MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Specific relative Entropy between continuous martingales.“

verfasst von / submitted by

Dott. Edoardo Kimani Bellotto

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

Master of Science (MSc)

Wien, 2024 / Vienna, 2024

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

Abstract

In this Master Thesis we deal with the concept of *specific relative entropy*, an important tool in stochastic analysis for comparing continuous martingales. In particular, we focus on the Brownian motion and its derivations, providing applications and examples in the interest of financial mathematics.
The first part of the thesis takes on a more theoretical aspect. We start describing the notion of *relative entropy* [Chapter 1] in order to elaborate the concept of *specific relative entropy* [Chapter 2], while the second part of the thesis is dedicated to the applications. Indeed, we will compute the specific relative entropy of Brownian motions and geometric Brownian motions [Chapter 3]. Finally, through some simulations on Python, we will study the specific relative entropy of the "Heston Model" and the "SABR-Model" in comparison to the geometric Brownian motion [Chapter 4].
This work can be seen as a widening of the Master Thesis of Clara Unterberger [1], since we will expand many results of it to the multivariate case. For this reason we will often refer to Unterberger's paper, altough many statements in [1] actually come from the PhD thesis of Nina Gantert [2].

Zusammenfassung

In der vorliegenden Masterarbeit steht die sogenannte *specific relative entropy* im Zentrum. Es handelt sich hierbei um ein wichtiges Werkzeug in der stochastischen Analyse zum Vergleich kontinuierlicher Martingale. Der Fokus liegt auf der „Brownian motion“ und ihren Ableitungen. Vorgestellt werden diese unter anderem mithilfe von Anwendungsbeispielen in der Finanzmathematik.
Der erste Teil der Masterarbeit beschäftigt sich mit der theoretischen Auseinandersetzung zur Thematik. Zunächst wird das Konzept der relativen Entropie herausgearbeitet [Kapitel 1], um dann anschließend das Modell der spezifischen relativen Entropie zu etablieren [Kapitel 2]. Der zweite wesentliche Teil dieser Arbeit konzentriert sich auf Anwendungsbeispiele [Kapitel 3], wobei relative Entropien der „Brownian motion“ und geometrische „Brownian motion“ berechnet werden. Schlussendlich wird die spezifische relative Entropie des „Heston-Model“ und des „SABR-Model“ in Vergleich zur geometrischen „Brownian motion“ mithilfe von Python thematisiert [Kapitel 4].
Diese Arbeit kann als Erweiterung der Masterarbeit von Clara Unterberger angesehen werden, da die Resultate aus [1] in mehrere Dimensionen verallgemeinert weren. Aus diesem Grund werden Bezüge zu Unterberger hergestellt, wobei einige Aspekte aus [1] ursprünglich der Doktorarbeit von Nina Gantert [2] entstammen.

To my girlfriend, my family and my friends, with a big thanks to my supervisor.

Contents

1 Relative Entropy

In this chapter we assume familiarity with the concept of relative entropy through examples and properties. For instance, we compute the relative entropy between two multivariate Gaussians.

Definition 1.1. [1] Let μ and ν be probability measures on the same measurable space $(\Omega, \mathcal{F})$. Then the relative entropy between μ and ν is defined as,

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

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

Remark 1. [1]. The relative entropy can be interpreted as a measure of "distance" between two measures defined on the same space. Indeed, it is always non negative and it is null when the two measures coincides. However, the triangle inequality does not hold, therefore the relative entropy does not define a metric.

We recall that the trace of a square matrix $A \in \mathbb{R}^{l\times l}$ is the sum of the diagonal elements of A, i.e. $tr(A) := \sum_{i=1}^{l} A_{ii}$. Furthermore, the trace is a linear map and it is commutative under the product of matrices. We also remind that if $x \in \mathbb{R}^l$ then $tr(Axx^T) = x^T Ax$, where x^T is the transpose of x.

