MASTERARBEIT | MASTER'S THESIS

Titel | Title

Specific Relative Entropy and the Gantert Conjecture

verfasst von | submitted by

Gregor Kirchner BSc

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

Master of Science (MSc)

Wien | Vienna, 2025

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:	Assoz. Prof. Dr. Julio Daniel Backhoff-Veraguas

Abstract

The notion of the specific relative entropy (SRE) was introduced by N. Gantert in her doctoral thesis (see [12]). This SRE provides a notion of distance between laws of continuous processes, which is finer than the well-established relative entropy. It is of particular interest in the context of continuous martingale laws, which typically have infinite relative entropy, while their specific relative entropy is often finite.

In the same thesis, Gantert provided a lower bound for the SRE between the law of a square-integrable, continuous martingale and the Wiener measure. This result is now known as the Gantert inequality. It was further conjectured that for every square-integrable, continuous martingale this lower bound would in fact be optimal. Over the recent years, this conjecture has been verified for several useful classes of martingales (see [16], [6], [7]). However, a counterexample was also found (see [3]). Thus, while the Gantert conjecture does not hold in full generality, it remains an open problem to characterize the class of square-integrable, continuous martingales for which it is valid.

The definition of the SRE requires a deliberate choice of a refining sequence of partitions of the unit interval. In her original definition, Gantert has chosen the dyadic partition sequence, which has since become the standard in the literature. Motivated by some concrete problems, we will propose in this thesis a more general definition of the SRE that allows the choice of partition sequences besides the dyadic one. For this general definition, we will then proceed to prove the Gantert inequality and we verify the Gantert conjecture for the classes for which it has already been verified in the dyadic case. Furthermore, we will verify the Gantert conjecture for a new class of processes, which we call invertible, iterative Bass martingales.

Zusammenfassung

Das Konzept der spezifischen relativen Entropie (SRE) wurde von N. Gantert in ihrer Doktorarbeit eingeführt (siehe [12]). Diese SRE stellt einen Abstandsbegriff für Verteilungen von stetigen Prozessen dar, der feiner ist jener der lange etablierten relativen Entropie. Sie ist von besonderem Interesse im Kontext von Verteilungen von stetigen Martingalprozessen, da diese typischerweise eine unendliche relative Entropie haben, während ihre spezifische relative Entropie oft endlich ist.

In ihrer Doktorarbeit wurde von Gantert auch eine untere Schranke für die SRE zwischen der Verteilung eines quadrat-integrierbaren, stetigen Martingals und dem Wiener Maß bewiesen. Dieses Resultat ist nun bekannt als Gantert-Ungleichung. Es wurde weiters die Vermutung aufgestellt, dass für jeden quadrat-integrierbaren, stetigen Martingalprozess diese untere Schranke sogar optimal sein soll. In den vergangen Jahren konnte diese Vermutung für ein paar nützliche Klassen von Martingalprozessen bewiesen werden (siehe [16], [6], [7]). Es wurde jedoch auch ein Gegenbeispiel entdeckt (siehe [3]). Während die Gantert-Vermutung also nicht in voller Allgemeinheit gültig sein kann, so verbleibt es dennoch ein offenes Problem die Klasse von quadrat-integrierbaren, stetigen Martingalprozessen zu charakterisieren für die sie korrekt ist.

Die Definition der SRE erfordert die Auswahl einer konkreten verfeinernden Folge von Partitionen des Einheitsintervalls. In der originalen Definition wurde von Gantert die dyadische Partitionenfolge verwendet und dies hat sich als Standard in der Literatur etabliert. Motiviert von konkreten Problemstellungen, möchten wir in dieser Masterarbeit eine allgemeinere Definition der SRE darlegen, welche die Auswahl von anderen Partitionsfolgen anstelle der Dyadischen erlaubt. Für diese allgemeinere Definition werden wir dann die Gantert-Ungleichung beweisen und die Gantert-Vermutung für alle jene Klassen verifizieren für die dies bereits im dyadischen Fall geschehen ist. Ebenso werden wir die Gantert-Vermutung für eine neue Klasse von Prozessen verifizieren, die wir als invertierbare, iterative Bass-Martingale bezeichnen.

Acknowledgements

My deepest thanks go to my supervisor, Julio Backhoff, for his guidance and support throughout the process of writing this thesis, and for the many hours of his time that he dedicated to me.

Contents

Section 1

Preliminaries

In this first section, we will establish some notation and review definitions and results that we will use throughout this text. We assume that the reader is familiar with the basic theory of stochastic processes in continuous time. In particular, we assume familiarity with martingales and Brownian motion. Some knowledge about stochastic integration and stochastic differential equations would be ideal, however, the definitions and results that we use in this regard can be read up on in the Appendix.

1.1 Notations and Conventions

To start, we will outline some notations and conventions.

We will regularly work with partitions of the unit interval $[0, 1]$. For us, a partition of $[0, 1]$ will be any finite tuple $(\pi_0, ..., \pi_k)$ with $0 = \pi_0 < \pi_1 < ... < \pi_k = 1$. Given an arbitrary partition π of $[0, 1]$, we use the following conventions.

- The partition points of π will be denoted by π_i where i ranges from 0 to $\#\pi$ where $\#\pi := \{\text{number of partition points}\} - 1$.
- $\Delta_i\pi$ denotes $\pi_i - \pi_{i-1}$ for any $i \in [\#\pi] := \{1, ..., \#\pi\}$.
- We define the *mesh size* of π by $\mathrm{mesh}(\pi) := \max\limits_{i\in[\#\pi]} \Delta_i\pi$

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. In general, we will use capital letters to denote $\mathbb{R}$-valued random variables and bold capital letters to denote $\mathbb{R}^d$-valued ($d > 1$) random variables (which we refer to as random vectors). Given a random variable Y and a random vector $\boldsymbol{X}$, we will use the following notations.

- $\mathcal{L}_\mathbb{P}Y$ denotes the law of Y under the measure $\mathbb{P}$.
- $\mathcal{L}_\mathbb{P}[Y \mid \boldsymbol{X}]$ denotes the conditional law of Y given $\boldsymbol{X}$ (as defined in Appendix A.2)
- $\mathcal{L}_\mathbb{P}[Y \mid \boldsymbol{X} = \boldsymbol{x}]$ denotes the conditional law of Y given $\boldsymbol{X} = \boldsymbol{x}$ for some $\boldsymbol{x} \in \mathbb{R}^d$ (as defined in Appendix A.2)
- $\mathbb{E}_\mathbb{P}[Y]$ denotes the expectation of Y with respect to the measure $\mathbb{P}$. Sometimes, we will use the more precise notation $\mathbb{E}^\omega_\mathbb{P}[Y(\omega)]$ to highlight that ω is the integration variable.

For example, for an appropriate $f : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, the expression $\mathbb{E}^{\omega}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[f(\boldsymbol{X}(\omega), Y)]]$ shall be read as

$$\mathbb{E}^{\omega}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[f(\boldsymbol{X}(\omega), Y)]] = \int_{\Omega}\int_{\Omega} f(\boldsymbol{X}(\omega), Y(\nu))\, d\mathbb{P}(\nu)\, d\mathbb{P}(\omega)$$

For a topological space T, we denote its Borel-σ-algebra by $\mathcal{B}(T)$ or just $\mathcal{B}$ if the space is clear from the context. In the common case, T will be $\mathbb{R}^d$ for some $d \geq 1$, naturally equipped with its Euclidean topology. We denote by m the Lebesgue measure on $(\mathbb{R}^d, \mathcal{B})$ (from the context it will always be clear that m is the d-dimensional Lebesgue measure). Further, we use the following notations.

- $p[Y]$ denotes the density of Y with respect to the Lebesgue measure if it exists.

- $p[Y \mid \boldsymbol{X} = \boldsymbol{x}]$ denotes the conditional density of Y given $\boldsymbol{X} = \boldsymbol{x}$ for some $\boldsymbol{x} \in \mathbb{R}^d$, which we define as the density of $\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]$ w.r.t. to the Lebesgue measure if it exists.

- $p[Y \mid \boldsymbol{X}]$ denotes the conditional density of Y given $\boldsymbol{X}$, which we define as the random density given by $p[Y \mid \boldsymbol{X}](\omega, y) := p[Y \mid \boldsymbol{X} = \boldsymbol{X}(\omega)](y)$.

Given $\mu \in \mathbb{R}$ and $\sigma^2 \in (0, \infty)$, we denote by

- $N[\mu, \sigma^2]$ the Gaussian density with mean μ and variance σ^2, i.e.

$$N[\mu, \sigma^2] : \mathbb{R} \to [0, \infty), \qquad N[\mu, \sigma^2](x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  In the case of $\mu = 0$, we sometimes shorten the notation to N_{σ^2}.

- $\mathcal{N}[\mu, \sigma^2]$ the Gaussian distribution with mean μ and variance σ^2, i.e. the measure on $\mathbb{R}$ that satisfies

$$N[\mu, \sigma^2] = \frac{d\mathcal{N}[\mu, \sigma^2]}{dm}$$

  In the case of $\mu = 0$, we sometimes shorten the notation to $\mathcal{N}_{\sigma^2}$.

Whenever we work with stochastic processes, we consider a filtered probability space $(\Omega, \mathcal{S}, (\mathcal{S}_t)_{t\in[0,1]}, \mathbb{S})$ as given. We implicitly assume that this space is rich enough, such that all our desired processes are definable. Our processes will, with few exceptions, always have a time frame of $[0, 1]$ and will in general always be $\mathbb{R}$-valued. Given such a process $M = (M_t)_{t\in[0,1]}$ and a partition π of $[0, 1]$, we will use for any $i \in [\#\pi]$ the following notations.

- $\boldsymbol{M}_{\pi_i}$ denotes the random vector $(M_{\pi_0}, ..., M_{\pi_i})$.

- $\boldsymbol{M}_{\pi_i} = f$ is short for $(M_{\pi_0}, ..., M_{\pi_i}) = (f(\pi_0), ..., f(\pi_i))$ for any function $f : [0, 1] \to \mathbb{R}$.

- $\Delta_i^{\pi} M$ denotes the random variable $M_{\pi_i} - M_{\pi_{i-1}}$.

1.2 Quadratic Variation

The notion of the quadratic variation plays a vital role in the theory of stochastic integration and also in our study of the specific relative entropy. We shall thus review some definitions and results connected to it. To fit our setting, we will formulate them in terms of continuous functions $[0,1] \to \mathbb{R}$ and continuous processes with time frame $[0,1]$.

Definition 1 (p-Variation)**.** Let $f : [0,1] \to \mathbb{R}$ be continuous, $p \geq 1$ and $t \in [0,1]$ be arbitrary. The *p-variation* of f over $[0,t]$ is given by

$$V_t^{(p)}(f) := \sup_{\pi} \sum_{i=1}^{\#\pi} \Big| f(\pi_i \wedge t) - f(\pi_{i-1} \wedge t) \Big|^p$$

where π ranges over all partitions of $[0,1]$. In the case of $p = 1$, the above is the *total variation* of f over $[0,t]$.

We shall say that a continuous function $f : [0,1] \to \mathbb{R}$ has *finite variation* if $V_1^{(1)}(f) < \infty$ (which in particular implies $V_t^{(1)}(f) < \infty$ for all $t \in [0,1]$). We say that a continuous process $X = (X_t)_{t \in [0,1]}$ is a *finite variation process* if $V_1^{(1)}(X) < \infty$ almost surely.

The total variation admits the following equivalent definition: Let $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ be a refining sequence of partitions of $[0,1]$ with mesh$(\pi^{(k)}) \to 0$, then the total variation of a continuous function $f : [0,1] \to \mathbb{R}$ over $[0,t]$ is given by

$$V_t^{(1)}(f) = \lim_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Big| f(\pi_i^{(k)} \wedge t) - f(\pi_{i-1}^{(k)} \wedge t) \Big|$$

In particular, the limit on the right side always exists and is independent of the choice of $\mathfrak{P}$. However, the corresponding statement is not true for p-variation where $p > 1$! In the case that we are interested in, that of $p = 2$, the limit

$$\lim_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Big(f(\pi_i^{(k)} \wedge t) - f(\pi_{i-1}^{(k)} \wedge t) \Big)^2$$

might not exist, and if it does, it might be distinct for different choices of $\mathfrak{P}$ (see Example 1 below). In particular, the limit need not coincide with the 2-variation of f over $[0,t]$.

Definition 2 (Quadratic Variation)**.** Let $f : [0,1] \to \mathbb{R}$ be continuous, $t \in [0,1]$ and $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ a refining partition sequence of $[0,1]$ with mesh size converging to 0. The *quadratic variation* of f over $[0,t]$ along the partition sequence $\mathfrak{P}$ is given by

$$[f]_t^{\mathfrak{P}} := \lim_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Big(f(\pi_i^{(k)} \wedge t) - f(\pi_{i-1}^{(k)} \wedge t) \Big)^2$$

if the limit exists.

Example 1. Let $B = (B_t)_{t\in[0,1]}$ be a Brownian motion, then for almost every $\omega \in \Omega$

$$\inf_{\mathfrak{P}} [B(\omega)]_1^{\mathfrak{P}} = 0 \quad \text{and} \quad \sup_{\mathfrak{P}} [B(\omega)]_1^{\mathfrak{P}} = \infty \tag{1.1}$$

where $\mathfrak{P}$ ranges over all refining partition sequences of $[0,1]$ with mesh size going to 0. A proof for this can be found in ([9], pp. 47-48). So for almost every sample path of Brownian motion, one can find partition sequences along which the quadratic variation is as small/large as one desires. However, it is well-known that regardless of what partition sequence $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ we choose, we have for every $t \in [0,1]$ that

$$\sum_{i=1}^{\#\pi^{(k)}} \left(B_{\pi_i^{(k)}\wedge t} - B_{\pi_{i-1}^{(k)}\wedge t} \right)^2 \xrightarrow{\mathbb{S}} t \tag{1.2}$$

This convergence even holds in L^2 (see [11], Prop. 2.16) and almost surely (see [15], Thm. 3.4.3).

It is a central result for the theory of stochastic integration that, as in (1.2), for every square-integrable, continuous martingale $M = (M_t)_{t\in[0,1]}$, $t \in [0,1]$ and refining partition sequence $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ with mesh size going to 0, there exists a random variable $[M]_t$ such that

$$\sum_{i=1}^{\#\pi^{(k)}} \left(M_{\pi_i^{(k)}\wedge t} - M_{\pi_{i-1}^{(k)}\wedge t} \right)^2 \xrightarrow{\mathbb{S}} [M]_t \tag{1.3}$$

Even stronger, as stated in the following Theorem, one can find a continuous, non-decreasing process $[M] = ([M]_t)_{t\in[0,1]}$ so that the convergence in (1.3) holds uniformly over all $t \in [0,1]$.

Theorem 1.1 (*Existence of Quadratic Variation*)**.** Let $M = (M_t)_{t\in[0,1]}$ be a square-integrable, continuous martingale, then there exists a (up to indistinguishability) unique, adapted, non-decreasing, continuous process $[M] = ([M]_t)_{t\in[0,1]}$ such that

- $[M]_0 = 0$ a.s.
- $M^2 - [M]$ is a continuous martingale

Furthermore, if $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ is any refining partition sequence of $[0,1]$ with mesh size converging to 0, then

$$\sup_{t\in[0,1]} \left| [M]_t - \sum_{i=1}^{\#\pi^{(k)}} \left(M_{\pi_i^{(k)}\wedge t} - M_{\pi_{i-1}^{(k)}\wedge t} \right)^2 \right| \xrightarrow{\mathbb{S}} 0$$

Proof. see for example ([11], Thm. 4.9 + Thm. 4.13 (ii))

□

Uniqueness up to indistinguishability in the above Theorem means that any two processes X and Y that have the properties we require of $[M]$ satisfy

$$\mathbb{S}[X_t = Y_t \text{ for all } t \in [0,1]] = 1$$

i.e. almost all sample paths of X and Y coincide.

Definition 3 (Quadratic Variation Process)**.** Let M be a square-integrable, continuous martingale. The unique process $[M]$ from Theorem 1.1 is the *quadratic variation process* of M.

Often, we will refer to $[M]$ as just the quadratic variation of M.

We now show how the quadratic variation can be evaluated for most of the martingales that we encounter throughout this text. For a Brownian motion B, we know that the quadratic variation process $[B]$ is given by

$$[B]_t = t$$

If M is a square-integrable, continuous martingale and $H \in L^2(M)$ (see Appendix A.3), then the quadratic variation of the stochastic integral $H.M$ is by Proposition A.9 given by

$$[H.M]_t = \left[\int_0^{\cdot} H_s \, dM_s\right]_t = \int_0^t H_s^2 \, d[M]_s \tag{1.4}$$

So in particular, if $H \in L^2(B)$, then

$$[H.B]_t = \int_0^t H_s^2 \, ds$$

For a function $f : [0,1] \times \mathbb{R} \to \mathbb{R}$, the process $(f(t, M_t))_{t\in[0,1]}$ might not be a martingale, however, if f is sufficiently smooth, then the process is at least a *semimartingale.*

Definition 4 (Semimartingale)**.** An adapted process $X = (X_t)_{t\in[0,1]}$ is a continuous *semimartingale* if there exists a decomposition

$$X = X_0 + A + N \tag{1.5}$$

where X_0 is a random variable, A an adapted, continuous, finite variation process with $A_0 = 0$ and N a continuous martingale with $N_0 = 0$.

The decomposition in (1.5) is referred to as the *Doob-Meyer decomposition* of X. It is unique up to indistinguishability. We sensibly extend the definition of the quadratic variation process to continuous semimartingales by defining it as the quadratic variation process of its martingale part, i.e. we define $[X] := [N]$.

If M is a square-integrable, continuous martingale and $f \in C^2([0,1] \times \mathbb{R}; \mathbb{R})$, then, by Itō's formula (see Theorem A.10), the process $X = (f(t, M_t))_{t \in [0,1]}$ is a semimartingale and we are provided with the Doob-Meyer decomposition

$$X_t = X_0 + \underbrace{\int_0^t \partial_t f(s, M_s)\, ds + \frac{1}{2} \int_0^t \partial_{xx} f(s, M_s)\, d[M]_s}_{\text{finite variation part}} + \underbrace{\int_0^t \partial_x f(s, M_s)\, dM_s}_{\text{martingale part}}$$

In accordance with (1.4), the quadratic variation of X is then given by

$$[X]_t = \int_0^t (\partial_x f(s, M_s))^2\, d[M]_s$$

1.3 Relative Entropy

The specific relative entropy builds upon the well-established notion of relative entropy, so in this section we will define the relative entropy and look at some of its basic properties.

Definition 5 (Relative Entropy)**.** Let $\mathbb{P}, \mathbb{Q}$ be probability measures on a common measurable space $(\Omega, \mathcal{F})$. The *relative entropy* of $\mathbb{Q}$ with respect to $\mathbb{P}$ is given by

$$H(\mathbb{Q} \,|\, \mathbb{P}) := \begin{cases} \int_\Omega \log \frac{d\mathbb{Q}}{d\mathbb{P}}\, d\mathbb{Q} = \mathbb{E}_\mathbb{Q}\Big[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\Big] & , \text{ if } \mathbb{Q} \ll \mathbb{P} \\ \infty & , \text{ else} \end{cases}$$

Lemma 1.2. Let $\mathbb{P}, \mathbb{Q}$ be probability measures on a measurable space $(\Omega, \mathcal{F})$, then

(1) $H(\mathbb{Q} \,|\, \mathbb{P}) \geq 0$

(2) $H(\mathbb{Q} \,|\, \mathbb{P}) = 0 \Leftrightarrow \mathbb{Q} = \mathbb{P}$

Proof. ad (1): If $\mathbb{Q} \not\ll \mathbb{P}$, then by definition $H(\mathbb{Q} \,|\, \mathbb{P}) = \infty$ and the claim holds. Assume now that $\mathbb{Q} \ll \mathbb{P}$. We consider the non-negative, convex function

$$\phi : [0, \infty) \to [0, \infty), \qquad \phi(x) = \begin{cases} x \log x - x + 1 & , \text{ if } x > 0 \\ 1 & , \text{ if } x = 0 \end{cases}$$

and note that

$$\mathbb{E}_\mathbb{P}\Big[\phi\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big)\Big] = \mathbb{E}_\mathbb{P}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}} \cdot \log\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big)\Big] - \mathbb{E}_\mathbb{P}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}}\Big] + 1 = \mathbb{E}_\mathbb{Q}\Big[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\Big] = H(\mathbb{Q} \,|\, \mathbb{P})$$

By Jensen's inequality, we then obtain

$$H(\mathbb{Q} \,|\, \mathbb{P}) = \mathbb{E}_\mathbb{P}\Big[\phi\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big)\Big] \geq \phi\Big(\mathbb{E}_\mathbb{P}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}}\Big]\Big) = \phi(1) = 0 \tag{1.6}$$

ad (2): We note that equality holds in (1.6) if and only if $\frac{d\mathbb{Q}}{d\mathbb{P}} = 1$ $\mathbb{P}$-a.s., i.e. if $\mathbb{P} = \mathbb{Q}$. $\square$

The above Lemma suggests that we interpret the relative entropy as a 'distance' between probability measures, i.e. a measure for how different they are. However, it does not define a metric, as it is neither symmetric nor does it satisfy the triangle inequality!

