On the specific relative entropy between martingale diffusions on the line

Julio Backhoff-Veraguas[*]   Clara Unterberger[†]

Abstract

The specific relative entropy, introduced in the Wiener space setting by N. Gantert, allows to quantify the discrepancy between the laws of potentially mutually singular measures. It appears naturally as the large deviations rate function in a randomized version of Donsker's invariance principle, as well as in a novel transport-information inequality recently derived by H. Föllmer. A conjecture, put forward by the aforementioned authors, concerns a closed form expression for the specific relative entropy between continuous martingale laws in terms of their quadratic variations. We provide a first partial result in this direction, by establishing this conjecture in the case of well-behaved martingale diffusions on the line.

1 Introduction

Let $(\mathcal{C}, \mathcal{F})$ be the Wiener space $\mathcal{C} = C([0,1], \mathbb{R})$ and $\mathcal{F} = \sigma(X_t : t \in [0,1])$ where $X_t \colon \mathcal{C} \to \mathbb{R}$ with $X_t(\omega) = \omega(t)$ is the canonical process. Define also

$$\mathcal{F}^n := \sigma(X_{\frac{k}{n}} : k = 0, 1, \ldots, n).$$

The central object of study in this article is the concept of specific relative entropy, introduced in the setting of Wiener space by N. Gantert in her PhD thesis [6] and recently studied by the works of H. Föllmer [4, 2].

Definition 1.1. *Given $\mathbb{Q}, \mathbb{P}$ probability measures on $(\mathcal{C}, \mathcal{F})$, the upper specific relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as*

$$h^u\left(\mathbb{Q} \,|\, \mathbb{P}\right) := \limsup_{n \to \infty} \frac{1}{n} H(\mathbb{Q} \,|\, \mathbb{P})|_{\mathcal{F}^n},$$

where $H(\mathbb{Q} \,|\, \mathbb{P})|_{\mathcal{F}^n}$ is the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ restricted to $\mathcal{F}^n$, namely

$$H(\mathbb{Q} \,|\, \mathbb{P})|_{\mathcal{F}^n} := \begin{cases} \int_{\mathcal{C}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}^n} \, d\mathbb{Q} & \text{if } \mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{P}|_{\mathcal{F}^n} \\ +\infty & \text{otherwise.} \end{cases}$$

---

[*]University of Vienna, Austria. E-mail: julio.backhoff@univie.ac.at
[†]University of Vienna, Austria. E-mail: clara.unterberger@gmail.com

Similarly the lower specific relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as

$$h^\ell\left(\mathbb{Q}\,|\,\mathbb{P}\right) := \liminf_{n\to\infty} \frac{1}{n} H(\mathbb{Q}\,|\,\mathbb{P})|_{\mathcal{F}^n}.$$

Finally, if $h^\ell(\mathbb{Q}\,|\,\mathbb{P}) = h^u(\mathbb{Q}\,|\,\mathbb{P})$, we denote this common value by $h(\mathbb{Q}\,|\,\mathbb{P})$ and call it the specific relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Remark 1.2.

1. The specific relative entropy is meaningful even in situations where measures singular to each other are being compared. This is the case of continuous martingale laws, which typically have infinite relative entropy but may still have a finite specific relative entropy when compared with each other. In [6, Kapitel II.4] it was shown that $h(\cdot|\mathbb{W})$, with $\mathbb{W}$ being Wiener measure, is the rate function in a large deviations principle associated to a randomized Donsker-type approximation of Brownian motion.

2. We took the liberty to introduce the terminology of "upper/lower specific relative entropies". In fact in [4] already h^ℓ is called specific relative entropy, whereas [6] restricts itself to the case when h^ℓ and h^u coincide. The works [4, 2] considers rather the subsequence of dyadic sigma-algebras $\mathcal{F}^{2^n}$ in the definitions of specific relative entropy.

3. Both [4, 2] and [6] mostly study the specific relative entropy for $\mathbb{Q}, \mathbb{P}$ being general martingale laws, or even more general stochastic processes than martingales. In the present work we will specialize the discussion to $\mathbb{Q}, \mathbb{P}$ being the laws of very well-behaved *martingale diffusions*.

4. The concept of specific relative entropy is older than the work [6]. For instance this concept appears in Donsker and Varadhan's study of large deviations for stationary Gaussian processes [1] and in Föllmer's study of lattice models [3].

Let $B = (B_t)_{t\in[0,1]}$ be a one-dimensional Brownian motion on some filtered probability space $(\Omega, \mathcal{S}, (\mathcal{S}_t)_{t\in[0,1]}, \$)$. Let $M^x = (M^x_t)_{t\in[0,1]}$ be the solution of

$$\begin{cases} dM^x_t = \sigma(M^x_t)\, dB_t \\ M^x_0 = x, \end{cases}$$

and likewise let $N^x = (N^x_t)_{t\in[0,1]}$ be the solution of

$$\begin{cases} dN^x_t = \eta(N^x_t)\, dB_t \\ N^x_0 = x. \end{cases}$$

Throughout this article we assume rather strong regularity assumptions on the diffusion coefficients σ, η. They will guarantee a unique strong solution to the above stochastic differential equations in particular, but more importantly, they will allow us to gauge the small-time behaviour of M^x and N^x. The precise assumptions are:

