MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

" On the Specific Relative Entropy
between Martingale Measures"

verfasst von / submitted by

Clara Unterberger, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	Ass.-Prof. Dr. Julio Daniel Backhoff-Veraguas

Abstract

The specific relative entropy was introduced by N. Gantert in her dissertation, [3], to measure the discrepancy between the laws of continuous processes. It arises as a refinement of the standard relative entropy when the laws in question are mutually singular. Recently, H. Föllmer has rekindled the interest on the subject in the work [1] where the specific relative entropy appears in a novel transport-information inequality.

We first introduce all necessary concepts from stochastic analysis and some important properties of relative entropy. With these at hand we summarize the fundamental aspects of the specific relative entropy by discussing and illustrating the contribution [3, Chapter 1] in detail. Here, the focus is on the specific relative entropy of the law of a continuous martingale with respect to Wiener measure. The two main results in that regard are: there exists a closed form expression for the specific relative entropy in terms of the quadratic variation in the Gaussian case; and at least an inequality holds in all generality.

This leads to the conjecture that this expression gives a formula for the specific relative entropy in the general case as well. The main contribution of this thesis is to verify this conjecture in two directions. Firstly, we consider a class of time-inhomogeneous diffusions, so-called monotone transformations of Brownian motion. Secondly, we consider time-homogeneous diffusions with rather regular coefficients. In both cases we show the validity of the conjectured formula for the specific relative entropy.

Zusammenfassung

Die spezifische relative Entropie wurde von N. Gantert in ihrer Dissertation, [3], eingeführt um die Diskrepanz zwischen zwei Verteilungen stetiger Prozesse zu messen. Sie tritt als Verfeinerung der gewöhnlichen relativen Entropie auf, wenn die betreffenden Verteilungen zueinander singulär sind. Kürzlich wurde durch die Arbeit [1] von H. Föllmer, in der die spezifische relative Entropie in einer Transport-Informations Ungleichung auftaucht, das Interesse an diesem Thema neu geweckt.

Wir beginnen damit alle nötigen Konzepte aus den Bereichen Stochastische Analysis und relative Entropie betreffend vorzustellen. Anschließend fassen wir die fundamentalen Aspekte der spezifisch relativen Entropie zusammen, indem wir den Beitrag [3, Chapter 1] detailliert erläutern. Dabei liegt der Fokus auf der spezifischen relativen Entropie zwischen der Verteilung eines stetigen Martingals und dem Wiener Maß. Die zwei Hauptresultate diesbezüglich sind, dass, im Falle eines Gaußprozesses, ein geschlossener Ausdruck für die spezifisch relative Entropie bezüglich der quadratischen Variation existiert und dass im allgemeinen Fall zumindest eine Ungleichung gilt.

Dies lässt vermuten, dass dieser Ausdruck auch im allgemeinen Fall eine Formel für die spezifische relative Entropie darstellt. Der Hauptbeitrag dieser Arbeit besteht darin, die Vermutung in zwei Richtungen zu verifizieren. Wir betrachten eine Klasse Zeit-inhomogener Diffusionsprozesse, sogenannte monotone Transformationen der brownschen Bewegung, ebenso wie Zeit-homogene Diffusionsprozesse mit recht starken Regularitätsanforderungen an die Koeffizienten, und zeigen die Gültigkeit der vermuteten Formel für die spezifische relative Entropie in diesen Fällen.

Acknowledgements

First and foremost, I would like to thank my advisor Julio Backhoff for all of his advice, ideas, time and positive encouragement. Also a big thank you to the people that made me feel home in Vienna: Hannah, Norbert, Chiara, Julia, Alex, Rajko and especially (against his explicit wish to never be mentioned in a thesis related to mathematics) to Severin!

Contents

1 Preliminaries

In this chapter we recall definitions and results needed in later chapters. We will assume familiarity with measure-theoretic probability theory and concepts involving discrete stochastic processes, Brownian motion and filtrations.

1.1 Relative Entropy

In the following we state some basic properties of the relative entropy between two probability measures, see for example [4, Chapter 7].

Definition 1. Let μ and ν be probability measures on the same measurable space $(\Omega, \mathcal{F})$.
Then the relative entropy between μ and ν is defined as,

$$\mathrm{H}(\mu \,|\, \nu) := \begin{cases} \int_\Omega \log \frac{d\mu}{d\nu}\, d\mu & \text{if } \mu \ll \nu \\ +\infty & \text{otherwise.} \end{cases}$$

where $\frac{d\mu}{d\nu}$ is the Radon-Nikodym density of μ with respect to ν.

Example 1. Let $m_1, m_2 \in \mathbb{R}$ and $\sigma_1, \sigma_2 > 0$. Then

$$\mathrm{H}\left(\mathcal{N}(m_1, \sigma_1^2) \,|\, \mathcal{N}(m_2, \sigma_2^2)\right) = \frac{1}{2}\left(\frac{\sigma_1^2}{\sigma_2^2} - 1 + \frac{(m_2 - m_1)^2}{\sigma_2^2} - \log \frac{\sigma_1^2}{\sigma_2^2}\right).$$

In the special case, where $m_1 = m_2$, we have,

$$\mathrm{H}\left(\mathcal{N}(m_1, \sigma_1^2) \,|\, \mathcal{N}(m_1, \sigma_2^2)\right) = F\left(\frac{\sigma_1^2}{\sigma_2^2}\right),$$

where $F(x) := \frac{1}{2}(x - 1 - \log x)$.

Remark 1. For two probability measures μ and ν on some measurable space $(\Omega, \mathcal{F})$ the relative entropy measures their 'distance' in the following sense: It always holds that $\mathrm{H}(\mu \,|\, \nu) \geq 0$ and equality holds if and only if $\mu = \nu$. However, the relative entropy is not a metric, as it is not symmetric and does not satisfy the triangle inequality.

Lemma 1. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces, μ and ν probability measures on $(X, \mathcal{A})$ such that $\mu \ll \nu$ and $T \colon X \to Y$ such that $\mathcal{A} = \sigma(T) \coloneqq \sigma(\{T^{-1}(B) : B \subset \mathbb{R}$ Borel-measurable $\})$. Then $T(\mu) \ll T(\nu)$ and*

$$\frac{dT(\mu)}{dT(\nu)}(T(x)) = \frac{d\mu}{d\nu}(x) \quad \text{for } \nu\text{-a.e. } x \in X$$

and

$$H(\mu \,|\, \nu) = H(T(\mu) \,|\, T(\nu)) \,.$$

Lemma 2. *For a random vector (V, W) we denote by $\mu_{V,W}$ its law and by $\mu_W^{V=v}$ a regular conditional distribution of W given $V = v$.*
Let $(X_1, ..., X_n)$ and $(Y_1, ..., Y_n)$ be random vectors. Then

$$\begin{aligned} &H(\mu_{X_1,...,X_n} \,|\, \mu_{Y_1,...,Y_n}) \\ &= \sum_{k=2}^{n} \int H\left(\mu_{X_k}^{X_1=x_1,...,X_{k-1}=x_{k-1}} \,|\, \mu_{Y_k}^{Y_1=x_1,...,Y_{k-1}=x_{k-1}}\right) \, d\mu_{X_1,...,X_{k-1}}(x_1, ..., x_{k-1}) \\ &\qquad + H(\mu_{X_1} \,|\, \mu_{Y_1}) \,. \end{aligned}$$

Lemma 3. *Let μ, ν be probability measures on $(X, \mathcal{A})$ such that $H(\mu \,|\, \nu) < \infty$. Let $\mathcal{B} \subset \mathcal{A}$ be a sub-sigma algebra and $\mu|_{\mathcal{B}}, \nu|_{\mathcal{B}}$ be the restrictions of μ and ν to $\mathcal{B}$. Then*

(i)

$$\frac{d\mu|_{\mathcal{B}}}{d\nu|_{\mathcal{B}}} = \mathbb{E}_\nu \left[\frac{d\mu}{d\nu} \,\middle|\, \mathcal{B} \right]$$

(ii)

$$H(\mu|_{\mathcal{B}} \mid \nu|_{\mathcal{B}}) \leq H(\mu \,|\, \nu)$$

1.2 Stochastic Analysis

Most concepts introduced here have more general analogues but we will restrict ourselves to the framework needed later on. All processes considered here are real-valued and we refer to processes with (almost sure) continuous sample paths as continuous processes.

We start by recalling some important properties of martingales, see for example [5].

Definition 2. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mathbb{P})$ be a filtered probability space, where $I \subset \mathbb{R}$. Let $X = (X_t)_{t\in I}$ be a process. Then X is a martingale (wrt. $(\mathcal{F}_t)_{t\in I}$) if

(i) X is adapted,

(ii) X is integrable, i.e. $\mathbb{E}[|X_t|] < \infty$ for all $t \in I$,

(iii) for all $s, t \in I$ such that $s < t$,

$$\mathbb{E}[X_t \,|\, \mathcal{F}_s] = X_s \quad \text{a.s..} \tag{1.1}$$

If instead of (1.1) we have, $\mathbb{E}[X_t \,|\, \mathcal{F}_s] \geq X_s$ $\big(\mathbb{E}[X_t \,|\, \mathcal{F}_s] \leq X_s\big)$, then X is called a submartingale (supermartingale).

Before we focus on continuous martingales, we recall the following important fact about convergence of discrete martingales.

Theorem 1 (Convergence theorem for uniformly integrable martingales)**.** *Let $I = \mathbb{N}_0$ and let $X = (X_n)_{n\in I}$ be a uniformly integrable martingale on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in I}, \mathbb{P})$. Then there exists some $\mathcal{F}_\infty := \sigma(\bigcup_{n\in I} \mathcal{F}_n)$-integrable random variable Z such that $X_n \xrightarrow{n\to\infty} Z$ a.s. and in L^1. Moreover, $X_n = \mathbb{E}[Z \,|\, \mathcal{F}_n]$ for all $n \in I$.*

Remark 2. A martingale $X = (X_n)_{n\in\mathbb{N}_0}$ of the form $X_n = \mathbb{E}[Y \,|\, \mathcal{F}_n]$ for some $Y \in L^1$ is called closed. In that case, X is uniformly integrable and $Z = \mathbb{E}[Y \,|\, \mathcal{F}_\infty]$ a.s., where Z is the limiting random variable from Theorem 1.

From now on we will mostly be interested in index sets of the form $[0, T]$, where $T \in (0, \infty)$. Throughout this and the following chapters we denote by $\lambda := \lambda|_{[0,T]}$ the restriction of Lebesgue measure to $[0, T]$.

Remark 3. We say a process $X = (X_t)_{t\in[0,T]}$ is square-integrable if $\mathbb{E}[X_t^2] < \infty$ for all $t \in [0, T]$. If X is a martingale, then X is square-integrable if and only if $\mathbb{E}[X_T^2] < \infty$.

Theorem 2 (Quadratic Variation)**.** *Let M be a continuous, square-integrable martingale. Then there exists a unique adapted, continuous, non-decreasing process $\langle M \rangle$ with $\langle M \rangle_0 = 0$ such that $M^2 - \langle M \rangle$ is a martingale. The process $\langle M \rangle$ is called the quadratic variation process of M.*

Theorem 3 (Doob's L^2-inequality)**.** *Let M be a continuous martingale. Then*

$$\mathbb{E}\left[\sup_{s\in[0,t]} M_s^2\right] \leq 4\mathbb{E}\left[M_t^2\right],$$

for all $t \in [0,T]$.

The goal now is to give a definition of the integral against Brownian motion. From now on we fix a Brownian motion B on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ and assume the filtration is generated by B completed by $\mathbb{P}$-Null sets.

Definition 3. Let $H = (H_t)_{t\in[0,T]}$ be a process. We call H predictable if the map $H : [0,T] \times \Omega \to \mathbb{R}$ is $\mathcal{P}$-measurable, where

$$\begin{aligned} \mathcal{P} &:= \sigma(X : \text{X is a left continuous and adapted process}) \\ &= \sigma(A \times (s,t] : s < t \in [0,T], A \in \mathcal{F}_s) \end{aligned}$$

is called the predictable sigma algebra.

Definition 4. We call a process simple, if it is of the form

$$H_t(\omega) = \sum_{k=1}^{n} h_{k-1}(\omega)\mathbb{1}_{(t_{k-1},t_k]}(t),$$

for some $n \in \mathbb{N}$, $0 = t_0 < ... < t_n = T$ and h_{k-1} bounded and $\mathcal{F}_{t_{k-1}}$-measurable. We denote the collection of all simple processes by $\mathcal{S}$.