Example 1.1. Given two l-multivariate normal distributions $\mathbb{P}_i \sim \mathcal{N}_l(\mu_i, \Sigma_i)$, where $i = 1, 2$, $l \in \mathbb{N}$, $\mu_i \in \mathbb{R}^l, \Sigma_i \in \mathbb{R}^{l\times l}$, then the relative entropy between the two variables is:

$$H\big(\mathbb{P}_1|\mathbb{P}_2\big) = \frac{1}{2}\big(tr(\Sigma_2^{-1}\Sigma_1) - l + (\mu_2 - \mu_1)^T\Sigma_2^{-1}(\mu_2 - \mu_1) - \log(\frac{\det \Sigma_1}{\det \Sigma_2})\big), \tag{1}$$

where $tr(.)$ indicates the trace of a matrix and $.^T$ its transpose.

Proof. By definition:

$$\begin{aligned} H(\mathbb{P}_1|\mathbb{P}_2) &= \int \log \frac{d\mathbb{P}_1}{d\mathbb{P}_2} d\mathbb{P}_1 \\ &= \mathbb{E}_{\mathbb{P}_1}[\log p_1 - \log p_2] \end{aligned} \tag{2}$$

where p_1 and p_2 are the distributions of the two variables, i.e $\forall x \in \mathbb{R}^l$, $i = 1, 2$:

$$p_i(x) = \frac{1}{\sqrt{(2\pi)^l \det \Sigma_i}} \exp\left(-\frac{1}{2}(x - \mu_i)^T\Sigma_i^{-1}(x - \mu_i)\right).$$

and the logarithms of the distributions are:

$$\log p_i(x) = -\frac{l}{2}\log 2\pi - \frac{1}{2}\log(\det \Sigma_i) - \frac{1}{2}(x - \mu_i)^T\Sigma_i^{-1}(x - \mu_i).$$

Thus, by some linear algebra and the properties of the logarithms, (2) is equal to:

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}_1}\big[-\log(\det\Sigma_1)-(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)+\log(\det\Sigma_2)+(x-\mu_2)^T\Sigma_2^{-1}(x-\mu_2)\big]$$
$$=-\frac{1}{2}\log\big(\frac{\det\Sigma_1}{\det\Sigma_2}\big)+\frac{1}{2}\mathbb{E}_{\mathbb{P}_1}\big[-(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)+(x-\mu_2)^T\Sigma_2^{-1}(x-\mu_2)\big]$$
$$=-\frac{1}{2}\log\big(\frac{\det\Sigma_1}{\det\Sigma_2}\big)+\frac{1}{2}\mathbb{E}_{\mathbb{P}_1}\big[-tr\big(\Sigma_1^{-1}(x-\mu_1)(x-\mu_1)^T\big)+tr\big(\Sigma_2^{-1}(x-\mu_2)(x-\mu_2)^T\big)\big].$$

Now using the following:

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}_1}[(x-\mu_1)(x-\mu_1)^T]=Var_{P_1}[x]=\Sigma_1\\&E_{\mathbb{P}_1}[xx^T]=Var_{\mathbb{P}_1}[x]+\mathbb{E}_{\mathbb{P}_1}[x]\mathbb{E}_{\mathbb{P}_1}[x]^T=\Sigma_1+\mu_1\mu_1^T\\&\mathbb{E}_{\mathbb{P}_1}[x]=\mu_1.\end{aligned}$$

We obtain:

$$=-\frac{1}{2}\log\big(\frac{\det\Sigma_1}{\det\Sigma_2}\big)+\frac{1}{2}\mathbb{E}_{\mathbb{P}_1}\big[-tr(\Sigma_1^{-1}\Sigma_1)+tr(\Sigma_2^{-1}(xx^T-x\mu_2^T-\mu_2x^T+\mu_2\mu_2^T))\big]$$
$$=-\frac{1}{2}\log\big(\frac{\det\Sigma_1}{\det\Sigma_2}\big)-\frac{1}{2}l+\frac{1}{2}tr\big(\Sigma_2^{-1}(\Sigma_1+\mu_1\mu_1^T-\mu_1\mu_2^T-\mu_2\mu_1^T+\mu_2\mu_2^T)\big).$$