Assumption 1.3. *The coefficients $\sigma, \eta\colon \mathbb{R} \to \mathbb{R}_+$ are twice continuously differentiable and such that*

$$\sigma(\cdot), \eta(\cdot) \in (\delta, 1/\delta),$$

for some $0 < \delta \leq 1$. Moreover, for some $L \in \mathbb{R}$ we have

$$\max\{\|\sigma'\|_\infty, \|\sigma''\|_\infty, \|\eta'\|_\infty, \|\eta''\|_\infty\} < L.$$

The main result of this article establishes the existence of the specific relative entropy between the laws of M^x and N^x, and provides a closed form expression for this quantity:

Theorem 1.4. *Under Assumption 1.3, let M^x and N^x be as above and call $\mathbb{Q}^x$ and $\mathbb{P}^x$ their respective laws in $(\mathcal{C}, \mathcal{F})$. Then*

  (i) the specific relative entropy of $\mathbb{Q}^x$ with respect to $\mathbb{P}^x$ exists, and

  (ii) it has the closed form

$$h(\mathbb{Q}^x \,|\, \mathbb{P}^x) = \frac{1}{2}\mathbb{E}\left[\int_0^1 \left\{ \frac{\sigma(M_s^x)^2}{\eta(M_s^x)^2} - 1 - \log\frac{\sigma(M_s^x)^2}{\eta(M_s^x)^2} \right\} ds \right]. \tag{1.1}$$

Recalling that X stood for the canonical process, Formula (1.1) becomes

$$h(\mathbb{Q}^x \,|\, \mathbb{P}^x) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}^x}\left[\int_0^1 \left\{ \frac{\sigma(X_s)^2}{\eta(X_s)^2} - 1 - \log\frac{\sigma(X_s)^2}{\eta(X_s)^2} \right\} ds \right].$$

A formula of this very kind had been conjectured in [6, Kapitel I.1] in the case that $\eta \equiv 1$. Specifically, calling $\mathbb{W}$ the Wiener measure and $\mathbb{Q}$ a martingale law (w.l.o.g. both started at zero), the conjecture asserted the equality

$$h(\mathbb{Q} \,|\, \mathbb{W}) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \left\{ \sigma(X_s)^2 - 1 - \log\sigma(X_s)^2 \right\} ds \right].$$

In fact, it was already established in [6, 4, 2] that the r.h.s. here is always the smaller one, and this fact has already found an application in [4, 2] were a novel transport-information inequality was derived. Yet to the best of our knowledge the conjecture (i.e. the equality of left- and right-hand sides) had only been confirmed in the following cases: for Gaussian martingales, i.e. when σ is deterministic (that is to say, not a function of space), for the case of geometric Brownian motion, and for the case of $M_t = \int_0^t B_s dB_s$. Our contribution is hence to provide a first corroboration of this conjecture beyond the Gaussian case in a rather systematic way. Furthermore we also allowed for η being non-constant. This seems relevant to us, as one could be possibly interested in other reference martingales than Brownian motion. The proof of Formula (1.1) is achieved by careful expansion of the transition densities of the involved martingales. This goes some way into explaining why we restrict ourselves to martingale diffusions with fairly regular coefficients.

As a corollary to Theorem 1.4 we are able to obtain a statistical estimator for the specific relative entropy based on independent samples. As in functional data analysis [9] (in the so-called asymptotic dense regime), the idea is that we have access to iid sample paths of a continuous process which we can observe only at mesh points, the size of the mesh going to zero as the number of sampled paths grows (in particular, no sampled path can be observed at more than a finite number of time instances). This is the content of our second main result:

Corollary 1.5. *Under Assumption 1.3, let M^x and N^x be as before and call $\mathbb{Q}^x$ and $\mathbb{P}^x$ their respective laws in $(\mathcal{C}, \mathcal{F})$. Let $\{M^{x,n}\}_{n \in \mathbb{N}}$ be iid samples from $\mathbb{Q}^x$. Then we have the almost sure convergence*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k \leq n} \frac{1}{k} \log\left(\frac{d\mathbb{Q}^x}{d\mathbb{P}^x}\Big|_{\mathcal{F}^k}\left(M_{\frac{0}{k}}^{x,k}, M_{\frac{1}{k}}^{x,k}, \ldots, M_{\frac{k}{k}}^{x,k} \right) \right) = h(\mathbb{Q}^x \,|\, \mathbb{P}^x). \tag{1.2}$$

The proof of this corollary is based on Theorem 1.4, a strong law of large numbers for independent (possibly non-identically distributed) samples, and the short time behaviour of the transition densities of the involved martingales.

In our last main result we obtain an almost-sure counterpart to Theorem 1.4 along dyadic partitions. For simplicity we restrict ourselves to the case when the reference martingale N^x is Brownian motion started at x.

Proposition 1.6. *Suppose that $\eta \equiv 1$ and that σ fulfills the conditions in Assumption 1.3. Then we have the almost-sure limit*

$$\lim_{k \to \infty} \frac{1}{2^k} \log \left(\frac{d\mathbb{Q}^x}{d\mathbb{P}^x} \Big|_{\mathcal{F}^{2^k}} \left(M^x_{\ell \cdot 2^{-k}} \right)^{2^k}_{\ell=0} \right) = \frac{1}{2} \int_0^1 \left\{ \sigma(M^x_s)^2 - 1 - \log \sigma(M^x_s)^2 \right\} ds \qquad (1.3)$$

This proposition establishes a conjecture in [6, p. 15] in the case when M is a martingale diffusion satisying our strong regularity assumptions. Incidentally, our proof reveals that the same result is true if we replace the role of 2^k by n_k, as long as $\sum_{k=0}^\infty n_k < \infty$.

