right) (1 + \log(1/q)) \right)^{s/2} q,
\]
where $q := \mathbb{P}(A)$ and $c_2 = c_2(s)$ is a constant depending on $s$ only. \quad \Box

**Lemma 5.5.** Suppose that $\hat{\varphi}^{(r)}$ ($r \geq 0$) are reconstructions for $\varphi$ such that
\[
\lim_{r \to \infty} \mathbb{E}[\|\varphi - \hat{\varphi}^{(r)}\|_{[0, 1]}^{2s}]^{1/2s} = 0
\]
for an $s \geq 1$. Then
\[
\mathbb{E}[\|W_{\varphi(\cdot)} - W_{\hat{\varphi}^{(r)}(\cdot)}\|_{[0, 1]}^{1/2s}]^{1/2s} = \mathcal{O} \left( \sqrt{d(r) \log(1/d(r))} \right),
\]
where $d(r) = \mathbb{E}[\|\varphi - \hat{\varphi}^{(r)}\|_{[0, 1]}^{2s}]^{1/2s}$.

**Proof.** Consider the r.v.’s $\varepsilon := \varepsilon(r) := \|\varphi - \hat{\varphi}^{(r)}\|$ and $\tau := \varphi(1)$. Notice that
\[
\|W_{\varphi(\cdot)} - W_{\hat{\varphi}^{(r)}(\cdot)}\| \leq \|W\|_{\varepsilon, \tau + \varepsilon}.
\]
Let now $\mathbb{I} := \{e^i : i \in \mathbb{N}_0\}$,
\[
\bar{\varepsilon} := \bar{\varepsilon}(r) := \min([\varepsilon, \infty) \cap d(r) \mathbb{I}) \quad \text{and} \quad \bar{\tau} := \bar{\tau}(r) := \min([\tau, \infty) \cap \mathbb{I}).
\]
\(\bar{\varepsilon}\) and $\bar{\tau}$ are discrete r.v.’s dominating $\varepsilon$ and $\tau$ and satisfying
\[
\bar{\varepsilon} \leq e\varepsilon + d(r) \quad \text{and} \quad \bar{\tau} \leq e\tau + 1. \quad (11)
\]
Denote by $(p_{\rho, t})$ the probability weights of $(\bar{\varepsilon}, \bar{\tau})$. Then Lemma 5.4 yields
\[
\mathbb{E}[\|W\|_{\varepsilon, \tau + \varepsilon}^{s}] \leq \sum_{\rho, t} \mathbb{E}[1_{\{\varepsilon, \tau\}=(\rho, t)} \|W_{\rho, t + \rho}\|^{s}]
\]
\[
\leq c_1 \mathbb{E} \left[ \left( \bar{\varepsilon} \left( 1 + \log \left( 1 + \frac{\bar{\tau}}{\bar{\varepsilon}} \right) \right) (1 + \log(1/p_{\bar{\varepsilon}, \bar{\tau}})) \right)^{s/2} \right]
\]
\[
\leq c_1 \mathbb{E} \left[ \bar{\varepsilon}^s \left( 1 + \log \left( 1 + \frac{\bar{\tau}}{\bar{\varepsilon}} \right) \right)^{s/2} \mathbb{E} \left[ \left( 1 + \log(1/p_{\bar{\varepsilon}, \bar{\tau}}) \right)^s \right]^{1/2} \right]
\]
\[
= : c_1 \Sigma_1 \cdot \Sigma_2 \quad (12)
\]
for some appropriate constant $c_1 = c_1(s)$. Notice that the second term can be controlled by
\[
\Sigma_2^2 \leq 2^s (H^s(\bar{\varepsilon}, \bar{\tau}) + 1),
\]
where $H^s$ denotes the generalized entropy
\[
H^s(\bar{\varepsilon}, \bar{\tau}) := \sum_{\rho, t} p_{\rho, t} \left( \log(1/p_{\rho, t}) \right)^s.
\]
Now choose \( \rho = \mathbb{E}[e^{2s}]^{1/2s} e^t \) and \( t = e^j (i, j \in \mathbb{N}_0) \). If \( i, j \in \mathbb{N} \), one obtains with (11) and the Cauchy–Schwarz inequality

\[
p_{\rho,t} \leq \frac{\mathbb{E}[\bar{\tau}]}{\mathbb{E}[e^{2s}]^{1/2s} e^{i+j-2}} \leq \frac{\mathbb{E}[(e\bar{\tau} + \mathbb{E}[e^{2s}]^{1/2s})(\tau + 1)]}{\mathbb{E}[e^{2s}]^{1/2s} e^{i+j-2}} \\
\leq \frac{(e + 1)\mathbb{E}[(\tau + 1)^{2}]^{1/2}}{e^{i+j-2}}.
\]