In the following example, we will evaluate the relative entropy between two Gaussian distributions. As we will use this result regularly, it will be useful to fix from now on the function

$$\Upsilon : [0,\infty) \to \overline{\mathbb{R}}, \qquad \Upsilon(x) = \begin{cases} \frac{1}{2}(x - 1 - \log(x)) & , \text{ if } x > 0 \\ \infty & , \text{ if } x = 0 \end{cases} \tag{1.7}$$

Example 2. Let $\mu_1, \mu_2 \in \mathbb{R}$ and $\sigma_1^2, \sigma_2^2 \in (0,\infty)$, then we obtain

$$\begin{aligned} H(\mathcal{N}[\mu_1,\sigma_1^2] \,|\, \mathcal{N}[\mu_2,\sigma_2^2]) &= \int_\Omega \log \frac{d\mathcal{N}[\mu_1,\sigma_1^2]}{d\mathcal{N}[\mu_2,\sigma_2^2]}\, d\mathcal{N}[\mu_1,\sigma_1^2] \\ &= \int_{\mathbb{R}} \log \frac{N[\mu_1,\sigma_1^2](x)}{N[\mu_2,\sigma_2^2](x)} \cdot N[\mu_1,\sigma_1^2](x)\, dx \\ &= \frac{1}{2}\left(\frac{\sigma_1^2}{\sigma_2^2} - 1 - \log\frac{\sigma_1^2}{\sigma_2^2} + \frac{(\mu_2-\mu_1)^2}{\sigma_2^2}\right) = \Upsilon\left(\frac{\sigma_1^2}{\sigma_2^2}\right) + \frac{(\mu_2-\mu_1)^2}{2\sigma_2^2} \end{aligned}$$

In particular if $\mu_1 = \mu_2 =: \mu$, we get

$$H(\mathcal{N}[\mu,\sigma_1^2] \,|\, \mathcal{N}[\mu,\sigma_2^2]) = \Upsilon\left(\frac{\sigma_1^2}{\sigma_2^2}\right)$$

We now list a few useful results regarding the relative entropy.

Lemma 1.3. Let $\mathbb{Q} \ll \mathbb{S} \ll \mathbb{P}$ be probability measures on a measurable space $(\Omega, \mathcal{F})$, then

$$H(\mathbb{Q} \,|\, \mathbb{P}) = H(\mathbb{Q} \,|\, \mathbb{S}) + \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{S}}{d\mathbb{P}}\right]$$

Proof. By the chain rule for Radon-Nikodym derivatives, we have

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \overset{\mathbb{P}\text{-a.s.}}{=} \frac{d\mathbb{Q}}{d\mathbb{S}} \cdot \frac{d\mathbb{S}}{d\mathbb{P}} \tag{1.8}$$

and thus obtain

$$H(\mathbb{Q} \,|\, \mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\right] \overset{(1.8)}{=} \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\mathbb{S}}\right] + \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{S}}{d\mathbb{P}}\right] = H(\mathbb{Q} \,|\, \mathbb{S}) + \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{S}}{d\mathbb{P}}\right]$$

(Note that the last summand is not necessarily equal to $H(\mathbb{S} \,|\, \mathbb{P})$ since the expectation is taken with respect to $\mathbb{Q}$ and not $\mathbb{S}$!).

□

Given probability measures $\mathbb{Q} \ll \mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ and $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-algebra, then $\mathbb{Q}|_{\mathcal{G}} \ll \mathbb{P}|_{\mathcal{G}}$ and we let

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}} \text{ denote } \frac{d\mathbb{Q}|_{\mathcal{G}}}{d\mathbb{P}|_{\mathcal{G}}} \qquad \text{and} \qquad H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{G}} \text{ denote } H(\mathbb{Q}|_{\mathcal{G}} \mid \mathbb{P}|_{\mathcal{G}})$$

So we have

$$H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{G}} = \int_{\Omega} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}} d\mathbb{Q}|_{\mathcal{G}} = \int_{\Omega} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}} d\mathbb{Q} \tag{1.9}$$

Lemma 1.4. Let $\mathbb{Q}, \mathbb{P}$ be probability measures on a measurable space $(\Omega, \mathcal{F})$.

(1) If $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ are sub-σ-algebras, then

$$H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{H}} \leq H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{G}}$$

(2) If $(\mathcal{F}_n)_{n \geq 0}$ is a sequence of σ-algebras increasing to $\mathcal{F}$, then

$$H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_n} \nearrow H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}}$$

Proof. ad (1): If $\mathbb{Q}|_{\mathcal{G}} \not\ll \mathbb{P}|_{\mathcal{G}}$, then the claim is trivial. Assume now that $\mathbb{Q}|_{\mathcal{G}} \ll \mathbb{P}|_{\mathcal{G}}$. Then we have $\mathbb{Q}|_{\mathcal{H}} \ll \mathbb{P}|_{\mathcal{H}}$ and in particular

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{H}} \overset{\mathbb{P}\text{-a.s.}}{=} \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}} \,\bigg|\, \mathcal{H}\right] \tag{1.10}$$

Thus, using Jensen's inequality and the tower property of conditional expectation, we obtain

$$\begin{aligned}
H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{H}} = \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{H}}\right] &= \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{H}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{H}}\right] \\
&\overset{(1.10)}{=} \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}}\bigg|\mathcal{H}\right] \log \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}}\bigg|\mathcal{H}\right]\right] \\
&\leq \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}}\bigg|\mathcal{H}\right]\right] \\
&= \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}}\right] = \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{G}}\right] = H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{G}}
\end{aligned}$$

ad (2): We first assume that $\mathbb{Q} \ll \mathbb{P}$. Let ϕ be the function from Lemma 1.2, then for every $n \geq 0$ we have

$$H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_n} = \mathbb{E}_{\mathbb{P}}\left[\phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_n}\right)\right] = \mathbb{E}_{\mathbb{P}}\left[\phi\left(\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\bigg|\, \mathcal{F}_n\right]\right)\right]$$

By Lévy's Upwards Theorem, we have

$$\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\bigg|\, \mathcal{F}_n\right] \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\bigg|\, \mathcal{F}\right] = \frac{d\mathbb{Q}}{d\mathbb{P}}$$

and by Fatou's Lemma, we obtain

$$\liminf_{n\to\infty} H(\mathbb{Q}\,|\,\mathbb{P})|_{\mathcal{F}_n} \geq \mathbb{E}_{\mathbb{P}}\left[\liminf_{n\to\infty} \phi\left(\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\,\middle|\, \mathcal{F}_n\right]\right)\right] = \mathbb{E}_{\mathbb{P}}\left[\phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = H(\mathbb{Q}\,|\,\mathbb{P})$$

which proves the claim since we know from part (1) that $H(\mathbb{Q}\,|\,\mathbb{P})|_{\mathcal{F}_n} \leq H(\mathbb{Q}\,|\,\mathbb{P})$.

We now consider the case where $\mathbb{Q} \not\ll \mathbb{P}$. If $\mathbb{Q}|_{\mathcal{F}_N} \not\ll \mathbb{P}|_{\mathcal{F}_N}$ for some $N \geq 0$, then $\mathbb{Q}|_{\mathcal{F}_n} \not\ll \mathbb{P}|_{\mathcal{F}_n}$ for all $n \geq N$ and the claim holds trivially. So assume now that $\mathbb{Q}|_{\mathcal{F}_n} \ll \mathbb{P}|_{\mathcal{F}_n}$ for all $n \geq 0$. By Lemma A.1 (1) there exists an $\mathcal{F}$-measurable Z such that

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_n} \overset{\mathbb{Q}-\text{a.s.}}{\to} Z$$

and which satisfies $\mathbb{Q}(Z = \infty) > 0$ (by assumption of $\mathbb{Q} \not\ll \mathbb{P}$) and also $Z > 0$ $\mathbb{Q}$-a.s.. We thus obtain by Fatou's Lemma that

$$\liminf_{n\to\infty} H(\mathbb{Q}\,|\,\mathbb{P})|_{\mathcal{F}_n} = \liminf_{n\to\infty} \int_\Omega \log \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_n} d\mathbb{Q} \geq \int_\Omega \log Z \, d\mathbb{Q} = \infty = H(\mathbb{Q}\,|\,\mathbb{P})$$

which concludes the claim.

□

Lemma 1.5. Let $(\Omega, \mathcal{F})$ and $(\Lambda, \mathcal{G})$ be measurable spaces, $\mathbb{Q}, \mathbb{P}$ probability measures on $(\Omega, \mathcal{F})$ and $T : \Omega \to \Lambda$ measurable such that $\sigma(T) = \mathcal{F}$, then

$$H(\mathbb{Q}\,|\,\mathbb{P}) = H(\mathcal{L}_{\mathbb{Q}}T\,|\,\mathcal{L}_{\mathbb{P}}T)$$

Proof. If $\mathbb{Q} \not\ll \mathbb{P}$, then $\mathcal{L}_{\mathbb{Q}}T \not\ll \mathcal{L}_{\mathbb{P}}T$ and the equality holds trivially. So assume now $\mathbb{Q} \ll \mathbb{P}$, which implies $\mathcal{L}_{\mathbb{Q}}T \ll \mathcal{L}_{\mathbb{P}}T$. For every $B \in \mathcal{G}$ we have

$$\mathbb{Q}(T^{-1}(B)) = \mathcal{L}_{\mathbb{Q}}T(B) = \int_B \frac{d\mathcal{L}_{\mathbb{Q}}T}{d\mathcal{L}_{\mathbb{P}}T}\, d\mathcal{L}_{\mathbb{P}}T = \int_{T^{-1}(B)} \frac{d\mathcal{L}_{\mathbb{Q}}T}{d\mathcal{L}_{\mathbb{P}}T} \circ T \, d\mathbb{P} \tag{1.11}$$

If $\sigma(T) = \mathcal{F}$, then for every $A \in \mathcal{F}$ there exists a $B \in \mathcal{G}$ such that $A = T^{-1}(B)$. So, by (1.11) we have

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \overset{\mathbb{P}\text{-a.s.}}{=} \frac{d\mathcal{L}_{\mathbb{Q}}T}{d\mathcal{L}_{\mathbb{P}}T} \circ T$$

and thus

$$\begin{aligned} H(\mathbb{Q}\,|\,\mathbb{P}) = \int_\Omega \log \frac{d\mathbb{Q}}{d\mathbb{P}}\, d\mathbb{Q} &= \int_\Omega \log \frac{d\mathcal{L}_{\mathbb{Q}}T}{d\mathcal{L}_{\mathbb{P}}T} \circ T \, d\mathbb{Q} \\ &= \int_\Lambda \log \frac{d\mathcal{L}_{\mathbb{Q}}T}{d\mathcal{L}_{\mathbb{P}}T}\, d\mathcal{L}_{\mathbb{Q}}T = H(\mathcal{L}_{\mathbb{Q}}T\,|\,\mathcal{L}_{\mathbb{P}}T) \end{aligned}$$

□

Lemma 1.6. Let $\mathcal{M}_1(\Omega)$ denote the convex set of probability measures on a measurable space $(\Omega, \mathcal{F})$. Then for every $\mathbb{P} \in \mathcal{M}_1(\Omega)$, the function

$$\mathcal{M}_1(\Omega) \to [0, \infty], \qquad \mathbb{Q} \mapsto H(\mathbb{Q} \mid \mathbb{P})$$

is convex.

Proof. Let $\mathbb{P} \in \mathcal{M}_1(\Omega)$ be arbitrary. Further, let $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}_1(\Omega)$ and $\lambda_1, \lambda_2 \in [0,1]$ such that $\lambda_1 + \lambda_2 = 1$. If $\mathbb{Q}_1 \not\ll \mathbb{P}$ or $\mathbb{Q}_2 \not\ll \mathbb{P}$, then we trivially have

$$H(\lambda_1\mathbb{Q}_1 + \lambda_2\mathbb{Q}_2 \mid \mathbb{P}) \leq \lambda_1 H(\mathbb{Q}_1 \mid \mathbb{P}) + \lambda_2 H(\mathbb{Q}_2 \mid \mathbb{P})$$

Assume now that $\mathbb{Q}_1 \ll \mathbb{P}$ and $\mathbb{Q}_2 \ll \mathbb{P}$. This implies $\lambda_1\mathbb{Q}_1 + \lambda_2\mathbb{Q}_2 \ll \mathbb{P}$ and by the convexity of the function $x \mapsto x \cdot \log x$ we obtain

$$\begin{aligned}
H(\lambda_1\mathbb{Q}_1 + \lambda_2\mathbb{Q}_2 \mid \mathbb{P}) &= \mathbb{E}_{\mathbb{P}}\left[\frac{d(\lambda_1\mathbb{Q}_1 + \lambda_2\mathbb{Q}_2)}{d\mathbb{P}} \cdot \log \frac{d(\lambda_1\mathbb{Q}_1 + \lambda_2\mathbb{Q}_2)}{d\mathbb{P}}\right] \\
&\leq \lambda_1 \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}_1}{d\mathbb{P}} \cdot \log \frac{d\mathbb{Q}_1}{d\mathbb{P}}\right] + \lambda_2 \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}_2}{d\mathbb{P}} \cdot \log \frac{d\mathbb{Q}_2}{d\mathbb{P}}\right] \\
&= \lambda_1 H(\mathbb{Q}_1 \mid \mathbb{P}) + \lambda_2 H(\mathbb{Q}_2 \mid \mathbb{P})
\end{aligned}$$

$\square$

Section 2

Specific Relative Entropy

The setting throughout the rest of this text will now be as follows: We consider the *Wiener space* $(\mathcal{C}, \mathcal{F})$, consisting of

$$\mathcal{C} := C([0,1];\mathbb{R}) = \{f : [0,1] \to \mathbb{R} \text{ continuous}\}$$

and $\mathcal{F}$ the Borel-σ-algebra of $\mathcal{C}$ equipped with the topology of uniform convergence. On the Wiener space we define the *canonical process* $X = (X_t)_{t\in[0,1]}$ via

$$X_t(\omega) := \omega(t)$$

and we note that $\mathcal{F} = \sigma(X_t \mid t \in [0,1])$. Further, we assume a filtered probability space $(\Omega, \mathcal{S}, (\mathcal{S}_t)_{t\in[0,1]}, \mathbb{S})$ as given. On this space, we will consider different continuous processes (with time frame $[0,1]$) and interpret them as random variables taking values in the Wiener space. For every $x \in \mathbb{R}$, we let B^x be a Brownian motion starting at x and we denote by $\mathbb{W}^x$ its law on the Wiener space. For $x = 0$, we will just write $\mathbb{W}$. This measure is called the *Wiener measure*, however, we will also use this term to refer to the measures $\mathbb{W}^x$ where $x \neq 0$.

2.1 Dyadic Specific Relative Entropy

The notion of the *specific relative entropy* was introduced by N.Gantert in her doctoral thesis (see [12]), which dealt, amongst other things, with large deviations of the quadratic variation of Brownian motion. This introduction was motivated by the fact that to measure deviations in the quadratic variation of continuous processes, one requires a finer notion of distance for their laws than the relative entropy. Indeed, if the canonical process X has a different quadratic variation under probability measures $\mathbb{P}$ and $\mathbb{Q}$ on the Wiener space, then the measures are mutually singular and the relative entropy $H(\mathbb{Q} \mid \mathbb{P})$ (and also $H(\mathbb{P} \mid \mathbb{Q})$) is simply equal to ∞.

To define the specific relative entropy we now proceed as follows. Let $\mathfrak{D} = (d^{(k)})_{k\geq 0}$ denote the *dyadic partition sequence* of $[0,1]$, i.e. we define

$$d^{(k)} := \left(0, \tfrac{1}{2^k}, \ldots, \tfrac{2^k-1}{2^k}, 1\right)$$

For every $k \geq 0$, we define the σ-algebra generated by the partition $d^{(k)}$ as

$$\mathcal{F}_{d^{(k)}} := \sigma(X_0, X_{\frac{1}{2^k}}, ..., X_1) \subseteq \mathcal{F}$$

By Lemma 1.5, we have for arbitrary probability measures $\mathbb{Q}, \mathbb{P}$ on the Wiener space that

$$H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_{d^{(k)}}} = H(\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1] \mid \mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1])$$

Since $\mathfrak{D}$ is refining, we know that the sequence $(\mathcal{F}_{d^{(k)}})_{k \geq 0}$ of σ-algebras generated by $\mathfrak{D}$ is increasing. In particular, as the dyadic points are dense in the unit interval, the sequence increases to the σ-algebra $\mathcal{F}$ and we get by Lemma 1.4 (2)

$$H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_{d^{(k)}}} \nearrow H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}} = H(\mathbb{Q} \mid \mathbb{P}) = H(\mathcal{L}_{\mathbb{Q}} X \mid \mathcal{L}_{\mathbb{P}} X)$$

If $\mathbb{Q} \not\ll \mathbb{P}$, then $H(\mathbb{Q} \mid \mathbb{P}) = \infty$, however, all the relative entropies $H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_{d^{(k)}}}$ are likely finite as they only take the distribution of X under $\mathbb{Q}$ and $\mathbb{P}$ at the points of $d^{(k)}$ into account. To obtain the specific relative entropy, we now take the scaling limit of $H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_{d^{(k)}}}$ scaled by 2^{-k} (the mesh size of $d^{(k)}$) for $k \to \infty$.

Definition 6 (Dyadic Specific Relative Entropy)**.** Let $\mathbb{Q}, \mathbb{P}$ be probability measures on $(\mathcal{C}, \mathcal{F})$. The *dyadic specific relative entropy* of $\mathbb{Q}$ with respect to $\mathbb{P}$ is given by

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) := \overline{\lim_{k \to \infty}} \frac{1}{2^k} H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_{d^{(k)}}}$$

Remarks.

- We refer to the above quantity as the *dyadic* specific relative entropy since later on we will generalize the notion to partition sequences besides $\mathfrak{D}$. This prefix is not used outside of the context of this text!
- The use of the limes superior in the definition is only for the convenience of $h^{\mathfrak{D}}$ being always defined. However, the quantity is only of interest to us if it is the actual limit of its defining sequence.

As mentioned before, the specific relative entropy is supposed to be a finer notion of distance than the relative entropy. So given probability measures $\mathbb{Q}, \mathbb{P}$ on the Wiener space, how does $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P})$ relate to $H(\mathbb{Q} \mid \mathbb{P})$?

- If $H(\mathbb{Q} \mid \mathbb{P}) < \infty$, then $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) = 0$. Indeed, by Lemma 1.4 (1) we have

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) = \overline{\lim_{k \to \infty}} \frac{1}{2^k} H(\mathbb{Q} \mid \mathbb{P})|_{\mathcal{F}_{d^{(k)}}} \leq \overline{\lim_{k \to \infty}} \frac{1}{2^k} H(\mathbb{Q} \mid \mathbb{P}) = 0$$

  In particular, we always have $h^{\mathfrak{D}}(\mathbb{Q} | \mathbb{P}) \leq H(\mathbb{Q} | \mathbb{P})$. Recall that for the relative entropy we observed that $H(\mathbb{Q} \mid \mathbb{P}) = 0$ if and only if $\mathbb{Q} = \mathbb{P}$. The same is not true for the specific relative entropy! If $\mathbb{Q} \neq \mathbb{P}$, then the measures may still satisfy $H(\mathbb{Q} \mid \mathbb{P}) < \infty$ and thus $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) = 0$.