Theorem 4 (Itô Isometry)**.** *There exists a unique linear map*

$$I : L^2([0,T] \times \Omega, \mathcal{P}, \lambda \otimes \mathbb{P}) \to L^2(\Omega, \mathcal{F}, \mathbb{P})$$

such that,

(i) for $H = \sum_{k=1}^{n} h_{k-1}\mathbb{1}_{(t_{k-1},t_k]} \in \mathcal{S}$,

$$I(H) = \sum_{k=1}^{n} h_{k-1}(B_{t_k} - B_{t_{k-1}})$$

and

(ii) for $H \in L^2([0,T] \times \Omega, \mathcal{P}, \lambda \otimes \mathbb{P})$,

$$\|H\|_{L^2([0,T]\times\Omega,\mathcal{P},\lambda\otimes\mathbb{P})} = \|I(H)\|_{L^2(\Omega,\mathcal{F},\mathbb{P})}.$$

Theorem 5. *Let* H *be a predictable process such that* $\mathbb{E}[\int_0^T H_t^2\, dt] < \infty$ *and* I *the map from Theorem 4. Then we can define a continuous process by* $I_t^H := I(H\mathbb{1}_{[0,t]})$*. This process* I^H *is a square-integrable martingale.*

Definition 5. Let H be a predictable process such that $\mathbb{E}[\int_0^T H_t^2\, dt] < \infty$. Then we define the Itô-integral of H wrt. B on the interval $[s,t] \subseteq [0,T]$ as,

$$\int_s^t H_u\, dB_u := I_t^H - I_s^H,$$

where $(I_t^H)_{t\in[0,T]}$ is as in Theorem 5.

Remark 4. The Itô-isometry for the integral of H wrt. B becomes,

$$\mathbb{E}\left[\int_0^t H_s^2\, ds\right] = \mathbb{E}\left[\left(\int_0^t H_u\, dB_u\right)^2\right].$$

Theorem 6. *Let* H *be a predictable process such that* $\mathbb{E}[\int_0^T H_t^2\, dt] < \infty$ *and* $M_t = \int_0^t H_u\, dB_u$*. Then the quadratic variation process of* M *is given by* $\langle M\rangle_t = \int_0^t H_s^2\, ds$*.*

Remark 5. The definition of the Itô integral can be extended to predictable integrands H such that $\int_0^T H_t^2\, dt < \infty$ a.s. in a consistent way. However the resulting process $\int_0^t H_s\, dB_s$ is then in general not a square-integrable martingale. We do not give the details here since we will later always be in the situation covered by Definition 5, but we do not care if we encounter integrands that only satisfy the weaker integrability condition above in the remainder of this chapter.

We now state the most general version of the Itô formula we will need.

Theorem 7 (Itô formula)**.** *Let* H *be a predictable, bounded process and* $M_t := \int_0^t H_u\, dB_u$ *and* $g \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R})$*. Then*

$$g(t, M_t) - g(0, M_0) = \int_0^t \partial_2 g(s, M_s) H_s\, dB_s + \int_0^t \partial_1 g(s, M_s) + \frac{1}{2}\partial_{22} g(s, M_s) H_s^2\, ds$$

Remark 6. In the context of Theorem 7, the process $g(t, M_t)$ is a square-integrable martingale if and only if $\mathbb{E}[\int_0^T \partial_2 g(s, M_s)^2 H_s^2 \, ds] < \infty$ and the ds-part vanishes.

We now introduce an example that combines several aspects of the theory so far and that will play a role again in Chapter 3.

Example 2. Let f be measurable and such that $\mathbb{E}[f(B_1)^2] < \infty$. We can define a martingale by

$$M_t := \mathbb{E}\left[f(B_1) \,|\, \mathcal{F}_t\right].$$

The martingale M has the special property that

$$M_t = f(t, B_t),$$

where $f(t,x) := \int_{\mathbb{R}} f(z+x) \; d\mathcal{N}(0, 1-t)(z)$. This follows from the fact that $B_1 - B_t$ is independent from $\mathcal{F}_t$ and has law $\mathcal{N}(0, 1-t)$, while B_t is $\mathcal{F}_t$ measurable. So we have,

$$M_t = \mathbb{E}\left[f(B_1) \mid \mathcal{F}_t\right] = \mathbb{E}\left[f(B_1 - B_t + B_t) \mid \mathcal{F}_t\right] = \int_{\mathbb{R}} f(z + B_t) \; d\mathcal{N}(0, 1-t)(z).$$

Note that, even if f is not differentiable, the convolution with the Gaussian is regularizing and therefore $f(\cdot, \cdot) \in \mathcal{C}^{1,2}([0,1) \times \mathbb{R})$. Now an application of Itô's formula, Theorem 7, yields,

$$M_t = M_0 + \int_0^t \partial_2 f(s, B_s) \; dB_s + \int_0^t \partial_1 f(s, Bs) + \frac{1}{2} \partial_{22} f(s, B_s) \, ds.$$

Since M is a square-integrable martingale by definition, by Remark 6 the finite variation parts vanish and

$$M_t = \mathbb{E}\left[f(B_1)\right] + \int_0^t \partial_2 f(s, B_s) \, dB_s.$$

In particular this means

$$\langle M \rangle_t = \int_0^t \partial_2 f(s, B_s)^2 \, ds.$$

We now turn to some basic results about stochastic differential equations. The rest of this chapter follows [2, Chapter 5] if not stated otherwise. We consider the stochastic differential equation (SDE) of the form,

$$\begin{cases} dX_t = \sigma(t, X_t)\, dB_t \\ X_0 = x_0, \end{cases} \tag{1.2}$$

where $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ is measurable and $x \in \mathbb{R}$.
By a solution to (1.2) we understand a continuous and adapted process X such that, for all $t \in [0, T]$,

$$X_t = x_0 + \int_0^t \sigma(s, X_s)\, dB_s \quad \text{a.s.}.$$

We will also refer to solutions of (1.2) as diffusions.

Definition 6. We call a solution $X = (X_t)_{t\in[0,T]}$ to (1.2) unique if any other solution $X' = (X'_t)_{t\in[0,T]}$ is such that, a.s.,

$$X_t = X'_t \text{ for all } t \in [0, T].$$

Theorem 8 (Existence and Uniqueness)**.** *Let $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ be measurable and assume there exists a constant $K > 0$ such that for all $t \in [0, T]$ and $x, y \in \mathbb{R}$,*

(i) $|\sigma(t, y) - \sigma(t, x)| \leq K|x - y|$

(ii) $|\sigma(t, x)| \leq K(1 + |x|)$.

Then equation (1.2) has a unique solution $X = (X_t)_{t\in[0,T]}$. Moreover, X is a square-integrable martingale.

Definition 7. We say that an adapted process X has the Markov property if, for all $t \in [0, T]$,

$$\mathcal{L}\big((X_{t+h})_{h\in(0,T-t]} \,|\, \mathcal{F}_t\big) = \mathcal{L}\big((X_{t+h})_{h\in(0,T-t]} \,|\, X_t\big).$$

Remark 7. Let σ be as in Theorem 8. Let $s \in [0, T)$. Denote by $X^{s,x}$ (for $X^{0,x}$ we drop the superscript and write X) the unique solution to

$$\begin{cases} dX_t = \sigma(t, X_t)\, dB_t \quad t \in [s, T] \\ X_s = x, \end{cases}$$

which we understand to be adapted to the filtration generated by $\tilde{B}_h := B_{s+h} - B_s$. For $s < t \in [0,1]$, $x \in \mathbb{R}$ and $A \subset \mathbb{R}$ measurable, define $P(t, A; s, x) := \mathbb{P}(X_t^{s,x} \in A)$. Then the solution X has the Markov property and $(x, A) \mapsto P(t, A; s, x)$ is a regular conditional distribution of X_t given X_s, in particular,

$$\mathbb{P}(X_t \in A \,|\, \mathcal{F}_s) = P(t, A; s, X_s) \quad \text{a.s..}$$

Definition 8. Let X be the solution to 1.2.

(i) We call the function P defined in Remark 7 a transition probability function of X.

(ii) We call X time-homogeneous if $P(t, A; s, x) = P(t-s, A; 0, x)$, i.e. its transition probability function depends on s, t only through $t - s$.

(iii) If the measure $A \mapsto P(t, A; s, x)$ has a density for all $x \in \mathbb{R}, s < t \in [0, T]$, i.e. there exists a function $p(t, y; s, x)$ defined for all $x, y \in \mathbb{R}, s < t \in [0, T]$ such that,

$$P(t, dy; s, x) = p(t, y; s, x)\, dy,$$

then we call p transition density function of X. In the time-homogeneous case we write $p(t - s, x, y)$.

Remark 8. If σ is as in Theorem 8 and additionally $\sigma(t, x) = \sigma(x)$, then X is time-homogeneous.

Example 3. For $\sigma(t, x) = 1$ we get $P(t - s, \cdot, x) = \mathcal{N}(x, t - s)$ which has the transition density

$$p(t - s, x, y) = \gamma_{t-s}(x, y) := \frac{1}{\sqrt{2\pi(t - s)}} e^{-\frac{(x-y)^2}{2(t-s)}}$$

Remark 9. If we strengthen the assumptions on σ from Theorem 8 such that

(i) there exists $\delta > 0$ such that $\delta \leq \sigma(t, x) \leq \frac{1}{\delta}$ for all $t \in [0, T], x \in \mathbb{R}$ and

(ii) there exists $L > 0$ such that $|\sigma(t, y) - \sigma(s, x)| \leq L(|y - x| + |t - s|)$ for all $s, t \in [0, T], x, y \in \mathbb{R}$, then the unique solution X to (1.2) has a transition density function p. Moreover, the density p satisfies that there exist constants $c, C > 0$ such that,

$$p(t, y; s, x) \leq \frac{C}{\sqrt{t-s}} e^{-c\frac{(x-y)^2}{t-s}} \text{ for all } x, y \in \mathbb{R}, s < t \in [0, T].$$

See [2, Chapter 6], Theorem 5.4 for existence of the transition density function and Theorem 4.5 for the Gaussian upper bound.

2 The specific relative Entropy

This chapter follows [3, Chapter 1] very closely. All concepts and statements can be found there unless stated otherwise.
Let $B = (B_t)_{t\in[0,1]}$ be a Brownian motion (starting at 0) on some filtered probability space $(\Omega, \mathcal{S}, \mathcal{S}_t, \mathbb{S})$. We consider the Wiener space $(C, \mathcal{F})$, that is

$$C = C([0,1]) = \{\omega\colon [0,1] \to \mathbb{R} \mid \omega \text{ is continuous } \}$$

and $\mathcal{F} \coloneqq \sigma(X_t : t \in [0,1])$, where $X = (X_t)_{t\in[0,1]}$ is the canonical process , i.e. $X_t(\omega) = \omega(t)$. Further we consider the filtration $\mathcal{F}_t = \sigma(X_s : s \leq t)$.
Let $n \in \mathbb{N}$. For $k = 1, ..., n$, we will often write $\Delta_k^n Y \coloneqq Y_{\frac{k}{n}} - Y_{\frac{k-1}{n}}$ for the k-th increment of a process Y and similarly $\Delta_k^n \omega \coloneqq \omega(\frac{k}{n}) - \omega(\frac{k-1}{n})$ for a function ω.

Definition 9. Let $\mathcal{M}_1(C)$ be the collection of all probability measures on $(C, \mathcal{F})$.

We will mostly be interested in laws of continuous martingales.

Definition 10. A probability measure $\mathbb{Q} \in \mathcal{M}_1(C)$ is called a martingale measure, if the canonical process X is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,1]}$ under $\mathbb{Q}$. We denote the collection of all square-integrable martingale measures as $\mathcal{M}^2$.

Remark 10. Recall that under $\mathbb{Q} \in \mathcal{M}^2$, the quadratic variation process of X is the unique adapted, continuous, non-decreasing process $\langle X \rangle$ with $\langle X \rangle_0 = 0$ that satisfies,

$$X^2 - \langle X \rangle \text{ is a martingale.}$$

Moreover, there exists a sequence $(n_l)_{l\in\mathbb{N}}$ such that $\mathbb{Q}$-a.s for all $t \in [0,1]$,

$$\sum_{k=1}^{n_l} (X_{\frac{k}{n_l}\wedge t} - X_{\frac{k-1}{n_l}\wedge t})^2 \xrightarrow{l\to\infty} \langle X \rangle_t.$$

Let $\mathbb{P}, \mathbb{Q} \in \mathcal{M}_1(C)$. Recall that the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as,

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

Obviously, the relative entropy as a measure of 'distance' is only interesting if $\mathbb{Q} \ll \mathbb{P}$.