2 Proof of Theorem 1.4

The proof of Theorem 1.4 relies on the following rather precise estimates on the transition density function of martingale diffusions as above. Throughout we denote

$$d_\sigma(a, b) := \left| \int_a^b \frac{du}{\sigma(u)} \right|.$$

The metric $d_\sigma(a, b)$ is the geodesic distance induced by σ (more precisely, by σ^2) on $\mathbb{R}$.

Lemma 2.1. *Let σ and M^x be as above and let $p(t, x, y)$ be the transition density function of M^x from x at time 0 to y at time t. Under Assumption 1.3 there are constants C_1 and C_2, depending only on δ and L, such that, for all $t \in (0, 1]$ and $x, y \in \mathbb{R}$:*

$$e^{-C_2 t} \frac{1}{\sqrt{2\pi t}} \sqrt{\frac{\sigma(x)}{\sigma(y)}} \frac{1}{\sigma(y)} e^{\left(-d_\sigma(x,y)^2/2t\right)} \leq p(t, x, y) \leq e^{C_1 t} \frac{1}{\sqrt{2\pi t}} \sqrt{\frac{\sigma(x)}{\sigma(y)}} \frac{1}{\sigma(y)} e^{\left(-d_\sigma(x,y)^2/2t\right)}.$$

The content of Lemma 2.1 is surely known to experts, as the study of asymptotic expansions for the heat kernel is a classical subject. On the other hand Lemma 2.1 is non-asymptotic and of global nature, which we find useful. It seems to us that this result should follow from the precise estimates in [7, Corollary 3.1 and Theorem 4.1], but we cannot guarantee this assertion as we lack geometric training. Hence we rather provide our own self-contained proof:

Proof of Lemma 2.1. We use throughout that B is a Brownian motion started at the origin. Let $g \colon \mathbb{R} \to \mathbb{R}$ be the strictly increasing and surjective function defined by

$$g(y) := \int_x^y \frac{1}{\sigma(u)} \, du,$$

and g^{-1} its inverse. By Itô's lemma we have

$$g(M^x_t) = B_t - \frac{1}{2} \int_0^t \sigma'(M^x_s) \, ds.$$

Hence $Y_t := g(M^x_t)$ satisfies $Y_t = B_t + \int_0^t b(Y_s) ds$, where $b := -\frac{1}{2} \sigma' \circ g^{-1}$.

For any $\Phi \geq 0$ measurable and bounded, we have

$$\mathbb{E}\left[\Phi(M^x_t)\right] = \mathbb{E}\left[\Phi(g^{-1}(Y_t))\right]$$

$$= \mathbb{E}\left[\Phi(g^{-1}(B_t)) \exp\left(\int_0^t b(B_s) \, dB_s - \frac{1}{2} \int_0^t b(B_s)^2 \, ds\right)\right]$$

$$= \mathbb{E}\left[\Phi(g^{-1}(B_t)) \exp\left(\int_0^{B_t} b(u) \, du - \frac{1}{2} \int_0^t b'(B_s) \, ds - \frac{1}{2} \int_0^t b(B_s)^2 \, ds\right)\right] \quad (2.1)$$

where the second equality is due to Girsanov's theorem (applicable as σ' is bounded), whereas the last equality follows from

$$\int_0^{B_t} b(u)\,du = \int_0^t b(B_s)\,dB_s + \frac{1}{2}\int_0^t b'(B_s)\,ds,$$

which is just Itô's lemma applied to the function $f(z) := \int_0^z b(u)\,du$.

To get an upper bound for (2.1) note that,

$$-b' - b^2 \leq -b' = \frac{1}{2}(\sigma \cdot \sigma'') \circ g^{-1} \leq \frac{L}{2\delta} \tag{2.2}$$

which implies,

$$\exp\left(-\frac{1}{2}\int_0^t b'(B_s)\,ds - \frac{1}{2}\int_0^t b(B_s)^2\,ds\right) \leq \exp\left(t\frac{L}{4\delta}\right).$$

Inserting in (2.1) with $C_1 := \frac{L}{4\delta}$ yields,

$$\mathbb{E}\left[\Phi(M_t^x)\right] \leq \exp\left(tC_1\right)\mathbb{E}\left[\Phi(g^{-1}(B_t))\exp\left(\int_0^{B_t} b(u)\,du\right)\right]. \tag{2.3}$$

Note that

$$b = -\frac{1}{2}\sigma' \circ g^{-1} = -\frac{1}{2}\frac{(\sigma \circ g^{-1})'}{\sigma \circ g^{-1}} = -\frac{1}{2}(\log \circ \sigma \circ g^{-1})'$$

and $g^{-1}(B_0) = g^{-1}(0) = x$. Therefore

$$\int_0^{B_t} b(u)\,du = -\frac{1}{2}\log\left(\frac{\sigma(g^{-1}(B_t))}{\sigma(g^{-1}(B_0))}\right) = \log\sqrt{\frac{\sigma(x)}{\sigma(g^{-1}(B_t))}}$$