Finally, by the linearity of the trace and some linear algebra:

$$=\frac{1}{2}\big(-\log\big(\frac{\det\Sigma_1}{\det\Sigma_2}\big)-l+tr(\Sigma_2^{-1}\Sigma_1)+tr(\mu_1^T\Sigma_2^{-1}\mu_2-2\mu_1^T\Sigma_2^{-1}\mu_2+\mu_2^T\Sigma_2^{-1}\mu_2)\big)$$
$$=\frac{1}{2}\big(-\log\big(\frac{\det\Sigma_1}{\det\Sigma_2}\big)-l+tr(\Sigma_2^{-1}\Sigma_1)+(\mu_2-\mu_1)^T\Sigma_2^{-1}(\mu_2-\mu_1)\big).$$

□

Remark 2. If the two normal variables have the same mean $\mu\in\mathbb{R}^l$ then:

$$H(\mathcal{N}_l(\mu,\Sigma_1)|\mathcal{N}_l(\mu,\Sigma_2))=\frac{1}{2}\Big(tr(\Sigma_2^{-1}\Sigma_1)-l-\log(\frac{\det\Sigma_1}{\det\Sigma_2})\Big).$$

Definition 1.2. The previous remark suggests us to define the function $F_l(X):=\frac{1}{2}(tr(X)-l-\log(\det X))$ for every positive definite matrix $X\in\mathbb{R}^{l\times l}$. In this way, we have that $H(\mathcal{N}_l(\mu,\Sigma_1)|\mathcal{N}_l(\mu,\Sigma_2))=F_l(\Sigma_2^{-1}\Sigma_1)$.

Example 1.2. [1]. If we restrict ourselves to the univariate case then, given $l=1$, $m_1,m_2\in\mathbb{R}$ and $\sigma_1^2,\sigma_2^2>0$, we have:

$$H(\mathcal{N}_1(m_1,\sigma_1^2)|\mathcal{N}_1(m_2,\sigma_2^2))=\frac{1}{2}\Big(\frac{\sigma_1^2}{\sigma_2^2}-1+\frac{(m_2-m_1)^2}{\sigma_2^2}-\log\frac{\sigma_1^2}{\sigma_2^2}\Big).$$

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

$$H(\mathcal{N}_1(m_1,\sigma_1^2)|\mathcal{N}_1(m_1,\sigma_2^2))=F_1(\frac{\sigma_1^2}{\sigma_2^2}),$$

where $F_1(x)=\frac{1}{2}(x-1-\log x)$.

For clarity, as $l \in \mathbb{N}$, we refer to the $l \times l$-identity matrix with the symbol $\mathcal{I}$, to the $l \times l$-null matrix with the symbol $\mathcal{O}_{l\times l}$ and to the l-null vector with the symbol 0_l.

Lemma 1.1. *1. Let $\alpha \in \mathbb{R}$ then $F_l(\alpha\mathcal{I}) = lF_1(\alpha)$, where $\mathcal{I}$ is the l-identity matrix.*

2. Let $D = diag(\sigma_i^2)$ be a diagonal matrix with $\sigma_i^2 \in \mathbb{R}$ and $i = 1, \ldots, l$. Then $F_l(D) = \sum_{i=1}^l F_1(\sigma_i^2)$.

3. Let $\Sigma \in \mathbb{R}^{l\times l}$ diagonalizable l-matrix, and let $D = A^{-1}\Sigma A = diag(eigenvalues\, of\, \Sigma)$. Then $F_l(\Sigma) = F_l(D) = \sum_{i=1}^l F_1(eigenvalues\, of\, \Sigma)$.