Remark 11. It is therefore clear that relative entropy is not a suitable notion if we consider martingale measures. Take as an example $\mathbb{Q} = \mathcal{L}((\sigma B_t)_{t\in[0,1]})$ and $\mathbb{P} = \mathcal{L}((\eta B_t)_{t\in[0,1]})$, where $\sigma, \eta \in \mathbb{R} \setminus \{0\}$ such that $\sigma^2 \neq \eta^2$. Then

$$\sum_{k=1}^{2^n} (X_{\frac{k}{2^n}\wedge t} - X_{\frac{k-1}{2^n}\wedge t})^2 \xrightarrow{n\to\infty} \sigma^2 t \text{ for all } t \in [0,1] \qquad \mathbb{Q}\text{-a.s}$$
$$\sum_{k=1}^{2^n} (X_{\frac{k}{2^n}\wedge t} - X_{\frac{k-1}{2^n}\wedge t})^2 \xrightarrow{n\to\infty} \eta^2 t \text{ for all } t \in [0,1] \qquad \mathbb{P}\text{-a.s},$$

so that neither $\mathbb{Q} \ll \mathbb{P}$ nor $\mathbb{P} \ll \mathbb{Q}$ holds.

To go beyond the absolute continuous case, we define for all $n \in \mathbb{N}$

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

Even if $\mathbb{Q}$ is not absolutely continuous with respect to $\mathbb{P}$ on $(C, \mathcal{F})$, we might have absolute continuity if we restrict the measures $\mathbb{Q}$ and $\mathbb{P}$ to the smaller sigma algebras $\mathcal{F}^n$ for all $n \in \mathbb{N}$.

Definition 11. Let $n \in \mathbb{N}$ and $\mathbb{P}, \mathbb{Q} \in \mathcal{M}_1(\mathcal{C})$ such that $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{P}|_{\mathcal{F}^n}$. We denote the Radon-Nikodym density of $\mathbb{Q}|_{\mathcal{F}^n}$ with respect to $\mathbb{P}|_{\mathcal{F}^n}$ on $(\mathcal{C}, \mathcal{F}^n)$ as $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}^n}$, and define

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

The following observation will be useful many times.

Lemma 4. *Let $n \in \mathbb{N}$. Let $\mathbb{P}, \mathbb{Q}$ be probability measures on $(C, \mathcal{F})$ such that $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{P}|_{\mathcal{F}^n}$. Then $\mathbb{P}$-a.s.,*

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}^n}(\omega) = \frac{d\mathcal{L}_{\mathbb{Q}}\big(X_0, \Delta_1^n X, ..., \Delta_n^n X\big)}{d\mathcal{L}_{\mathbb{P}}\big(X_0, \Delta_1^n X, ..., \Delta_n^n X\big)}\left(\omega(0), \Delta_1^n \omega, ..., \Delta_n^n \omega\right).$$

In particular,

$$H(\mathbb{Q}\,|\,\mathbb{P})\,|_{\mathcal{F}^n} = H(\mathcal{L}_{\mathbb{Q}}(X_0, \Delta_1^n X, ..., \Delta_n^n X)\,|\,\mathcal{L}_{\mathbb{P}}(X_0, \Delta_1^n X, ..., \Delta_n^n X)).$$

Moreover, if $\mathbb{P}$ and $\mathbb{Q}$ are such that X has independent increments. Then

$$H(\mathbb{Q}\,|\,\mathbb{P})\,|_{\mathcal{F}^n} = H(\mathcal{L}_{\mathbb{Q}}(X_0)\,|\,\mathcal{L}_{\mathbb{P}}(X_0)) + \sum_{k=1}^{n} H(\mathcal{L}_{\mathbb{Q}}(\Delta_k^n X)\,|\,\mathcal{L}_{\mathbb{P}}(\Delta_k^n X))\,.$$

Proof. The important observation is, that the map

$$\pi^n \colon (C, \mathcal{F}^n) \to \mathbb{R}^{n+1}$$
$$\pi^n\,(\omega) = \begin{pmatrix} \omega(0) \\ \omega(\frac{1}{n}) - \omega(0) \\ \vdots \\ \omega(1) - \omega(\frac{n-1}{n}) \end{pmatrix}$$

is measurable and $\pi^n(\mathbb{Q}|_{\mathcal{F}^n}) = \mathcal{L}_{\mathbb{Q}}\big(X_0, \Delta_1^n X, ..., \Delta_n^n X\big)$. Now the statement follows from Lemma 1. The additivity property then follows directly from Lemma 2. □

Remark 12. Let $\mathbb{Q}, \mathbb{P} \in \mathcal{M}_1(C)$ be such that $X_0 = x$ for some $x \in \mathbb{R}$ both $\mathbb{Q}$-a.s. and $\mathbb{P}$-a.s.. Then we can disregard the slot for time 0 when computing the relative entropy on the restrictions to $\mathcal{F}^n$. Indeed, we have $\mathbb{P}$-a.s.,

$$\frac{d\mathcal{L}_{\mathbb{Q}}\big(X_0, \Delta_1^n X, ..., \Delta_n^n X\big)}{d\mathcal{L}_{\mathbb{P}}\big(X_0, \Delta_1^n X, ..., \Delta_n^n X\big)}\,(\omega(0), \Delta_1^n\omega, ..., \Delta_n^n\omega) = \frac{d\mathcal{L}_{\mathbb{Q}}\big(\Delta_1^n X, ..., \Delta_n^n X\big)}{d\mathcal{L}_{\mathbb{P}}\big(\Delta_1^n X, ..., \Delta_n^n X\big)}\,(\Delta_1^n\omega, ..., \Delta_n^n\omega)\,,$$

and therefore,

$$\mathrm{H}\,(\mathbb{Q}\,|\,\mathbb{P})\,|_{\mathcal{F}^n} = \mathrm{H}\,(\mathcal{L}_{\mathbb{Q}}(\Delta_1^n X, ..., \Delta_n^n X)\,|\,\mathcal{L}_{\mathbb{P}}(\Delta_1^n X, ..., \Delta_n^n X))\,.$$

Definition 12. Let $\mathbb{P}, \mathbb{Q} \in \mathcal{M}_1(\mathcal{C})$. Then the specific relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as,

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

if the limit exists in $[0, +\infty]$.

Example 4. Going back to the measures in Remark 11, $\mathbb{Q} = \mathcal{L}((\sigma B_t)_{t\in[0,1]})$ and $\mathbb{P} = \mathcal{L}((\eta B_t)_{t\in[0,1]})$ we can compute with Lemma 4 and the formula for the relative entropy between two Normal random variables from Example 1,

$$\begin{aligned}
\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{P}\right)|_{\mathcal{F}^n} &= \sum_{k=1}^{n} \mathrm{H}\left(\mathcal{L}(\Delta_k^n \sigma B)\,|\,\mathcal{L}(\Delta_k^n \eta B)\right) \\
&= \sum_{k=1}^{n} \mathrm{H}\left(\mathcal{N}(0, \tfrac{\sigma^2}{n})\,|\,\mathcal{N}(0, \tfrac{\eta^2}{n})\right) \\
&= n\frac{1}{2}\left(\frac{\sigma^2}{\eta^2} - 1 - \log\frac{\sigma^2}{\eta^2}\right).
\end{aligned}$$

For the specific relative entropy we therefore get,

$$h(\mathbb{Q}\,|\,\mathbb{P}) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{P}\right)|_{\mathcal{F}^n} = \frac{1}{2}\left(\frac{\sigma^2}{\eta^2} - 1 - \log\frac{\sigma^2}{\eta^2}\right).$$

From now on we always take the reference measure to be Wiener measure $\mathbb{W}$, i.e. $\mathbb{W} = \mathcal{L}(B)$ and assume $\mathbb{Q} \in \mathcal{M}_1(C)$ is such that $X_0 = 0$ a.s.. In other words, $\mathbb{Q}$ is concentrated on $C_0 := \{\omega\colon [0,1] \to \mathbb{R} \mid \omega \text{ is continuous}, \omega(0) = 0\}$. Since the same is true for $\mathbb{W}$, we are in the situation of Remark 12.

Remark 13. Recalling Lemma 3 we note that, if $\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{W}\right) < \infty$, then $h(\mathbb{Q}\,|\,\mathbb{W}) = 0$. See Proposition 3 in [1] for the following observation: Assume $\mathbb{Q} \ll \mathbb{W}$. Then $\mathbb{Q}$ is the law of a Brownian motion with absolute continuous drift. More concretely, $\mathbb{Q}$ is such that $X_t = W_t + \int_0^t b(s,X)\,ds$, where W is a Brownian motion under $\mathbb{Q}$ and b is a predictable process such that $\int_0^1 b(s,X)^2\,ds < \infty$ $\mathbb{Q}$-a.s. In that case, we have,

$$\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{W}\right) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 b(s,X)^2\,ds\right].$$

We note that the relative entropy is finite if and only if $\mathbb{Q} \ll \mathbb{W}$ and additionally $\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 b(s,X)^2\,ds\right] < \infty$.

We will consider an example that should emphasize the following: The implication in Remark 13 cannot be reversed, i.e. the relative entropy being finite, is not necessary for the specific relative entropy to vanish.

Example 5. Let $\mathbb{Q}$ be the law of $(B_t + b(t))_{t\in[0,1]}$, where $b \in C_0$.
Note that the increments of X under $\mathbb{Q}$ are independent and for all $s < t$,

$$X_t - X_s \sim_{\mathbb{Q}} \mathcal{N}(b(t) - b(s), t - s)$$

We can compute the relative entropy on $\mathcal{F}^n$ as before using Lemma 4 and the formula for the relative entropy between Normal random variables,

$$\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{W}\right)|_{\mathcal{F}^n} = \sum_{k=1}^{n} \mathrm{H}\left(\mathcal{N}(\Delta_k^n b, \tfrac{1}{n})\,|\,\mathcal{N}(0, \tfrac{1}{n})\right) = \frac{n}{2}\sum_{k=1}^{n}(\Delta_k^n b)^2.$$

So for the specific relative entropy, we have,

$$h(\mathbb{Q}\,|\,\mathbb{W}) = \lim_{n\to\infty} \frac{1}{2}\sum_{k=1}^{n}(\Delta_k^n b)^2.$$

We see that if b is of bounded variation, then the specific relative entropy vanishes, even if b is not absolutely continuous and thus $\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{W}\right) = \infty$ by Remark 13.

The specific relative entropy can even be finite if we consider laws of processes that do not have continuous sample paths.

Example 6. Let $x \in [0,1]$ be any irrational number. Let $\mathbb{Q}$ be the law of the process, defined as,

$$Y_t = B_t + \mathbb{1}_{(x,1]}(t).$$

Then Y has càdlàg sample paths and $\mathbb{Q} \in \mathcal{M}_1(\mathcal{D})$, where $\mathcal{D} = \mathcal{D}([0,1])$ is the Skorokhod space. Note that for all $n \in \mathbb{N}$ and $k = 1, ..., n$,

$$\Delta_k^n Y = \begin{cases} 1 + \Delta_k^n B & \text{if } x \in (\frac{k-1}{n}, \frac{k}{n}], \\ \Delta_k^n B & \text{else.} \end{cases}$$

We can compute,

$$\mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{W}\right)|_{\mathcal{F}^n} = \mathrm{H}\left(\mathcal{N}(1, \tfrac{1}{n})\,|\,\mathcal{N}(0, \tfrac{1}{n})\right) = \frac{n}{2}$$

and so $h(\mathbb{Q}\,|\,\mathbb{W}) = \frac{1}{2}$.

From now on we focus on martingale measures. First we consider a particular class of measures in $\mathcal{M}^2$, for which the specific relative entropy with respect to Wiener measure exists in $[0, +\infty]$ and is given by an explicit formula in terms of the quadratic variation. For every $a \in C_0$ that is non-decreasing we can define $\mathbb{Q}^a := \mathcal{L}(B_{a(\cdot)})$. Then $\mathbb{Q}^a \in \mathcal{M}^2$ is Gaussian with independent increments and $\langle X \rangle_t = a(t)$.