If \( i = 0 \) and \( j \in \mathbb{N} \), then

\[
p_{\rho,t} \leq \frac{\mathbb{E}[\bar{\tau}]}{e^{j-1}} \leq \frac{\mathbb{E}[e^{\bar{\tau}+1}]}{e^{j-2}},
\]

whereas for \( i \in \mathbb{N} \) and \( j = 0 \), one obtains

\[
p_{\rho,t} \leq \frac{\mathbb{E}[\bar{\tau}]}{\mathbb{E}[e^{2s}]^{1/2s} e^{i-1}} \leq \frac{\mathbb{E}[e\tau + \mathbb{E}[e^{2s}]^{1/2s}]}{\mathbb{E}[e^{2s}]^{1/2s} e^{i-1}} \leq e + 1.
\]

Note that the above estimates for \( p_{\rho,t} \) do not depend on the rate \( r \geq 0 \) and decrease sufficiently fast to zero in order to provide the finiteness of \( H^s(\bar{\tau}, \tau) \). Moreover, \( H^s(\bar{\tau}, \tau) \) is uniformly bounded for all \( r \geq 0 \) by some constant \( c_3 < \infty \) depending on \( \mathbb{E}[(\tau + 1)^{2}]^{1/2} \) only. Consequently, \( \Sigma_2 \) is uniformly bounded.

It remains to analyze the first expression \( \Sigma_1 \). Recall that \( \bar{\tau} \geq 1 \). Hence, (11) and the Cauchy–Schwarz inequality yield

\[
\Sigma_1^{2/s} = \mathbb{E}\left[\bar{\tau} \left(1 + \log\left(1 + \frac{\bar{\tau}}{\bar{\tau}}\right)\right)^s\right]^{1/s} \leq \mathbb{E}\left[\bar{\tau} \left(1 + \log \bar{\tau} + \log\left(1 + \frac{1}{\bar{\tau}}\right)\right)\right]^{s/s} \\
\leq \mathbb{E}\left[\bar{\tau}^{2s}\right]^{1/2s} \mathbb{E}\left[1 + \log(1 + \varepsilon) + \log\left(1 + \frac{1}{d(r)}\right)\right]^{2s} \\
\leq (e + 1)d(r) \mathbb{E}\left[1 + \log(1 + \varepsilon) + \log\left(1 + \frac{1}{d(r)}\right)\right]^{2s}.
\]

Consequently, with (12) we arrive at

\[
\mathbb{E}[\|W_{\psi(\cdot)} - W_{\psi(r(\cdot))}\|^s]^{1/s} = O\left(\sqrt{d(r) \log(1/d(r))}\right).
\]

6. Decoupling \( (W_{t})_{t \in [0,\tau]} \) and the approximate time-change \( \hat{\psi} \)

We need some more notation. A function \( f \in C_0[0,\infty) \) admitting a representation \( f(t) = \int_0^t \hat{f}(u)du \) with a locally integrable \( \hat{f} : [0, \infty) \rightarrow \mathbb{R} \) will be called weakly differentiable, and we let

\[
\|f\|_{\mathcal{H}} := \left\{ \begin{array}{ll} \|\hat{f}\|_{L^2[0,\infty)} & \text{if } f \text{ is weakly differentiable with differential } \hat{f} \\ \infty & \text{else}. \end{array} \right.
\]

In analogy to the above, let for \( T > 0 \) and \( f \in C_0[0, T], \|f\|_{\mathcal{H}_T} = \|\hat{f}\|_{L^2[0,T]} \) if \( f \) is weakly differentiable and \( \|f\|_{\mathcal{H}_T} = \infty \), otherwise. The corresponding Hilbert spaces are denoted by \( \mathcal{H} \) and \( \mathcal{H}_T \).