We can now rewrite the expectation on the right-hand side of (2.3) by using this observation and then expressing it in terms of the density of B_t. The desired representation then follows from a change of variables with $y = g^{-1}(z)$, namely:

$$\mathbb{E}\left[\Phi(g^{-1}(B_t))\exp\left(\int_0^{B_t} b(u)\,du\right)\right]$$

$$= \mathbb{E}\left[\Phi(g^{-1}(B_t))\sqrt{\frac{\sigma(x)}{\sigma(g^{-1}(B_t))}}\right]$$

$$= \int_{\mathbb{R}} \Phi(g^{-1}(z))\sqrt{\frac{\sigma(x)}{\sigma(g^{-1}(z))}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{z^2}{2t}\right)dz$$

$$= \int_{\mathbb{R}} \Phi(y)\sqrt{\frac{\sigma(x)}{\sigma(y)}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{g(y)^2}{2t}\right)g'(y)\,dy$$

$$= \int_{\mathbb{R}} \Phi(y)\sqrt{\frac{\sigma(x)}{\sigma(y)}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{1}{2t}\left(\int_x^y \frac{1}{\sigma(u)}\,du\right)^2\right)\frac{1}{\sigma(y)}\,dy.$$

We conclude that

$$\mathbb{E}\left[\Phi(M_t^x)\right] \leq \exp\left(tC_1\right)\int_{\mathbb{R}} \Phi(y)\sqrt{\frac{\sigma(x)}{\sigma(y)}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{1}{2t}d_\sigma(x,y)^2\right)\frac{1}{\sigma(y)}\,dy.$$

As this holds for all $\Phi \geq 0$ measurable and bounded, this proves the desired upper bound.

The lower bound follows analogously by replacing (2.2) with

$$-b' - b^2 = \frac{1}{2}(\sigma\sigma'') \circ g^{-1} - (\frac{1}{2}\sigma' \circ g^{-1})^2 \geq -\left(\frac{L\delta}{2} + \frac{L^2}{4}\right).$$

This implies

$$\exp\left(-\frac{1}{2}\int_0^t b'(B_s)\,ds - \frac{1}{2}\int_0^t b(B_s)^2\,ds\right) \geq \exp\left(-tC_2\right),$$

where $C_2 := \frac{L\delta}{4} + \frac{L^2}{8}$. □

We are now in a position to prove Theorem 1.4.

Proof of Theorem 1.4. Let $q(t,x,y)$ and $p(t,x,y)$ be the transition density functions of M^x and N^x respectively. To ease notation we now drop the superscript x from M and N. Since M and N are Markov and time-homogeneous, we have by the tensorization property of the relative entropy that

$$\begin{aligned}
\mathrm{H}(\mathbb{Q}^x \,|\, \mathbb{P}^x)|_{\mathcal{F}^n} &= \mathrm{H}\left(\mathcal{L}(M_0, M_{\frac{1}{n}}, \ldots, M_1) \,|\, \mathcal{L}(N_0, N_{\frac{1}{n}}, \ldots, N_1)\right) \\
&= \sum_{k=1}^n \int_{\mathbb{R}^2} \log \frac{d\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = x_{k-1})(x_k)}{d\mathcal{L}(N_{\frac{k}{n}}|N_{\frac{k-1}{n}} = x_{k-1})(x_k)} \, d\mathcal{L}(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})(x_{k-1}, x_k) \\
&= \sum_{k=1}^n \int_{\mathbb{R}^2} \log \frac{q(\frac{1}{n}, x_{k-1}, x_k)}{p(\frac{1}{n}, x_{k-1}, x_k)} \, d\mathcal{L}(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})(x_{k-1}, x_k) \\
&= \sum_{k=1}^n \mathbb{E}\left[\log \frac{q(\frac{1}{n}, M_{\frac{k-1}{n}}, M_{\frac{k}{n}})}{p(\frac{1}{n}, M_{\frac{k-1}{n}}, M_{\frac{k}{n}})}\right]
\end{aligned}$$

By Lemma 2.1 for the upper bound of q and the lower bound of p, we derive the existence of a constant C such that for all $x, y \in \mathbb{R}, t \in (0, 1]$:

$$\log \frac{q(t,x,y)}{p(t,x,y)} \leq Ct + \frac{1}{2}\log\frac{\sigma(x)\eta(y)}{\sigma(y)\eta(x)} - \log\frac{\sigma(y)}{\eta(y)} - \frac{1}{2t}d_\sigma(x,y)^2 + \frac{1}{2t}d_\eta(x,y)^2.$$

Therefore, for all $n \in \mathbb{N}$ and any $k = 1, \ldots, n$:

$$\begin{aligned}
&\mathbb{E}\left[\log \frac{q(\frac{1}{n}, M_{\frac{k-1}{n}}, M_{\frac{k}{n}})}{p(\frac{1}{n}, M_{\frac{k-1}{n}}, M_{\frac{k}{n}})}\right] \\
&\leq \mathbb{E}\left[\frac{C}{n} + \frac{1}{2}\log\frac{\sigma(M_{\frac{k-1}{n}})\eta(M_{\frac{k}{n}})}{\sigma(M_{\frac{k}{n}})\eta(M_{\frac{k-1}{n}})} - \log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})} - \frac{n}{2}d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 + \frac{n}{2}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right].
\end{aligned}$$
(2.4)