Proof. By Remark 2 and some linear algebra we get the three properties:

$$\begin{aligned} F_l(\alpha\mathcal{I}) &= \frac{1}{2}\big(tr(\alpha\mathcal{I}) - l - \log(\det(\alpha\mathcal{I}))\big) \\ &= \frac{1}{2}(l\alpha - l - l\log\alpha) \\ &= \frac{l}{2}(\alpha - 1 - l\log\alpha) \\ &= lF_1(\alpha). \end{aligned}$$

This shows *1*, while for *2*:

$$\begin{aligned} F_l(D) &= \frac{1}{2}\big(tr(D) - l - \log(\det(D))\big) \\ &= \frac{1}{2}(\sum_{i=1}^l \sigma_i^2 - \sum_{i=1}^l 1 - \sum_{i=1}^l \log\sigma_i^2) \\ &= \sum_{i=1}^l \frac{1}{2}(\sigma_i^2 - 1 - \log\sigma_i^2) \\ &= \sum_{i=1}^l F_1(\sigma_i^2). \end{aligned}$$

Finally, we get *3*:

$$\begin{aligned} F_l(\Sigma) &= \frac{1}{2}\big(tr(\Sigma) - l - \log(det\Sigma)\big) \\ &= \frac{1}{2}(tr(D) - l - \log(detD)) \\ &= F_l(D) = \sum_{i=1}^l F_1((eigenvalues\, of\, \Sigma). \end{aligned}$$

□

For a later use we look at other examples and we recall the following definition.

Definition 1.3. A stochastic process M_t is said to be a standard geometric Brownian motion if $dM_t := M_t dB_t$, i.e. if $M_t := e^{B_t - \frac{t}{2}}$ with B_t the usual Brownian motion. In particular, M_t has distribution $lognormal(-\frac{t}{2}, t)$.

Remark 3. We recall that if $X \sim lognormal(\mu, \sigma^2)$ then $\log(X) \sim \mathcal{N}_1(\mu, \sigma^2)$. Moreover, $\mathbb{E}[X] = e^{\mu + \frac{\sigma^2}{2}}$ and $Var[X] = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$.

Example 1.3. The relative entropy between a standard geometric Brownian motion M_t and a Brownian motion B_t, both starting at 1, is:

$$H(\mathcal{L}(M_t)|\mathcal{L}(B_t)) = \frac{1}{2}(t - 1 + \frac{e^t - 1}{t}).$$

Proof. Let $B_t \sim \mathcal{N}_1(1, t)$ be the Brownian motion with $B_0 = 1$ and let b_t be its density function. Let $M_t := e^{B_t - 1 - \frac{t}{2}} \sim lognormal(-\frac{t}{2}, t)$ be the standard geometric Brownian motion with $M_0 = e^{B_0 - 1 - 0} = 1$ and let m_t be its densitiy function. We first look at the logarithms of the respective densities:

$$\begin{aligned}
\log(m_t(x)) &= \log(\frac{1}{\sqrt{2\pi t x^2}} \exp(-(\frac{(\log x + \frac{t}{2})^2}{2t})) \\
&= -\frac{1}{2}\log(2\pi t) - \log x - \frac{1}{2t}(\log^2 x + \frac{t^2}{4} + t\log x)
\end{aligned}$$

and,

$$\log(b_t(x)) = \log(\frac{1}{\sqrt{2\pi t}} \exp(\frac{-(x-1)^2}{2t})) = -\frac{1}{2}\log(2\pi t) - \frac{1}{2t}(x^2 + 1 - 2x).$$

For the expectations we need from Remark 3:

$$\begin{aligned}
&\mathbb{E}\big[M_t\big] = e^{-\frac{t}{2} + \frac{t}{2}} = 1 \\
&Var\big[M_t\big] = (e^t - 1)e^{t-t} = e^t - 1 \\
&\mathbb{E}\big[M_t^2\big] = e^t - 1 + 1 = e^t \\
&\log(M_t) \sim \mathcal{N}_1(-\frac{t}{2}, t).
\end{aligned}$$