- If $H(\mathbb{Q}\,|\,\mathbb{P}) = \infty$, then $h^{\mathfrak{D}}(\mathbb{Q}\,|\,\mathbb{P})$ can take on any value from 0 to ∞, as we will see in Examples later on.

Let us look at some equivalent definitions of the dyadic specific relative entropy. Depending on the context, some of them might be more convenient to use than the original one.

Proposition 2.1. Let $\mathbb{Q}, \mathbb{P}$ be probability measures on $(\mathcal{C}, \mathcal{F})$, then $h^{\mathfrak{D}}(\mathbb{Q}\,|\,\mathbb{P})$ is equal to

(1) (*discretized form*)

$$\overline{\lim_{k\to\infty}} \frac{1}{2^k} H(\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1] \mid \mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1])$$

(2) (*increment form*)

$$\overline{\lim_{k\to\infty}} \frac{1}{2^k} H(\mathcal{L}_{\mathbb{Q}}[X_0, \Delta_1^{d^{(k)}} X, ..., \Delta_{2^k}^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[X_0, \Delta_1^{d^{(k)}} X, ..., \Delta_{2^k}^{d^{(k)}} X])$$

(3) (*conditional discretized form*)

$$\overline{\lim_{k\to\infty}} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega])\right]$$

under the assumption that $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) < \infty$.

(4) (*conditional increment form*)

$$\overline{\lim_{k\to\infty}} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X | \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X | \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega])\right]$$

under the assumption that $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) < \infty$.

Proof. ad (1): For every $k \geq 0$, the function $T_k := (X_0, X_{\frac{1}{2^k}}, ..., X_1)$ satisfies $\sigma(T_k) = \mathcal{F}_{d^{(k)}}$. So we may apply Lemma 1.5 to obtain

$$H(\mathbb{Q}\,|\,\mathbb{P})|_{\mathcal{F}_{d^{(k)}}} = H(\mathcal{L}_{\mathbb{Q}} T_k \mid \mathcal{L}_{\mathbb{P}} T_k) = H(\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1] \mid \mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1])$$

which we plug into the original definition to conclude the claim.

ad (2): For every $k \geq 0$, the function $S_k = (X_0, \Delta_1^{d^{(k)}} X, ..., \Delta_{2^k}^{d^{(k)}} X)$ satisfies $\sigma(S_k) = \mathcal{F}_{d^{(k)}}$. So we may apply Lemma 1.5 to obtain

$$\begin{aligned} H(\mathbb{Q}\,|\,\mathbb{P})|_{\mathcal{F}_{d^{(k)}}} &= H(\mathcal{L}_{\mathbb{Q}} S_k \mid \mathcal{L}_{\mathbb{P}} S_k) \\ &= H(\mathcal{L}_{\mathbb{Q}}[X_0, \Delta_1^{d^{(k)}} X, ..., \Delta_{2^k}^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[X_0, \Delta_1^{d^{(k)}} X, ..., \Delta_{2^k}^{d^{(k)}} X]) \end{aligned}$$

which we plug into the original definition to conclude the claim.

ad (3): Let $k \geq 0$ be arbitrary. By Corollary A.6, we have

$$\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1] \ll \mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]$$

if and only if $\mathcal{L}_{\mathbb{Q}}[X_0] \ll \mathcal{L}_{\mathbb{P}}[X_0]$ and for $\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]$-a.e. $\boldsymbol{x} = (x_0, x_1, ..., x_{2^k}) \in \mathbb{R}^{2^k+1}$

$$\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = (x_0, ..., x_{i-1})] \ll \mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = (x_0, ..., x_{i-1})], \quad \forall i \in [2^k]$$

in which case, we have

$$\frac{d\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]}{d\mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]}(\boldsymbol{x}) = \frac{d\mathcal{L}_{\mathbb{Q}}X_0}{d\mathcal{L}_{\mathbb{P}}X_0}(x_0) \cdot \prod_{i=1}^{2^k} \frac{d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = (x_0, ..., x_{i-1})]}{d\mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = (x_0, ..., x_{i-1})]}(x_i) \tag{2.1}$$

So under this assumption of absolute continuity, we obtain

$$\begin{aligned}
&H(\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1] \mid \mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]) \\
&= \int_{\mathbb{R}^{2^k+1}} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]}{d\mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1]}(\boldsymbol{x})\, d\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}} ..., X_1](\boldsymbol{x}) \\
&\overset{(2.1)}{=} \int_{\mathbb{R}^{2^k+1}} \log \frac{d\mathcal{L}_{\mathbb{Q}}X_0}{d\mathcal{L}_{\mathbb{P}}X_0}(x_0) \\
&\qquad + \sum_{i=1}^{2^k} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = (x_0, ..., x_{i-1})]}{d\mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = (x_0, ..., x_{i-1})]}(x_i)\, d\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1](\boldsymbol{x}) \\
&\overset{\substack{\text{Lem} \\ \text{A.5}}}{=} \int_{\mathbb{R}} \log \frac{d\mathcal{L}_{\mathbb{Q}}X_0}{d\mathcal{L}_{\mathbb{P}}X_0}\, d\mathcal{L}_{\mathbb{Q}}X_0 \\
&\qquad + \sum_{i=1}^{2^k} \int_{\mathbb{R}^i} \int_{\mathbb{R}} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \boldsymbol{x}]}{d\mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \boldsymbol{x}]}\, d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \boldsymbol{x}]\, d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}_{d_{i-1}^{(k)}}](\boldsymbol{x}) \\
&= H(\mathcal{L}_{\mathbb{Q}}X_0 \mid \mathcal{L}_{\mathbb{P}}X_0) \\
&\qquad + \sum_{i=1}^{2^k} \int_{\mathbb{R}^i} H(\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \boldsymbol{x}] \mid \mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \boldsymbol{x}])\, d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}_{d_{i-1}^{(k)}}](\boldsymbol{x}) \\
&= H(\mathcal{L}_{\mathbb{Q}}X_0 \mid \mathcal{L}_{\mathbb{P}}X_0) + \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} | \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega])\right]
\end{aligned}$$

(*Recall that, as introduced in Section 1, the expression* $\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega$ *is short-hand for* $(X_{d_0^{(k)}}, X_{d_1^{(k)}}, ..., X_{d_{i-1}^{(k)}}) = (\omega(d_0^{(k)}), \omega(d_1^{(k)}), ..., \omega(d_{i-1}^{(k)}))$.

Note that also without the assumption of absolute continuity, the first term of the above equality is equal to the last term (as both are ∞). We now plug the above into the discretized form derived in (1). If $H(\mathcal{L}_{\mathbb{Q}}X_0 \mid \mathcal{L}_{\mathbb{P}}X_0) < \infty$, then we have for the first summand

$$\frac{1}{2^k} H(\mathcal{L}_{\mathbb{Q}}X_0 \mid \mathcal{L}_{\mathbb{P}}X_0) \to 0$$

and we can remove it, yielding the conditional discretized form.

ad (4): We note that for every $\omega \in \mathcal{C}$, $k \geq 0$ and $i \in [2^n]$, if

$$\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \ll \mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}} \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]$$

then we have

$$\begin{aligned}
&H(\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}}|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}}|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]) \\
&\qquad = \int_{\mathbb{R}} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}}|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]}{d\mathcal{L}_{\mathbb{P}}[X_{d_i^{(k)}}|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]}(\nu)\, d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}}|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega](\nu) = \\
&\qquad = \int_{\mathbb{R}} \log \frac{d\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]}{d\mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]}(\nu - X_{d_{i-1}^{(k)}}(\omega))\, d\mathcal{L}_{\mathbb{Q}}[X_{d_i^{(k)}}|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega](\nu) \\
&\qquad = \int_{\mathbb{R}} \log \frac{d\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]}{d\mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]}(\nu)\, d\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega](\nu) \\
&\qquad = H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega])
\end{aligned}$$

We note that even without the assumption of absolute continuity, the first term of the above equality is trivially equal to the last term (as both are ∞). We now apply this equality to the conditional discretized form derived in (3) to conclude the claim.

□

Remark. The above proof shows that if the *starting point entropy* $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0)$ is finite, then it will contribute nothing to the specific relative entropy $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P})$. On the other hand, if $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) = \infty$, then necessarily $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) = \infty$. Going forward, the measures $\mathbb{Q}$ and $\mathbb{P}$ will often be assumed to be concentrated on $\{X_0 = x_0\}$ for some $x_0 \in \mathbb{R}$, so that $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) = 0$ and we can freely use the conditional discretized/increment form.

If the canonical process has independent increments under measures $\mathbb{Q}, \mathbb{P}$, then the specific relative entropy $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P})$ can be easily evaluated as the following Lemma shows.

Lemma 2.2. Let $\mathbb{Q}, \mathbb{P}$ be probability measures on $(\mathcal{C}, \mathcal{F})$ with $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) < \infty$.

(1) If X has independent increments under $\mathbb{P}$, then

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) \geq \overline{\lim_{k\to\infty}} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X])$$

(2) If X has independent increments under both $\mathbb{Q}$ and $\mathbb{P}$, then

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) = \overline{\lim_{k\to\infty}} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X])$$

Proof. In the following, we assume that

$$\mathcal{L}_{\mathbb{Q}}[X_0, X_{\frac{1}{2^k}}, ..., X_1] \ll \mathcal{L}_{\mathbb{P}}[X_0, X_{\frac{1}{2^k}}, ..., X_1], \quad \forall k \geq 0$$

as without this assumption both claims are trivially true.

ad (1): Since X has independent increments under $\mathbb{P}$, we have for every $k \geq 0$ and $\mathbb{Q}$-a.e. $\omega \in \mathcal{C}$ that

$$\mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] = \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X], \qquad \forall i \in [2^k] \tag{2.2}$$

Since the starting point entropy $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0)$ is finite, we may use the conditional increment form of $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P})$ and apply the convexity of the relative entropy (established in Lemma 1.6) to obtain

$$\begin{aligned}
h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) &= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega])\right] \\
&\overset{(2.2)}{=} \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X])\right] \\
&\overset{\substack{\text{Lem} \\ 1.6}}{\geq} \varlimsup_{k\to\infty} \frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathbb{E}_{\mathbb{Q}}^{\omega}[\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega]] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X]) \\
&= \varlimsup_{k\to\infty} \frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X])
\end{aligned}$$

ad (2): If X has independent increments under $\mathbb{P}$ and $\mathbb{Q}$, then additionally to (2.2), we also have for every $k \geq 0$ and $\mathbb{Q}$-a.e. $\omega \in \mathcal{C}$ that

$$\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X \mid \boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] = \mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X], \quad \forall i \in [2^k]$$

So we obtain

$$\begin{aligned}
h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) &= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X|\boldsymbol{X}_{d_{i-1}^{(k)}} = \omega])\right] \\
&= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X])\right] \\
&= \varlimsup_{k\to\infty} \frac{1}{2^k}\sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d^{(k)}} X])
\end{aligned}$$

□

With the help of this Lemma, we will now evaluate the specific relative entropy for a few simple examples. Examples 3, 4 and 6 are hereby taken from Gantert's doctoral thesis (see [12]).

Example 3. Let $b \in C([0,1];\mathbb{R})$ be arbitrary and denote $x_0 := b(0)$ and $\mathbb{Q} := \mathcal{L}_{\mathbb{S}}[B+b]$. The canonical process X has independent increments under $\mathbb{Q}$ and the Wiener measure $\mathbb{W}^{x_0}$, so we have by Lemma 2.2 (2) that

$$\begin{aligned} h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d(k)} X] \mid \mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{d(k)} X]) = \\ &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{N}[b(\tfrac{i}{2^k}) - b(\tfrac{i-1}{2^k}), \tfrac{1}{2^k}] \mid \mathcal{N}[0, \tfrac{1}{2^k}]) \\ &= \frac{1}{2} \varlimsup_{k\to\infty} \sum_{i=1}^{2^k} \Big(b(\tfrac{i}{2^k}) - b(\tfrac{i-1}{2^k})\Big)^2 \end{aligned}$$

where for the last step we used the formula for the relative entropy between Gaussian distributions which we established in Example 2. So under the assumption that the quadratic variation of b over $[0,1]$ along the dyadic partition sequence exists, we have

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \frac{[b]_1^{\mathfrak{D}}}{2}$$

Example 4. Given $\alpha \neq \beta > 0$, let $\mathbb{Q} := \mathcal{L}_{\mathbb{S}}[\sqrt{\alpha} B]$ and $\mathbb{P} := \mathcal{L}_{\mathbb{S}}[\sqrt{\beta} B]$. Using Lemma 2.2 (2) we can calculate

$$\begin{aligned} h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d(k)} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{d(k)} X]) = \\ &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{N}[0, \tfrac{\alpha}{2^k}] \mid \mathcal{N}[0, \tfrac{\beta}{2^k}]) = \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} \Upsilon\Big(\frac{\alpha}{\beta}\Big) = \Upsilon\Big(\frac{\alpha}{\beta}\Big) \end{aligned}$$

Example 5. Let $(\hat{B}_t)_{t\in[0,1]}$ be defined by $\hat{B}_t := B_{t^2}$ and $\mathbb{Q} := \mathcal{L}_{\mathbb{S}}\hat{B}$. By Lemma 2.2 (2) we obtain

$$\begin{aligned} h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}) &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d(k)} X] \mid \mathcal{L}_{\mathbb{W}}[\Delta_i^{d(k)} X]) \\ &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{N}[0, (\tfrac{i}{2^k})^2 - (\tfrac{i-1}{2^k})^2] \mid \mathcal{N}[0, \tfrac{1}{2^k}]) \\ &= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} \Upsilon(\tfrac{2i-1}{2^k}) = \int_0^1 \Upsilon(2s)\, ds = \frac{1-\log(2)}{2} \end{aligned}$$

Example 6. Let $T^H = (T^H_t)_{t\in[0,1]}$ be a fractional Brownian motion with parameter $H \in (0,1)$. That is, T^H is the unique Gaussian process that satisfies for all $s,t \in [0,1]$

$$\mathbb{E}_{\mathbb{S}}[T^H_t] = 0, \qquad \operatorname{Cov}(T^H_t, T^H_s) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$$

Note that for $H = \frac{1}{2}$, this is just a standard Brownian motion. For $H \neq \frac{1}{2}$, the process does not have independent increments, however, they are stationary with

$$T^H_t - T^H_s \sim_{\mathbb{S}} T^H_{t-s} \sim_{\mathbb{S}} \mathcal{N}[0, |t-s|^{2H}]$$

Let $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} T^H$ then we by Lemma 2.2 (1) we get

$$\begin{aligned}
h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}) &\geq \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{d^{(k)}} X] \mid \mathcal{L}_{\mathbb{W}}[\Delta_i^{d^{(k)}} X]) \\
&= \varlimsup_{k\to\infty} \frac{1}{2^k} \sum_{i=1}^{2^k} H(\mathcal{N}[0, \tfrac{1}{2^{2kH}}] \mid \mathcal{N}[0, \tfrac{1}{2^k}]) \\
&= \varlimsup_{k\to\infty} \Upsilon(2^{k(1-2H)}) = \begin{cases} \infty & , \text{ if } H \neq \frac{1}{2} \\ 0 & , \text{ if } H = \frac{1}{2} \end{cases}
\end{aligned}$$

2.2 Generalized Specific Relative Entropy

For the dyadic specific relative entropy $h^{\mathfrak{D}}$ we selected the dyadic partition sequence $\mathfrak{D}$ as our choice of subdivision. However, there is nothing that makes $\mathfrak{D}$ exceptional amongst all subdivisions that we could have chosen. We thus want to establish a definition of the specific relative entropy that allows us to choose partition sequences besides the dyadic one. Ideally, we would hope that every choice of partition sequence leads to the same specific relative entropy, allowing us to freely choose the subdivision that we situationally find most convenient to work with.

Let us start by outlining the partition sequences that we will allow. The relevant property of the dyadic partition sequence that we used was that the corresponding sequence of generated σ-algebras increases to $\mathcal{F}$. This is true by virtue of $\mathfrak{D}$ being refining with mesh size converging to 0. So these are properties we want to require of our admissible partition sequences. Furthermore, we want to take the entire unit interval into account, so every partition in an admissible sequence should contain the points 0 and 1 (In Section 1 we defined partitions of $[0,1]$ to include 0 and 1, so we do not mention this explicitly in the following definition.)

Definition 7 (Admissible Partition Sequence)**.** A partition sequence $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ of the interval $[0,1]$ is

- *admissible*, if it satisfies
  - $\mathfrak{P}$ is refining, i.e. $\pi^{(k+1)}$ refines $\pi^{(k)}$ for all $k \geq 0$
  - $\mathrm{mesh}(\pi^{(k)}) \to 0$
- *equidistant*, if $\Delta_i \pi^{(k)} = \mathrm{mesh}(\pi^{(k)})$ for all $k \geq 0$ and $i \in [\#\pi^{(k)}]$

Clearly, the dyadic partition sequence $\mathfrak{D}$ is admissible and equidistant. Let now $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ be an arbitrary admissible partition sequence. As in the dyadic case, we define for every $k \geq 0$ the sub-σ-algebra generated by $\pi^{(k)}$ via

$$\mathcal{F}_{\pi^{(k)}} := \sigma(X_i \mid i \in \pi^{(k)}) \subseteq \mathcal{F}$$

By our assumptions, these σ-algebras increase to $\mathcal{F}$ for $k \to \infty$. Suppose now that $\mathbb{Q}$ and $\mathbb{P}$ are probability measures on the Wiener space. If $\mathfrak{P}$ is equidistant, then, in accordance with Definition 6, it would be natural to define the specific relative entropy of $\mathbb{Q}$ w.r.t. $\mathbb{P}$ with subdivision $\mathfrak{P}$ via

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) := \varlimsup_{k \to \infty} \mathrm{mesh}(\pi^{(k)}) H(\mathbb{Q} \mid \mathbb{P})\big|_{\mathcal{F}_{\pi^{(k)}}} \tag{2.3}$$

With the obvious adaptations made, the equivalent definitions provided by Proposition 2.1 would then also be valid.

The case where $\mathfrak{P}$ is not equidistant is a bit more subtle. If we were to use (2.3) as our definition, then, written in the conditional increment form, $h^{\mathfrak{P}}$ would be given by

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) = \varlimsup_{k \to \infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\mathrm{mesh}(\pi^{(k)}) \sum_{i=1}^{\#\pi^{(k)}} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega])\right]$$

(under the assumption of $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) < \infty$). So the summand corresponding to the increment $\Delta_i^{\pi^{(k)}} X$ would be scaled by the mesh size of the partition $\pi^{(k)}$ regardless of how large the interval $[\pi_{i-1}^{(k)}, \pi_i^{(k)}]$ actually is. It will become quite apparent as we progress through the text that the more sensible definition is to scale each summand by the length of the corresponding interval instead of just the mesh size[1]. We are thus led to the following generalized definition of the specific relative entropy.

Definition 8 (Specific Relative Entropy)**.** Let $\mathbb{Q}, \mathbb{P}$ be probability measures on $(\mathcal{C}, \mathcal{F})$ and $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ admissible. The *specific relative entropy of* $\mathbb{Q}$ *w.r.t.* $\mathbb{P}$ *with subdivision* $\mathfrak{P}$ is given by

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) := \varlimsup_{k \to \infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega])\right]$$

if $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) < \infty$ and is set equal to ∞ if $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) = \infty$.

[1]We give credit to Daniel Lacker, who suggested this definition to the supervisor in private communication.

Remarks.

- For equidistant partition sequences, the above definition coincides with our suggested definition in (2.3). In particular, if we choose the dyadic partition sequence in the above definition, we obtain precisely the dyadic specific relative entropy as defined in Definition 6. For non-equidistant partition sequences, however, there is no equivalent definition of the form of (2.3), so we have to use the conditional increment form.
- Of course, we could have equivalently defined $h^{\mathfrak{P}}$ using the conditional discretized form. In this case we would just have to replace $\Delta_i^{\pi^{(k)}} X$ by $X_{\pi_i^{(k)}}$ in the definition. Throughout the text, we will use both forms.