We recall some results of the theory of enlargements of filtrations (see [13,1]). Let \( (\mathcal{F}^W_t) \) be the filtration generated by the Wiener process \( (W_t) \) and denote by \( G \) a discrete random variable
with probability weights \((p_g)\). We consider the enlarged filtration \((\mathcal{G}_t)_{t \geq 0} = (\mathcal{F}_t^W \vee \sigma(G))_{t \geq 0}\) and assume that for some fixed \(s \geq 1\) the generalized entropy

\[
H^s(G) := \mathbb{E} \left[ \left( \log \frac{1}{p_G} \right)^s \right]
\]

is finite. Then the process \((W_t)\) is a \((\mathcal{G}_t)\)-semimartingale, and its Doob–Meyer decomposition \(W_t = \bar{W}_t + \bar{Y}_t\) comprises a \((\mathcal{G}_t)\)-Wiener process \((\bar{W}_t)\) and a process of bounded variation \((\bar{Y}_t)\) (which we call information drift) satisfying

\[
\mathbb{E}[\|\bar{Y}\|_{H^s_T}^{2s}] \leq \kappa_s(H^s(G) + 1). \tag{13}
\]

Here, the constant \(\kappa_s\) depends on \(s\) only.

We recall that \(H^1\) is compactly embedded into \(C[0, 1]\), and that its entropy numbers satisfy

\[
e_n(H^1, C[0, 1]) \approx \frac{1}{n}, \quad n \to \infty.
\]

**Lemma 6.1.** Let \(\tilde{s} > s > 0\). There exists a constant \(c = c(s, \tilde{s})\) such that

\[
D^{(q)}(r|Y, T, s) \leq c\sqrt{T}\mathbb{E}[\|Y\|_{H^s_T}^{\tilde{s}}]^{1/\tilde{s}} \frac{1}{r + 1}
\]

for all \(T > 0, r \geq 0\) and \(H_T\)-valued r.v.'s \(Y\).

**Proof.** By Lemma 2.1, the statement holds for fixed time \(T = 1\) for an appropriate constant \(c > 0\). Notice that for \(T > 0\) the maps

\[
\pi_T^{(1)} : H_T \to H^1, \quad f \mapsto \frac{1}{\sqrt{T}} f(T \cdot) \quad \text{and}
\]

\[
\pi_T^{(2)} : C[0, T] \to C[0, 1], \quad f \mapsto f(T \cdot)
\]

are isometric isomorphisms. Consequently,

\[
D^{(q)}(r|Y, T, s) = D^{(q)}(r|\pi_T^{(2)} Y, 1, s) \\
\leq c\mathbb{E}[\|\pi_T^{(2)}(Y)\|_{H^1}^{\tilde{s}}]^{1/\tilde{s}} \frac{1}{r + 1} = c\mathbb{E}[\|\sqrt{T} \pi_T^{(1)}(Y)\|_{H^1}^{\tilde{s}}]^{1/\tilde{s}} \frac{1}{r + 1} = c\sqrt{T}\mathbb{E}[\|Y\|_{H^s_T}^{\tilde{s}}]^{1/\tilde{s}} \frac{1}{r + 1}. \quad \square
\]

**Lemma 6.2.** For any \(s \geq 1\), there exists a constant \(c < \infty\) such that

\[
D^{(q)}(r|\bar{Y}, \tau, s) \leq c(H^s(G) + 1)^{1/2s} \left[ \sqrt{T} \frac{1}{r} + T \tau^{1-\frac{q}{r}} \mathbb{E}[\tau^q]^{1/2s} \right]
\]

for all \(q \geq s, T > 0, r \geq 1\), all \([0, \infty)\)-valued r.v.'s \(\tau\) and arbitrary side information \(G\).

**Proof.** Fix \(T > 0\) and \(r \geq 1\). The previous lemma and Eq. (13) imply
\[ D^{(q)}(r|\tilde{Y}, T, s) \leq c_1 \sqrt{T} \mathbb{E}[\|\tilde{Y}\|_{\mathcal{H}_T}^{2s}]^{1/2s} \frac{1}{r + 1} \]
\[ \leq c_1 s^{1/2s} \sqrt{T} (H^s(G) + 1)^{1/2s} \frac{1}{r + 1} \]
\[ = \frac{c_2}{2} \sqrt{T} (H^s(G) + 1)^{1/2s} \frac{1}{r + 1}, \]

for some appropriate constants \(c_1\) and \(c_2\) depending on \(s\) only. Consequently, there exists a codebook \(C \subseteq C(0, \infty)\) of size \([e']^s\) containing the constant function \(0\) and satisfying

\[ \mathbb{E}[\min_{\tilde{Y} \in C} \|\tilde{Y} - \hat{Y}\|_{[0, T]}^{1/s}] \leq c_2 \sqrt{T} (H^s(G) + 1)^{1/2s} \frac{1}{r}. \quad (14) \]