Summing over $k = 1, \ldots, n$ the $\log \frac{\sigma(M_{\frac{k-1}{n}})\eta(M_{\frac{k}{n}})}{\sigma(M_{\frac{k}{n}})\eta(M_{\frac{k-1}{n}})}$ terms form a telescopic sum. So for all $n \in \mathbb{N}$:

$$\begin{aligned}
&\frac{1}{n}\mathrm{H}(\mathbb{Q}^x \,|\, \mathbb{P}^x)|_{\mathcal{F}^n} \\
&\leq \frac{1}{n}\sum_{k=1}^n \mathbb{E}\left[\frac{C}{n} + \frac{1}{2}\log\frac{\sigma(M_{\frac{k-1}{n}})\eta(M_{\frac{k}{n}})}{\sigma(M_{\frac{k}{n}})\eta(M_{\frac{k-1}{n}})} - \log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})} - \frac{n}{2}d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 \right. \\
&\qquad\qquad \left. + \frac{n}{2}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right]
\end{aligned}$$

$$= \frac{C}{n} + \mathbb{E}\left[\frac{1}{2n}\log\frac{\sigma(M_0)\eta(M_1)}{\sigma(M_1)\eta(M_0)} - \frac{1}{n}\sum_{k=1}^{n}\log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})} - \frac{1}{2}\sum_{k=1}^{n}d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right.$$
$$\left. + \frac{1}{2}\sum_{k=1}^{n}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right] \tag{2.5}$$

We now show that (2.5) converges to the right-hand side of (1.1). Firstly, since σ and η are bounded away from 0 and bounded from above, it is clear that,

$$\lim_{n\to\infty}\mathbb{E}\left[\frac{1}{2n}\log\frac{\sigma(M_0)\eta(M_1)}{\sigma(M_1)\eta(M_0)}\right] = 0$$

Furthermore for almost every $\omega \in \Omega$ the map $t \mapsto \log(\frac{\sigma(M_t(\omega))}{\eta(M_t(\omega))})$ is continuous, so that

$$\sum_{k=1}^{n}\log\frac{\sigma(M_{\frac{k}{n}}(\omega))}{\eta(M_{\frac{k}{n}}(\omega))}\mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t) \xrightarrow{n\to\infty} \log\frac{\sigma(M_t(\omega))}{\eta(M_t(\omega))} \qquad \text{Leb} \otimes \mathbb{P} - a.e.$$

Since the left-hand side above is bounded uniformly in n, we have by dominated convergence

$$\lim_{n\to\infty}\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})}\right] = \lim_{n\to\infty}\mathbb{E}\left[\int_0^1\sum_{k=1}^{n}\log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})}\mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)\,ds\right]$$
$$= \mathbb{E}\left[\int_0^1\log\frac{\sigma(M_s)}{\eta(M_s)}\,ds\right].$$

Next, Itô's lemma applied to the function $F(z) := \int_0^z\frac{1}{\eta(u)}\,du$ yields,

$$F(M_t) = \int_0^{M_t}\frac{1}{\eta(u)}\,du = \int_0^t\frac{\sigma(M_u)}{\eta(M_u)}\,dB_u - \frac{1}{2}\int_0^t\frac{\eta'(M_u)}{\eta(M_u)^2}\sigma(M_u)^2\,du.$$

Thus we have

$$\mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right] = \mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}\left(\int_{M_{\frac{k-1}{n}}}^{M_{\frac{k}{n}}}\frac{1}{\eta(u)}\,du\right)^2\right]$$
$$= \mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}\left(F(M_{\frac{k}{n}}) - F(M_{\frac{k-1}{n}})\right)^2\right].$$

At this point one can directly deduce

$$\lim_{n\to\infty}\mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right] = \mathbb{E}\left[\int_0^1\frac{1}{2}\left(\frac{\sigma(M_u)}{\eta(M_u)}\right)^2\,du\right], \tag{2.6}$$

from the fact that

$$\mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}\left(F(M_{\frac{k}{n}}) - F(M_{\frac{k-1}{n}})\right)^2\right] \to \mathbb{E}[\langle F(M)\rangle_1],$$

as follows directly per properties of the quadratic variation of Itô processes with bounded coefficients. To be self-contained we provide now a detailed argument:

$$\mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}\left(F(M_{\frac{k}{n}}) - F(M_{\frac{k-1}{n}})\right)^2\right]$$

$$= \frac{1}{2} \sum_{k=1}^{n} \mathbb{E}\left[\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \right)^2 - \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2} \sigma(M_u)^2 \, du \right.$$

$$\left. + \frac{1}{4} \left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2} \sigma(M_u)^2 \, du \right)^2 \right]$$

$$= \mathbb{E}\left[\frac{1}{2} \int_0^1 \left(\frac{\sigma(M_u)}{\eta(M_u)} \right)^2 \, du - \frac{1}{2} \sum_{k=1}^{n} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2} \sigma(M_u)^2 \, du \right.$$

$$\left. + \frac{1}{2} \sum_{k=1}^{n} \frac{1}{4} \left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2} \sigma(M_u)^2 \, du \right)^2 \right].$$

The expression above converges to $\mathbb{E}\left[\frac{1}{2} \int_0^1 \left(\frac{\sigma(M_u)}{\eta(M_u)} \right)^2 \, du \right]$ as $n \to \infty$. Indeed, since $\|\eta'\|_\infty < L$ and $\sigma, \eta \in (\delta, \frac{1}{\delta})$, we have that

$$\mathbb{E}\left[\left| \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \right| \right] \leq \left(\mathbb{E}\left[\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \right)^2 \right] \right)^{\frac{1}{2}}$$

$$= \left(\mathbb{E}\left[\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)^2}{\eta(M_u)^2} \, du \right] \right)^{\frac{1}{2}}$$

$$\leq \sqrt{\frac{1}{n}} \frac{1}{\delta^2},$$

and therefore as $n \to \infty$

$$\left| \sum_{k=1}^{n} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2} \sigma(M_u)^2 \, du \right| \leq \frac{L}{\delta^4} \frac{1}{n} \sum_{k=1}^{n} \left| \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)} \, dB_u \right| \xrightarrow{L^1} 0.$$