Hence we obtain:

$$\begin{aligned}
\mathbb{E}_{M_t}\big[\log(m_t(x))\big] &= -\frac{1}{2}\log 2\pi t + \frac{t}{2} - \frac{1}{2t}(t + \frac{t^2}{4} + \frac{t^2}{4} - \frac{t^2}{2}) \\
&= -\frac{1}{2}\log 2\pi t + \frac{t}{2} - \frac{1}{2}
\end{aligned}$$

and,

$$\mathbb{E}_{M_t}\big[\log(b_t(x))\big] = -\frac{1}{2}\log(2\pi t) - \frac{e^t - 1}{2t}.$$

Finally, we get the desired relative entropy:

$$\begin{aligned}
H(\mathcal{L}(M_t)|\mathcal{L}(B_t)) &= \mathbb{E}_{M_t}\big[\log(m_t(x)) - \log(b_t(x))\big] \\
&= \mathbb{E}_{M_t}[\log(m_t(x))] - \mathbb{E}_{M_t}[\log(b_t(x))] \\
&= \frac{t}{2} - \frac{1}{2} + \frac{e^t - 1}{2t} \\
&= \frac{1}{2}(t - 1 + \frac{e^t - 1}{t}).
\end{aligned}$$

Remark 4. Note that if $t \to 0$ we expect that $H(\mathcal{L}(M_t)|\mathcal{L}(B_t)) \to H(\mathcal{L}(M_0)|\mathcal{L}(B_0)) = 0$ since $B_0 = M_0 = 1$. Indeed, in Example 1.3 we have $H(\mathcal{L}(M_t)|\mathcal{L}(B_t)) = \frac{1}{2}(t - 1 + \frac{e^t-1}{t})$ which vanishes as $t \to 0$.

□

We will also use often the following example.

Example 1.4. Let $l \in \mathbb{R}$. Given two l-multivariate lognormal distributions $\mathbb{Q}_i \sim lognormal(\mu_i, \Sigma_i)$ where $i = 1, 2$, $l \in \mathbb{N}$, $\mu_i \in \mathbb{R}^l$, $\Sigma_i \in \mathbb{R}^{l\times l}$ then the relative entropy between the two variables coincides with the relative entropy between two l-multivariate normal variables with the same parameters, i.e.:

$$H(lognormal(\mu_1, \Sigma_1)|lognormal(\mu_2, \Sigma_2)) = H(\mathcal{N}_l(\mu_1, \Sigma_1)|\mathcal{N}_l(\mu_2, \Sigma_2)).$$

To show this, we consider the lognormal densities as $x \in \mathbb{R}^l$:

$$q_i(x) := \frac{1}{\sqrt{(2\pi)^l \det \Sigma_i}} \exp\big(-\frac{1}{2}(x - \mu_i)^T \Sigma^{-1}(x - \mu_i)\big)$$

and taking their logarithms:

$$\begin{aligned}
\log(q_i(x)) &= -\frac{1}{2}\log\big((2\pi)^l \det \Sigma_i\big) - \frac{1}{2}(x - \mu_i)^T \Sigma^{-1}(x - \mu_i) \\
&= -\frac{1}{2}\log\big((2\pi)^l \det \Sigma_i\big) - \frac{1}{2} tr(\Sigma_i^{-1}(x - \mu_i)(x - \mu_i)^T).
\end{aligned}$$

After noticing that:

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_1}[(x - \mu_1)(x - \mu_1)^T] = Var_{\mathbb{Q}_1}[x] = \Sigma_1 \\
&\mathbb{E}_{\mathbb{Q}_1}[\log x] = \mu_1
\end{aligned}$$

we can look at the expectations:

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_1}[\log q_1(x)] = -\frac{1}{2}\log\big((2\pi)^l \det \Sigma_1\big) - \mu_1 - \frac{1}{2} tr(\mathcal{I}) \\
&\mathbb{E}_{\mathbb{Q}_1}[\log q_2(x))] = -\frac{1}{2}\log\big((2\pi)^l \det \Sigma_2\big) - \mu_1 - \frac{1}{2}(tr(\Sigma_1\Sigma_2^{-1}) - \frac{1}{2}(\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1).
\end{aligned}$$