Next, denote by \(\hat{Y}\) a \(C\)-valued r.v. minimizing the distance \(\|\cdot - Y\|_{[0, T]}\), and observe that

\[ \mathbb{E}[\|\tilde{Y} - \hat{Y}\|_{[0, T]}^{1/s}] \leq \mathbb{E}[1_{\{\tau \leq T\}} \|\tilde{Y} - \hat{Y}\|_{[0, T]}^{1/s}] + \mathbb{E}[1_{\{\tau > T\}} \|\tilde{Y}\|_{[0, T]}^{1/s}] =: I_1 + I_2. \]

It remains to analyze \(I_2\). The natural embedding from \(\mathcal{H}_\tau\) to \(C[0, \tau]\) has norm \(\sqrt{\tau}\) so that

\[ I_2 \leq \mathbb{E}[1_{\{\tau > T\}} \sqrt{T} s \|\tilde{Y}\|_{\mathcal{H}_\tau}^{1/s}] \leq \mathbb{E}[\mathbb{E}[1_{\{\tau > T\}} \tau \|\tilde{Y}\|_{\mathcal{H}_\tau}^{2s}]]^{1/2s} \]
\[ \leq (T^{s-q} \mathbb{E}[\tau^q])^{1/2s} \left(\frac{1}{2s} \left(\frac{1}{2s} \left(H^s(G) + 1\right)^{1/2s}\right) \right). \]

Combining this with estimate (14) leads to the assertion. \(\square\)

Now we apply the above results to the case where the enlarging random variable is \(G = \hat{\phi}^{(n)}\), with \(\hat{\phi}^{(n)}\) defined as in Section 4. Recall that the definition of \(\hat{\phi}^{(n)}\) depends on a parameter \(\alpha \in (0, \beta/2)\). For fixed \(\alpha\) we can now control the coding complexity of the related information drift \(\tilde{Y}^{(n)}\):

**Proposition 6.3.** Let \(\gamma_1, \gamma_2 > 0\), relate \(r > 0\) and \(n \in \mathbb{N}\) via \(n = n(r) = [r^{\gamma_1}]\) and denote by \(\tilde{Y}^{(n)}\) the information drift of \(W\) under the enlarged filtration \(\mathcal{F}^W_t \vee \sigma(\hat{\phi}^{(n)})\). Then for any \(\varepsilon > 0\)

\[ D^{(q)}(r^{\gamma_2} | \tilde{Y}^{(n)}, \tau^{(n)}, s) = O(r^{\frac{\gamma_1}{2} - \gamma_2 + \varepsilon}), \]

where \(\tau^{(n)} := \hat{\phi}^{(n)}(1)\).

**Proof.** It suffices to consider \(s \geq 2\). Note that one can choose \(c_1 \geq 0\) such that the function

\[ f : [1, \infty) \to [0, \infty), \quad x \mapsto (\log x)^s + c_1 \log x \]

is concave. Consequently, it follows that for any \(Z\) with finite range

\[ H^s(Z) = \mathbb{E}[(\log 1/pZ)^s] \leq \mathbb{E} f(1/pZ) \leq f(\mathbb{E}[1/pZ]) = f(|\text{range}(Z)|). \]

Due to Proposition 4.1 there exists a constant \(C\) such that for all \(n\)

\[ \log |\text{range}(\hat{\phi}^{(n)})| \leq Cn(1 + \log n). \]

Hence,

\[ (H^s(\hat{\phi}^{(n)}) + 1)^{1/2s} = O(\sqrt{r^{\gamma_1} \log r}) \]
as $r \to \infty$. Fix $\epsilon > 0$ and consider $T = T(r) = r^\epsilon$ and $q = \frac{2}{\epsilon} \gamma_2 s$. Then

$$\frac{\sqrt{T}}{r^{\gamma_2}} + T^{\frac{1}{2}}(1-\frac{2}{\epsilon}) \mathbb{E}\left[\left(\tau^{(n)}\right)^q\right]\left(1 + \mathbb{E}\left[\left(\tau^{(n)}\right)^q\right]\right) = \frac{1}{r^{\gamma_2 - \frac{2}{\epsilon}}} \left(1 + \mathbb{E}\left[\left(\tau^{(n)}\right)^q\right]\right).$$  \hspace{1cm} (15)

Since $\sup_{n\in\mathbb{N}} \mathbb{E}[\left(\tau^{(n)}\right)^q] < \infty$ we conclude that (15) is of the order $\mathcal{O}(r^{-\gamma_2 + \epsilon/2})$, so that due to Lemma 6.2

$$D(q)(r^{\gamma_2} \tilde{Y}^{(n)} , \tau^{(n)} , s) = \mathcal{O}\left(r^{\frac{2}{\epsilon} - \gamma_2 + \epsilon}\right). \quad \Box$$

7. Main representation of the diffusion

In this section we make use of all the previous results and establish the main representation of the diffusion.