The natural question that arises from our generalized definition is whether the choice of partition sequence actually matters or whether any partition sequence simply leads to the same quantity. To get a first idea, we will revisit our examples from before and see whether we obtain different specific relative entropies for different partition sequences. Before doing so, let us record the analogous statement to Lemma 2.2 which we require for the examples.

Lemma 2.3. Let $\mathbb{Q}, \mathbb{P}$ be probability measures on $(\mathcal{C}, \mathcal{F})$ such that $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0)$ is finite and $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ admissible.

(1) If X has independent increments under $\mathbb{P}$, then

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) \geq \varlimsup_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{\pi^{(k)}} X])$$

(2) If X has independent increments under both $\mathbb{Q}$ and $\mathbb{P}$, then

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) = \varlimsup_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{\pi^{(k)}} X])$$

Proof. Identical to Lemma 2.2. □

Example 3 continued. Let $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ be an admissible partition sequence, then by Lemma 2.3 (2) we obtain

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) &= \varlimsup_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X] \mid \mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}} X]) \\ &= \varlimsup_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{N}[b(\pi_i^{(k)}) - b(\pi_{i-1}^{(k)}), \Delta_i \pi^{(k)}] \mid \mathcal{N}[0, \Delta_i \pi^{(k)}]) \\ &= \varlimsup_{k \to \infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \frac{\left(b(\pi_i^{(k)}) - b(\pi_{i-1}^{(k)})\right)^2}{2 \Delta_i \pi^{(k)}} = \varlimsup_{n \to \infty} \frac{1}{2} \sum_{i=1}^{\#\pi^{(k)}} \left(b(\pi_i^{(k)}) - b(\pi_{i-1}^{(k)})\right)^2 \end{aligned}$$

So if for the function b the quadratic variation over $[0,1]$ along the partition sequence

$\mathfrak{P}$ exists, then

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) = \frac{[b]_1^{\mathfrak{P}}}{2}$$

Suppose b is a sample path of the Brownian motion B which satisfies (1.1), then, by choosing $\mathfrak{P}$ appropriately, we can make $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P})$ as small/large as we want.

Example 4 continued. Let $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ be an admissible partition sequence, then by Lemma 2.3 (2) we obtain

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) &= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X] \mid \mathcal{L}_{\mathbb{P}}[\Delta_i^{\pi^{(k)}} X]) \\ &= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{N}[0, \tfrac{\alpha}{\Delta_i \pi^{(k)}}] \mid \mathcal{N}[0, \tfrac{\beta}{\Delta_i \pi^{(k)}}]) \\ &= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \Upsilon\Big(\frac{\alpha}{\beta}\Big) = \Upsilon\Big(\frac{\alpha}{\beta}\Big) \end{aligned}$$

So $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P})$ coincides for all admissible $\mathfrak{P}$.

Example 5 continued. Let $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ be an admissible partition sequence, then by Lemma 2.3 (2) we obtain

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}) &= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X] \mid \mathcal{L}_{\mathbb{W}}[\Delta_i^{\pi^{(k)}} X]) \\ &= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{N}[0, (\pi_i^{(k)})^2 - (\pi_{i-1}^{(k)})^2] \mid \mathcal{N}[0, \Delta_i \pi^{(k)}]) \\ &= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \Upsilon(\pi_{i-1}^{(k)} + \pi_i^{(k)}) \\ &= \overline{\lim_{k\to\infty}} -\frac{1}{2} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log(\pi_i^{(k)} + \pi_{i-1}^{(k)}) = -\frac{1}{2} \int_0^1 \log(2x)\, dx = \frac{1 - \log(2)}{2} \end{aligned}$$

So $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W})$ coincides for all admissible $\mathfrak{P}$.

Example 6 continued. Let $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ be an admissible partition sequence, then by Lemma 2.3 (1) we obtain

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}) &\geq \underline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X] \mid \mathcal{L}_{\mathbb{W}}[\Delta_i^{\pi^{(k)}} X]) \\ &= \underline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{N}[0, (\Delta_i \pi^{(k)})^{2H}] \mid \mathcal{N}[0, \Delta_i \pi^{(k)}]) = ... \end{aligned}$$

$$... = \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \Upsilon((\Delta_i \pi^{(k)})^{2H-1}) = \begin{cases} \infty & , \text{ if } H \neq \frac{1}{2} \\ 0 & , \text{ if } H = \frac{1}{2} \end{cases}$$

So $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W})$ coincides for all admissible $\mathfrak{P}$.

So, as we can see from Example 3, different partition sequences can indeed lead to a different specific relative entropy. However, as the other examples suggest, this might not happen if we are working with laws of a 'nice' class of processes. The result we establish in the following section will hint at the fact that the class of square-integrable, continuous martingales might be such a nice class.

Before moving on, let us record a simple Lemma that we will use a few times throughout the remainder of this text.

Lemma 2.4. Let $\mathbb{Q}, \mathbb{P}$ be probability measures on $(\mathcal{C}, \mathcal{F})$ and $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ admissible. Under the assumption that

(a) $H(\mathcal{L}_{\mathbb{Q}} X_0 \mid \mathcal{L}_{\mathbb{P}} X_0) < \infty$

(b) $\mathcal{L}_{\mathbb{Q}}[X_{\pi_0^{(k)}}, ...X_{\pi_{\#\pi}^{(k)}}] \ll \mathcal{L}_{\mathbb{P}}[X_{\pi_0^{(k)}}, ...X_{\pi_{\#\pi}^{(k)}}]$ for every $k \geq 0$

we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) = \overline{\lim_{k\to\infty}} \, \mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{P}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]} \circ X_{\pi_i^{(k)}} \right]$$

Proof. Under the stated assumptions, we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) = \overline{\lim_{k\to\infty}} \, \mathbb{E}^{\omega}_{\mathbb{Q}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{P}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \right]$$

$$= \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \int_{\mathcal{C}} \int_{\mathbb{R}} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega]}{d\mathcal{L}_{\mathbb{P}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega]}(\nu) \, d\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega](\nu) \, d\mathbb{Q}(\omega)$$

$$= \overline{\lim_{k\to\infty}} \, \mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{P}}[X_{\pi_i^{(k)}} | \boldsymbol{X}_{\pi_{i-1}^{(k)}}]} \circ X_{\pi_i^{(k)}} \right]$$

Note that we need assumption (a) for the first equality (otherwise $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{P}) = \infty$ by definition). Meanwhile, assumption (b) implies $\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \ll \mathcal{L}_{\mathbb{P}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega]$ for $\mathbb{Q}$-a.e. $\omega \in \mathcal{C}$, making the second equality valid.

$\square$

2.3 Gantert Inequality

In the following, we let $\mathcal{M}_c^2$ denote the class of all square-integrable, continuous martingale measures on the Wiener space. That is, all the measures under which the canonical process is a square-integrable martingale. This class contains, for example, all the measures from Example 4 and 5 and the Wiener measure $\mathbb{W}^x$ for every $x \in \mathbb{R}$.

In this section we study the quantity $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0})$ where $\mathbb{Q} \in \mathcal{M}_c^2$ is assumed to be concentrated on $\{X_0 = x_0\}$ for some $x_0 \in \mathbb{R}$ (so that the starting point entropy is 0) and $\mathfrak{P}$ is an arbitrary admissible partition sequence. We will establish the *Gantert inequality* which provides us with a neat lower bound for this type of specific relative entropy in terms of the quadratic variation of X under the measure $\mathbb{Q}$.

So let $\mathbb{Q} \in \mathcal{M}_c^2$ be arbitrary such that $\mathbb{Q}$ is concentrated on $\{X_0 = x_0\}$ for some $x_0 \in \mathbb{R}$. By Theorem 1.1, the canonical process X has a quadratic variation process $[X]$ (under the measure $\mathbb{Q}$). Let $\nu_{[X]}$ denote the <u>random</u> measure on $([0,1], \mathcal{B})$ whose cumulative function is given by $[X]$. Note that this is well defined since $[X]$ is a continuous, non-decreasing process with $[X]_0 = 0$. Using Lebesgue decomposition, we can write

$$\nu_{[X]} = \nu_\perp + \nu_\ll$$

for some random measures $\nu_\perp$ and $\nu_\ll$, where $\nu_\perp$ is singular and $\nu_\ll$ absolutely continuous w.r.t. to the Lebesgue measure m. On $[0,1] \times \mathcal{C}$ we define the *previsible σ-algebra* $\mathcal{P}$ via

$$\mathcal{P} := \sigma(\{(t_1, t_2] \times A \mid t_1 < t_2,\ A \in \sigma(X_\tau \mid \tau \leq t_1)\})$$

The fact that $\nu_\ll \ll m$ implies that $(\nu_\ll \otimes \mathbb{Q})|_\mathcal{P} \ll (m \otimes \mathbb{Q})|_\mathcal{P}$ and we define

$$\sigma^2 := \frac{d\nu_\ll \otimes \mathbb{Q}}{dm \otimes \mathbb{Q}}\bigg|_\mathcal{P}$$

So for every $\omega \in \mathcal{C}$ and $t_1 < t_2 \in [0,1]$ we have

$$\nu_{[X](\omega)}((t_1, t_2]) = \nu_{\perp(\omega)}((t_1, t_2]) + \underbrace{\int_{t_1}^{t_2} \sigma^2(s, \omega)\, ds}_{=\nu_{\ll(\omega)}((t_1,t_2])}$$

It was proven by N. Gantert (see [12], Satz 1.3) that in the case of $\nu_{[X]}$ being absolutely continuous, i.e. $\nu_{[X]} = \nu_\ll$, one obtains the following lower bound for the dyadic specific relative entropy of $\mathbb{Q}$ with respect to $\mathbb{W}^{x_0}$

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \mathbb{E}_\mathbb{Q}\Big[\int_0^1 \Upsilon(\sigma^2(s, X))\, ds\Big]$$

Note that the above inequality is also true, but fairly uninteresting, if $\mathbb{Q}$ is not concentrated on $\{X_0 = x_0\}$, as in this case the starting point entropy is ∞ and thus $h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \infty$.

The inequality was later generalized by H. Föllmer (see [8], Thm. 17) to the case where $\nu_{[X]}$ is not necessarily absolutely continuous. In this case one obtains the following inequality

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \Upsilon(\sigma^2(s,X))\,ds\right]$$

We will refer to this inequality as the *Gantert inequality.* In the following Theorem, we will show that the Gantert inequality remains valid when we replace the dyadic partition sequence by any other admissible partition sequence. The proof we provide will be a combination of the proofs of Gantert and Föllmer adapted to our generalized definition of the specific relative entropy.

Theorem 2.5 (*Gantert Inequality*)**.** Let $\mathbb{Q} \in \mathcal{M}_c^2$ be concentrated on $\{X_0 = x_0\}$ for some $x_0 \in \mathbb{R}$, then for any admissible partition sequence $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \Upsilon(\sigma^2(s,X))\,ds\right]$$

Proof. **1.** For every $k \geq 0$ we define an auxilliary process $Y^{(k)} = (Y_t^{(k)})_{t\in[0,1]}$ as follows. For $t \in [0, \pi_1^{(k)}]$ we let

$$Y_t := B_t\sqrt{\frac{\mathbb{E}_{\mathbb{Q}}[(\Delta_1^{\pi^{(k)}}X)^2]}{\pi_1^{(k)}}}$$

and assuming that the process is defined up to time $\pi_{i-1}^{(k)}$, we define it for $t \in (\pi_{i-1}^{(k)}, \pi_i^{(k)}]$ via

$$Y_t = Y_{\pi_{i-1}^{(k)}} + (B_t - B_{\pi_{i-1}^{(k)}})\sqrt{\frac{\mathbb{E}_{\mathbb{Q}}[(\Delta_i^{\pi^{(k)}}X)^2 \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{\Delta_i\pi^{(k)}}}$$

We denote by $\mathbb{Q}^{(k)} := \mathcal{L}_{\mathbb{S}}Y^{(k)}$ the law of $Y^{(k)}$ on the Wiener space and we note that by construction this measure satisfies

$$\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}] = \mathcal{N}[0, \mathbb{E}_{\mathbb{Q}}[(\Delta_i^{\pi^{(k)}}X)^2 \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]], \quad \forall i \in [\#\pi^{(k)}] \tag{2.4}$$

Due to this property, we will refer to $\mathbb{Q}^{(k)}$ as the *Gaussian approximation measure* of $\mathbb{Q}$ with respect to the partition $\pi^{(k)}$.

2. Since the starting point entropy is 0, the specific relative entropy is given by

$$h^{\mathfrak{P}}(\mathbb{Q}\,|\,\mathbb{W}^{x_0}) = \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega]\,|\,\mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega])\right]$$

and we have

$$\mathbb{E}_{\mathbb{Q}}^{\omega}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega]\mid \mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega])\right] = \ldots$$

$$
\begin{aligned}
\dots &\overset{\text{Lem}\ 1.3}{=} \mathbb{E}^{\omega}_{\mathbb{Q}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]\mid \mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega])\Bigg]\\
&\quad + \mathbb{E}^{\omega}_{\mathbb{Q}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \int_{\mathbb{R}} \log\frac{d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]}{d\mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]}\, d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]\Bigg]\\
&\geq \mathbb{E}^{\omega}_{\mathbb{Q}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \int_{\mathbb{R}} \log\frac{d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]}{d\mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]}\, d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}=\omega]\Bigg]\\
&\overset{\text{Lem}\ 4.5}{=} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)}\mathbb{E}_{\mathbb{Q}}\Bigg[\log\frac{d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}}X|\boldsymbol{X}_{\pi_{i-1}^{(k)}}]}\circ \Delta_i^{\pi^{(k)}}X\Bigg]
\end{aligned}
$$

Note that we obtained the inequality, since the first summand of the left-hand side is non-negative. Our goal is to show that the right-hand side of the above converges to the right side of the Gantert inequality for $k\to\infty$.

3. For any $k\geq 0$, we consider on $[0,1]\times\mathcal{C}$ the σ-algebra

$$\mathcal{P}_{\pi^{(k)}} := \sigma(\{(t_1,t_2]\times A \mid t_1<t_2\in\pi^{(k)},\ A\in\sigma(X_\tau\mid \tau\leq t_1,\tau\in\pi^{(k)})\})\subseteq\mathcal{P}$$

We note that $(\nu_{[X]}\otimes\mathbb{Q})|_{\mathcal{P}_{\pi^{(k)}}}\ll(m\otimes\mathbb{Q})|_{\mathcal{P}_{\pi^{(k)}}}$ for every $k\geq 0$ (regardless of whether or not we have $\nu_{[X]}\ll m$) and we define

$$\sigma_k^2 := \frac{d\nu_{[X]}\otimes\mathbb{Q}}{dm\otimes\mathbb{Q}}\Big|_{\mathcal{P}_{\pi^{(k)}}}$$

For every $i\in[\#\pi^{(k)}]$ and $t\in(\pi_{i-1}^{(k)},\pi_i^{(k)}]$, the function σ_k^2 satisfies

$$\Delta_i\pi^{(k)}\cdot\sigma_k^2(t,X)=\mathbb{E}_{\mathbb{Q}}\Big[\nu_{[X]}((\pi_{i-1}^{(k)},\pi_i^{(k)}])\mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}\Big]=\mathbb{E}_{\mathbb{Q}}\Big[(\Delta_i^{\pi^{(k)}}X)^2\mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}\Big]\tag{2.5}$$

which we plug into (2.4) to obtain

$$\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X\mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]=\mathcal{N}(0,\Delta_i\pi^{(k)}\cdot\sigma_k^2(\pi_i^{(k)},X))\tag{2.6}$$

4. Using (2.5) and (2.6) we can show that

$$
\begin{aligned}
&\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)}\mathbb{E}_{\mathbb{Q}}\Bigg[\log\frac{d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}}X\mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}}X\mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}\circ \Delta_i^{\pi^{(k)}}X\Bigg]\\
&=\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)}\mathbb{E}_{\mathbb{Q}}\Bigg[\log\frac{N[0,\Delta_i\pi^{(k)}\cdot\sigma_k^2(\pi_i^{(k)},X)]}{N[0,\Delta_i\pi^{(k)}]}\circ\Delta_i^{\pi^{(k)}}X\Bigg]\\
&=\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)}\mathbb{E}_{\mathbb{Q}}\Bigg[\frac{(\Delta_i^{\pi^{(k)}}X)^2}{2\Delta_i\pi^{(k)}}-\frac{(\Delta_i^{\pi^{(k)}}X)^2}{2\Delta_i\pi^{(k)}\cdot\sigma_k^2(\pi_i^{(k)},X)}-\frac{1}{2}\log\sigma_k^2(\pi_i^{(k)},X)\Bigg]=\dots
\end{aligned}
$$

$$\begin{aligned}
\dots &= \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)}\mathbb{E}_{\mathbb{Q}}\Big[\frac{1}{2}\left(\sigma_k^2(\pi_i^{(k)},X) - 1 - \log(\sigma_k^2(\pi_i^{(k)},X))\right)\Big] \\
&= \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)}\mathbb{E}_{\mathbb{Q}}\Big[\Upsilon(\sigma_k^2(\pi_i^{(k)},X))\Big] \\
&= \mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 \Upsilon(\sigma_k^2(s,X))\,ds\Big]
\end{aligned}$$

We have thus shown that

$$h^{\mathfrak{P}}(\mathbb{Q}\mid\mathbb{W}^{x_0}) \geq \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 \Upsilon(\sigma_k^2(s,X))\,ds\Big]$$

5. To conclude, we note that since $\mathfrak{P}$ is refining with mesh size converging to 0, the σ-algebras $\mathcal{P}_{\pi^{(k)}}$ increase to $\mathcal{P}$. We thus have by A.1 (1) that

$$\sigma_k^2 = \frac{d\nu_{[X]}\otimes\mathbb{Q}}{dm\otimes\mathbb{Q}}\bigg|_{\mathcal{P}_{\pi^{(k)}}} \overset{m\otimes\mathbb{Q}\text{-a.s.}}{\to} \frac{d\nu_{\ll}\otimes\mathbb{Q}}{dm\otimes\mathbb{Q}}\bigg|_{\mathcal{P}} = \sigma^2$$

and by Lemma A.1 (2) we obtain

$$\lim_{k\to\infty} \mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 \Upsilon(\sigma_k^2(s,X))\,ds\Big] = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 \Upsilon(\sigma^2(s,X))\,ds\Big]$$

from which the claim follows.

□

Remarks.

- Note that we actually proved a slightly stronger statement than what is formulated. Namely, that the inequality still holds if we replace the limes superior in the definition of $h^{\mathfrak{P}}(\mathbb{Q}\mid\mathbb{W}^{x_0})$ by the limes inferior. So if we can show the opposite inequality

$$h^{\mathfrak{P}}(\mathbb{Q}\mid\mathbb{W}^{x_0}) \leq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 \Upsilon(\sigma^2(s,X))\,ds\Big]$$

  then we can conclude that the defining sequence of $h^{\mathfrak{P}}(\mathbb{Q}\mid\mathbb{W}^{x_0})$ actually converges to the right side of the Gantert inequality (instead of just its limes superior being equal to the right side).

- We may write the right side of the Gantert inequality in terms of relative entropy as follows

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 \Upsilon(\sigma^2(s,X))\,ds\Big] \\
&\qquad = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\Big[\nu_{\perp}([0,1]) + \int_0^1 \sigma^2(s,X)\,ds - 1 - \int_0^1 \log\sigma^2(s,X)\,ds\Big] \\
&\qquad = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\Big[\nu_{[X]}([0,1]) - 1 - H(m\mid\nu_{[X]})\Big]
\end{aligned}$$

- We see immediately that $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) < \infty$ implies $\sigma^2 > 0$ $m \otimes \mathbb{Q}$-a.s.
- The right side of the Gantert inequality is the same regardless of what partition sequence we are considering!