We will restrict ourselves to the case that a is absolutely continuous. Equivalently, there exists some $\sigma \in L^2(\lambda)$ such that $a(t) = \int_0^t \sigma(s)^2\, ds$ and the measure $\mathbb{Q}^a$ is the law of the process $Y = (Y_t)_{t\in[0,1]}$ defined by,

$$Y_t = \int_0^t \sigma(s)\, dB_s.$$

For the general case see Theorem 1 in [3, Chapter 1].

Theorem 9. *Let $a(t) = \int_0^t \sigma(s)^2\, ds$ for some $\sigma \in L^2(\lambda)$. Let $\mathbb{Q}^a \in \mathcal{M}^2$ be as defined above. Then,*

$$\begin{aligned} h(\mathbb{Q}^a \,|\, \mathbb{W}) &= \frac{1}{2}\int_0^1 \sigma(s)^2 - 1 - \log(\sigma(s)^2)\, ds \\ &= \sup_{n\in\mathbb{N}} \frac{1}{n} H(\mathbb{Q}^a \,|\, \mathbb{W})\,|_{\mathcal{F}^n} \in [0,\infty]. \end{aligned}$$

Moreover,

$$\frac{1}{n}\log \frac{d\mathbb{Q}^a}{d\mathbb{W}}\Big|_{\mathcal{F}^n} \xrightarrow{L^1(\mathbb{Q}^a)} \frac{1}{2}\int_0^1 \sigma(s)^2 - 1 - \log\sigma(s)^2\, ds \quad \text{as } n\to\infty$$

if the right-hand side is finite.

Remark 14. Note that we obtain again Example 4 with $\eta = 1$ as a special case by taking $\sigma(s) = \sigma$.

Proof. Under $\mathbb{Q}^a$ the process X is Gaussian with independent increments and for all $s < t$,

$$X_t - X_s \sim_{\mathbb{Q}^a} \mathcal{N}(0,\, a(t) - a(s))$$

Recall that $\mathrm{H}\left(\mathcal{N}(0,\sigma^2)\,|\,\mathcal{N}(0,\eta^2)\right) = F(\frac{\sigma^2}{\eta^2})$, where $F(x) = \frac{1}{2}(x - 1 - \log(x))$. By Lemma 4 and independence of increments under $\mathbb{Q}^a$ as well as under $\mathbb{W}$, we can write,

$$\begin{aligned} \mathrm{H}\left(\mathbb{Q}^a \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} &= \sum_{k=1}^n \mathrm{H}\left(\mathcal{L}_{\mathbb{Q}^a}(\Delta_k^n X) \,|\, \mathcal{L}_{\mathbb{W}}(\Delta_k^n X)\right) \\ &= \sum_{k=1}^n \mathrm{H}\left(\mathcal{N}(0,\, \Delta_k^n a) \,|\, \mathcal{N}(0,\, \tfrac{1}{n})\right) \\ &= \sum_{k=1}^n F(n\Delta_k^n a). \end{aligned} \tag{2.1}$$

Now let $\mathcal{B}_n := \sigma\left((\frac{k-1}{n}, \frac{k}{n}] : k = 1, ..., n\right)$. Since we have, $a(t) = \int_0^t \sigma(s)^2\, ds$, where

$\sigma^2 \in L^1(\lambda)$,

$$\mathbb{E}_\lambda[\sigma^2 \mid \mathcal{B}_n](t) = n \sum_{k=1}^{n} \Delta_k^n a \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t).$$

We get in (2.1),

$$\begin{aligned} \sum_{k=1}^{n} F(n\Delta_k^n a) &= n \int_0^1 F(n\Delta_k^n a) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)\, ds \\ &= n \int_0^1 F(\mathbb{E}_\lambda[\sigma^2 \mid \mathcal{B}_n](s)) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)\, ds \end{aligned}$$

and dividing by n, we obtain,

$$\frac{1}{n} \mathrm{H}\left(\mathbb{Q}^a \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} = \int_0^1 F(\mathbb{E}_\lambda[\sigma^2 \mid \mathcal{B}_n](s)) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)\, ds.$$

We now use that the function F is convex and Jensen's conditional inequality and the tower property of conditional expectation to obtain an upper bound for all $n \in \mathbb{N}$,

$$\begin{aligned} \frac{1}{n} \mathrm{H}\left(\mathbb{Q}^a \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} &= \int_0^1 F(\mathbb{E}_\lambda[\sigma^2 \mid \mathcal{B}_n](s))\, ds \\ &\leq \int_0^1 \mathbb{E}_\lambda[F(\sigma^2) \mid \mathcal{B}_n](s)\, ds \\ &= \int_0^1 F(\sigma(s)^2)\, ds. \end{aligned}$$

Note that $\mathbb{E}_\lambda[\sigma^2 \mid \mathcal{B}_n] \to \sigma^2$ λ-a.e., by the Lebesgue differentiation theorem. Using the upper bound and the a.s. convergence we now get,

$$\begin{aligned} \int_0^1 F\left(\sigma(s)^2\right)\, ds &= \int_0^1 \liminf_{n\to\infty} F\left(\mathbb{E}_\lambda\left[\sigma^2 \mid \mathcal{B}_n\right](s)\right)\, ds \\ &\leq \liminf_{n\to\infty} \int_0^1 F\left(\mathbb{E}_\lambda\left[\sigma^2 \mid \mathcal{B}_n\right](s)\right)\, ds \\ &\leq \limsup_{n\to\infty} \int_0^1 F\left(\mathbb{E}_\lambda\left[\sigma^2 \mid \mathcal{B}_n\right](s)\right)\, ds \\ &\leq \int_0^1 F(\sigma(s)^2)\, ds, \end{aligned}$$

where the first inequality is by Fatou's lemma and $F \geq 0$ and the last inequality by the upper bound from above. We sum up,

$$\begin{aligned} h(\mathbb{Q}^a \,|\, \mathbb{W}) &= \lim_{n\to\infty} \frac{1}{n} \mathrm{H}\left(\mathbb{Q}^a \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} \\ &= \lim_{n\to\infty} \int_0^1 F(\mathbb{E}_\lambda[\sigma^2 \mid \mathcal{B}_n](s))\, ds \\ &= \sup_{n\geq 1} \frac{1}{n} \mathrm{H}\left(\mathbb{Q}^a \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} \\ &= \int_0^1 F(\sigma(s)^2)\, ds \end{aligned}$$

which proves the first part of the theorem.

Now we assume that $\int_0^1 \sigma(s)^2 - 1 - \log \sigma(s)^2 \, ds < \infty$ and in particular $\Delta_k^n a > 0$ for all $n \in \mathbb{N}$ and $k = 1, ..., n$. To show L^1 convergence, note that by Lemma 4,

$$\begin{aligned} \frac{1}{n} \log \frac{d\mathbb{Q}^a}{d\mathbb{W}}\Big|_{\mathcal{F}^n}(X) &= \frac{1}{n} \sum_{k=1}^n \log \frac{d\mathcal{N}(0,\, \Delta_k^n a)}{d\mathcal{N}(0,\, \frac{1}{n}))} (\Delta_k^n X) \\ &= \frac{1}{2n} \sum_{k=1}^n \Big\{ n(\Delta_k^n X)^2 - \frac{(\Delta_k^n X)^2}{\Delta_k^n a} - \log\left(n\Delta_k^n a\right) \Big\}. \end{aligned}$$

We will consider each of the three terms in the sum above individually. For the first one, we show that $\sum_{k=1}^n (\Delta_k^n X)^2 \xrightarrow{L^2(\mathbb{Q}^a)} a(1)$ as $n \to \infty$. For all $n \in \mathbb{N}$ we can write,

$$\mathbb{E}_{\mathbb{Q}^a}\left[\left(\sum_{k=1}^n (\Delta_k^n X)^2 - a(1)\right)^2\right] = \mathbb{E}_{\mathbb{Q}^a}\left[\left(\sum_{k=1}^n \Big((\Delta_k^n X)^2 - \Delta_k^n a\Big)\right)^2\right].$$

Expanding the square, the cross-terms vanish, since $\Delta_k^n X$ and $\Delta_i^n X$ are independent for $k \neq i$ and $\mathbb{E}_{\mathbb{Q}^a}\left[(\Delta_k^n X)^2 - \Delta_k^n a\right] = 0$ and we get,

$$\mathbb{E}_{\mathbb{Q}^a}\left[\left(\sum_{k=1}^n \Big((\Delta_k^n X)^2 - \Delta_k^n a\Big)\right)^2\right] = \sum_{k=1}^n \mathbb{E}_{\mathbb{Q}^a}\left[\Big((\Delta_k^n X)^2 - \Delta_k^n a\Big)^2\right].$$

Factoring out $(\Delta_k^n a)^2$ and since $\frac{\Delta_k^n X}{\sqrt{\Delta_k^n a}} \sim_{\mathbb{Q}^a} \mathcal{N}(0,1)$ we have,

$$\begin{aligned} \sum_{k=1}^n \mathbb{E}_{\mathbb{Q}^a}\left[\Big((\Delta_k^n X)^2 - \Delta_k^n a\Big)^2\right] &= \sum_{k=1}^n (\Delta_k^n a)^2 \mathbb{E}_{\mathbb{Q}^a}\left[\left(\frac{(\Delta_k^n X)^2}{\Delta_k^n a} - 1\right)^2\right] \\ &= \mathbb{E}_{\mathbb{Q}^a}\left[\Big(\mathcal{N}(0,1)^2 - 1\Big)^2\right] \sum_{k=1}^n (\Delta_k^n a)^2. \end{aligned}$$

Summing up we have for all $n \in \mathbb{N}$,

$$\mathbb{E}_{\mathbb{Q}^a}\left[\left(\sum_{k=1}^{n}(\Delta_k^n X)^2 - a(1)\right)^2\right] = \mathbb{E}_{\mathbb{Q}^a}\left[\left(\mathcal{N}(0,1)^2 - 1\right)^2\right]\sum_{k=1}^{n}(\Delta_k^n a)^2$$

and $\sum_{k=1}^{n}(\Delta_k^n a)^2 \to 0$ as $n \to \infty$ since a is non-decreasing.

Next, for every $n \in \mathbb{N}$, for $k = 1, ..., n$,

$$\frac{\Delta_k^n X}{\sqrt{\Delta_k^n a}} \sim_{\mathbb{Q}^a} \mathcal{N}(0,1) \quad iid$$

and so by the Law of Large Numbers,

$$\frac{1}{n}\sum_{k=1}^{n}\frac{(\Delta_k^n X)^2}{\Delta_k^n a} \xrightarrow{L^1(\mathbb{Q}^a)} 1 \quad \text{as } n \to \infty.$$

Since we already know

$$\lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^a}\left[\frac{1}{2n}\sum_{k=1}^{n} n(\Delta_k^n X)^2 - \frac{(\Delta_k^n X)^2}{\Delta_k^n a} - \log(n\Delta_k^n a)\right] = \int_0^1 \sigma(s)^2 - 1 - \log\sigma(s)^2\, ds$$

from the first part of the proof, we have for the deterministic term,

$$\frac{1}{2n}\sum_{k=1}^{n}\log(n\Delta_k^n a) \xrightarrow{n\to\infty} \int_0^1 \frac{1}{2}\log\sigma(s)^2\, ds,$$

which finishes the proof. □

Next we consider more general martingale measures, again with absolutely continuous quadratic variation. In that case, we obtain a lower bound for the specific relative entropy along the dyadic subsequence.

Theorem 10. *Let $\mathbb{Q} \in \mathcal{M}^2$ be such that $X_0 = 0$ a.s. and $\langle X\rangle_t = \int_0^t \sigma(s,X)^2\, ds$, with $\sigma \in L^2(\lambda \otimes \mathbb{Q})$ predictable. Then we have,*

$$\liminf_{n\to\infty} \frac{1}{2^n} H(\mathbb{Q}\,|\,\mathbb{W})\,|_{\mathcal{F}^{2^n}} \geq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \sigma(s,X)^2 - 1 - \log(\sigma(s,X)^2)\, ds\right]. \qquad (2.2)$$

Proof. Let $n \in \mathbb{N}$. We start by defining a new measure on $(C, \mathcal{F}^n)$.