**Theorem 7.1.** Fix $\alpha \in (0, \beta/2)$ and $\gamma_1 \in ((1 + \alpha)^{-1}, 1)$. Moreover, let $\varphi^{(n)}$ be as in Section 4, relate $n$ and $r > 0$ via $n = n(r) = [r^{\gamma_1}]$, and let $W_t = \tilde{W}^{(n)}_t + \tilde{Y}^{(n)}_t$ denote the $(\mathcal{F}^W_t \vee \sigma(\varphi^{(n)}))$-Doob–Meyer decomposition of $W$.

For fixed $s > 0$ there exist $C[0, 1]$-valued r.v.’s $\tilde{R}^{(n)}$ and $\hat{R}^{(r)}$ such that

- $X = \tilde{W}^{(n)}_{\hat{\varphi}^{(n)}(\cdot)} + \tilde{R}^{(n)}$,
- $\tilde{W}^{(n)}$ is a Wiener process that is independent of $\hat{\varphi}^{(n)}$,
- $\mathbb{E}[\|\tilde{R}^{(n)} - \hat{R}^{(r)}\|_{[0, 1]}^s]^{1/s} = \mathcal{O}(r^{-\frac{1}{2} - \delta})$, for some $\delta > 0$,
- $\log|\text{range}(\hat{R}^{(r)}, \hat{\varphi}^{(n)})| = \mathcal{O}(r^\gamma)$, for some $\gamma \in (0, 1)$.

**Proof.** We consider $s \geq 1$ only. Due to Proposition 4.1,

$$\mathbb{E}[\|\varphi - \hat{\varphi}^{(n)}\|^2_{[0, 1]}]^{1/2r} = \mathcal{O}(n^{-(1+\alpha)}) = \mathcal{O}(r^{-(1+\alpha)\gamma_1}).$$

Now Lemma 5.5 implies the existence of a constant $\delta_1 > 0$ with

$$\mathbb{E}[\|W_{\hat{\varphi}^{(n)}(\cdot)} - W_{\hat{\varphi}^{(n)}(\cdot)}\|^s_{[0, 1]}]^{1/s} = \mathcal{O}(r^{-\frac{1}{2} - \delta_1}).$$  \hspace{1cm} (16)

Next, choose $\gamma_2 \in (0, 1)$ with $\frac{\gamma_1}{2} - \gamma_2 < -\frac{1}{2}$. Due to Proposition 6.3, there exist $C[0, \infty)$-valued reconstructions $\hat{Y}^{(r)}$ satisfying

$$\log|\text{range}( \hat{Y}^{(r)} ) | \leq r^{\gamma_2} \quad \text{and} \quad \mathbb{E}[\|\hat{Y}^{(n)}(\cdot) - \hat{Y}^{(r)}(\cdot)\|^s_{[0, \hat{\varphi}^{(n)}(\cdot)]}]^{1/s} = \mathcal{O}(r^{-\frac{1}{2} - \delta_2})$$

for a fixed $\delta_2 \in (0, - (\frac{\gamma_1}{2} - \gamma_2) - \frac{1}{2})$.

The bounded variation part $A$ of the diffusion $X$ satisfies $\mathbb{E}[\|A\|^2_{\mathcal{H}_1}] < \infty$, so that Lemma 6.1 yields the existence of reconstructions $\hat{A}^{(r)}$ for which

$$\log|\text{range}( \hat{A}^{(r)} ) | \leq r^{2/3} \quad \text{and} \quad \mathbb{E}[\|A - \hat{A}^{(r)}\|^s_{[0, 1]}]^{1/s} = \mathcal{O}(r^{-2/3}).$$  \hspace{1cm} (18)

Now we rewrite $X$ in terms of the newly introduced r.v.’s:

$$X_t = A_t + M_t = A_t + W_{\hat{\varphi}(t)} = A_t + (W_{\varphi(t)} - W_{\hat{\varphi}^{(n)}(t)}) + W_{\hat{\varphi}^{(n)}(t)}$$

$$= A_t + (W_{\varphi(t)} - W_{\hat{\varphi}^{(n)}(t)} + \tilde{Y}_{\hat{\varphi}^{(n)}(t)} + \tilde{W}_{\hat{\varphi}^{(n)}(t)}).$$
Due to (16)–(18) it follows that the process \( \hat{R}_t^{(r)} := \hat{A}_t^{(r)} + \hat{Y}_t^{(r)}(\phi^{(n)}(t)) \) satisfies for \( \delta := \min(\delta_1, \delta_2, 1/6) > 0 \)