Let us take a closer look at the proof of the Gantert inequality. We introduced the Gaussian approximation measures $\mathbb{Q}^{(k)}$ under which each increment $\Delta_i^{\pi^{(k)}} X$ conditioned on the preceeding values of X at times $\pi_0^{(k)}, ..., \pi_{i-1}^{(k)}$ has Gaussian distribution with the same mean and variance as it would have under the measure $\mathbb{Q}$. We then decomposed the quantity

$$\mathbb{E}_{\mathbb{Q}}^{\omega}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{W}^{x_0}}[\Delta_i^{\pi^{(k)}} X \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega]\right]$$

into

$$T_{\pi^{(k)}}^{\mathbb{Q}} + G_{\pi^{(k)}}^{\mathbb{Q}}$$

where $T_{\pi^{(k)}}^{\mathbb{Q}}$ is the *non-Gaussian part*

$$T_{\pi^{(k)}}^{\mathbb{Q}} := \mathbb{E}_{\mathbb{Q}}^{\omega}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} H(\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega] \mid \mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}} = \omega])\right]$$

which we know to be non-negative, and $G_{\pi^{(k)}}^{\mathbb{Q}}$ is the *Gaussian part*

$$G_{\pi^{(k)}}^{\mathbb{Q}} := \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathcal{L}_{\mathbb{Q}^{(k)}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{W}^{x}}[\Delta_i^{\pi^{(k)}} X | \boldsymbol{X}_{\pi_{i-1}^{(k)}}]} \circ \Delta_i^{\pi^{(k)}} X\right]$$

We then showed that the Gaussian part converges for $k \to \infty$ to the right side of the Gantert inequality to conclude that

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}) = \varlimsup_{k\to\infty} T_{\pi^{(k)}}^{\mathbb{Q}} + \frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \Upsilon(\sigma^2(s,X))\, ds\right] \tag{2.7}$$

So we have equality in the Gantert inequality if and only if $\mathbb{Q}$ is 'asymptotically Gaussian', in the sense that for $k \to \infty$, i.e. by considering increasingly finer partitions of the unit interval, the non-Gaussian part $T_{\pi^{(k)}}^{\mathbb{Q}}$ converges to 0.

Example 4 continued. The quadratic variation of X under $\mathbb{Q}$ is given by $[X]_t = \alpha t = \int_0^t \alpha\, ds$ and we already know that upon setting $\beta = 1$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}) = \Upsilon(\alpha) = \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \Upsilon(\alpha)\, ds\right]$$

So for this example, we have equality in the Gantert inequality for every admissible $\mathfrak{P}$.

Example 5 continued. The quadratic variation of X under $\mathbb{Q}$ is given by $[X]_t = t^2 = \int_0^t 2s\,ds$ and we know that

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}) = \frac{1-\log(2)}{2} = \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \Upsilon(2s)\,ds\right]$$

So for this example, we have equality in the Gantert inequality for every admissible $\mathfrak{P}$.

In the above examples, the lower bound for the specific relative entropy that the Gantert inequality provides is in fact optimal. This is fairly unsurprising, as in both examples the Gaussian approximation measures simply coincide with the measure $\mathbb{Q}$ and thus the non-Gaussian part is 0 for every partition.

It was however conjectured by Gantert that one would obtain equality for every measure in $\mathcal{M}_c^2$. Let us state this conjecture adapted to our generalized definition.

Gantert Conjecture

For every $\mathbb{Q} \in \mathcal{M}_c^2$ concentrated on $\{X_0 = x_0\}$ for some $x_0 \in \mathbb{R}$ and $\mathfrak{P}$ admissible, we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}}[\nu_{\perp}([0,1])] + \mathbb{E}_{\mathbb{Q}}\left[\int_0^1 \Upsilon(\sigma^2(s,X))\,ds\right]$$

where $\nu_{\perp}$ and σ^2 are as defined in the beginning of this section. In particular, the quantity $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0})$ coincides for all admissible $\mathfrak{P}$.

Over the past years, this conjecture has been verified (in the dyadic case) for a few useful subclasses of $\mathcal{M}_c^2$. In Section 3, we will show that for these subclasses, the conjecture remains valid under our general definition of the specific relative entropy. We will also establish its validity for a new class. However, we will also present a counterexample to the conjecture in Section 3.2. Thus, while the conjecture does not hold in the generality we stated it, it appears to remain true for a substantial subclass of measures in $\mathcal{M}_c^2$, and characterizing this class remains an open problem.

Section 3

Martingale Measures with Gantert Equality

3.1 Deterministic Time-Transformations of Brownian Motion

We begin with the simple class of deterministic time-transformations of Brownian motion. Suppose $a : [0,1] \to [0,1]$ is continuous and increasing with $a(0) = 0$. We consider for some starting point $x_0 \in \mathbb{R}$ the Brownian motion B^{x_0} time-transformed by a. That is, we consider the process $\hat{B}^{x_0} = (\hat{B}^{x_0}_t)_{t \in [0,1]}$ defined by

$$\hat{B}^{x_0}_t := B^{x_0}_{a(t)}$$

and we let $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} \hat{B}^{x_0}$. The quadratic variation of the canonical process X under the measure $\mathbb{Q}$ is then given by $[X]_t = a(t)$. Suppose $\nu_\perp$ and σ^2 are such that for all $t_1 < t_2 \in [0,1]$

$$a(t_2) - a(t_1) = \nu_\perp([t_1, t_2]) + \int_{t_1}^{t_2} \sigma(s)^2 \, ds$$

then the Gantert inequality tells us that for any admissible partition sequence $\mathfrak{P}$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \frac{1}{2} \nu_\perp([0,1]) + \int_0^1 \Upsilon(\sigma^2(s)) \, ds$$

It was shown by Gantert (see [12], Chapter 1) that in the dyadic case and under the assumption that a is absolutely continuous, equality holds in the above. We now show that this is also true in the general setting. Given the observations that we made at the end of the previous section, this is very straight-forward.

Proposition 3.1. Let $\mathbb{Q}$ be as above and $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ admissible, then

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \frac{1}{2} \nu_\perp([0,1]) + \int_0^1 \Upsilon(\sigma^2(s)) \, ds$$

Proof. We observe that for every $k \geq 0$ and every $i \in [\#\pi^{(k)}]$ we have

$$\mathcal{L}_{\mathbb{Q}}[\Delta_i^{\pi^{(k)}} X \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}] = \mathcal{N}[0, a(\pi_i^{(k)}) - a(\pi_{i-1}^{(k)})] = \mathcal{N}[0, \mathbb{E}_{\mathbb{Q}}[(\Delta_i^{\pi^{(k)}} X)^2 \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]]$$

So for every $k \geq 0$, $\mathbb{Q}$ coincides with the Gaussian approximation measure $\mathbb{Q}^{(k)}$. In particular, for every $k \geq 0$, the non-Gaussian part $T^{\mathbb{Q}}_{\pi^{(k)}}$ is equal to 0 and we can conclude from (2.7). $\square$

3.2 Martingale Diffusions

In this section, we verify the Gantert conjecture for two classes of Martingale diffusions. For the first class, we consider solutions of stochastic differential equations of the form

$$\begin{cases} dM_t = \sigma(M_t)\, dB_t \\ M_0 = x_0 \end{cases} \tag{$\star$}$$

where $x_0 \in \mathbb{R}$ is arbitrary and the diffusion coefficient σ is assumed to satisfy

(D1) $\sigma \in C^2(\mathbb{R};\mathbb{R})$

(D2) $\sigma \in (\delta, \frac{1}{\delta})$ for some $\delta \in (0,1]$

(D3) $\|\sigma'\|_\infty, \|\sigma''\|_\infty < L$ for some $L \geq 0$

Note that these assumptions guarantee that ($\star$) has a unique strong solution $M = (M_t)_{t\in[0,1]}$ and that for M there exists a transition probability density (see Appendix A.4 for details). Furthermore, M is a square-integrable, continuous martingale with absolutely continuous quadratic variation given by

$$[M]_t = \int_0^t \sigma(M_s)^2\, ds$$

Thus, upon denoting $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} M$, the Gantert inequality tells us that for any admissible partition sequence $\mathfrak{P}$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(\sigma(M_s)^2)\, ds\right] \tag{3.1}$$

It was proven by J. Backhoff-Veraguas and C. Unterberger (see [6], Theorem 1.4) that equality holds in the above for the dyadic case. Following their proof, we will show that this is the case for every admissible partition sequence.

Remark. The statement of ([6], Theorem 1.4) is actually more general than what we just mentioned. Namely, if N is the solution of

$$\begin{cases} dN_t = \eta(N_t)\, dB_t \\ N_0 = x_0 \end{cases}$$

where η also satisfies (D1)-(D3), then upon denoting $\mathbb{P} := \mathcal{L}_{\mathbb{S}} N$ one obtains

$$h^{\mathfrak{D}}(\mathbb{Q} \mid \mathbb{P}) = \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon\left(\frac{\sigma(M_s)^2}{\eta(M_s)^2}\right) ds\right]$$

However, in the following, we will restrict ourselves to the case of $\mathbb{P} = \mathbb{W}^{x_0}$, i.e. where $\eta \equiv 1$.

The main ingredient to prove equality in (3.1) is the following precise estimate for the transition probability density of M.

Lemma 3.2. Let M be the solution of $(\star)$, then there exist a constant $C > 0$ such that

$$p[M_t \mid M_s = x](y) \leq e^{Ct} \frac{1}{\sqrt{2\pi(t-s)}} \sqrt{\frac{\sigma(x)}{\sigma(y)}} \frac{1}{\sigma(y)} e^{-\frac{d_\sigma(x,y)^2}{2(t-s)}}$$

where $d_\sigma(x,y) := \int_x^y \frac{1}{\sigma(s)}\, ds$.

Proof. see ([6], Lemma 2.1)

□

Theorem 3.3. Let M be the solution of $(\star)$ and $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} M$, then for any admissable $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(\sigma(M_s)^2)\, ds\right]$$

Proof. By Lemma 2.4, we know that the specific relative entropy is given by

$$\begin{aligned}
h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) &\overset{\text{Lem 2.4}}{=} \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{W}^{x_0}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]} \circ X_{\pi_i^{(k)}}\right] \\
&= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{p[M_{\pi_i^{(k)}} \mid \boldsymbol{M}_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i \pi^{(k)}] \circ M_{\pi_i^{(k)}}}\right] \\
&= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i \pi^{(k)}] \circ M_{\pi_i^{(k)}}}\right]
\end{aligned}$$

Note that for the last equality, we used the fact that M is Markovian, so the conditional density of $M_{\pi_i^{(k)}}$ given $\boldsymbol{M}_{\pi_{i-1}^{(k)}}$ coincides with the conditional density of $M_{\pi_i^{(k)}}$ given $M_{\pi_{i-1}^{(k)}}$.

Using Lemma 3.2, we obtain the following bound for the logarithm in the RHS of the above

$$\begin{aligned}
\log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i \pi^{(k)}] \circ M_{\pi_i^{(k)}}} \leq & C\Delta_i \pi^{(k)} + \frac{1}{2} \log \frac{\sigma(M_{\pi_{i-1}^{(k)}})}{\sigma(M_{\pi_i^{(k)}})} + \frac{(\Delta_i^{\pi^{(k)}} M)^2}{2\Delta_i \pi^{(k)}} \\
& - \frac{d_\sigma(M_{\pi_{i-1}^{(k)}}, M_{\pi_i^{(k)}})^2}{2\Delta_i \pi^{(k)}} - \log \sigma(M_{\pi_i^{(k)}})
\end{aligned}$$

We thus obtain

$$\begin{aligned}
&\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i \pi^{(k)}] \circ M_{\pi_i^{(k)}}}\right] \\
&\quad \leq \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} C(\Delta_i \pi^{(k)})^2\right] + \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{\sigma(M_{\pi_{i-1}^{(k)}})}{\sigma(M_{\pi_i^{(k)}})}\right] + \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} M)^2\right] \\
&\qquad - \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{k=1}^{\#\pi^{(k)}} d_\sigma(M_{\pi_{i-1}^{(k)}}, M_{\pi_i^{(k)}})^2\right] - \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \sigma(M_{\pi_i^{(k)}})\right]
\end{aligned}$$

and we shall proceed by checking that the limes superior for $k \to \infty$ of the right side of the above is less or equal to the right side of (3.1). For the first summand, we have

$$\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} C(\Delta_i \pi^{(k)})^2\right] \leq C \cdot \mathrm{mesh}(\pi^{(k)}) \to 0$$

For the second summand, we observe that the function $t \mapsto \mathbb{E}_{\mathbb{S}}[\log \sigma(M_t)]$ is uniformly continuous by virtue of σ being bounded away from 0 and ∞ and thus

$$\frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{\sigma(M_{\pi_{i-1}^{(k)}})}{\sigma(M_{\pi_i^{(k)}})}\right] \leq \frac{1}{2} \max_{i \in [\#\pi^{(k)}]} \left\{\mathbb{E}_{\mathbb{S}}\big[\log \sigma(M_{\pi_{i-1}^{(k)}})\big] - \mathbb{E}_{\mathbb{S}}\big[\log \sigma(M_{\pi_i^{(k)}})\big]\right\} \to 0$$

For the third summand we have

$$\begin{aligned}\frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} M)^2\right] &= \frac{1}{2}\sum_{i=1}^{\#\pi^{(k)}} \mathbb{E}_{\mathbb{S}}\left[\left(\int_{\pi_{i-1}^{(k)}}^{\pi_i^{(k)}} \sigma(M_s)\, dB_s\right)^2\right] \\ &= \frac{1}{2}\sum_{i=1}^{\#\pi^{(k)}} \mathbb{E}_{\mathbb{S}}\left[\int_{\pi_{i-1}^{(k)}}^{\pi_i^{(k)}} \sigma(M_s)^2\, ds\right] = \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 \sigma(M_s)^2\, ds\right]\end{aligned}$$

For the fourth summand, we apply Itō's formula to the function $F(z) = \int_0^z \frac{1}{\sigma(s)}\, ds$ to obtain

$$F(M_t) = \int_0^t \frac{1}{\sigma(M_s)}\, dM_s - \frac{1}{2}\int_0^t \frac{\sigma'(M_s)}{\sigma(M_s)^2}\, d[M]_s = \int_0^t 1\, dB_s - \frac{1}{2}\int_0^t \sigma'(M_s)\, ds \tag{3.2}$$

which tells us that the quadratic variation of the process $F(M)$ is given by $[F(M)]_t = t$ and we conclude that

$$\frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{k=1}^{\#\pi^{(k)}} d_\sigma(M_{\pi_{i-1}^{(k)}}, M_{\pi_i^{(k)}})^2\right] = \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\sum_{k=1}^{\#\pi^{(k)}} \Big(F(M_{\pi_i^{(k)}}) - F(M_{\pi_{i-1}^{(k)}})\Big)^2\right] \to \frac{1}{2}\mathbb{E}_{\mathbb{S}}[[F(M)]_1] = \frac{1}{2}$$

For the fifth summand, we note that by continuity of $t \mapsto \log \sigma(M_t)$ we have

$$\sum_{i=1}^{\#\pi^{(k)}} \log \sigma(M_{\pi_i^{(k)}}) 1_{(\pi_{i-1}^{(k)}, \pi_i^{(k)}]} \overset{m \otimes \mathbb{Q}\text{-a.s.}}{\to} \log \sigma(M_{\pi_i^{(k)}})$$

Now since the left-hand side is uniformly bounded (since σ is bounded), we obtain by the Dominated Convergence Theorem that

$$\begin{aligned}\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \sigma(M_{\pi_i^{(k)}})\right] &= \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \sum_{i=1}^{\#\pi^{(k)}} \log \sigma(M_{\pi_i^{(k)}}) 1_{(\pi_{i-1}^{(k)}, \pi_i^{(k)}]}(s)\, ds\right] \\ &\to \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \log \sigma(M_s)\, ds\right] = \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 \log \sigma(M_s)^2\, ds\right]\end{aligned}$$

In total we have thus shown that

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) &= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i\pi^{(k)}] \circ M_{\pi_i^{(k)}}}\right] \\ &\leq \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 \sigma(M_s)^2\, ds\right] - \frac{1}{2} - \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 \log \sigma(M_s)^2\, ds\right] = \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(\sigma(M_s)^2)\, ds\right] \end{aligned}$$

Since the Gantert inequality provides us with the opposite inequality, the result follows. □

For the second class of martingale diffusions for which we want to verify the Gantert conjecture, we look at solutions of SDEs of the form

$$\begin{cases} dM_t = \sigma(t, M_t)\, dB_t \\ M_0 = x_0 \end{cases} \tag{$\diamond$}$$

where $x_0 \in \mathbb{R}$ is arbitrary and the diffusion coefficient σ is assumed to satisfy

(T1) $\sigma \in C^2([0,1] \times \mathbb{R}; \mathbb{R})$

(T2) $\sigma \in (\delta, \frac{1}{\delta})$ for some $\delta \in (0,1]$

(T3) there is a constant $L > 0$ such that

$$|\sigma(s,x) - \sigma(t,y)| < L(|t-s| + |y-x|), \quad \forall x, y \in \mathbb{R},\ s < t \in [0,1]$$

As for the previous class we considered, the assumptions (T1)-(T3) guarantee the existence of a strong solution $M = (M_t)_{t\in[0,1]}$ and that M has a transition probability density. Further, M is a square-integrable martingale with a quadratic variation given by

$$[M]_t = \int_0^t \sigma(s, M_s)^2\, ds$$

Thus, upon defining $\mathbb{Q} := \mathcal{L}_{\mathbb{S}}M$, the Gantert inequality tells us that for every admissible partition sequence $\mathfrak{P}$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(\sigma(s, M_s)^2)\, ds\right]$$

It was proven by J.-D. Benamou et al. (see [7], Proposition 2) that equality holds in the above in the dyadic case. Taking inspiration from their proof, we will show that equality holds for every admissible partition sequence with an argument that has a closer resemblance to the proof we gave for Theorem 3.3.

To prove the claim, we will use the following two Lemmas regarding the transition probability density of M.

Lemma 3.4. Let M be the solution of ($\diamond$). For every compact set $K \subseteq \mathbb{R}$, there exists a $\tau \geq 0$ such that for all $x, y \in K,\ s \leq t \in [0,1]$ with $|x-y|, |t-s| < \tau$ we have

$$p[M_t \mid M_s = x](y) = \frac{1}{\sqrt{2\pi(t-s)\sigma(s,x)^2}} e^{-\frac{(y-x)^2}{2(t-s)\sigma(s,x)^2}} (1 + R(s,t,x,y)(t-s))$$

where R is a bounded function on the set of all (s,t,x,y) that satisfy the above.

Proof. see ([2], Theorem 1.2.)

□

Lemma 3.5. Let M be the solution of ($\diamond$), then there are constants $c_{\pm}, C_{\pm} > 0$ such that for all $s \leq t \in [0,1]$ and $x, y \in \mathbb{R}$

$$\frac{C_-}{\sqrt{t-s}} e^{-c_- \frac{|y-x|^2}{t-s}} \leq p[M_t \mid M_s = x](y) \leq \frac{C_+}{\sqrt{t-s}} e^{-c_+ \frac{|y-x|^2}{t-s}}$$

Proof. see ([1], Theorem 1)

□

Theorem 3.6. Let M be the solution of ($\diamond$) and $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} M$, then for any admissible $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \mathbb{E}_{\mathbb{S}} \left[\int_0^1 \Upsilon(\sigma(s, M_s)^2)\, ds \right]$$

Proof. **1.** As in the proof of Theorem 3.3, we know that the specific relative entropy is given by

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \varlimsup_{k \to \infty} \mathbb{E}_{\mathbb{S}} \left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i \pi^{(k)}] \circ M_{\pi_i^{(k)}}} \right] \tag{3.3}$$

Let $\varepsilon > 0$ be arbitrary. By Lemma 3.4 we find a $\tau > 0$ such that for all $x, y \in [-\frac{2}{\varepsilon}, \frac{2}{\varepsilon}]$ and $s \leq t \in [0,1]$ with $|y-x|, |t-s| \leq \tau$ we have

$$p[M_t \mid M_s = x](y) = \underbrace{\frac{1}{\sqrt{2\pi(t-s)\sigma(s,x)^2}} e^{-\frac{(y-x)^2}{2(t-s)\sigma(s,x)^2}}}_{=:q(s,t,x,y)} (1 + R(s,t,x,y)(t-s)) \tag{3.4}$$

for some function R which is bounded on the set of all (s,t,x,y) that satisfy the above. W.l.o.g. we shall assume that $\tau \leq \frac{1}{\varepsilon}$. For every $k \geq 0$ and $i \in [\#\pi^{(k)}]$, we want to split the expectation

$$\mathbb{E}_{\mathbb{S}} \left[\log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i \pi^{(k)}] \circ M_{\pi_i^{(k)}}} \right]$$

into two expectations over the sets

- $A_{i,k} := \{|M_{\pi^{(k)}_{i-1}}| \leq \frac{1}{\varepsilon},\ |\Delta_i^{\pi^{(k)}} M| \leq \tau\}$
- $B_{i,k} := \{|M_{\pi^{(k)}_{i-1}}| > \frac{1}{\varepsilon},\ |\Delta_i^{\pi^{(k)}} M| \leq \tau\} \cup \{|\Delta_i^{\pi^{(k)}} M| > \tau\}$

Since we are interested in $k \to \infty$, we may w.l.o.g. assume that $\Delta_i \pi^{(k)} \leq \tau$, so that by construction, if $\omega \in A_{i,k}$, then (3.4) is valid for $p[M_{\pi^{(k)}_i} \mid M_{\pi^{(k)}_{i-1}} = M_{\pi^{(k)}_{i-1}}(\omega)](M_{\pi^{(k)}_i}(\omega))$.