Thus, we get the relative entropy:

$$\begin{aligned}
H(\mathbb{Q}_1|\mathbb{Q}_2) &= \mathbb{E}_{\mathbb{Q}_1}[\log q_1(x) - \log q_2(x)] \\
&= \frac{1}{2}\big(-\log(\frac{\det \Sigma_1}{\det \Sigma_2}) - l + \frac{1}{2} tr(\Sigma_2^{-1}\Sigma_1) + (\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1)\big).
\end{aligned}$$

Exactly as in (1). Indeed, this is the consequence of a more general result, given by the following Lemma with $T(x) = e^x$.

Lemma 1.2. *[1]. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces, μ and ν probability measures on $(X, \mathcal{A})$ such that $\mu \ll \nu$ and $T : X \to Y$ such that $\mathcal{A} = \sigma(T) := \sigma(\{T^{-1}(B) : B \in \mathcal{B}\})$. Then $T(\mu) \ll T(\nu)$ and for ν-a.e. $x \in X$:*

$$\frac{dT(\mu)}{dT(\nu)}(T(x)) = \frac{d\mu}{dv}(x)$$

and

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

Moreover, in the applications we will often use the following two lemmas.

Lemma 1.3. *[1]. For a random vector (V, W) we denote by $\mu_{V,W}$ its law and by $\mu_W^{V=v}$ a regular conditional distribution of W given $V = v$. Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ be random vectors. Then:*

$$H(\mu_{X_1,\ldots,X_n}|\mu_{Y_1,\ldots,Y_n}) = H(\mu_{X_1}|\mu_{Y_1}) +$$
$$+ \sum_{k=2}^{n} \int H\big(\mu_{X_k}^{X_1=x_1,\ldots,X_{k-1}=x_{k-1}}|\mu_{Y_k}^{Y_1=x_1,\ldots,Y_{k-1}=x_{k-1}}\big) d\mu_{X_1,\ldots,X_{k-1}}(x_1, \ldots, x_{k-1}),$$

which for a later use we can write as:

$$\begin{aligned}
&H\big(\mathcal{L}((X_1, \ldots, X_n))|\mathcal{L}((Y_1, \ldots, Y_n))\big) = \\
&H(\mathcal{L}(X_1)|\mathcal{L}(Y_1)) + \sum_{k=2}^{n} \int H\Big(\mathcal{L}(X_k|X_1 = x_1, \ldots, X_{k-1} = x_{k-1}) \\
&|\, \mathcal{L}(Y_k|Y_1 = x_1, \ldots, Y_{k-1} = x_{k-1})\Big) d\mathcal{L}(X_1, \ldots, X_{k-1}) dx_1 \ldots dx_{k-1}.
\end{aligned}$$

Lemma 1.4. *[1]. Let μ, ν be probability measures on $(X, \mathcal{A})$ such that $H(\mu|\nu) < \infty$. Let $\mathcal{B} \subset \mathcal{A}$ be a sub-sigma algebra and $\mu|_{\mathcal{B}}, \nu|_{\mathcal{B}}$ be the restrictions of μ and ν to $\mathcal{B}$.Then*
(i)

$$\frac{d\mu|_{\mathcal{B}}}{d\nu|_{\mathcal{B}}} = \mathbb{E}_\nu\Big[(\frac{d\mu}{d\nu})|\mathcal{B}\Big].$$

(ii)

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

2 Specific Relative Entropy

In this chapter we assume familiarity with the concept of *specific relative entropy*, through examples and properties. In particular, we show the multidimensional case of Theorem 9 and Theorem 10 of [1].

To do so we consider, as in [1], the Wiener space $(C, \mathcal{F})$, such that $C := C([0,1]) = \{\omega : [0,1] \to \mathbb{R} \,|\, \omega \text{ is continuous}\}$ and $\mathcal{F} := \sigma(X_t : t \in [0,1])$, where $X = (X_t)_{t\in[0,1]}$ is the canonical process, i.e. $X_t(\omega) = \omega(t)$. Further we consider the filtration $\mathcal{F}_t = \sigma(X_s : s \le t)$. Finally, since we are interested to work with laws of continuous martingales we recall the following definitions.