\[
\mathbb{E}[\| \hat{R}^{(n)}(t) - \hat{R}_t^{(r)} \|^s]^{1/s} \leq \mathbb{E}[\| A - \hat{A}_t^{(r)} \|^s]^{1/s} + \mathbb{E}[\| W_{\phi^{(n)}} - W_{\phi^{(n)}(t)} \|^s]^{1/s} + \mathbb{E}[\| \hat{Y}^{(n)}_t(\phi^{(n)}(t)) - \hat{Y}_t^{(r)}(\phi^{(n)}(t)) \|^s]^{1/s} = \mathcal{O}(r^{-1/2 - \delta}).
\]

Moreover, Proposition 4.1 implies that

\[
\log |\text{range } (\phi^{(n)})| = \mathcal{O}(n \log n) = \mathcal{O}(r^{\gamma_1} \log r) = \mathcal{O}(r^{(1+\gamma_1)/2}).
\]

Combining this with the range estimates for \( \hat{Y}^{(r)} \) and \( \hat{A}^{(r)} \), we obtain

\[
\log |\text{range } (\hat{R}^{(r)}, \phi^{(n)})| = \mathcal{O}(r^{\gamma}),
\]

where \( \gamma = \max((1 + \gamma_1)/2, \gamma_2, 2/3) < 1. \)

8. The coding complexity of \( X \) in \( C[0, 1] \)

In this section we prove Theorem 1.1. We will need the notion of \textit{conditional entropy}. For two discrete r.v.’s \( Z \) and \( G \), let

\[
H(Z|G = g) = \mathbb{E}[\log 1/p_{Z|G} | G = g] \quad \text{and} \quad H(Z|G) = \mathbb{E}[\log 1/p_{Z|G}],
\]

where \( p_{Z|G} \) denotes the conditional probability \( \mathbb{P}(Z = z | G = g) \), which is well defined for \( \mathbb{P}_G \)-a.a. \( g \). For basic properties of the conditional entropy one might consult [12].

In the rest of this section \( s > 0, \alpha \in (0, \beta/2) \) and \( \gamma_1 \in ((1 + \alpha)^{-1}, 1) \) are fixed. Moreover, relate \( n \) and \( r > 0 \) via \( n = [r^{\gamma_1}] \) and let \( \hat{\phi} = \hat{\phi}^{(n)}, \hat{W} = \hat{W}^{(n)}, \hat{R} = \hat{R}^{(n)}, \hat{R} = \hat{R}^{(r)} \) be as in Theorem 7.1. We also let \( \tau = \tau^{(n)} = \hat{\phi}^{(n)}(1) \), and for simplicity we omit the parameters \( n \) and \( r \) in the notation for the stochastic processes.

We first turn to the proof of the upper bounds.

\textbf{Proof of the upper bounds.} We start by proving the upper asymptotic bound for the quantization formula. For each possible realization \( t \) of \( \tau \) we choose a codebook \( C(t) \) of size \([e^r] \) with entries in \( C[0, \infty) \) such that

\[
\mathbb{E}[\min \limits_{\hat{w} \in C(t)} \| \hat{W} - \hat{w} \|^s_{[0, r]}]^{1/s} \leq (1 + 1/r)\sqrt{\mathcal{D}(q)}(r | W, s).
\]

Now let

\[
\hat{W} = \text{argmin } \| \hat{W} - \hat{w} \|^s_{[0, r]}
\]

and observe that due to the independence of \( \tau \) and \( \hat{W} \) (Theorem 7.1), we have

\[
\mathbb{E}[\| \hat{W} - \hat{W}^{(n)}_{\phi^{(n)}} \|^s_{[0, r]}]^{1/s} \leq (1 + 1/r)\mathbb{E}[\sqrt{\mathcal{D}(q)}(r | W, s)].
\]