2. As an auxiliary result, let us note that for every $k \geq 0$ we have

$$\overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \sqrt{\mathbb{S}[B_{i,k}]} \leq 2\varepsilon \|M_1\|_{L^2(\mathbb{S})} \tag{3.5}$$

Indeed, we have that

$$\begin{aligned}
\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \sqrt{\mathbb{S}[B_{i,k}]} &\leq \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \sqrt{\mathbb{S}[|\Delta_i^{\pi^{(k)}} M| > \tau] + \mathbb{S}\left[\max_{0\leq t\leq 1} |M_t| > \tfrac{1}{\varepsilon}\right]} \\
&\leq \max_{i\in[\#\pi^{(k)}]} \sqrt{\mathbb{S}[|\Delta_i^{\pi^{(k)}} M| > \tau]} + \sqrt{\mathbb{S}\left[\max_{0\leq t\leq 1} |M_t| > \tfrac{1}{\varepsilon}\right]} \\
&\leq \max_{i\in[\#\pi^{(k)}]} \sqrt{\mathbb{S}[|\Delta_i^{\pi^{(k)}} M| > \tau]} + 2\varepsilon \|M_1\|_{L^2(\mathbb{S})} \overset{k\to\infty}{\to} 2\varepsilon \|M_1\|_{L^2(\mathbb{S})}
\end{aligned}$$

where in the last step we used Lemma 3.5 to conclude that

$$\begin{aligned}
\mathbb{S}[|\Delta_i^{\pi^{(k)}} M| > \tau] &\overset{\substack{\text{Lem} \\ 3.5}}{\leq} 2 \int_\tau^\infty \frac{C_+}{\sqrt{\Delta_i \pi^{(k)}}} e^{-c_+ \frac{s^2}{\Delta_i \pi^{(k)}}} \, ds \\
&= 2C_+ \sqrt{\frac{\pi}{c_+}} \cdot \mathcal{N}[0, \tfrac{\Delta_i \pi^{(k)}}{2c_+}]([\tau, \infty)) \\
&\overset{\substack{\text{Lem} \\ A.3}}{\leq} \frac{C_+ \sqrt{\Delta_i \pi^{(k)}}}{c_+ \tau} \exp\left(-\frac{c_+ \tau^2}{\Delta_i \pi^{(k)}}\right) \leq \frac{C_+}{c_+ \tau} \exp\left(-\frac{c_+ \tau^2}{\operatorname{mesh}(\pi^{(k)})}\right) \to 0
\end{aligned}$$

3. For the expectation on the right-hand side of (3.3) taken over the A-sets, we obtain

$$\begin{aligned}
&\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{p[M_{\pi^{(k)}_i} \mid M_{\pi^{(k)}_{i-1}}] \circ M_{\pi^{(k)}_i}}{N[M_{\pi^{(k)}_{i-1}}, \Delta_i \pi^{(k)}] \circ M_{\pi^{(k)}_i}} \cdot 1_{A_{i,k}}\right] \\
&\quad \overset{(3.4)}{=} \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \frac{q(\pi^{(k)}_{i-1}, \pi^{(k)}_i, M_{\pi^{(k)}_{i-1}}, M_{\pi^{(k)}_i})}{N[M_{\pi^{(k)}_{i-1}}, \Delta_i \pi^{(k)}] \circ M_{\pi^{(k)}_i}} \cdot 1_{A_{i,k}}\right] \\
&\qquad + \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log(1 + \Delta_i \pi^{(k)} R(\pi^{(k)}_{i-1}, \pi^{(k)}_i, M_{\pi^{(k)}_{i-1}}, M_{\pi^{(k)}_i})) \cdot 1_{A_{i,k}}\right] = \ldots
\end{aligned}$$

$$\begin{aligned}
... = \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} M)^2 \cdot 1_{A_{i,k}}\Bigg] &- \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \frac{(\Delta_i^{\pi^{(k)}} M)^2}{\sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})^2} \cdot 1_{A_{i,k}}\Bigg] \\
&- \mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}}) \cdot 1_{A_{i,k}}\Bigg] \\
&+ \mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log(1 + \Delta_i \pi^{(k)} R(\pi_i^{(k)}, \pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}}, M_{\pi_i^{(k)}})) \cdot 1_{A_{i,k}}\Bigg]
\end{aligned}$$

and we now want to find a good upper bound for the limes superior for $k \to \infty$ of the right side. For the first summand, we have

$$\frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} M)^2 \cdot 1_{A_{i,k}}\Bigg] \leq \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} M)^2\Bigg] \to \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Big[\int_0^1 \sigma(s, M_s)^2 \, ds\Big]$$

For the second summand, we first note that

$$\frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \frac{(\Delta_i^{\pi^{(k)}} M)^2}{\sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})^2}\Bigg] \to \frac{1}{2}$$

and that along the B-sets we have

$$\begin{aligned}
\overline{\lim_{k\to\infty}} &\left|\frac{1}{2}\mathbb{E}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \frac{(\Delta_i^{\pi^{(k)}} M)^2}{\sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})^2} \cdot 1_{B_{i,k}}\Bigg]\right| \\
&\leq \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \left|\mathbb{E}_{\mathbb{S}}\Bigg[\frac{(\Delta_i^{\pi^{(k)}} M)^2}{\sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})^2} \cdot 1_{B_{i,k}}\Bigg]\right| \\
&\leq \overline{\lim_{k\to\infty}} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \underbrace{\sqrt{\mathbb{E}_{\mathbb{S}}\Bigg[\frac{(\Delta_i^{\pi^{(k)}} M)^4}{\sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})^4}\Bigg]}}_{\leq C_1 \text{ (independent of } i \text{ and } k)} \sqrt{\mathbb{S}[B_{i,k}]} \\
&\leq C_1 \overline{\lim_{k\to\infty}} \Delta_i \pi^{(k)} \sqrt{\mathbb{S}[B_{i,k}]} \leq 2\varepsilon C_1 \|M_1\|_{L^2(\mathbb{S})}
\end{aligned}$$

So we can conclude that, over the A-sets, we have

$$\overline{\lim_{k\to\infty}} -\frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \frac{(\Delta_i^{\pi^{(k)}} M)^2}{\sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})^2} \cdot 1_{A_{i,k}}\Bigg] \leq -\frac{1}{2} + 2\varepsilon C_1 \|M_1\|_{L^2(\mathbb{S})}$$

For the third summand, we first note that (by the same argument as in the proof of Theorem 3.3) we have

$$\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log \sigma(\pi_{i-1}^{(k)}, M_{\pi_{i-1}^{(k)}})\Bigg] \to \mathbb{E}_{\mathbb{S}}\Big[\int_0^1 \log \sigma(s, M_s) \, ds\Big] = \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Big[\int_0^1 \log \sigma(s, M_s)^2 \, ds\Big]$$

and that along the B-sets we have

$$\begin{aligned}
&\varlimsup_{k\to\infty}\left|\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\log\sigma(\pi_{i-1}^{(k)},M_{\pi_{i-1}^{(k)}})\cdot 1_{B_{i,k}}\right]\right| \\
&\qquad\le \varlimsup_{k\to\infty}\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\left|\mathbb{E}_{\mathbb{S}}\left[\log\sigma(\pi_{i-1}^{(k)},M_{\pi_{i-1}^{(k)}})\cdot 1_{B_{i,k}}\right]\right| \\
&\qquad\le \varlimsup_{k\to\infty}\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\underbrace{\sqrt{\mathbb{E}_{\mathbb{S}}\left[\left(\log\sigma(\pi_{i-1}^{(k)},M_{\pi_{i-1}^{(k)}})\right)^2\right]}}_{\le C_2\text{ (indepedent of } i \text{ and } k)}\sqrt{\mathbb{S}[B_{i,k}]} \\
&\qquad\le C_2\varlimsup_{k\to\infty}\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\sqrt{\mathbb{S}[B_{i,k}]}\le 2\varepsilon C_2\|M_1\|_{L^2(\mathbb{S})}
\end{aligned}$$

So we can conclude that, over the A-sets, we have

$$\begin{aligned}
\varlimsup_{k\to\infty}-\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\log\sigma(\pi_{i-1}^{(k)},M_{\pi_{i-1}^{(k)}})\cdot 1_{A_{i,k}}\right]&\le-\frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1\log\sigma(s,M_s)^2\,ds\right] \\
&\quad+2\varepsilon C_2\|M_1\|_{L^2(\mathbb{S})}
\end{aligned}$$

For the fourth summand, we have by virtue of R being bounded that

$$\begin{aligned}
&\left|\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\log(1+\Delta_i\pi^{(k)}R(\pi_{i-1}^{(k)},\pi_i^{(k)},M_{\pi_{i-1}^{(k)}},M_{\pi_i^{(k)}}))\cdot 1_{A_{i,k}}\right]\right| \\
&\qquad\le\mathbb{E}_{\mathbb{S}}\left[\max_{i\in[\#\pi^{(k)}]}\left|\log(1+\Delta_i\pi^{(k)}R(\pi_{i-1}^{(k)},\pi_i^{(k)},M_{\pi_{i-1}^{(k)}},M_{\pi_i^{(k)}}))\cdot 1_{A_{i,k}}\right|\right]\to 0
\end{aligned}$$

In total, we have now shown that

$$\begin{aligned}
&\varlimsup_{k\to\infty}\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\log\frac{p[M_{\pi_i^{(k)}}\mid M_{\pi_{i-1}^{(k)}}]\circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}},\Delta_i\pi^{(k)}]\circ M_{\pi_i^{(k)}}}\cdot 1_{A_{i,k}}\right] \\
&\qquad\le\mathbb{E}_{\mathbb{S}}\left[\int_0^1\Upsilon(\sigma^2(s,M_s))\,ds\right]+2\varepsilon(C_1+C_2)\|M_1\|_{L^2(\mathbb{S})}
\end{aligned}$$

4. For the expectation over the B-sets, we use Lemma 3.5 to conclude that for all $x,y\in\mathbb{R}, s\le t\in[0,1]$ we have

$$\left|\log\frac{p[M_t\mid M_s=x](y)}{N[x,t-s](y)}\right|\le\underbrace{\log\sqrt{2\pi}(C_+\vee C_-)}_{=:C}+\underbrace{(1-2(c_+\vee c_-))}_{=:c}\frac{(y-x)^2}{t-s}=C+c\frac{(y-x)^2}{t-s}$$

We therefore get

$$\varlimsup_{k\to\infty}\left|\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}}\Delta_i\pi^{(k)}\log\frac{p[M_{\pi_i^{(k)}}\mid M_{\pi_{i-1}^{(k)}}]\circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}},\Delta_i\pi^{(k)}]\circ M_{\pi_i^{(k)}}}\cdot 1_{B_{i,k}}\right]\right|\le\ldots$$

$$
\begin{aligned}
\ldots &\leq \varlimsup_{k\to\infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \left| \mathbb{E}_{\mathbb{S}}\left[\left(C + c\frac{(\Delta_i^{\pi^{(k)}} M)^2}{\Delta_i \pi^{(k)}} \right) \cdot 1_{B_{i,k}} \right] \right| \\
&\leq \varlimsup_{k\to\infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \underbrace{\sqrt{\mathbb{E}_{\mathbb{S}}\left[C^2 + 2cC\frac{(\Delta_i^{\pi^{(k)}} M)^2}{\Delta_i \pi^{(k)}} + c^2\frac{(\Delta_i^{\pi^{(k)}} M)^4}{(\Delta_i \pi^{(k)})^2} \right]}}_{\leq C_3 \text{ (independent of } i \text{ and } k)} \sqrt{\mathbb{S}[B_{i,k}]} \\
&\leq C_3 \varlimsup_{k\to\infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \sqrt{\mathbb{S}[B_{i,k}]} \leq 2\varepsilon C_3 \|M_1\|_{L^2(\mathbb{S})}
\end{aligned}
$$

5. Putting our results together we obtain

$$
\begin{aligned}
h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^x) &= \varlimsup_{k\to\infty} \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \log \frac{p[M_{\pi_i^{(k)}} \mid M_{\pi_{i-1}^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i\pi^{(k)}] \circ M_{\pi_i^{(k)}}} \right] \\
&\leq \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(\sigma(s, M_s)^2)\, ds \right] + 2\varepsilon(C_1 + C_2 + C_3)|M_1\|_{L^2(\mathbb{S})}
\end{aligned}
$$

Since ε was chosen arbitrarily, we can conclude the claim.

$\square$

A natural question that arises is how far the assumptions (T1)-(T3) can be relaxed with Theorem 3.6 remaining valid. It seems that at least the assumption of σ being bounded away from 0 and ∞ is essential, as the following example shows.

Example 7. We consider the *Aldous martingale* studied by J. Backhoff-Veraguas and M. Beiglböck in [3]. This martingale is the solution to the SDE

$$
\begin{cases} dM_t = \frac{\sin(\pi M_t)}{\pi\sqrt{1-t}}\, dB_t \\ M_0 = x_0 \end{cases}
$$

for $x_0 \in (0, 1)$. Upon denoting $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} M$, we have by the Gantert inequality for every admissible $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ that

$$
h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon\left(\frac{\sin^2(\pi M_s)}{\pi^2(1-s)} \right) ds \right] \tag{3.6}
$$

and we would expect equality to hold in the above. However, this is not the case! The Aldous martingale has the property that $M_1 \in \{0, 1\}$ almost surely. So for every $k \geq 0$

$$
H(\mathcal{L}_{\mathbb{Q}}[X_0, X_{\pi_1^{(k)}} \ldots, X_1] \mid \mathcal{L}_{\mathbb{W}^{x_0}}[X_0, X_{\pi_1^{(k)}} \ldots, X_1]) = \infty
$$

by virtue of the two laws being mutually singular. This implies that for every admissible $\mathfrak{P}$ we have $h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \infty$, while the right-hand side of (3.6) is finite. So we are provided with a counterexample for the Gantert conjecture!

3.3 Iterative Bass Martingales

In this section, we want to prove the Gantert conjecture for a class of *iterative Bass martingales*. We call a martingale $M = (M_t)_{t\in[0,1]}$ a *Bass martingale* if it satisfies

$$M_1 = f(B_1)$$

for some $f : \mathbb{R} \to \mathbb{R}$ such that $f(B_1)$ is square-integrable[1]. In this case, M satisfies for every $t \in [0, 1]$ that

$$M_t = \mathbb{E}_{\mathbb{S}}[f(B_1) \mid \mathcal{S}_t]$$

We shall consider the more general class of martingales, where f can be a function of the Brownian motion B at a finite number of times in $[0, 1]$.

> **Definition 9** (Iterative Bass Martingale)**.** Let $\tau = (0 = \tau_0, \tau_1, ..., \tau_m = 1)$ be a partition of $[0, 1]$ and $f : \mathbb{R}^m \to \mathbb{R}$ such that $f(B_{\tau_1}, ..., B_{\tau_m})$ is square-integrable. A martingale $(M_t)_{t\in[0,1]}$ is an *iterative Bass martingale* with underlying partition τ and function f if
>
> $$M_1 = f(B_{\tau_1}, ..., B_{\tau_m})$$
>
> (The requirement of $\tau_0 = 0$ has no impact on the definition apart from easing the notation throughout this section.)

Suppose M is an iterative Bass martingale with underlying partition $\tau = (\tau_0, ..., \tau_m)$ and function f. For every $i \in [m]$ and $t \in (\tau_{i-1}, \tau_i]$ we have

$$\begin{aligned} M_t &= \mathbb{E}_{\mathbb{S}}[f(B_{\tau_1}, ..., B_{\tau_m}) \mid \mathcal{S}_t] \\ &= \mathbb{E}_{\mathbb{S}}[f(B_{\tau_1}, ..., B_{\tau_{i-1}}, B_t + (B_{\tau_i} - B_t), \dots, B_t + (B_{\tau_i} - B_t) + \Delta^\tau_{i+1}B ... + \Delta^\tau_m B) \mid \mathcal{S}_t] \end{aligned}$$

Since $B_{\tau_1}, ..., B_{\tau_{i-1}}$ and B_t are $\mathcal{S}_t$-measurable and the jointly independent increments $B_{\tau_i} - B_t, \Delta^\tau_{i+1}B, ..., \Delta^\tau_m B$ are independent of $\mathcal{S}_t$, we have

$$M_t \overset{\mathbb{S}\text{-a.s.}}{=} \hat{f}_i(t, B_{\tau_1}, ..., B_{\tau_{i-1}}, B_t) \tag{3.7}$$

where the function $\hat{f}_i : (\tau_{i-1}, \tau_i] \times \mathbb{R}^{i-1} \times \mathbb{R} \to \mathbb{R}$ is defined by

$$\hat{f}_i(t, \boldsymbol{x}, x_t) = \int_{\mathbb{R}} ... \int_{\mathbb{R}} f(\boldsymbol{x}, x_t + \delta_i, \dots, x_t + \delta_i + ... + \delta_m)\, d\mathcal{N}_{\tau_i - t}(\delta_i) ... d\mathcal{N}_{\Delta_m \tau}(\delta_m)$$

In particular, the starting point of M is given by

$$M_0 \overset{\mathbb{S}\text{-a.s.}}{=} \mathbb{E}_{\mathbb{S}}[f(B_{\tau_1}, ..., B_{\tau_m})] \tag{3.8}$$

In the following, if we write $M = \mathrm{IBM}[\tau, f]$ we shall mean that M is the concrete iterative Bass martingale with underlying partition τ and function f defined by (3.7) and (3.8).

[1] Here we follow the nomenclature used in the theory of martingale optimal transport (see for example [4] or [5]). However, compared to these references, we do <u>not</u> require f to be the derivative of a convex function.

Example 8. Let us look at two simple examples.

(1) For $M = \mathrm{IBM}[(0, \frac{1}{2}, 1), x_{\frac{1}{2}} + x_1]$ it is straightforward to evaluate that

$$\hat{f}_1(t, x_t) = 2x_t, \qquad \hat{f}_2(t, x_{\frac{1}{2}}, x_t) = x_{\frac{1}{2}} + x_t$$

So the martingale M is explicitly given by

$$M_t = \begin{cases} 2B_t & , t \in [0, \frac{1}{2}] \\ B_{\frac{1}{2}} + B_t & , t \in (\frac{1}{2}, 1] \end{cases}$$

(2) For $M = \mathrm{IBM}[(0, 1), x_1^2]$ we have

$$\hat{f}_1(t, x_t) = x_t^2 + (1 - t)$$

So the martingale M is explicitly given by

$$M_t = B_t^2 + (1 - t)$$

We now want to outline a class of 'nice' iterative Bass martingales that we know to have continuous sample paths and for which we will be able to verify the Gantert conjecture.

Definition 10 (Invertible Iterative Bass Martingale)**.** The iterative Bass martingale $M = \mathrm{IBM}[(\tau_0, ..., \tau_m), f]$ is *invertible*, if

(I1) $f \in C^1(\mathbb{R}^m, \mathbb{R})$ and $\partial_i f(B_{\tau_1}, ..., B_{\tau_m})$ is integrable for every $i \in [m]$

(I2) for all $i \in [m]$, $t \in (\tau_{i-1}, \tau_i]$ and $x_{\tau_1}, ..., x_{\tau_{i-1}} \in \mathbb{R}$, the function

$$x_t \mapsto \hat{f}_i(t, x_{\tau_1}, ..., x_{\tau_{i-1}}, x_t)$$

is a diffeomorphism on $\mathbb{R}$.