Moreover,

$$\sum_{k=1}^{n} \left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2} \sigma(M_u)^2 \, du \right)^2 \leq \sum_{k=1}^{n} \frac{L^2}{\delta^8} \frac{1}{n^2} = \frac{L^2}{\delta^8} \frac{1}{n} \xrightarrow{L^1} 0 \quad \text{as } n \to \infty.$$

All in all, we have established (2.6). Applying the same arguments with η replaced by σ, shows

$$\lim_{n \to \infty} \mathbb{E}\left[\frac{1}{2} \sum_{k=1}^{n} d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 \right] = \frac{1}{2}.$$

Finally we conclude that

$$h^u(\mathbb{Q}^x | \mathbb{P}^x) := \limsup_{n \to \infty} \frac{1}{n} \mathrm{H}(\mathbb{Q}^x \, | \, \mathbb{P}^x)|_{\mathcal{F}^n} \leq \frac{1}{2} \mathbb{E}\left[\int_0^1 \frac{\sigma(M_s)^2}{\eta(M_s)^2} - 1 - \log \frac{\sigma(M_s)^2}{\eta(M_s)^2} \, ds \right].$$

Analogously, by the lower bound for q and the upper bound for p from Lemma 2.1, applied in (2.4), we get:

$$h^\ell(\mathbb{Q}^x | \mathbb{P}^x) := \liminf_{n \to \infty} \frac{1}{n} \mathrm{H}(\mathbb{Q}^x \, | \, \mathbb{P}^x)|_{\mathcal{F}^n} \geq \frac{1}{2} \mathbb{E}\left[\int_0^1 \frac{\sigma(M_s)^2}{\eta(M_s)^2} - 1 - \log \frac{\sigma(M_s)^2}{\eta(M_s)^2} \, ds \right],$$

which finishes the proof. $\square$

Remark 2.2. We find the limit in (2.6), equivalently

$$\lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^x}\left[\sum_{k=1}^n d_\eta(X_{\frac{k-1}{n}}, X_{\frac{k}{n}})^2\right] = \mathbb{E}_{\mathbb{Q}^x}\left[\int_0^1 \eta(X_u)^{-2} d\langle X\rangle_u\right],$$

curious and worth highlighting. We suspect that there is a geometric interpretation to this result, but we do not pursue this line of reasoning here.

3 Proof of Corollary 1.5

In this part it will be convenient to apply a weaker version of Lemma 2.1, namely:

Lemma 3.1. *Let σ, M^x and $p(t,x,y)$ as in Lemma 2.1. Under Assumption 1.3 there is a constant C, depending only on δ and L, such that, for all $t \in (0,1]$ and $x,y \in \mathbb{R}$:*

$$e^{-Ct}\frac{\delta^2}{\sqrt{2\pi t}}e^{\left(-(x-y)^2/2t\delta^2\right)} \le p(t,x,y) \le e^{Ct}\frac{1}{\delta^2\sqrt{2\pi t}}e^{\left(-\delta^2(x-y)^2/2t\right)}.$$

This follows directly from Lemma 2.1 and Assumption 1.3 on σ, so we omit the proof. We remark that at least the upper bound in Lemma 3.1 can be obtained in far more generality with PDE methods (see Theorems 4.5 and 5.4 in [5, Chapter 6]).

We put ourselves now in the setting of Corollary 1.5, and introduce some notation here: If $\mathbb{Q}$ is a probability measure on $\mathcal{C}$ and if $\mathbb{Q}|_{\mathcal{F}^n}$ possesses a density with respect to n-dimensional Lebesgue measure,[1] we denote this density by the lower-case symbol q_n. If $X \in \mathcal{C}$ we denote $X_{(n)} := (X_{i/n})_{i=0}^n$. Written in this terms, our aim is to establish the almost sure convergence

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k\le n}\frac{1}{k}\log\left(\frac{q_k^x}{p_k^x}(M_{(k)}^{x,k})\right) = h(\mathbb{Q}^x \mid \mathbb{P}^x). \tag{3.1}$$

We now omit the x-dependence everywhere, in order to keep notation at bay.

Proof of Corollary 1.5. We will use the following Kolmogorov's version of the strong law of large numbers (KSLLN) for independent non-identically distributed sequences: If $(Y_k)_k$ are independent, with finite second moments, and such that $\sum_k \frac{1}{k^2}\mathrm{Var}(Y_k) < \infty$, then $\frac{1}{n}\sum_{k\le n}(Y_k - \mathbb{E}[Y_k]) \to 0$ almost surely as $n \to \infty$. This can be found in [8, Theorem 2.3.10]. For us

$$Y_k := \frac{1}{k}\log\left(\frac{q_k}{p_k}(M_{(k)}^k)\right),$$

so by definition the $(Y_k)_k$ are independent and $\mathbb{E}[Y_k] = \frac{1}{k}\mathrm{H}(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}^k}$. Per Theorem 1.4 we know $\frac{1}{k}\mathrm{H}(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}^k} \to h(\mathbb{Q} \mid \mathbb{P})$, hence $\frac{1}{n}\sum_{k\le n}\mathbb{E}[Y_k] \to h(\mathbb{Q} \mid \mathbb{P})$. Thus if we can apply KSLLN then we will have derived (3.1) as desired. Thus it remains to verify that $\sum_k \frac{1}{k^2}\mathrm{Var}(Y_k) < \infty$, but as $\sum_k \frac{1}{k^2}\mathbb{E}[Y_k]^2 < \infty$ in our case, it is enough to verify

$$\sum_k \frac{1}{k^2}\mathbb{E}[Y_k^2] < \infty. \tag{3.2}$$

To this end remark that $Y_k^2 \le \frac{2}{k^2}\left\{\log(q_k(M_{(k)}^k))^2 + \log(p_k(M_{(k)}^k))^2\right\}$. We now establish

$$\sum_k \frac{1}{k^4}\mathbb{E}[\log(\ell_k(M_{(k)}))^2] < \infty, \tag{3.3}$$

---

[1]To be fully precise: the first marginal of $\mathbb{Q}|_{\mathcal{F}^n}$ is a Dirac measure. But as all martingales we consider are started at a fixed x we may "drop" the first marginal and consider the density of the remaining arguments w.r.t. Lebesgue.