We consider \( \hat{X} := \hat{X}^{(r)} := \hat{W}^{(n)}_{\phi^{(n)}} + \hat{R} \) as a reconstruction for \( X \) and observe that for \( s \geq 1 \)

\[
\mathbb{E}[\| X - \hat{X} \|^s]^{1/s} \leq \mathbb{E}[\| \hat{W} - \hat{W}^{(n)} \|^s_{[0, r]}]^{1/s} + \mathbb{E}[\| \hat{R} - \hat{R}^{(n)} \|^s]^{1/s} \\
\lesssim K \mathbb{E}[\| \sigma_{t} \|^s_{L^2[0, 1]}]^{1/s} \frac{1}{\sqrt{r}}
\]
and for $s < 1$

$$
\mathbb{E}[\|X - \hat{X}\|^{s}]^{1/s} \leq \left(\mathbb{E}[\|\hat{W} - \hat{W}\|^{s}_{[0, \tau]}] + \mathbb{E}[\|\hat{R} - \hat{R}\|^{s}]\right)^{1/s} 
\leq K \mathbb{E}[\|\sigma.\|_{L^{2}[0, 1]}^{s}]^{1/s} \left(\frac{1}{\sqrt{r}}\right).
$$

In the latter computations we have used that $\lim_{r \to \infty} \mathbb{E}[\sqrt{\tau^{s}}]^{1/s} = \mathbb{E}[\|\sigma.\|_{L^{2}[0, 1]}^{s}]^{1/s}$.

It remains to analyze the size of the range of $\hat{X}$; recall that $\log |\text{range } \hat{X}(\tau)|$ is of order $o(r)$ so that $\log |\text{range } (\hat{W})| = (1 + o(1))r$. Consequently, $\hat{X}$ has range of size $e^{(1+o(1))r}$ which completes the proof of the upper bound for the quantization error.

Now we focus on the entropy coding problem. Let $r_{t} = r_{t}^{(r)} = r^{s/(s+2)}/\mathbb{E}[r^{s/(s+2)}] + \sqrt{r}$ $(t \geq 0)$ and let $\hat{W} = \hat{W}^{(r)}$ denote a reconstruction for $\hat{W}$ such that for any $t \in \text{range } (\tau)$ one has $H(\hat{W}|\tau = t) \leq r_{t}$ and

$$
\mathbb{E}[\|\hat{W} - \hat{W}\|^{s}_{[0, \tau]}|\tau = t]^{1/s} \leq (1 + 1/r)\sqrt{tD(\phi)}(r_{t}|W, s).
$$

Since $r$. converges uniformly to $\infty$ on range $(\tau)$, we conclude that

$$
\mathbb{E}[\|\hat{W} - \hat{W}\|^{s}_{[0, \tau]}]^{1/s} \leq K \mathbb{E}[r^{s/2}r_{r_{t}}^{s/2}]^{1/s} \sim K \mathbb{E}[r^{s/(s+2)}r_{r_{t}}^{s/(s+2)}]^{1/2} \sim K \|\sigma.\|_{L^{2}[0, 1]}^{s} \left(\frac{1}{\sqrt{r}}\right).
$$

Analogously to the above, the same estimate remains valid for $\mathbb{E}[\|X - \hat{X}\|^{s}_{[0, 1]}]^{1/s}$. On the other hand the entropy of $\hat{X}$ can be controlled by

$$
H(\hat{X}) \leq H(\hat{W}, \hat{R}, \hat{\phi}) \leq H(\hat{R}) + H(\hat{\phi}) + H(\hat{W}|\tau).
$$

Note that $H(\hat{R})$ and $H(\hat{\phi})$ are of order $o(r)$ and

$$
H(\hat{W}|\tau) \leq \mathbb{E}[r_{\tau}] = r + \sqrt{r} \sim r,
$$

and the second assertion follows. $\square$

Now we turn to the proof of the lower bounds.

**Proof of the lower bounds.** We only treat the case $s \geq 1$ here. In order to obtain the estimate for $s < 1$ one only has to adjust the arguments that use the triangle inequality. First consider the quantization setting. Denote by $\hat{X} = \hat{X}^{(r)}$ arbitrary reconstructions for $X$ that have range of size $[e^{t}]$, and let

$$
\hat{W}_{t} := \hat{W}^{(r)}_{t} := X_{\hat{\phi}^{-1}(t)} - \hat{R}_{\hat{\phi}^{-1}(t)} \quad (t \in [0, \tau]),
$$

where $\hat{\phi}^{-1}(t) := \inf\{s \geq 0 : \hat{\phi}(s) \geq t\}$. Due to the continuity of $\hat{\phi}$, one has $\hat{\phi} \circ \hat{\phi}^{-1} = \text{id}$ on $[0, \tau]$, so that $\hat{W}_{t} = X_{\hat{\phi}^{-1}(t)} - \hat{R}_{\hat{\phi}^{-1}(t)}$ for $t \in [0, \tau]$. Therefore,