Suppose the iterative Bass martingale $M = \mathrm{IBM}[(\tau_0, ..., \tau_m), f]$ is invertible. The condition (I1) ensures that for every $i \in [m]$ the function $\hat{f}_i$ is C^∞-smooth. In particular, we can conclude that M has continuous sample paths[2]. We associate with M the process $F = (F_t)_{t \in (0,1]}$, which we define for $t \in (\tau_{i-1}, \tau_i]$ via

$$F_t := \partial_{i+1} \hat{f}_i(t, \boldsymbol{B}_{\tau_{i-1}}, B_t)$$

[2]Since B is an $(\mathcal{S}_t)$-Brownian motion, we have $\mathbb{E}_{\mathbb{S}}[f(B_{\tau_1}, ..., B_{\tau_m}) \mid \mathcal{S}_t] = \mathbb{E}_{\mathbb{S}}[f(B_{\tau_1}, ..., B_{\tau_m}) \mid \mathcal{F}_t^B]$ where $(\mathcal{F}_t^B)_{t \in [0,1]}$ is the filtration generated by the Brownian motion B. This implies that even without any extra assumptions on f, there exists an IBM with underlying function f and partition τ which has continuous sample paths.

We note that the process F restricted to each time frame $(\tau_{i-1}, \tau_i]$ is a martingale. Indeed, for $t \in (\tau_{i-1}, \tau_i]$ we have

$$\begin{aligned}
F_t &= \partial_{i+1}\hat{f}_i(t, \boldsymbol{B}_{\tau_{i-1}}, B_t) \\
&= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \sum_{k=i}^{m} \partial_k f(\boldsymbol{B}_{\tau_{i-1}}, B_t + \delta_i, \dots, B_t + \delta_i + \dots + \delta_m)\, d\mathcal{N}_{\tau_i - t}(\delta_i) \dots d\mathcal{N}_{\Delta_m \tau}(\delta_m) \\
&= \mathbb{E}_{\mathbb{S}}\left[\sum_{k=i}^{m} \partial_k f(B_{\tau_1}, \dots, B_{\tau_m}) \,\middle|\, \mathcal{S}_t \right]
\end{aligned}$$

On the interval $(\tau_{i-1}, \tau_i]$, we apply Itō's formula to $M_t = \hat{f}_i(t, \boldsymbol{B}_{\tau_{i-1}}, B_t)$ to obtain for $t \in (\tau_{i-1}, \tau_i]$ that

$$M_t = M_{\tau_{i-1}} + \int_{\tau_{i-1}}^{t} (\partial_1 + \tfrac{1}{2}\partial_{i+1}^2)\hat{f}_i(s, \boldsymbol{B}_{\tau_{i-1}}, B_s)\, ds + \int_{\tau_{i-1}}^{t} \partial_{i+1}\hat{f}_i(s, \boldsymbol{B}_{\tau_{i-1}}, B_s)\, dB_s$$

The quadratic variation of M is thus given by

$$[M]_t = \sum_{i=1}^{m} \int_{\tau_{i-1}}^{\tau_i} \Big(\partial_{i+1}\hat{f}_i(s, \boldsymbol{B}_{\tau_{i-1}}, B_s)\Big)^2\, ds = \int_0^t F_s^2\, ds$$

So M has an absolutely continuous quadratic variation with density F^2. Thus, if we denote by $\mathbb{Q} := \mathcal{L}_{\mathbb{S}} M$ the law of M on the Wiener space and by $x_0 := \mathbb{E}_{\mathbb{S}}[f(B_{\tau_1}, \dots, B_{\tau_m})]$ the starting point of M, then by the Gantert inequality we have for every admissible partition sequence $\mathfrak{P}$ that

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) \geq \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(F_s^2)\, ds\right] \tag{3.9}$$

and we want to show that equality holds. However, we will only be able to show this for admissible partition sequences that refine the underlying partition τ. That is, for $\mathfrak{P} = (\pi^{(k)})_{k \geq 0}$ we require the existence of a $K \geq 0$ such that for all $k \geq K$ the partition $\pi^{(k)}$ refines τ.

To show the desired equality, we make use of condition (I2). This condition ensures that for every $t \in (\tau_{i-1}, \tau_i]$, the values of $M_{\tau_1}, \dots, M_{\tau_{i-1}}, M_t$ determine the values of $B_{\tau_1}, \dots, B_{\tau_{i-1}}, B_t$. Indeed, let us denote for every $j \in [m]$ by $\hat{f}_j^{-1}(s, x_{\tau_1}, \dots, x_{\tau_{j-1}}, .)$ the inverse of the function

$$x_t \mapsto \hat{f}_j(s, x_{\tau_1}, \dots, x_{\tau_{j-1}}, x_t)$$

for fixed $s \in (\tau_{j-1}, \tau_j]$ and $x_{\tau_1}, \dots, x_{\tau_{j-1}} \in \mathbb{R}$. We then have

$$\begin{aligned}
B_{\tau_1} &= \hat{f}_1^{-1}(\tau_1, M_{\tau_1}) \\
B_{\tau_2} &= \hat{f}_2^{-1}(\tau_2, B_{\tau_1}, M_{\tau_2}) = \hat{f}_2^{-1}(\tau_2, \hat{f}_1^{-1}(\tau_1, M_{\tau_1}), M_{\tau_2}) \\
B_{\tau_3} &= \hat{f}_3^{-1}(\tau_3, B_{\tau_1}, B_{\tau_2}, M_{\tau_3}) = \hat{f}_3^{-1}(\tau_3, \hat{f}_1^{-1}(\tau_1, M_{\tau_1}), \hat{f}_2^{-1}(\tau_2, \hat{f}_1^{-1}(\tau_1, M_{\tau_1}), M_{\tau_2}), M_{\tau_3}) \\
&\ \vdots \\
B_t &= \hat{f}_i^{-1}(t, B_{\tau_1}, \dots, B_{\tau_{i-1}}, M_t) = \hat{f}_i^{-1}(t, \hat{f}_1^{-1}(\tau_1, M_{\tau_1}), \dots, M_t)
\end{aligned}$$

We shall use this idea to evaluate in Lemma 3.8 the conditional density of M_t given $\boldsymbol{M}_{\tau_{i-1}}$.

Before we state this Lemma, let us make a small remark. To check whether $M = \mathrm{IBM}[\tau, f]$ satisfies condition (I2), we are required to evaluate the functions $\hat{f}_i$. Ideally, we would like to determine whether or not condition (I2) holds directly from the underlying function f. Unfortunately, there appears to be no simple way to characterize all the functions for which this is the case. However, the following Lemma provides at least a simple sufficient condition.

Lemma 3.7. Let $M = \mathrm{IBM}[\tau, f]$ be such that (I1) is satisfied. If there exists a $K > 0$ such that for every $i \in [\#\tau]$ we have

$$\sum_{k=i}^{m} \partial_k f(x_1, ..., x_m) > K, \quad \forall (x_1, ..., x_m) \in \mathbb{R}^m$$

then M is invertible.

Proof. Let $i \in [\#\tau]$, $t \in (\tau_{i-1}, \tau_i]$ and $\boldsymbol{x} \in \mathbb{R}^{i-1}$ be arbitary, then

$$\partial_{i+1} \hat{f}_i(t, \boldsymbol{x}, x_t) = \int_{\mathbb{R}} ... \int_{\mathbb{R}} \sum_{k=i}^{m} \partial_k f(\boldsymbol{x}, x_t + \delta_i, ..., x_t + \delta_i + ... + \delta_m) d\mathcal{N}_{\tau_i - t}(\delta_i) ... d\mathcal{N}_{\Delta_m \tau}(\delta_m) > K$$

Thus, the function $x_t \mapsto \hat{f}_i(t, \boldsymbol{x}, x_t)$ is a diffeomorphism and condition (I2) holds, making M invertible.

$\square$

Lemma 3.8. Let $M = \mathrm{IBM}[\tau, f]$ be invertible and π a partition that refines τ, then for every $i \in [\#\pi]$

$$p[M_{\pi_i} \mid \boldsymbol{M}_{\pi_{i-1}}] \circ M_{\pi_i} = \frac{1}{|F_{\pi_i}|} \cdot N[B_{\pi_{i-1}}, \Delta_i \pi] \circ B_{\pi_i}$$

Proof. Let $i \in [\#\pi]$ be arbitrary and j such that $\pi_i \in (\tau_{j-1}, \tau_j]$. Due to the invertibility of M and the fact that π refines τ, the value of $B_{\pi_{i-1}}$ is determined by the values of $M_{\pi_0}, ..., M_{\pi_{i-1}}$ and we have

$$p[B_{\pi_i} \mid \boldsymbol{M}_{\pi_{i-1}}] = N[B_{\pi_{i-1}}, \Delta_i \pi]$$

Further, the function $\phi(x) := \hat{f}_j(\pi_i, \boldsymbol{B}_{\tau_{j-1}}, x)$ is a diffeomorphism that satisfies $\phi(B_{\pi_i}) = M_{\pi_i}$ and

$$|(\phi^{-1})'(M_{\pi_i})| = \frac{1}{|\phi'(B_{\pi_i})|} = \frac{1}{|\partial_{j+1} \hat{f}_j(\pi_i, \boldsymbol{B}_{\tau_{j-1}}, B_{\pi_i})|} = \frac{1}{|F_{\pi_i}|}$$

We thus obtain by a change of variable that

$$\begin{aligned} p[M_{\pi_i} \mid \boldsymbol{M}_{\pi_{i-1}}] \circ M_{\pi_i} &= |(\phi^{-1})'(M_{\pi_i})| \cdot p[B_{\pi_i} \mid \boldsymbol{M}_{\pi_i}] \circ \phi^{-1}(M_{\pi_i}) \\ &= \frac{1}{|F_{\pi_i}|} \cdot p[B_{\pi_i} \mid \boldsymbol{M}_{\pi_i}] \circ B_{\pi_i} \\ &= \frac{1}{|F_{\pi_i}|} \cdot N[B_{\pi_{i-1}}, \Delta_i \pi] \circ B_{\pi_i} \end{aligned}$$

$\square$

Note that the assumption of π refining τ is essential for the above proof. Indeed, we used the previously established fact that $B_{\pi_{i-1}}$ is determined by $M_{\pi_{i-1}}$ and $M_{\tau_1}, ..., M_{\tau_{j-1}}$ (where τ_{j-1} is the last partition point of τ that preceeds π_{i-1}), which are amongst $M_{\pi_0}, ..., M_{\pi_{i-1}}$ by assumption of π refining τ.

Let us now proof the main result.

Theorem 3.9. Let $M = \mathrm{IBM}[\tau, f]$ be invertible with starting point x_0. Let $\mathbb{Q} := \mathcal{L}_{\mathbb{S}}M$, then for every admissible partition sequence $\mathfrak{P} = (\pi^{(k)})_{k\geq 0}$ that refines τ we have

$$h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) = \mathbb{E}_{\mathbb{S}}\Big[\int_0^1 \Upsilon(F_s^2)\, ds\Big]$$

Proof. By Lemma 2.4 we know that the specific relative entropy is given by

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) &\overset{\text{Lem}\,2.4}{=} \varlimsup_{k\to\infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \mathbb{E}_{\mathbb{Q}}\Bigg[\log \frac{d\mathcal{L}_{\mathbb{Q}}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]}{d\mathcal{L}_{\mathbb{W}^x}[X_{\pi_i^{(k)}} \mid \boldsymbol{X}_{\pi_{i-1}^{(k)}}]} \circ X_{\pi_i^{(k)}}\Bigg] \\ &= \varlimsup_{k\to\infty} \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \mathbb{E}_{\mathbb{S}}\Bigg[\log \frac{p[M_{\pi_i^{(k)}} \mid \boldsymbol{M}_{\pi_i^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i\pi^{(k)}] \circ M_{\pi_i^{(k)}}}\Bigg] \end{aligned}$$

Since $\mathfrak{P}$ refines the underlying partition τ, we find by definition a $K \geq 0$ such that for all $k \geq K$ the partition $\pi^{(k)}$ refines τ. So for sufficiently large k we have by Lemma 3.8 that

$$\begin{aligned} &\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \mathbb{E}_{\mathbb{S}}\Bigg[\log \frac{p[M_{\pi_i^{(k)}} \mid \boldsymbol{M}_{\pi_i^{(k)}}] \circ M_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i\pi^{(k)}] \circ M_{\pi_i^{(k)}}}\Bigg] \\ &\quad= \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \mathbb{E}_{\mathbb{S}}\Bigg[\log \frac{N[B_{\pi_{i-1}^{(k)}}, \Delta_i\pi^{(k)}] \circ B_{\pi_i^{(k)}}}{N[M_{\pi_{i-1}^{(k)}}, \Delta_i\pi^{(k)}] \circ M_{\pi_i^{(k)}}} - \log |F_{\pi_i^{(k)}}|\Bigg] \\ &\quad= \sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \mathbb{E}_{\mathbb{S}}\Bigg[\frac{(\Delta_i^{\pi^{(k)}} M)^2}{2\Delta_i\pi^{(k)}} - \frac{(\Delta_i^{\pi^{(k)}} B)^2}{2\Delta_i\pi^{(k)}} - \log |F_{\pi_i^{(k)}}|\Bigg] \\ &\quad= \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} M)^2\Bigg] - \frac{1}{2}\mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} (\Delta_i^{\pi^{(k)}} B)^2\Bigg] - \mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \log |F_{\pi_i^{(k)}}|\Bigg] \\ &\quad= \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 F_s^2\, ds\right] - \frac{1}{2} - \mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \log |F_{\pi_i^{(k)}}|\Bigg] \end{aligned}$$

We recall that the process F restricted to each time frame $(\tau_{i-1}, \tau_i]$ is a martingale. The same is true $|F|$, as by condition (I2) F is either non-negative or non-positive. So we can invoke Lemma A.2 to conclude that

$$\begin{aligned} h^{\mathfrak{P}}(\mathbb{Q} \mid \mathbb{W}^{x_0}) &= \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 F_s^2\, ds\right] - \frac{1}{2} - \liminf_{k\to\infty} \mathbb{E}_{\mathbb{S}}\Bigg[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i\pi^{(k)} \log |F_{\pi_i^{(k)}}|\Bigg] \\ &\leq \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 F_s^2\, ds\right] - \frac{1}{2} - \frac{1}{2}\mathbb{E}_{\mathbb{S}}\left[\int_0^1 \log F_s^2\, ds\right] = \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \Upsilon(F_s^2)\, ds\right] \end{aligned}$$

□

Appendix A

Appendix

A.1 Auxilliary Results

Lemma A.1. Let ν, μ be two probability measures on a space $(\Omega, \mathcal{F})$ and $(\mathcal{F}_n)_{n\geq 0}$ a sequence of sub-σ-algebras increasing to $\mathcal{F}_\infty$. If for every $n \geq 0$ we have $\nu|_{\mathcal{F}_n} \ll \mu|_{\mathcal{F}_n}$ with density ϕ_n, then

(1) $\phi_\infty := \lim_{n\to\infty} \phi_n$ exists ν-a.s. and μ-a.s. with

$$\phi_\infty \in [0, \infty)\ \mu\text{-a.s.} \quad \text{and} \quad \phi_\infty \in (0, \infty]\ \nu\text{-a.s.}$$

Furthermore, the Lebesgue decomposition $\nu = \nu_\perp + \nu_\ll$ of ν with respect to μ on $\mathcal{F}_\infty$ is given by

$$\nu_\perp(A) = \nu(A \cap \{\phi_\infty = \infty\}) \quad \text{and} \quad \nu_\ll(A) = \int_A \phi_\infty \, d\mu$$

(2) Let Υ be as defined in (1.7), then

$$\lim_{n\to\infty} \int_\Omega \Upsilon(\phi_n)\, d\mu = \frac{1}{2}\nu_\perp(\Omega) + \int_\Omega \Upsilon(\phi_\infty)\, d\mu$$

and this equation extends to the case where ν is a non-negative finite measure.

Proof. see ([8], Lemma 19) □

Lemma A.2. Let $(M_t)_{t\in[0,1]}$ be a non-negative, square-integrable, continuous martingale such that $\log M_1$ is integrable and $(\pi^{(k)})_{k\geq 0}$ an admissible partition sequence, then

$$\liminf_{k\to\infty} \mathbb{E}_\mathbb{S}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log M_{\pi_i^{(k)}}\right] \geq \mathbb{E}_\mathbb{S}\left[\int_0^1 \log M_t \, dt\right]$$

Proof. We first prove that

$$\liminf_{k\to\infty} \mathbb{E}_\mathbb{S}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log M_{\pi_{i-1}^{(k)}}\right] \geq \mathbb{E}_\mathbb{S}\left[\int_0^1 \log M_t \, dt\right] \tag{A.1}$$

For every $k \geq 0$, let $\mathcal{P}_{\pi^{(k)}}$ be the σ-algebra defined by

$$\mathcal{P}_{\pi^{(k)}} := \sigma(\{(t_1, t_2] \times A \mid t_1 < t_2 \in \pi^{(k)},\ A \in \mathcal{S}_{t_1}\})$$

then for every $t \in [0,1]$, we have

$$\mathbb{E}_{m\otimes\mathbb{S}}[M_t \mid \mathcal{P}_{\pi^{(k)}}] = \sum_{i=1}^{\#\pi^{(k)}} M_{\pi^{(k)}_{i-1}} 1_{(\pi^{(k)}_{i-1}, \pi^{(k)}_i]} \tag{A.2}$$

We thus obtain

$$\begin{aligned}
\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\pi^{(k)}} \Delta_i \pi^{(k)} \log M_{\pi^{(k)}_{i-1}}\right] &= \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \sum_{i=1}^{\#\pi^{(k)}} \log M_{\pi^{(k)}_{i-1}} 1_{(\pi^{(k)}_{i-1} \pi^{(k)}_i]}(t)\, dt\right] \\
&\overset{(A.2)}{=} \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \log \mathbb{E}_{m\otimes\mathbb{S}}[M_t \mid \mathcal{P}_{\pi^{(k)}}]\, dt\right] \\
&\geq \mathbb{E}_{\mathbb{S}}\left[\int_0^1 \log M_t\, dt\right]
\end{aligned} \tag{A.3}$$

To conlude, we note that $t \mapsto \mathbb{E}_{\mathbb{S}}[\log M_t]$ is uniformly continuous on $[0,1]$ so that

$$\begin{aligned}
&\mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log M_{\pi^{(k)}_i}\right] - \mathbb{E}_{\mathbb{S}}\left[\sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} \log M_{\pi^{(k)}_{i-1}}\right] \\
&\qquad = \sum_{i=1}^{\#\pi^{(k)}} \Delta_i \pi^{(k)} (\mathbb{E}_{\mathbb{S}}[\log M_{\pi^{(k)}_i}] - \mathbb{E}_{\mathbb{S}}[\log M_{\pi^{(k)}_{i-1}}]) \\
&\qquad \leq \max_{i \in [\#\pi^{(k)}]} (\mathbb{E}_{\mathbb{S}}[\log M_{\pi^{(k)}_i}] - \mathbb{E}_{\mathbb{S}}[\log M_{\pi^{(k)}_{i-1}}]) \to 0
\end{aligned}$$

□

Lemma A.3 (*Mill's Inequality*)**.** Let $X \sim_{\mathbb{S}} \mathcal{N}(0, \sigma^2)$, then for any $\tau > 0$

$$\mathbb{S}[X > \tau] \leq \frac{\sigma}{\sqrt{2\pi\tau^2}} \cdot \exp\left(-\frac{\tau^2}{2\sigma^2}\right)$$

Proof. We have

$$\begin{aligned}
\tau \cdot \mathbb{S}[X > \tau] &= \tau \int_\tau^\infty d\mathcal{N}[0, \sigma^2] \\
&\leq \int_\tau^\infty x\, d\mathcal{N}[0, \sigma^2](x) \\
&= \int_\tau^\infty \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\tau^2}{2\sigma^2}\right)
\end{aligned}$$

We then divide by τ to conclude the claim.