Definition 2.1. We define $\mathcal{M}_1(C)$ as the collection of all probability measures on $(C, \mathcal{F})$. Moreover, a martingale measure is a probability measure $\mathbb{Q} \in \mathcal{M}_1(C)$ such that the canonical process X is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,1]}$ under $\mathbb{Q}$.

Remark 5. The relative entropy as a measure of "distance" is only interesting when there is absolute continuity. Hence, it is not a useful tool for martingale measures, as we can see in Remark 11 of [1]. To go over this case, we define for all $n \in \mathbb{N}$

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

In this way, it might happen that $\mathbb{Q}$ is not absolutely continuous to $\mathbb{P}$ in the whole sigma-algebra $(\mathcal{C}, \mathcal{F})$, while it is in smaller sigma-algebras $\mathcal{F}^n$ for all $n \in \mathbb{N}$.

Definition 2.2. [1]. Let $n \in \mathbb{N}$ and $\mathbb{P}, \mathbb{Q} \in \mathcal{M}_1(\mathcal{C})$ martingale measures such that $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{P}|_{\mathcal{F}^n}$. We denote the Radon-Nikodym density of $\mathbb{Q}|_{\mathcal{F}^n}$ with respect to $\mathbb{P}|_{\mathcal{F}^n}$ on $(\mathcal{C}, \mathcal{F}^n)$ as $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}^n}$ and define:

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

Let $n \in \mathbb{N}$. For $k = 1, \dots, n$, we use the following notation. For the k-th increment of a process X we write $\Delta^n_k X := X_{\frac{k}{n}} - X_{\frac{k-1}{n}}$, while for a function ξ we use the notation $\Delta^n_k \xi := \xi(\frac{k}{n}) - \xi(\frac{k-1}{n})$.
From [1] we get the following and very useful lemma.

Lemma 2.1. *[1]. Let $n \in \mathbb{N}$ and $\mathbb{P}, \mathbb{Q}$ be probability measures on $(C, \mathcal{F})$ such that $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{P}|_{\mathcal{F}^n}$. Then $\mathbb{P}$-a.s.,*

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}^n}(\omega) = \frac{d\mathcal{L}_{\mathbb{Q}}(X_0, \Delta^n_1 X, \dots, \Delta^n_n X)}{d\mathcal{L}_{\mathbb{P}}(X_0, \Delta^n_1 X, \dots, \Delta^n_n X)}(\omega(0), \Delta^n_1 \omega, \dots, \Delta^n_n \omega).$$

In particular,

$$H(\mathbb{Q}|\mathbb{P})|_{\mathcal{F}^n} = H(\mathcal{L}_{\mathbb{Q}}(X_0, \Delta^n_1 X, \dots, \Delta^n_n X)|\mathcal{L}_{\mathbb{P}}(X_0, \Delta^n_1 X, \dots, \Delta^n_n X)).$$

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

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

Remark 6. [1]. Let $\mathbb{Q}, \mathbb{P} \in \mathcal{M}_1(C)$ be martingale measures such that $X_0 = x$ for some $x \in \mathbb{R}$ both $\mathbb{Q}$-a.s. and $\mathbb{P}$-a.s. Then we can disregard the slot for time 0 when computing the relative entropy on the restrictions to $\mathcal{F}^n$. Indeed, we have $\mathbb{P}$-a.s.,

$$\begin{aligned}
&\frac{d\mathcal{L}_{\mathbb{Q}}(X_0, \Delta_1^n X, \ldots, \Delta_n^n X)}{d\mathcal{L}_{\mathbb{P}}(X_0, \Delta_1^n X, \ldots, \Delta_n^n X)}(\omega(0), \Delta_1^n \omega, \ldots, \Delta_n^n \omega) = \\
&= \frac{d\mathcal{L}_{\mathbb{Q}}(X_0, \Delta_1^n X, \ldots, \Delta_n^n X)}{d\mathcal{L}_{\mathbb{P}}(X_0, \Delta_1^n X, \ldots, \Delta_n^n X)}(\Delta_1^n \omega, \ldots, \Delta_n^n \omega),
\end{aligned}$$

and therefore,

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

Definition 2.3. [1]. Let $\mathbb{Q}, \mathbb{P} \in \mathcal{M}_1(C)$ be martingale measures. Then the specific relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as,