$$
\mathbb{E}[\|\hat{W} - \hat{W}\|^{s}_{[0, \tau]}]^{1/s} \leq \mathbb{E}[\|X_{\hat{\phi}^{-1}(\cdot)} - \hat{X}_{\hat{\phi}^{-1}(\cdot)}\|^{s}_{[0, \tau]}]^{1/s} + \mathbb{E}[\|\hat{R}_{\hat{\phi}^{-1}(\cdot)} - \hat{R}_{\hat{\phi}^{-1}(\cdot)}\|^{s}_{[0, \tau]}]^{1/s} 
\leq \mathbb{E}[\|X - \hat{X}\|^{s}_{[0, 1]}]^{1/s} + \mathbb{E}[\|\hat{R} - \hat{R}\|^{s}_{[0, 1]}]^{1/s}.
$$

(19)
On the other hand, $\hat{W}$ satisfies the range constraint
\[
\log |\text{range } (\hat{W})| \leq \log |\text{range } (\hat{X}, \hat{R}, \hat{\phi})| \leq r + O(r^\gamma) \sim r.
\]
Using the independence of $\hat{W}$ and $\hat{\phi}$ we conclude that
\[
\mathbb{E}[\| \hat{W} - \hat{W}^s_{[0,\tau]} \|_{[0,\tau]}^s | \hat{\phi}]^{1/s} \geq \sqrt{r} D(q) (\log |\text{range } (\hat{W})||W, s).
\]
This gives
\[
\mathbb{E}[\| \hat{W} - \hat{W}^s_{[0,\tau]} \|_{[0,\tau]}^s ]^{1/s} \geq K \mathbb{E}[\tau^{s/2}]^{1/s} \frac{1}{\sqrt{r}} \sim K \mathbb{E}[\| \sigma \|_{L^2[0,1]}^s ]^{1/s} \frac{1}{\sqrt{r}}.
\]
Since $\mathbb{E}[\| \hat{R} - \hat{R}^s \|^{1/s} ]$ is of order $o(1/\sqrt{r})$ the assertion is a consequence of (19).

Let now $\hat{X} = \hat{X}(r)$ denote arbitrary reconstructions for $X$ with $H(\hat{X}) \leq r$ and define $\hat{W} = \hat{W}(r)$ as above. For $t \in \text{range } (\tau)$ let $r_t = r_t^{(r)} = H(\hat{W}|\tau = t) + \sqrt{r}$. Since
\[
H(\hat{W}|\tau) \leq H(\hat{X}) + H(\hat{R}) + H(\hat{\phi}) \leq r + o(r) \sim r
\]
we have
\[
\mathbb{E}[r_t] = H(\hat{W}|\tau) + \sqrt{r} \lesssim r.
\]
Note that $r$ converges on range $(\tau)$ uniformly to infinity. Combining this property with Eq. (1) and the estimate
\[
\mathbb{E}[\| \hat{W} - \hat{W}^s_{[0,\tau]} \|_{[0,\tau]}^s | \tau = t]^{1/s} \geq \sqrt{t} D(q)(r_t|W, s)
\]
leads to
\[
\mathbb{E}[\| \hat{W} - \hat{W}^s_{[0,\tau]} \|_{[0,\tau]}^s ]^{1/s} \geq K \mathbb{E} \left[ \frac{\tau^{s/2}}{\sqrt{r}^{s/2}} \right]^{1/s} = \langle \tau^{s/2}, r_t^{-s/2}, t \rangle_{L^2(D)}.
\]
Applying Hölder’s inequality (for exponents less than one) with $q = -2/s$ and adjoint coefficient $q^* = 2/(s + 2)$ gives
\[
\langle \tau^{s/2}, r_t^{-s/2}, t \rangle_{L^2(D)} \geq \mathbb{E}[\tau^{s/(s+2)}]^{(s+2)/2} \mathbb{E}[r_t]^{-s/2}.
\]
Consequently,
\[
\mathbb{E}[\| \hat{W} - \hat{W}^s_{[0,\tau]} \|_{[0,\tau]}^s ]^{1/s} \geq K \mathbb{E}[\tau^{s/(s+2)}]^{(s+2)/2} \mathbb{E}[r_t]^{-1/2}
\]
\[
\geq K \mathbb{E}[\| \sigma \|_{L^2[0,1]}^{2s/(s+2)}]^{(s+2)/2} \frac{1}{\sqrt{F}}.
\]
Due to Eq. (19) the proof is complete on recalling that $\mathbb{E}[\| \hat{R} - \hat{R}^s \|^{1/s} ]$ is of order $o(1/\sqrt{r})$. \qed

References