Let $\mathbb{Q}^n$ be such that for $k = 1, \dots, n$

$$X_{\frac{k}{n}} - X_{\frac{k-1}{n}} \mid X_0, ..., X_{\frac{k-1}{n}} \sim_{\mathbb{Q}^n} \mathcal{N}\left(0,\, \mathbb{E}_{\mathbb{Q}}\big[(\Delta_k^n X)^2 \mid X_0, ..., X_{\frac{k-1}{n}}\big]\right),$$

i.e. the new measure is defined such that the increments of X have the same conditional variances as under $\mathbb{Q}$.

Assume that $\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} < \infty$. Then it follows that, $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{W}|_{\mathcal{F}^n}$, which implies, $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{Q}^n$ and $\mathbb{Q}^n \ll \mathbb{W}|_{\mathcal{F}^n}$, and thus we have,

$$\frac{d\mathbb{Q}}{d\mathbb{W}}\Big|_{\mathcal{F}^n} = \frac{d\mathbb{Q}}{d\mathbb{Q}^n}\Big|_{\mathcal{F}^n} \frac{d\mathbb{Q}^n}{d\mathbb{W}}\Big|_{\mathcal{F}^n}.$$

Now taking the logarithm and integrating with respect to $\mathbb{Q}$ yields,

$$\begin{aligned} \mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}\right)|_{\mathcal{F}^n} &= \int_C \log \frac{d\mathbb{Q}}{d\mathbb{Q}^n}\Big|_{\mathcal{F}^n} d\mathbb{Q} + \int_C \log \frac{d\mathbb{Q}^n}{d\mathbb{W}}\Big|_{\mathcal{F}^n} d\mathbb{Q} \\ &= \mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{Q}^n\right)|_{\mathcal{F}^n} + \int_C \log \frac{d\mathbb{Q}^n}{d\mathbb{W}}\Big|_{\mathcal{F}^n} d\mathbb{Q}. \end{aligned} \tag{2.3}$$

In the remainder we will show that,

$$\frac{1}{2^n} \int_C \log \frac{d\mathbb{Q}^{2^n}}{d\mathbb{W}}\Big|_{\mathcal{F}^{2^n}} d\mathbb{Q}$$

converges to the right-hand side of (2.2).
The statement then follows, since $\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{Q}^n\right)|_{\mathcal{F}^n} \geq 0$ for all $n \in \mathbb{N}$. We will assume $\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}\right)|_{\mathcal{F}^{2^n}} < \infty$ for all $n \in \mathbb{N}$ and thus (2.3) holds for all $n \in \mathbb{N}$. The assumption is justified, since otherwise the left-hand side in (2.2) is infinite and the inequality trivially true. Note that for this argument we use that the dyadic numbers are nested.

Let $\mathcal{P}_n \coloneqq \sigma\big(A \times (s,t] \mid s < t \in \{0, \frac{1}{n}, \dots, 1\}, A \in \sigma(X_0, X_{\frac{1}{n}}, ..., X_s)\big)$. Since σ^2 is $\lambda \otimes \mathbb{Q}$-integrable, we can define,

$$\sigma_n^2 \coloneqq \mathbb{E}_{\lambda \otimes \mathbb{Q}}[\sigma^2 \mid \mathcal{P}_n].$$

We have for every $n \in \mathbb{N}$

$$\sigma_n^2(t, X) = n \sum_{k=1}^{n} \mathbb{E}_{\mathbb{Q}} \left[\int_{\frac{k-1}{n}}^{\frac{k}{n}} \sigma(s, X)^2 \, ds \mid X_0, ..., X_{\frac{k-1}{n}} \right] (X) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t).$$

Note that this means,

$$X_{\frac{k}{n}} - X_{\frac{k-1}{n}} \mid X_0, ..., X_{\frac{k-1}{n}} \sim_{\mathbb{Q}^n} \mathcal{N}\Big(0, \tfrac{1}{n}\sigma_n^2(\tfrac{k}{n}, X)\Big).$$

Indeed, this follows easily from,

$$\begin{aligned}
\frac{1}{n}\sigma_n^2(\tfrac{k}{n}, X) &= \mathbb{E}_{\mathbb{Q}}\left[\int_{\frac{k-1}{n}}^{\frac{k}{n}} \sigma(s, X)^2\, ds \mid X_0, ..., X_{\frac{k-1}{n}}\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}\left[\int_{\frac{k-1}{n}}^{\frac{k}{n}} \sigma(s, X)^2\, ds \mid \mathcal{F}_{\frac{k-1}{n}}\right] \mid X_0, ..., X_{\frac{k-1}{n}}\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}\left[(\Delta_k^n X)^2 \mid \mathcal{F}_{\frac{k-1}{n}}\right] \mid X_0, ..., X_{\frac{k-1}{n}}\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[(\Delta_k^n X)^2 \mid X_0, ..., X_{\frac{k-1}{n}}\right]
\end{aligned}$$

and since we have assumed $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{W}|_{\mathcal{F}^n}$, we have $\sigma_n^2(\frac{k}{n}, X) > 0$ $\mathbb{Q}$-a.s.. We can now compute the Radon-Nikodym density of $\mathbb{Q}^n$ with respect to $\mathbb{W}|_{\mathcal{F}^n}$,

$$\begin{aligned}
\frac{d\mathbb{Q}^n}{d\mathbb{W}}\Big|_{\mathcal{F}^n}(X) &= \prod_{k=1}^{n} \frac{d\mathcal{L}_{\mathbb{Q}^n}(\Delta_k^n X \mid X_0, ..., X_{\frac{k-1}{n}})}{d\mathcal{L}_{\mathbb{W}}(\Delta_k^n X \mid X_0, ..., X_{\frac{k-1}{n}})}(\Delta_k^n X) \\
&= \prod_{k=1}^{n} \frac{d\mathcal{N}\big(0, \frac{1}{n}\sigma_n^2(\frac{k}{n}, X)\big)}{d\mathcal{N}\big(0, \frac{1}{n}\big)}(\Delta_k^n X) \\
&= \prod_{k=1}^{n} \frac{1}{\sqrt{\sigma_n^2(\frac{k}{n}, X)}} \exp\left(-\frac{n}{2}\frac{(\Delta_k^n X)^2}{\sigma_n^2(\frac{k}{n}, X)} + \frac{n}{2}(\Delta_k^n X)^2\right).
\end{aligned}$$

Taking the logarithm and integrating with respect to $\mathbb{Q}$ and dividing by n, we get

$$\frac{1}{n}\int_C \log \frac{d\mathbb{Q}^n}{d\mathbb{W}}\Big|_{\mathcal{F}^n} d\mathbb{Q} = \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}_{\mathbb{Q}}\left[\frac{n}{2}(\Delta_k^n X)^2 - \frac{n}{2}\frac{(\Delta_k^n X)^2}{\sigma_n^2(\frac{k}{n}, X)} - \frac{1}{2}\log \sigma_n^2(\tfrac{k}{n}, X)\right]. \tag{2.4}$$

Recall that for all $n \geq 1$ and all $k = 1, ..., n$,

$$\frac{1}{n}\sigma_n^2(\tfrac{k}{n}, X) = \mathbb{E}_{\mathbb{Q}}\left[(\Delta_k^n X)^2 \mid X_0, ..., X_{\frac{k-1}{n}}\right]$$

and therefore with the tower property of conditional expectation,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[n\frac{(\Delta_k^n X)^2}{\sigma_n^2(\frac{k}{n},X)}\right] &= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}\left[n\frac{(\Delta_k^n X)^2}{\sigma_n^2(\frac{k}{n},X)} \mid X_0,...,X_{\frac{k-1}{n}}\right]\right]\\ &= \mathbb{E}_{\mathbb{Q}}\left[\frac{n}{\sigma_n^2(\frac{k}{n},X)}\mathbb{E}_{\mathbb{Q}}\left[(\Delta_k^n X)^2 \mid X_0,...,X_{\frac{k-1}{n}}\right]\right]\\ &= 1.\end{aligned}$$

We continue in (2.4) with $F(x) \coloneqq \frac{1}{2}(x-1-\log x)$ as before,

$$\begin{aligned}\frac{1}{n}\int_C \log\frac{d\mathbb{Q}^n}{d\mathbb{W}}\Big|_{\mathcal{F}^n} d\mathbb{Q} &= \frac{1}{n}\sum_{k=1}^n \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{2}\sigma_n^2(\tfrac{k}{n},X) - \frac{1}{2} - \frac{1}{2}\log\sigma_n^2(\tfrac{k}{n},X)\right]\\ &= \frac{1}{n}\sum_{k=1}^n \mathbb{E}_{\mathbb{Q}}\left[F(\sigma_n^2(\tfrac{k}{n},X))\right]\\ &= \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma_n^2(s,X))\,ds\right]\end{aligned}$$

where we have used, that $t\mapsto \sigma_n^2(t,X)$ is constant on intervals of the form $(\frac{k-1}{n},\frac{k}{n}]$. We will now use that F is convex and apply Jensen's inequality to obtain an upper bound for all $n\in\mathbb{N}$,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma_n^2(s,X))\,ds\right] &= \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\mathbb{E}_{\lambda\otimes\mathbb{Q}}[\sigma^2 \mid \mathcal{P}_n](s,X))\,ds\right]\\ &\le \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \mathbb{E}_{\lambda\otimes\mathbb{Q}}[F(\sigma^2) \mid \mathcal{P}_n](s,X)\,ds\right]\\ &= \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma(s,X)^2)\,ds\right],\end{aligned}$$

the last equality is due to Fubini's theorem for non-negative functions and the tower property.

Now we restrict ourselves to the subsequence of dyadic numbers. Note that $(\mathcal{P}_{2^n})_{n\ge 1}$ is a sequence of sigma algebras that increases to the predictable sigma algebra $\mathcal{P}$. Therefore, $\sigma_{2^n}^2$ is a closed martingale with respect to $(\mathcal{P}_{2^n})_{n\in\mathbb{N}}$ and from the convergence theorem for uniformly integrable martingales, Theorem 1, it follows that,

$$\sigma_{2^n}^2(t,X) \xrightarrow{n\to\infty} \sigma(t,X)^2 \;\; \lambda\otimes\mathbb{Q}\text{-a.s.}.$$

With Fatou's lemma and $F \geq 0$ and the upper bound from before we get,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma(s,X)^2)\,ds\right] &= \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \liminf_{n\to\infty} F(\sigma_{2^n}^2(s,X))\,ds\right] \\ &\leq \liminf_{n\to\infty} \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma_{2^n}^2(s,X))\,ds\right] \\ &\leq \limsup_{n\to\infty} \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma_{2^n}^2(s,X))\,ds\right] \\ &\leq \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma(s,X)^2)\,ds\right].\end{aligned}$$

We conclude,

$$\lim_{n\to\infty} \frac{1}{2^n} \int_C \log \frac{d\mathbb{Q}^{2^n}}{d\mathbb{W}}\Big|_{\mathcal{F}^{2^n}} d\mathbb{Q} = \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 F(\sigma(s,X)^2)\,ds\right].$$

□

Remark 15. Let $\mathbb{Q} \in \mathcal{M}^2$ as before and such that $\mathrm{H}(\mathbb{Q}\,|\,\mathbb{W})\,|_{\mathcal{F}^{2^n}} < \infty$ for all $n \in \mathbb{N}$. Let $\mathbb{Q}^{2^n}$ be as in the proof above, then we recall,

$$\frac{1}{2^n}\mathrm{H}(\mathbb{Q}\,|\,\mathbb{W})\,|_{\mathcal{F}^{2^n}} = \frac{1}{2^n}H(\mathbb{Q}\,|\,\mathbb{Q}^{2^n})|_{\mathcal{F}^{2^n}} + \frac{1}{2^n}\mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}^{2^n}}{d\mathbb{W}}\Big|_{\mathcal{F}^{2^n}}(X)\right].$$

The question that remains open is, if

$$\frac{1}{2^n}H(\mathbb{Q}\,|\,\mathbb{Q}^{2^n})|_{\mathcal{F}^{2^n}} \xrightarrow{n\to\infty} 0.$$

If so, then it follows from the proof that, at least along the subsequence of dyadic numbers, the specific relative entropy is given by,

$$\lim_{n\to\infty} \frac{1}{2^n}\mathrm{H}(\mathbb{Q}\,|\,\mathbb{W})\,|_{\mathcal{F}^{2^n}} = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \sigma(s,X)^2 - 1 - \log\sigma(s,X)^2\,ds\right]. \tag{2.5}$$

Remark 16. Note that all arguments in the proof of Theorem 10 and in Remark 15 still hold, if the sequence $(2^n)_{n\in\mathbb{N}}$ is replaced by $(p^n)_{n\in\mathbb{N}}$ for any $p \in \mathbb{N}$ and $p \geq 2$.