□

A.2 Regular Conditional Distributions

We follow Chapter 8 of [13]. Throughout this section, assume that we are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 11 (Regular Conditional Distribution)**.** Let $Y : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Lambda, \mathcal{G})$ be a random variable and $\mathcal{H} \subseteq \mathcal{F}$ a sub-σ-algebra. A map $\kappa_{Y,\mathcal{H}} : \Omega \times \mathcal{G} \to [0, 1]$ is a *regular conditional distribution* of Y given $\mathcal{H}$ if

(R1) $A \mapsto \kappa_{Y,\mathcal{H}}(\omega, A)$ is a probability measure on $(\Lambda, \mathcal{G})$ for every $\omega \in \Omega$

(R2) $\omega \mapsto \kappa_{Y,\mathcal{H}}(\omega, A)$ is version of $\mathbb{E}_{\mathbb{P}}[1_{Y \in A} \mid \mathcal{H}]$ for every $A \in \mathcal{G}$

In general, a regular conditional distribution does not have to exist. However, we are interested in the case where Y is $\mathbb{R}$-valued, where the existence is ensured by the following Theorem.

Theorem A.4 (*Existence of Regular Conditional Distribution*)**.** Let Y be an $\mathbb{R}$-valued random variable and $\mathcal{H} \subseteq \mathcal{F}$ a sub-σ-algebra. There exists a regular conditional distribution $\kappa_{Y,\mathcal{H}}$ of Y given $\mathcal{H}$.

Proof. see ([13], Theorem 8.37).

□

Suppose $\boldsymbol{X}$ is an $\mathbb{R}^n$-valued random variable and Y is an $\mathbb{R}$-valued random variable. By Theorem A.4 there exists a regular conditional distribution of Y given $\sigma(\boldsymbol{X})$. By virtue of $\mathcal{B}(\mathbb{R})$ being countably generated, this regular conditional distribution uniquely determined up to null sets. That is, if κ and $\hat{\kappa}$ are regular conditional distribution of Y given $\sigma(\boldsymbol{X})$, then for almost every $\omega \in \Omega$ we have

$$\kappa(\omega, A) = \hat{\kappa}(\omega, A), \qquad \forall A \in \mathcal{B}(\mathbb{R})$$

In this setting, we refer to the regular conditional distribution as the *conditional law of* Y *given* $\boldsymbol{X}$ and we denote it by $\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}]$.

By construction, for every $A \in \mathcal{B}(\mathbb{R})$ the function $\omega \mapsto \mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}](\omega, A)$ is $\sigma(\boldsymbol{X})$-measurable, so by the Doob-Dynkin Lemma there exists a measurable map $\phi_A : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}](\omega, A) = \phi_A \circ \boldsymbol{X}(\omega) \tag{A.4}$$

For fixed $\boldsymbol{x} \in \mathbb{R}^n$, we define the *conditional law of* Y *given* $\boldsymbol{X} = \boldsymbol{x}$ as the measure

$$\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](A) := \phi_A(x)$$

In this case (A.4) becomes

$$\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}](\omega, A) = \mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{X}(\omega)](A)$$

(We note that the ϕ-functions are not necessarily unique if $\boldsymbol{X}$ is not surjective. So in that case, the definition of the family $(\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}])_{\boldsymbol{x}\in\mathbb{R}^n}$ depends on the choice of these functions. However, for all $\boldsymbol{x}$ in the image of $\boldsymbol{X}$, so for all $\mathcal{L}_{\mathbb{P}}\boldsymbol{X}$-a.e. $\boldsymbol{x}$, the obtained measures $\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]$ will coincide. The following two results are therefore valid regardless of this choice.)

Theorem A.5 (*Disintegration Theorem*)**.** In the above setting, let $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be measurable such that $f(\boldsymbol{X}, Y)$ is integrable, then

$$\mathbb{E}[f(\boldsymbol{X}, Y)] = \int_{\mathbb{R}^n} \int_{\mathbb{R}} f(\boldsymbol{x}, y)\, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y)\, d\mathcal{L}_{\mathbb{P}}\boldsymbol{X}(\boldsymbol{x})$$

Proof. Follows from ([13], Theorem 8.38). $\square$

Corollary A.6. In the above setting, suppose $\mathbb{Q}$ is another probability measure on $(\Omega, \mathcal{F})$, then the following are equivalent

(a) $\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y] \ll \mathcal{L}_{\mathbb{P}}[\boldsymbol{X}, Y]$

(b) $\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}] \ll \mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]$ and $\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}] \ll \mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]$ for $\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]$-a.e. $\boldsymbol{x}$

If (a) or (b) holds, then for $\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}, Y]$-a.e. $(\boldsymbol{x}, y) \in \mathbb{R}^{n+1}$ we have

$$\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}, Y]}(\boldsymbol{x}, y) = \frac{d\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]}{d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]}(y) \cdot \frac{d\mathcal{L}_{\mathbb{Q}}\boldsymbol{X}}{d\mathcal{L}_{\mathbb{P}}\boldsymbol{X}}(\boldsymbol{x})$$

Proof. ad (a)$\Rightarrow$(b): Suppose $\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y] \ll \mathcal{L}_{\mathbb{P}}[\boldsymbol{X}, Y]$, then clearly $\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}] \ll \mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]$. We want to show that the function

$$(\omega, A) \mapsto \int_A \frac{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X},Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X},Y]}(\boldsymbol{X}(\omega), y)}{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]}(\boldsymbol{X}(\omega))}\, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}](\omega, y) \tag{A.5}$$

is a regular conditional distribution of Y given $\sigma(\boldsymbol{X})$ w.r.t. to the measure $\mathbb{Q}$. Let $A \in \mathcal{B}(\mathbb{R})$ be fixed, then for every $T \in \sigma(\boldsymbol{X})$ we have

$$\begin{aligned}
&\int_T \int_A \frac{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X},Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X},Y]}(\boldsymbol{X}(\omega), y)}{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]}(\boldsymbol{X}(\omega))}\, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}](\omega, y)\, d\mathbb{Q}(\omega) \\
&\qquad = \int_{\boldsymbol{X}(T)} \int_A \frac{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X},Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X},Y]}(\boldsymbol{x}, y)}{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]}(\boldsymbol{x})}\, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y)\, d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}](\boldsymbol{x}) \\
&\qquad = \int_{\boldsymbol{X}(T)} \int_A \frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}, Y]}(\boldsymbol{x}, y)\, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y)\, d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}](\boldsymbol{x}) \\
&\qquad = \int_{\boldsymbol{X}(T)} \int_A d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y] \\
&\qquad = \mathbb{E}_{\mathbb{Q}}[1_A \cdot 1_T]
\end{aligned}$$

So we have shown that for fixed $A \in \mathcal{B}(\mathbb{R})$ the function in (A.5) is a version of $\mathbb{E}_{\mathbb{Q}}[1_A \,|\, \sigma(\boldsymbol{X})]$ and for every fixed ω it defines a probability measure. We thus know by the uniqueness of the regular conditional distribution of Y given $\sigma(\boldsymbol{X})$ that for $\mathbb{Q}$-a.e. ω we have

$$\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X}](\omega, A) = \int_A \frac{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X},Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X},Y]}(\boldsymbol{X}(\omega), y)}{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]}(\boldsymbol{X}(\omega))} \, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X}](\omega, y), \quad \forall A \in \mathcal{B}(\mathbb{R})$$

which implies that for $\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]$-a.e. $\boldsymbol{x} \in \mathbb{R}^n$ we have

$$\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](A) = \int_A \frac{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X},Y]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X},Y]}(\boldsymbol{x}, y)}{\frac{d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}]}{d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}]}(\boldsymbol{x})} \, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y), \quad \forall A \in \mathcal{B}(\mathbb{R})$$

ad (b)$\Rightarrow$(a): We have for every $A \in \mathcal{B}(\mathbb{R}^{n+1})$ that

$$\begin{aligned}
\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y](A) &= \int_{\mathbb{R}^{n+1}} 1_{(\boldsymbol{x},y)\in A} \, d\mathcal{L}_{\mathbb{Q}}[\boldsymbol{X}, Y](\boldsymbol{x}, y) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}} 1_{(\boldsymbol{x},y)\in A} d\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y) \, d\mathcal{L}_{\mathbb{Q}}\boldsymbol{X}(\boldsymbol{x}) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}} 1_{(\boldsymbol{x},y)\in A} d\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y) \, d\mathcal{L}_{\mathbb{Q}}\boldsymbol{X}(\boldsymbol{x}) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}} 1_{(\boldsymbol{x},y)\in A} \frac{d\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]}{d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]}(y) \cdot \frac{d\mathcal{L}_{\mathbb{Q}}\boldsymbol{X}}{d\mathcal{L}_{\mathbb{P}}\boldsymbol{X}}(\boldsymbol{x}) \, d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}](y) \, d\mathcal{L}_{\mathbb{P}}\boldsymbol{X}(\boldsymbol{x}) \\
&= \int_A \frac{d\mathcal{L}_{\mathbb{Q}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]}{d\mathcal{L}_{\mathbb{P}}[Y \mid \boldsymbol{X} = \boldsymbol{x}]}(y) \cdot \frac{d\mathcal{L}_{\mathbb{Q}}\boldsymbol{X}}{d\mathcal{L}_{\mathbb{P}}\boldsymbol{X}}(\boldsymbol{x}) \, d\mathcal{L}_{\mathbb{P}}[\boldsymbol{X}, Y](\boldsymbol{x}, y)
\end{aligned}$$

which proves the claim.

$\square$

A.3 Stochastic Integration

We establish the notion of stochastic integration against square-integrable, continuous martingales (in particular Brownian motion), which we use throughout in the text. For this, we will follow Chapter 5 of [11].

Throughout this section, let $(\Omega, \mathcal{S}, (\mathcal{S}_t)_{t\in[0,1]}, \mathbb{S})$ be a filtered probability space. We denote by $\mathcal{M}^2$ the set of all càdlàg, square-integrable martingales and $\mathcal{M}_c^2$ the set of all square-integrable, continuous martingales. We shall consider $\mathcal{M}^2$ equipped with the norm $\|.\|_{\mathcal{M}^2}$ given by

$$\|M\|_{\mathcal{M}^2} := \sqrt{\mathbb{E}[M_1^2]}$$

and we note that $(\mathcal{M}^2, \|.\|_{\mathcal{M}^2})$ forms a Hilbert space containing $\mathcal{M}_c^2$ as a closed subspace.

Definition 12 (Simple Process)**.** A process $H = (H_t)_{t\in[0,1]}$ is *simple* if it is of the form

$$H_t = \sum_{i=1}^{\#\pi} Z_i \cdot 1_{(\pi_{i-1},\pi_i]}(t)$$

for π a partition of $[0,1]$ and Z_i a bounded $\mathcal{S}_{\pi_{i-1}}$-measurable random variable.

We shall denote the set of simple processes by $\mathfrak{S}$. Given $H = \sum_{i=1}^{\#\pi} Z_i \cdot 1_{(\pi_{i-1},\pi_i]} \in \mathfrak{S}$ and $M \in \mathcal{M}^2$ we define the *stochastic integral* of H against M as the process $H.M = ((H.M)_t)_{t\in[0,1]}$ given by

$$(H.M)_t = \int_0^t H_s \, dM_s := \sum_{i=1}^{\#\pi^{(k)}} Z_i \cdot (M_{\pi_i \wedge t} - M_{\pi_{i-1}\wedge t})$$

One can show that $H.M \in \mathcal{M}^2$. We now want to extend this definition to allow for a broader class of integrands.

Definition 13 (Previsible σ-algebra)**.** The *previsible σ-algebra* $\mathcal{P}$ on $[0,1]\times\Omega$ is given by

$$\mathcal{P} := \sigma(\{(t_1,t_2]\times E \mid t_1 < t_2,\ E \in \mathcal{S}_{t_1}\})$$

A process (interpreted as a map $[0,1]\times\Omega \to \mathbb{R}$) is *previsible* if it is $\mathcal{P}$-measurable.

It is easy to see that the simple processes are previsible. In particular, $\mathcal{P}$ is the smallest σ-algebra such that all adapted, left-continuous processes are measurable.

Given any $M \in \mathcal{M}_c^2$, we let μ_M be the measure on $([0,1]\times\Omega, \mathcal{P})$ defined by

$$\mu_M((t_1,t_2]\times A) := \mathbb{E}[1_A([M]_{t_2} - [M]_{t_1})]$$

This measure is finite, allowing us to invoke the following Lemma.

Lemma A.7. Let μ be finite measure on $([0,1] \times \Omega, \mathcal{P})$, then $\mathfrak{S}$ is dense in $L^2(\mu)$.

We will write $L^2(M)$ as short-hand for $L^2([0,1] \times \Omega, \mathcal{P}, \mu_M)$ and $\|.\|_{L^2(M)}$ for its norm which is given by

$$\|H\|_{L^2(M)} = \sqrt{\mathbb{E}\left[\int_0^1 H_s^2 \, d[M]_s\right]}$$

So $L^2(M)$ consists of all processes for which the above expectation is finite.

The following Theorem now shows us that the linear operator mapping a simple process to its stochastic integral against M describes an isometry from $\mathfrak{S}$ to $\mathcal{M}_c^2$.

Theorem A.8 (*Itō Isometry*)**.** Let $M \in \mathcal{M}_c^2$, then the linear operator

$$I : (\mathfrak{S}, \|.\|_{L^2(M)}) \to (\mathcal{M}_c^2, \|.\|_{\mathcal{M}^2}), \quad H \mapsto H.M = \int_0^{\cdot} H_s \, dM_s$$

is an isometry. That is

$$\mathbb{E}\left[\int_0^1 H_s^2 \, d[M]_s\right] = \mathbb{E}\left[\left(\int_0^1 H_s \, dM_s\right)^2\right]$$

Since we know that the set of simple processes $\mathfrak{S}$ is dense in $L^2(M)$, the above implies that the operator I has a unique isometric extension $\hat{I}$ to the entire $L^2(M)$.

Definition 14 (Stochastic Integral)**.** Let $H \in L^2(M)$ and $M \in \mathcal{M}_c^2$, then the *stochastic integral* of H against M is the process

$$H.M = ((H.M)_t)_{t\in[0,1]} = \left(\int_0^t H_s \, dM_s\right)_{t\in[0,1]}$$

defined by

$$(H.M)_t = \int_0^t H_s \, dM_s := \hat{I}(H)_t$$

Since $H.M$ is a square-integrable, continuous martingale, it has a quadratic variation process, which is given by the following useful formula.

Proposition A.9. Let $H \in L^2(M)$ and $M \in \mathcal{M}_c^2$, then

$$[H.M]_t = \int_0^t H_s^2 \, d[M]_s$$

One of our main tools is the seminal Itō formula, which (in the formulation that we are providing) tells us that for $M \in \mathcal{M}_c^2$ and $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ sufficiently smooth, the process $(f(t, M_t))_{t\in[0,1]}$ is a continuous semimartingale and we are provided with its Doob-Meyer decomposition.

Theorem A.10 (*Itō's Formula*)**.** Let $M \in \mathcal{M}_c^2$ and $f \in C^2([0,1] \times \mathbb{R}, \mathbb{R})$, then for all $t \in [0,1]$ we have

$$f(t, M_t) = f(0, M_0) + \int_0^t \partial_x f(s, M_s)\, dM_s + \int_0^t \partial_t f(s, M_s)\, ds + \frac{1}{2} \int_0^t \partial_{xx} f(s, M_s)\, d[M]_s$$

We immediately see that the semimartingale $F(M) = (f(t, M_t))_{t \in [0,1]}$ admits the decomposition

$$F(M) = F(M)_0 + A + N$$

where A is the finite variation part given by

$$A_t = \int_0^t \partial_t f(s, M_s)\, ds + \frac{1}{2} \int_0^t \partial_{xx} f(s, M_s)\, d[M]_s$$

and $N \in \mathcal{M}_c^2$ is the martingale part given by

$$N_t = \int_0^t \partial_x f(s, M_s)\, dM_s$$

In particular, we know that the quadratic variation of $F(M)$ is

$$[F(M)]_t = \int_0^t \left(\partial_x f(s, M_s)\right)^2\, d[M]_s$$

A.4 Stochastic Differential Equations

Given a Brownian motion B on some filtered probability space, $x \in \mathbb{R}$ and measurable functions $\sigma, b : [0,1] \times \mathbb{R} \to \mathbb{R}$, we consider the stochastic differential equation (SDE)

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

A continuous, adapted process $X = (X_t)_{t\in[0,1]}$ for which the above equations hold a.s. shall be referred to as a *(strong) solution* of $(\star)$. The existence and uniqueness of a solution is guaranteed under some basic assumptions on the coefficients as stated by the following Theorem.

Theorem A.11 (*Existence and Uniqueness Theorem for SDEs*)**.** If there are constants $C, L > 0$ such that

(A1) $|\sigma(t,x)| \le C(1 + |x|)$ for all $x \in \mathbb{R}$, $t \in [0,1]$

(A2) $|\sigma(t,x) - \sigma(t,y)| \le L|x-y|$ for all $x, y \in \mathbb{R}$, $t \in [0,1]$

then there exists a unique strong solution to $(\star)$.

Proof. see ([14], Theorem 5.2.1).

□

Given a solution X to $(\star)$, we say that X has a *transition probability density* if for every $s < t \in [0,1]$ and $x \in \mathbb{R}$ the density $p[X_t \mid X_s = x]$ exists. The following Theorem shows how assumption (A1) and (A2) can be strengthened to ensure the existence of a transition probability density of the solution.

Theorem A.12 (*Existence of Transition Probability Density*)**.** If there are constants $\delta \in (0,1]$ and $L > 0$ such that

(B1) $\delta \le \sigma(t,x) \le \frac{1}{\delta}$ for all $x \in \mathbb{R}$, $t \in [0,1]$

(B2) $|\sigma(s,x) - \sigma(t,y)| \le L(|t-s| + |y-x|)$ for all $x, y \in \mathbb{R}$ and $s, t \in [0,1]$

then the unique strong solution to $(\star)$ has a transition probability density.

Proof. see ([10], Chapter 6, Theorem 5.4)

□

Bibliography

[1] D. Aronson. “Bounds for the fundamental solution of a parabolic equation”. In: *Bulletin of the American Mathematical Society* 73 (1967).

[2] R. Azencott. “Densité des diffusions en temps petit développements asymptotiques”. In: *Séminaire de Probabilités XVIII*. Springer, 1984, pp. 402–498.

[3] J. Backhoff-Veraguas and M. Beiglböck. “The most exciting game”. In: *Electronic Communications in Probability* (2024).

[4] J. Backhoff-Veraguas, M. Beiglböck, W. Schachermayer, and Bertram Tschiderer. *Existence of Bass martingales and the martingale Benamou−Brenier problem in* $\mathbb{R}^d$. 2025. arXiv: 2306.11019 [math.PR].

[5] J. Backhoff-Veraguas, W. Schachermayer, and B. Tschiderer. *The Bass functional of martingale transport*. 2023. arXiv: 2309.11181 [math.PR].

[6] J. Backhoff-Veraguas and C. Unterberger. “On the specific relative entropy between martingale diffusions on the line”. In: *Electronic Communications in Probability* (2023).

[7] J.-D. Benamou, G. Chazareix, M. Hoffmann, G. Loeper, and F.-X. Vialard. *Entropic Semi-Martingale Optimal Transport*. 2024. arXiv: 2408.09361 [math.OC].

[8] H. Föllmer. “Optimal Couplings on Wiener Space and an Extension of Talagrand’s Transport Inequality”. In: *Stochastic Analysis, Filtering, and Stochastic Optimization*. Springer, 2022, pp. 147–175.

[9] D. Freedman. *Brownian Motion and Diffusion*. Springer, 1983.

[10] A. Friedman. *Stochastic Differential Equations and Applications, Volume 1*. Academic Press, 1975.

[11] J.-F. Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer, 2016.

[12] N. Gantert. “Einige große Abweichungen der Brownschen Bewegung”. PhD thesis. Universität Bonn, 1991.

[13] A. Klenke. *Probability Theory: A Comprehensive Course (2nd Edition)*. Springer, 2013.

[14] B. Øksendal. *Stochastic Differential Equations*. Springer, 2003.

[15] S. Shreve. *Stochastic Calculus for Finance II*. Springer, 2004.

[16] C. Unterberger. “On the Specific Relative Entropy between Martingale Measures”. MA thesis. University of Vienna, 2022.