where we let the symbol ℓ stand for either p or q, implying then (3.2). Applying the Markov property we have

$$\ell_k(M_{(k)}) = \prod_{i=1}^k \ell(1/k, M_{(i-1)/k}, M_{i/k}),$$

so taking logarithms, bounding the squared sum by a sum of squares, and applying Lemma 3.1, we obtain

$$\log(\ell_k(M_{(k)}))^2 \leq 4k \sum_{i \leq k} \left\{ \frac{C^2}{k^2} + \log(\delta^2)^2 + \frac{1}{4}\log(k/2\pi)^2 + \frac{k^2}{4\delta^4}(M_{(i-1)/k} - M_{i/k})^4 \right\}$$

$$= 4C^2 + 4k^2\log(\delta^2)^2 + k^2\log(k/2\pi)^2 + \frac{k^3}{\delta^4}\sum_{i \leq k}(M_{(i-1)/k} - M_{i/k})^4.$$

To check (3.3), first observe $\sum_k (4C^2 + 4k^2\log(\delta^2)^2 + k^2\log(k/2\pi)^2)/k^4 < \infty$, as $\log(x) \leq x^r/r$ for all $x \geq 1$ and $r \in (0,1)$. Second, by the BDG inequality, we have

$$\mathbb{E}[(M_{(i-1)/k} - M_{i/k})^4] \leq c\mathbb{E}\left[\left(\int_{(i-1)/k}^{i/k} \sigma(M_t)^2\, dt\right)^2\right] \leq \frac{c}{\delta^4 k^2},$$

so $\sum_{i \leq k} \mathbb{E}[(M_{(i-1)/k} - M_{i/k})^4] \leq \frac{c}{\delta^4 k}$ and $\sum_k \frac{1}{k^4}\sum_{i \leq k}\frac{k^3}{\delta^4}\mathbb{E}[(M_{(i-1)/k} - M_{i/k})^4] \leq c'\sum_k \frac{1}{k^2} < \infty$. □

4   Proof of Proposition 1.6

Proof of Proposition 1.6. To ease notation we now drop x as a superscript. Let $q(t,x,y)$ and $p(t,x,y)$ denote the transition density functions of respectively M and $N = B$. As in the proof of Theorem 1.4 we start with the inequality

$$\log \frac{q(\frac{1}{n}, M_{\frac{\ell-1}{n}}, M_{\frac{\ell}{n}})}{p(\frac{1}{n}, M_{\frac{\ell-1}{n}}, M_{\frac{\ell}{n}})}$$
$$\leq \frac{C}{n} + \frac{1}{2}\log\frac{\sigma(M_{\frac{\ell-1}{n}})}{\sigma(M_{\frac{\ell}{n}})} - \log\sigma(M_{\frac{\ell}{n}}) - \frac{n}{2}d_\sigma(M_{\frac{\ell-1}{n}}, M_{\frac{\ell}{n}})^2 + \frac{n}{2}|M_{\frac{\ell-1}{n}} - M_{\frac{\ell}{n}}|^2. \qquad (4.1)$$

By the time-homogeneous Markov property we have

$$\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}^{2^k}}\left(M_{\ell \cdot 2^{-k}}\right)_{\ell=0}^{2^k}\right) = \sum_{\ell=1}^{2^k}\log\frac{q(2^{-k}, M_{(\ell-1)\cdot 2^{-k}}, M_{\ell \cdot 2^{-k}})}{p(2^{-k}, M_{(\ell-1)\cdot 2^{-k}}, M_{\ell \cdot 2^{-k}})},$$

so our pre-limit quantity of interest,

$$\frac{1}{2^k}\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}^{2^k}}\left(M_{\ell \cdot 2^{-k}}\right)_{\ell=0}^{2^k}\right),$$

is bounded from above by

$$C2^{-k} + 2^{-k-1}\log\frac{\sigma(M_1)}{\sigma(M_0)} - \alpha_k - \beta_k + \delta_k,$$

where

$$\alpha_k := 2^{-k}\sum_{\ell=1}^{2^k}\log\sigma(M_{\ell \cdot 2^{-k}}),$$

$$\beta_k := \frac{1}{2} \sum_{\ell=1}^{2^k} d_\sigma(M_{(\ell-1)\cdot 2^{-k}}, M_{\ell\cdot 2^{-k}})^2,$$

$$\delta_k := \frac{1}{2} \sum_{\ell=1}^{2^k} |M_{(\ell-1)\cdot 2^{-k}} - M_{\ell\cdot 2^{-k}}|^2.$$