Remark 17. Let $\mathbb{Q} \in \mathcal{M}^2$ and $\nu_{\langle X\rangle}$ be the random measure on $[0,1]$ with distribution function $\langle X\rangle$. In the above theorem, the assumption $\langle X\rangle_t$ being absolutely continuous means $\nu_{\langle X\rangle} \ll \lambda$ and $\frac{d\nu_{\langle X\rangle}}{d\lambda}(t) = \sigma(t,X)^2$. So (2.2) takes the alternative form,

$$\liminf_{n\to\infty} \frac{1}{2^n}\mathrm{H}(\mathbb{Q}\,|\,\mathbb{W})\,|_{\mathcal{F}^{2^n}} \geq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\nu_{\langle X\rangle}([0,1]) - 1 - \mathrm{H}\left(\lambda\,|\,\nu_{\langle X\rangle}\right)\right].$$

Formulated this way, (2.2) also stays true, if the measure $\mathbb{Q}$ is such that $\langle X\rangle$ is not absolutely continuous. This extension of Theorem 10 can be found in [1], Theorem 17. In that case, the measure $\nu_{\langle X\rangle}$ has Lebesgue decomposition $\nu_{\langle X\rangle} = \nu_s^X + \nu_{ac}^X$, where ν_s^X denotes the singular part of $\nu_{\langle X\rangle}$ and $\sigma(\cdot, X)^2 := \frac{d\nu_{ac}^X}{d\lambda}$. Then,

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{2^n} \mathrm{H}\left(\mathbb{Q}\,|\,\mathbb{W}\right)|_{\mathcal{F}^{2^n}} &\geq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\nu_{\langle X\rangle}([0,1]) - 1 - \mathrm{H}\left(\lambda\,|\,\nu_{\langle X\rangle}\right)\right] \\ &= \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \sigma(s,X)^2 - 1 - \log\sigma(s,X)^2\,ds + \nu_s^X([0,1])\right].\end{aligned}$$

Remark 18. It immediately follows that if $\mathbb{Q} \in \mathcal{M}^2$ is such that $h(\mathbb{Q}\,|\,\mathbb{W}) < \infty$, then

$$\sigma(\cdot, X)^2 > 0 \quad \lambda \otimes \mathbb{Q}\text{-a.s.},$$

where σ^2 is the Radon-Nikodym density of the the absolutely continuous part of the measure $\nu_{\langle X\rangle}$ as in Remark 17.

3 The specific relative Entropy for two classes of Diffusion Processes

As in the previous chapter, we consider a Brownian motion $B = (B_t)_{t\in[0,1]}$ on some filtered probability space $(\Omega, \mathcal{S}, \mathcal{S}_t, \mathbb{S})$ and $(C, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]})$, where again

$$C = C([0,1]) = \{\omega\colon [0,1] \to \mathbb{R} \mid \omega \text{ is continuous}\}$$

The canonical process X, the filtration $(\mathcal{F}_t)_{t\in[0,1]}$ and sigma-algebras $\mathcal{F}^n$ are as in Chapter 2. Now we will also consider laws of processes not concentrated on $\{\omega \in C \mid \omega(0) = 0\}$, but instead we introduce the measures $\mathbb{W}^x := \mathcal{L}(B^x)$, where $B^x := B + x$ for $x \in \mathbb{R}$ as reference measures. This will allow us to still always neglect time point 0 when computing the relative entropy on the restriction to $\mathcal{F}^n$.

We start by verifying Formula (2.5) in the special case of the law of the martingale defined as in Example 2 under certain constraints on the defining function f.

Theorem 11. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a function such that:*

(i) $\mathbb{E}[f(B_1)^2] < \infty$

(ii) *f is continuously differentiable and $f' > 0$*

(iii) *f' satisfies,* $\mathbb{E}[f'(B_1)^2] < \infty$ *and* $\mathbb{E}[-\log f'(B_1)] < \infty$.

Let $M = (M_t)_{t\in[0,1]}$ be the martingale defined by

$$M_t := \mathbb{E}\left[f(B_1) \mid \mathcal{F}_t\right]$$

and let $\mathbb{Q}$ be the law of M on $(\mathcal{C}, \mathcal{F})$.

Then

$$h(\mathbb{Q} \,|\, \mathbb{W}^x) = \mathbb{E}\left[\int_0^1 \frac{1}{2}\left\{\partial_2 f(t, B_t)^2 - 1 - \log \partial_2 f(t, B_t)^2\right\} dt\right] \tag{3.1}$$

where $f(t,y) := \int_{\mathbb{R}} f(z+y)\; d\mathcal{N}(0, 1-t)(z)$ and $x := \mathbb{E}[f(B_1)]$.

Remark 19. From the discussion in Example 2 we know, $M_t = f(t, B_t)$ and $\langle M \rangle_t = \int_0^t \partial_2 f(s, B_s)^2 \, ds$. Here we additionally require f to be strictly increasing, so that for all $t \in [0,1]$ the function $f(t, \cdot)$ is also strictly increasing and as a consequence has an inverse $f^{-1}(t, \cdot)$ defined on $f(t, \mathbb{R})$. This means also B_t can be recovered from M_t, concretely $f^{-1}(t, M_t) = B_t$. Define for all $x \in f(t, \mathbb{R})$,

$$\sigma(t, x) \coloneqq \partial_2 f(t, \cdot) \circ f^{-1}(t, x)$$

Then M satisfies the SDE,

$$dM_t = \sigma(t, M_t) \, dB_t$$

and we can write (3.1) as,

$$\begin{aligned} h(\mathbb{Q} \,|\, \mathbb{W}^x) &= \mathbb{E}\left[\int_0^1 \frac{1}{2}\Big\{\sigma(t, M_t)^2 - 1 - \log \sigma(t, M_t)^2\Big\} \, dt\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \frac{1}{2}\Big\{\sigma(t, X_t)^2 - 1 - \log \sigma(t, X_t)^2\Big\} \, dt\right] \end{aligned}$$

Note that the integrability assumptions on f' capture that σ should not be 'too large' or 'too close to 0'.

Proof. In the first part of the proof, we will compute $\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n}$ for any $n \in \mathbb{N}$. In the second part, we will show convergence of $\frac{1}{n}\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n}$.

(i) Fix $n \in \mathbb{N}$. In the following denote $M^n \coloneqq (M_{\frac{1}{n}}, ..., M_1)$ and B^n analogously. Since $\mathbb{Q}$ and $\mathbb{W}^x$ are both concentrated on $\{X_0 = x\}$, we have,

$$\begin{aligned} \mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n} &= \mathrm{H}\left(\mathcal{L}_{\mathbb{Q}}(X_0, X_{\frac{1}{n}}, ..., X_1) \,|\, \mathcal{L}_{\mathbb{W}^x}(X_0, X_{\frac{1}{n}}, ..., X_1)\right) \\ &= \mathrm{H}\left(\mathcal{L}_{\mathbb{Q}}(X_{\frac{1}{n}}, ..., X_1 \,|\, X_0 = x) \,|\, \mathcal{L}_{\mathbb{W}^x}(X_{\frac{1}{n}}, ..., X_1 \,|\, X_0 = x)\right) \\ &= \mathrm{H}\left(\mathcal{L}(M_{\frac{1}{n}}, ..., M_1) \,|\, \mathcal{L}(B^x_{\frac{1}{n}}, ..., B^x_1)\right), \end{aligned} \tag{3.2}$$

where $B^x = B + x$. The key to computing the expression above is that the random vector M^n has an explicit density with respect to Lebesgue measure.

Define the map $\phi_n \colon \mathbb{R}^n \to \mathbb{R}^n$ as,

$$\phi_n\left(x_1, \ldots, x_n\right) = \begin{pmatrix} f(\frac{1}{n}, x_1) \\ \vdots \\ f(1, x_n) \end{pmatrix}$$

Recall that $f(t,\cdot)$ has an inverse $f^{-1}(t,\cdot)$ defined on $f(t,\mathbb{R})$ for every $t \in [0,1]$. Note that this means that ϕ_n is invertible on $\phi_n(\mathbb{R}^n)$ and

$$\phi_n^{-1}(x_1,\dots,x_n) = \begin{pmatrix} f^{-1}(\frac{1}{n},x_1) \\ \vdots \\ f^{-1}(1,x_n) \end{pmatrix}.$$

Moreover, $f(t,\cdot)$ is continuously differentiable with strictly positive derivative for every $t \in [0,1]$, ie. $\partial_2 f(t,\cdot) > 0$. The k-th entry of ϕ_n only depends on x_k and therefore we have for the Jacobian of ϕ_n,

$$J_{\phi_n}(x_1,...,x_n) = \Big(\partial_{x_i} f(\tfrac{j}{n},x_j) \colon i,j=1,...,n\Big) = \operatorname{diag}\big(\partial_2 f(\tfrac{i}{n},x_i) \colon i=1,...,n\big)$$

and for the Jacobian determinant,

$$\det(J_{\phi_n}(x_1,...,x_n)) = \prod_{k=1}^{n} \partial_2 f(\tfrac{k}{n},x_k) > 0$$

In particular ϕ_n and ϕ_n^{-1} are continuously differentiable on their respective domains. Since $M^n = \phi_n(B^n)$, where $\phi_n \colon \mathbb{R}^n \to \phi_n(\mathbb{R}^n)$ is a diffeomorphism and B^n has a density p_n^0, also the vector M^n has a density q_n^x, which is given by,

$$q_n^x(\cdot) = \mathbb{1}_{\phi_n(\mathbb{R}^n)}(\cdot) p_n^0 \circ \phi_n^{-1}(\cdot) \frac{1}{\det(J_{\phi_n} \circ \phi_n^{-1}(\cdot))}. \tag{3.3}$$

Indeed, we have for every $A \subset \mathbb{R}^n$ measurable,

$$\begin{aligned}
\mathbb{P}(M^n \in A) &= \mathbb{P}\left(B^n \in \phi_n^{-1}(A)\right) \\
&= \int_{\mathbb{R}^n} \mathbb{1}_{\phi_n^{-1}(A)}(\mathbf{y}) p_n^0(\mathbf{y})\, d\mathbf{y} \\
&= \int_{\phi_n(\mathbb{R}^n)} \mathbb{1}_A(\mathbf{x}) p_n^0(\phi_n^{-1}(\mathbf{x})) |\det(J_{\phi_n^{-1}}(\mathbf{x}))|\, d\mathbf{x}
\end{aligned}$$

where the last equality follows from a change of variables $\mathbf{y} = \phi_n^{-1}(\mathbf{x})$ and together with $\det(J_{\phi_n^{-1}}) = (\det(J_{\phi_n} \circ \phi_n^{-1}))^{-1}$ we get (3.3).

Note that $\mathbb{1}_{\phi_n(\mathbb{R}^n)} = 1$, $\mathcal{L}(M^n)$-a.e. and explicitly writing out (3.3) we get with $x_0 = x$,

$$q_n^x(\mathbf{x}) \\ = \left(\frac{2\pi}{n}\right)^{-\frac{n}{2}} \exp\left(-\frac{n}{2}\sum_{k=1}^n (f^{-1}(\tfrac{k}{n}, x_k) - f^{-1}(\tfrac{k-1}{n}, x_{k-1}))^2\right) \prod_{k=1}^n \frac{1}{\partial_2 f(\frac{k}{n}, f^{-1}(\frac{k}{n}, x_k))},$$

for $\mathcal{L}(M^n)$-a.e. $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$.

We continue in (3.2) and write p_n^x for the density of $(B^x_{\frac{1}{n}}, ..., B^x_1)$,

$$\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n} = \int_{\mathbb{R}^n} \log \frac{q_n^x(\mathbf{x})}{p_n^x(\mathbf{x})}\, d\mathcal{L}(M^n)(\mathbf{x})$$

Note that, $\mathcal{L}(M^n)$-a.e.,

$$\log \frac{q_n^x(\mathbf{x})}{p_n^x(\mathbf{x})} \\ = \frac{n}{2}\sum_{k=1}^n (\Delta_k^n x)^2 - \frac{n}{2}\sum_{k=1}^n (f^{-1}(\tfrac{k}{n}, x_k) - f^{-1}(\tfrac{k-1}{n}, x_{k-1}))^2 - \sum_{k=1}^n \log \partial_2 f(\tfrac{k}{n}, f^{-1}(\tfrac{k}{n}, x_k)).$$

Recall that $f^{-1}(t, M_t) = B_t$ and so integrating with respect to $\mathcal{L}(M^n)$ yields,

$$\begin{aligned}
&\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n} \\
&= \mathbb{E}\left[\frac{n}{2}\sum_{k=1}^n (M_{\frac{k}{n}} - M_{\frac{k-1}{n}})^2 - \frac{n}{2}\sum_{k=1}^n (B_{\frac{k}{n}} - B_{\frac{k-1}{n}})^2 - \sum_{k=1}^n \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}})\right] \\
&= \mathbb{E}\left[\frac{n}{2}\int_0^1 \partial_2 f(t, B_t)^2\, dt - \frac{n}{2} - \sum_{k=1}^n \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}})\right],
\end{aligned}$$

which finishes (i).

For (ii), first observe that for all $n \in \mathbb{N}$,

$$\frac{1}{n}\mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n} = \mathbb{E}\left[\int_0^1 \frac{1}{2}\partial_2 f(t, B_t)^2\, dt - \frac{1}{2} - \frac{1}{2n}\sum_{k=1}^n \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}})^2\right]. \tag{3.4}$$

In the remainder we will show that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^n \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}})\right] = \mathbb{E}\left[\int_0^1 \log \partial_2 f(t, B_t)\, dt\right]. \tag{3.5}$$

We will use that $\partial_2 f(t, B_t)$ is itself a continuous martingale. Indeed, it is of the form $\partial_2 f(t, B_t) = \mathbb{E}[f'(B_1)|\mathcal{F}_t]$ a.s., where $\mathbb{E}[f'(B_1)^2] < \infty$ by assumption.

We have,

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\log\partial_2 f(\tfrac{k}{n},B_{\frac{k}{n}})\right]=\mathbb{E}\left[\int_0^1\sum_{k=1}^{n}\log\partial_2 f(\tfrac{k}{n},B_{\frac{k}{n}})\mathbb{1}_{(\frac{k-1}{n},\frac{k}{n}]}(t)\,dt\right]$$

To use the martingale property of $\partial_2 f(t,B_t)$, we first need to assign every interval the value on its left endpoint,

$$\begin{aligned}\mathbb{E}&\left[\int_0^1\sum_{k=1}^{n}\log\partial_2 f(\tfrac{k}{n},B_{\frac{k}{n}})\mathbb{1}_{(\frac{k-1}{n},\frac{k}{n}]}(t)\,dt\right]\\&=\mathbb{E}\left[\int_0^1\sum_{k=1}^{n}\log\partial_2 f(\tfrac{k-1}{n},B_{\frac{k-1}{n}})\mathbb{1}_{(\frac{k-1}{n},\frac{k}{n}]}(t)\,dt\right]-\frac{C}{n},\end{aligned}$$

where

$$C:=\mathbb{E}\left[\log\partial_2 f(0,B_0)-\log\partial_2 f(1,B_1)\right]=\log\big(\mathbb{E}\left[f'(B_1)\right]\big)-\mathbb{E}\left[\log f'(B_1)\right].$$

We have that $C\geq 0$ by Jensen's inequality, and $C<\infty$ due to the assumptions $\mathbb{E}[f'(B_1)^2]<\infty$ and $\mathbb{E}[-\log f'(B_1)]<\infty$.

Define the sigma algebras $\mathcal{P}_n:=\sigma\left((\frac{k-1}{n},\frac{k}{n}]\times A : k=1,...,n, A\in\mathcal{F}_{\frac{k-1}{n}}\right)$ and note that,

$$\begin{aligned}\mathbb{E}_{\lambda\otimes\mathbb{P}}\left[\partial_2 f(\cdot,B.)\,|\,\mathcal{P}_n\right]&=\sum_{k=1}^{n}n\int_{\frac{k-1}{n}}^{\frac{k}{n}}\mathbb{E}\left[\partial_2 f(s,B_s)\,|\,\mathcal{F}_{\frac{k-1}{n}}\right]\,ds\,\mathbb{1}_{(\frac{k-1}{n},\frac{k}{n}]}\\&=\sum_{k=1}^{n}\partial_2 f(\tfrac{k-1}{n},B_{\frac{k-1}{n}})\mathbb{1}_{(\frac{k-1}{n},\frac{k}{n}]},\end{aligned}$$

by the martingale property.

Using that log is concave, we get for all $n\in\mathbb{N}$,

$$\begin{aligned}\mathbb{E}&\left[\int_0^1\sum_{k=1}^{n}\log\partial_2 f(\tfrac{k-1}{n},B_{\frac{k-1}{n}})\mathbb{1}_{(\frac{k-1}{n},\frac{k}{n}]}(t)\,dt\right]\\&=\mathbb{E}\left[\int_0^1\log\mathbb{E}_{\lambda\otimes\mathbb{P}}\left[\partial_2 f(\cdot,B.)\,|\,\mathcal{P}_n\right]\,dt\right]\\&\geq\mathbb{E}\left[\int_0^1\mathbb{E}_{\lambda\otimes\mathbb{P}}\left[\log\partial_2 f(\cdot,B.)\,|\,\mathcal{P}_n\right]\,dt\right]\\&=\mathbb{E}\left[\int_0^1\log\partial_2 f(t,B_t)\,dt\right],\end{aligned}$$

where the last equality is due to Fubini and the tower property of conditional expectation. The integrals appearing above are all well defined since $\log \partial_2 f(t, B_t)$ is $\lambda \otimes \mathbb{P}$ integrable. Indeed, the positive part of $\log \partial_2 f(t, B_t)$ is dominated by $\partial_2 f(t, B_t)$, which is square-integrable, and for the negative part, $(\log \partial_2 f(t, B_t))^-$, this follows from the fact that $(\log(\cdot))^-$ is convex and therefore $t \mapsto \mathbb{E}\left[(\log \partial_2 f(t, B_t))^-\right]$ is non-decreasing with finite value for $t = 1$. This also justifies the application of Fubini's theorem.

By continuity,

$$\sum_{k=1}^{n} \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}}) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t) \xrightarrow{n \to \infty} \log \partial_2 f(t, B_t) \qquad \lambda \otimes \mathbb{P}\text{-a.e.},$$

and additionaly, since $\log(x) \leq x$, for all $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}}) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t) \leq \sum_{k=1}^{n} \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}}) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t) \leq \sup_{t \in [0,1]} \partial_2 f(t, B_t)$$

and the expression on the right-hand side is $\lambda \otimes \mathbb{P}$-integrable by Doob's L^2-inequality and therefore with Fatou's lemma and the lower bound from above, we get

$$\begin{aligned}
\mathbb{E}\left[\int_0^1 \log \partial_2 f(t, B_t)\, dt\right] &\geq \limsup_{n \to \infty} \mathbb{E}\left[\int_0^1 \sum_{k=1}^{n} \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}})^2 \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t)\, dt\right] \\
&\geq \liminf_{n \to \infty} \mathbb{E}\left[\int_0^1 \sum_{k=1}^{n} \log \partial_2 f(\tfrac{k}{n}, B_{\frac{k}{n}})^2 \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t)\, dt\right] \\
&= \liminf_{n \to \infty} \mathbb{E}\left[\int_0^1 \sum_{k=1}^{n} \log \partial_2 f(\tfrac{k-1}{n}, B_{\frac{k-1}{n}}) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t)\, dt\right] - \frac{C}{n} \\
&\geq \liminf_{n \to \infty} \mathbb{E}\left[\int_0^1 \log \partial_2 f(t, B_t)\, dt\right] - \frac{C}{n}.
\end{aligned}$$

This proves (3.5) and we get in (3.4),

$$\lim_{n \to \infty} \frac{1}{n} \mathrm{H}\left(\mathbb{Q} \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n} = \frac{1}{2} \mathbb{E}\left[\int_0^1 \partial_2 f(t, B_t)^2 - 1 - \log \partial_2 f(t, B_t)^2\, dt\right].$$

□

Next, we consider the stochastic differential equation,

$$\begin{cases} dM_t = \sigma(t, M_t)\, dB_t \\ M_0 = x, \end{cases} \tag{3.6}$$

where $\sigma\colon [0,1] \times \mathbb{R} \to \mathbb{R}$ measurable and $x \in \mathbb{R}$. Under the assumptions on the diffusion coefficient σ from Remark 6 (in particular a unique solution exists) the specific relative entropy with respect to Wiener measure (if it exists) is finite.

Lemma 5. *Let $\sigma\colon [0,1] \times \mathbb{R} \to \mathbb{R}$ satisfy that*

(i) there exists $\delta \in (0,1]$ such that $\delta \leq \sigma(t,x) \leq \frac{1}{\delta}$ for all $t \in [0,1]$ and $x \in \mathbb{R}$

(ii) there exists a constant $L > 0$ such that $|\sigma(t,y) - \sigma(s,x)| \leq L(|y-x| + |t-s|)$ for all $x, y \in \mathbb{R}$ and $s, t \in [0,1]$.

Let $M^x = (M^x_t)_{t\in[0,1]}$ be the solution to (3.6) with law $\mathbb{Q}^x$ on $(\mathcal{C}, \mathcal{F})$. Then there exists a constant $C \geq 0$ such that,

$$\frac{1}{n} H(\mathbb{Q}^x \,|\, \mathbb{W}^x)\,|_{\mathcal{F}^n} < C \text{ for all } n \in \mathbb{N}.$$

In particular,

$$\limsup_{n\to\infty} \frac{1}{n} H(\mathbb{Q}^x \,|\, \mathbb{W}^x)\,|_{\mathcal{F}^n} < \infty.$$

Proof. To ease notation we write M for M^x. Let $\gamma_t(\cdot, x)$ be the Gaussian density with mean x and variance t. Denote by Q the probability transition function of M from Remark 7 and P the probability transition function with density $\gamma_\cdot(\cdot,\cdot)$. From Remark 9 we know M has a transition density function $q(t,y;s,x)$ that satisfies: There exist $C, c > 0$ such that,

$$q(t,y;s,x) \leq \frac{C}{\sqrt{t-s}} e^{-c\frac{(x-y)^2}{t-s}} \text{ for all } s < t \in [0,1] \text{ and } x, y \in \mathbb{R}.$$

Similar as in Lemma 4, we have by Lemma 2 and the Markov property,

$$\begin{aligned}
&\mathrm{H}\left(\mathcal{L}(M_0, M_{\frac{1}{n}}, \ldots, M_1) \mid \mathcal{L}(B^x_0, B^x_{\frac{1}{n}}, \ldots, B^x_1)\right) \\
&= \sum_{k=1}^{n} \int_{\mathbb{R}^2} \log \frac{Q(\frac{k}{n}, dx_{\frac{k}{n}}; \frac{k-1}{n}, x_{\frac{k-1}{n}})}{P(\frac{k}{n}, dx_{\frac{k}{n}}; \frac{k-1}{n}, x_{\frac{k-1}{n}})} \, d\mathcal{L}(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})(x_{k-1}, x_k) \\
&= \sum_{k=1}^{n} \int_{\mathbb{R}^2} \log \frac{q(\frac{k}{n}, x_k; \frac{k-1}{n}, x_{k-1})}{\gamma_{\frac{1}{n}}(x_k; , x_{k-1})} \, d\mathcal{L}(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})(x_{k-1}, x_k)
\end{aligned}$$

$$= \sum_{k=1}^{n} \mathbb{E}\left[\log \frac{q(\frac{k}{n}, M_{\frac{k}{n}}; \frac{k-1}{n}, M_{\frac{k-1}{n}})}{\gamma_{\frac{1}{n}}(M_{\frac{k}{n}}; M_{\frac{k-1}{n}})}\right].$$

Together with the Gaussian upper bound and since

$$\sum_{k=1}^{n} \mathbb{E}\left[(M_{\frac{k}{n}} - M_{\frac{k-1}{n}})^2\right] = \mathbb{E}\left[\int_0^1 \sigma(t, M_t)^2 \, dt\right] \in [\delta^2, \frac{1}{\delta^2}]$$

we get,

$$\begin{aligned}
\frac{1}{n}\mathrm{H}\left(\mathbb{Q}^x \,|\, \mathbb{W}^x\right)|_{\mathcal{F}^n} &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[\log \frac{q(\frac{k}{n}, M_{\frac{k}{n}}; \frac{k-1}{n}, M_{\frac{k-1}{n}})}{\gamma_{\frac{1}{n}}(M_{\frac{k}{n}}, M_{\frac{k-1}{n}})}\right] \\
&\leq \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[\log(C\sqrt{2\pi}) - cn(M_{\frac{k}{n}} - M_{\frac{k-1}{n}})^2 + \frac{n}{2}(M_{\frac{k}{n}} - M_{\frac{k-1}{n}})^2\right] \\
&\leq \log(C\sqrt{2\pi}) - c\delta^2 + \frac{1}{2\delta^2},
\end{aligned}$$

for all $n \in \mathbb{N}$. □

Now we restrict ourselves to diffusion coefficients that do not depend on the time variable t. In particular, a solution is time-homogeneous. Let $\sigma\colon \mathbb{R} \to \mathbb{R}$ be measurable and consider the SDE,

$$\begin{cases} dM_t = \sigma(M_t)\, dB_t \\ M_0 = x, \end{cases} \tag{3.7}$$

for $x \in \mathbb{R}$.