We have the almost-sure limit $\lim_k \alpha_k = \int_0^1 \log\sigma(M_t)dt$, by convergence of Riemann sums, the function $t \mapsto \log\sigma(M_t)$ being continuous and bounded. We observe also that

$$d_\sigma(M_{(\ell-1)\cdot 2^{-k}}, M_{\ell\cdot 2^{-k}})^2 = \left| B_{(\ell-1)\cdot 2^{-k}} - B_{\ell\cdot 2^{-k}} - \frac{1}{2}\int_{(\ell-1)\cdot 2^{-k}}^{\ell\cdot 2^{-k}} \sigma'(M_t)dt \right|^2,$$

by applying Itô formula to the process $\int_0^{M_t} \frac{dy}{\sigma(y)}$. Hence

$$2\beta_k = \sum_{\ell=1}^{2^k} \left\{ |B_{(\ell-1)\cdot 2^{-k}} - B_{\ell\cdot 2^{-k}}|^2 + \frac{1}{4}\left| \int_{(\ell-1)\cdot 2^{-k}}^{\ell\cdot 2^{-k}} \sigma'(M_t)dt \right|^2 - (B_{(\ell-1)\cdot 2^{-k}} \right.$$
$$\left. -B_{\ell\cdot 2^{-k}}) \int_{(\ell-1)\cdot 2^{-k}}^{\ell\cdot 2^{-k}} \sigma'(M_t)dt \right\}.$$

Now we observe the almost-sure limits $\sum_\ell |B_{(\ell-1)\cdot 2^{-k}} - B_{\ell\cdot 2^{-k}}|^2 \to 1$ and, as σ' is bounded, also $\sum_\ell |\int_{(\ell-1)\cdot 2^{-k}}^{\ell\cdot 2^{-k}} \sigma'(M_t)dt|^2 = 0$. This and Cauchy-Schwarz establish that $\beta_k \to 1/2$ almost-surely. Finally we remark that $2\delta_k \to \int_0^1 \sigma(M_t)^2 dt$ almost surely.[2]

All in all, we have proved the almost-sure bound

$$\limsup_k \frac{1}{2^k} \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}^{2^k}} \left(M_{\ell\cdot 2^{-k}}\right)_{\ell=0}^{2^k} \right) \le \frac{1}{2}\int_0^1 \left\{ \sigma(M_s)^2 - 1 - \log\sigma(M_s)^2 \right\} ds.$$

Since similar arguments deal with the lower bound and the limes inferior, we conclude. $\qquad\square$

References

[1] Monroe Donsker and Srinivasa Varadhan. Large deviations for stationary gaussian processes. *Communications in Mathematical Physics*, 97(1):187–210, 1985. MR0782966

[2] Hans Föllmer. Doob decomposition, Dirichlet processes and entropies on Wiener space. arXiv preprint 2022. MR4508843

[3] Hans Föllmer. Random fields and diffusion processes. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, pages 101–203. Springer, 1988. MR0983373

[4] Hans Föllmer. Optimal couplings on Wiener space and an extension of Talagrand's transport inequality. In *Stochastic Analysis, Filtering, and Stochastic Optimization*, pages 147–175. Springer, 2022. MR4433813

---

[2]In case that the reader is not aware of this fact we next provide a self-contained argument: First we estimate for $r < t$ that $\mathbb{E}[|(M_t - M_r)^2 - \int_r^t \sigma(M_s)^2 ds|^2] = \mathbb{E}[|\int_r^t (M_s - M_r)dM_s|^2] \le c|t-r|^2$, with c only depending on $\|\sigma\|_\infty$. Then we observe that $\mathbb{E}[\{(M_t - M_r)^2 - \int_r^t \sigma(M_s)^2 ds\} \,|\, \mathcal{F}_r] = \mathbb{E}[\int_r^t (M_s - M_r)dM_s \,|\, \mathcal{F}_r] = 0$. Hence developing squares and using the tower property we have

$$\mathbb{E}\left[\left| \sum_{\ell=1}^n (M_{\frac{\ell}{n}} - M_{\frac{(\ell-1)}{n}})^2 - \int_0^1 \sigma(M_t)^2 dt \right|^2 \right] = \sum_{\ell=1}^n \mathbb{E}\left[\left| (M_{\frac{\ell}{n}} - M_{\frac{(\ell-1)}{n}})^2 - \int_{\frac{(\ell-1)}{n}}^{\frac{\ell}{n}} \sigma(M_s)^2 ds \right|^2 \right] \le \frac{c}{n}.$$

A Borel-Cantelli argument yields that $\sum_{\ell=1}^{2^k} |M_{(\ell-1)\cdot 2^{-k}} - M_{\ell\cdot 2^{-k}}|^2 \to \int_0^1 \sigma(M_t)^2 dt$ almost-surely, since the sequence $\{2^{-k}\}_k$ is summable.

[5] Avner Friedman. Stochastic differential equations and applications. In *Stochastic Differential Equations*, pages 75–148. Springer, 2010. MR2830391

[6] Nina Gantert. *Einige große Abweichungen der Brownschen Bewegung*. Number 224. Universität Bonn, 1991. MR1179159

[7] Peter Li and Shing Tung Yau. On the parabolic kernel of the Schrödinger operator. *Acta Mathematica*, 156:153–201, 1986. MR0834612

[8] Pranab K Sen and Julio M Singer. *Large Sample Methods in Statistics: An Introduction with Applications*, volume 25. CRC Press, 1994. MR1293125

[9] Jane-Ling Wang, Jeng-Min Chiou, and Hans-Georg Müller. Functional data analysis. *Annual Review of Statistics and Its Application*, 3:257–295, 2016.