In the next lemma our assumptions on σ guarantee that a unique solution exists and that it has a transition density function. But more importantly they allow us to derive estimates on the transition density. These will then be the key to computing the specific relative entropy between the laws of two martingales arising as solutions of such SDEs.

Lemma 6. *Let M^x be the solution to (3.7) and let σ satisfy,*

(i) $\sigma\colon \mathbb{R} \to \mathbb{R}_+$ is twice continuously differentiable

(ii) for some $L \in \mathbb{R}$ we have $\max\{\|\sigma'\|_\infty, \|\sigma''\|_\infty\} < L$

(iii) there exists $\delta \in (0, 1]$ such that $\delta \leq \sigma(x) \leq \frac{1}{\delta}$ for all $x \in \mathbb{R}$.

Let $p(t,x,y)$ be the transition density function of M^x. Then there exist constants C_1 and C_2, depending only on δ and L, such that, for all $t \in (0,1]$ and $x, y \in \mathbb{R}$:

(i)

$$p(t,x,y) \geq e^{-C_2 t} \frac{1}{\sqrt{2\pi t}} \sqrt{\frac{\sigma(x)}{\sigma(y)}} \frac{1}{\sigma(y)} e^{-\frac{1}{2t}\left(\int_x^y \frac{1}{\sigma(u)}\, du\right)^2}$$

(ii)

$$p(t,x,y) \leq e^{C_1 t} \frac{1}{\sqrt{2\pi t}} \sqrt{\frac{\sigma(x)}{\sigma(y)}} \frac{1}{\sigma(y)} e^{-\frac{1}{2t}\left(\int_x^y \frac{1}{\sigma(u)}\, du\right)^2}$$

Proof. 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

$$g(M_t^x) = B_t - \frac{1}{2} \int_0^t \sigma'(M_s^x)\, ds.$$

Hence $Y_t := g(M_t^x)$ 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

$$\begin{aligned}
&\mathbb{E}\left[\Phi(M_t^x)\right] = \mathbb{E}\left[\Phi(g^{-1}(Y_t))\right] \\
&= \mathbb{E}\left[\Phi(g^{-1}(B_t)) \exp\left(\int_0^t b(B_s)\, dB_s - \frac{1}{2}\int_0^t b(B_s)^2\, ds\right)\right] \\
&= \mathbb{E}\left[\Phi(g^{-1}(B_t)) \exp\left(\int_0^{B_t} b(u)\, du - \frac{1}{2}\int_0^t b'(B_s)\, ds - \frac{1}{2}\int_0^t b(B_s)^2\, ds\right)\right] \qquad (3.8)
\end{aligned}$$

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 (3.8) note that,

$$-b' - b^2 \leq -b' = \frac{1}{2}(\sigma \cdot \sigma'') \circ g^{-1} \leq \frac{L}{2\delta} \qquad (3.9)$$

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

$$\mathbb{E}\left[\Phi(M_t^x)\right] \leq \exp(tC_1)\mathbb{E}\left[\Phi(g^{-1}(B_t))\exp\left(\int_0^{B_t} b(u)\,du\right)\right]. \tag{3.10}$$

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 (3.10) 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:

$$\begin{aligned}
\mathbb{E}\left[\Phi(g^{-1}(B_t))\exp\left(\int_0^{B_t} b(u)\,du\right)\right] &= \mathbb{E}\left[\Phi(g^{-1}(B_t))\sqrt{\frac{\sigma(x)}{\sigma(g^{-1}(B_t))}}\right] \\
&= \int_{\mathbb{R}} \Phi(g^{-1}(z))\sqrt{\frac{\sigma(x)}{\sigma(g^{-1}(z))}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{z^2}{2t}\right)dz \\
&= \int_{\mathbb{R}} \Phi(y)\sqrt{\frac{\sigma(x)}{\sigma(y)}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{g(y)^2}{2t}\right)g'(y)\,dy \\
&= \int_{\mathbb{R}} \Phi(y)\sqrt{\frac{\sigma(x)}{\sigma(y)}}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{1}{2t}\left(\int_x^y \frac{1}{\sigma(u)}\,du\right)^2\right)\frac{1}{\sigma(y)}\,dy.
\end{aligned}$$

We conclude that

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

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

The lower bound follows analogously by replacing (3.9) 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 $C_2 \coloneqq \frac{L\delta}{4} + \frac{L^2}{8}$. □

Remark 20. In Lemma 5 the estimate from Remark 9 lead to finiteness of the specific relative entropy. Now, Lemma 6 allows us to gauge small-time behaviour of the solution M^x more precisely and will lead to a closed-form expression for the specific relative entropy.

Remark 21. The two-sided estimate on the transition density function in Lemma 6 allows us to consider more general reference measures than Wiener measure $\mathbb{W}$. We take the new reference measure to be the law of N^x, which is solution to the SDE,

$$\begin{cases} dN_t = \eta(N_t)\,dB_t \\ N_0 = x, \end{cases} \tag{3.11}$$

where $\eta\colon \mathbb{R} \to \mathbb{R}$ measurable.

Theorem 12. *Assume the coefficients $\sigma, \eta\colon \mathbb{R} \to \mathbb{R}_+$ both satisfy the assumptions of Lemma 6.*
Let M^x be the solution to (3.7) and N^x the solution to (3.11) 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

$$h(\mathbb{Q}^x \,|\, \mathbb{P}^x) = \frac{1}{2}\mathbb{E}\left[\int_0^1 \left\{\frac{\sigma(M_s^x)^2}{\eta(M_s^x)^2} - 1 - \log\frac{\sigma(M_s^x)^2}{\eta(M_s^x)^2}\right\} ds\right]. \tag{3.12}$$

Remark 22. Recalling that X stood for the canonical process, Formula (3.12) becomes

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

Note that this is consistent with Example 4 in Chapter 2 by taking $\sigma(x) = \sigma$ and $\eta(x) = \eta$.

Remark 23. By taking $\eta(x) = 1$, we recover Formula 2.5 from Chapter 2, as a special case.

Proof. 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. As in the proof of Lemma 5, we have,

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

Define,

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

analogously for η. By Lemma 6 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 $k = 1, \ldots, n$:

$$\begin{aligned}&\mathbb{E}\left[\log \frac{q(\frac{1}{n}, M_{\frac{k-1}{n}}, M_{\frac{k}{n}})}{p(\frac{1}{n}, M_{\frac{k-1}{n}}, M_{\frac{k}{n}})}\right] \\ &\leq \mathbb{E}\Bigg[\frac{C}{n} + \frac{1}{2}\log\frac{\sigma(M_{\frac{k-1}{n}})\eta(M_{\frac{k}{n}})}{\sigma(M_{\frac{k}{n}})\eta(M_{\frac{k-1}{n}})} - \log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})} \\ &\qquad\qquad - \frac{n}{2}d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 + \frac{n}{2}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\Bigg]. \end{aligned} \tag{3.13}$$

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 $n \in \mathbb{N}$:

$$
\begin{aligned}
&\frac{1}{n}\mathrm{H}\left(\mathbb{Q}^x \,|\, \mathbb{P}^x\right) \\
&\leq \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\Bigg[\frac{C}{n} + \frac{1}{2}\log\frac{\sigma(M_{\frac{k-1}{n}})\eta(M_{\frac{k}{n}})}{\sigma(M_{\frac{k}{n}})\eta(M_{\frac{k-1}{n}})} - \log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})} \\
&\qquad\qquad - \frac{n}{2}d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 + \frac{n}{2}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\Bigg] \\
&= \frac{C}{n} + \mathbb{E}\Bigg[\frac{1}{2n}\log\frac{\sigma(M_0)\eta(M_1)}{\sigma(M_1)\eta(M_0)} - \frac{1}{n}\sum_{k=1}^{n}\log\frac{\sigma(M_{\frac{k}{n}})}{\eta(M_{\frac{k}{n}})} \\
&\qquad\qquad - \frac{1}{2}\sum_{k=1}^{n}d_\sigma(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 + \frac{1}{2}\sum_{k=1}^{n}d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\Bigg]
\end{aligned}
\tag{3.14}
$$

We now show that (3.14) converges to the right-hand side of (3.12). 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

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

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

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

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

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

Thus we have,

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n} d_\eta(M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2\right] \\
&\quad = \mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}\left(\int_{M_{\frac{k-1}{n}}}^{M_{\frac{k}{n}}} \frac{1}{\eta(u)}\,du\right)^2\right] \\
&\quad = \mathbb{E}\left[\frac{1}{2}\sum_{k=1}^{n}\Big(F(M_{\frac{k}{n}}) - F(M_{\frac{k-1}{n}})\Big)^2\right] \\
&\quad = \frac{1}{2}\sum_{k=1}^{n}\mathbb{E}\Bigg[\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)}\,dB_u\right)^2 - \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)}\,dB_u \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2}\sigma(M_u)^2\,du \\
&\qquad\qquad + \frac{1}{4}\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2}\sigma(M_u)^2\,du\right)^2\Bigg] \\
&\quad = \mathbb{E}\Bigg[\frac{1}{2}\int_0^1 \left(\frac{\sigma(M_u)}{\eta(M_u)}\right)^2 du - \frac{1}{2}\sum_{k=1}^{n}\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)}\,dB_u \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2}\sigma(M_u)^2\,du \\
&\qquad\qquad + \frac{1}{2}\sum_{k=1}^{n}\frac{1}{4}\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\eta'(M_u)}{\eta(M_u)^2}\sigma(M_u)^2\,du\right)^2\Bigg].
\end{aligned}
$$

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 for $k = 1, ..., n$,

$$
\begin{aligned}
\mathbb{E}\left[\left|\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)}\,dB_u\right|\right] &\leq \left(\mathbb{E}\left[\left(\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)}{\eta(M_u)}\,dB_u\right)^2\right]\right)^{\frac{1}{2}} \\
&= \left(\mathbb{E}\left[\int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\sigma(M_u)^2}{\eta(M_u)^2}\,du\right]\right)^{\frac{1}{2}} \\
&\leq \sqrt{\frac{1}{n}\frac{1}{\delta^2}},
\end{aligned}
$$

and therefore

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

as $n \to \infty$. Moreover,

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

which yields

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{2} \sum_{k=1}^{n} d_\eta (M_{\frac{k-1}{n}}, M_{\frac{k}{n}})^2 \right] = \mathbb{E}\left[\int_0^1 \frac{1}{2} \left(\frac{\sigma(M_u)}{\eta(M_u)} \right)^2 du \right]. \tag{3.15}$$

Lastly, the same arguments with η replaced by σ, show

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

We conclude that

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

Analogously, by the lower bound for q and the upper bound for p from Lemma 6, applied in (3.13), we get:

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

which finishes the proof. □

References

[1] 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.

[2] Avner Friedman. *Stochastic differential equations and applications 1*. Acad. Press, 1975.

[3] Nina Gantert. *Einige grosse Abweichungen der Brownschen Bewegung: Inaugural-Dissertation zur Erlangung der Doktorwürde*. Mathematisches Institut der Universität Bonn, 1991.

[4] Robert M Gray. *Entropy and Information Theory*. Springer Science & Business Media, 2011.

[5] Achim Klenke. *Probability theory: A comprehensive course*. Springer Science & Business Media, 2013.