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

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

Example 2.1. By Lemma 2.1 we can compute the specific relative entropy between two l-dimensional Brownian motions: $\mathbb{Q} = \mathcal{L}((\sigma B_t)_{t \in [0,1]})$ and $\mathbb{P} = \mathcal{L}((\eta B_t)_{t \in [0,1]})$ with $\sigma, \eta \in \mathbb{R}$.

$$H(\mathbb{Q}|\mathbb{P})|_{\mathcal{F}^n} = \sum_{k=1}^{n} H(\mathcal{L}(\Delta_k^n \sigma B)|\mathcal{L}(\Delta_k^n \eta B)).$$

Now, given the law of the increments:

$$\sigma \mathcal{N}_l(0_l, \frac{k}{n}\mathcal{I}) - \sigma \mathcal{N}_l(0_l, \frac{k-1}{n}\mathcal{I}) \sim \mathcal{N}_l\big(0_l, \frac{\sigma^2}{n}\mathcal{I}\big).$$

Then we have:

$$= \sum_{i=1}^{n} H\big(\mathcal{N}_l(0_l, \frac{\sigma^2}{n}\mathcal{I})|\mathcal{N}_l(0_l, \frac{\eta^2}{n}\mathcal{I})\big)$$

and by Remark 2:

$$= \sum_{i=1}^{n} F_l\big((\frac{\eta^2}{n}\mathcal{I})^{-1}(\frac{\sigma^2}{n}\mathcal{I})\big) = \sum_{i=1}^{n} F_l\big(\frac{\sigma^2}{\eta^2}\mathcal{I}\big).$$

Explicitly:

$$\begin{aligned}
&= \frac{1}{2}\sum_{k=1}^{n} tr(\frac{\sigma^2}{\eta^2}\mathcal{I}) - l - \log\big(\det \frac{\sigma^2}{\eta^2}\mathcal{I}\big) \\
&= \frac{1}{2}\sum_{k=1}^{n} l\frac{\sigma^2}{\eta^2} - l - l\log(\frac{\sigma^2}{\eta^2}) \\
&= n\frac{l}{2}(\frac{\sigma^2}{\eta^2} - 1 - \log\frac{\sigma^2}{\eta^2}).
\end{aligned}$$

Finally, by definition of the specific relative entropy:

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

Note that in the 1-dimensional case $h(\sigma B|\eta B) = F_1(\frac{\sigma^2}{\eta^2}) = \frac{1}{2}(\frac{\sigma^2}{\eta^2} - 1 - \log\frac{\sigma^2}{\eta^2})$.

In the next part of the chapter we want to extend to the multidimensional case the Theorems 9 and 10 of [1]. To do so we consider the same setup used in [1]. In particular, we work with the collection of all the square integrable martingale measures, which we define as $\mathcal{M}_l^2$.

Definition 2.4. A l-dimensional martingale measure $X(.) = (X^1(.), \dots, X^l(.))$ belongs to $\mathcal{M}_l^2$ if $\mathbb{E}\big[(X^i(.))^2\big] < \infty \ \forall i = 1, \dots, l$.

Defined $a(t) := \Sigma(t) \in \mathbb{R}^{l\times l}$ a symmetric positive definite and differentiable matrix such that for times $t \geq s$ we have $a(t) \geq a(s)$ in the sense of positive matrices. Then there exists a symmetric and positive definite matrix $G(s) \in \mathbb{R}^{l\times l}$ such that $a(t) = \int_0^t G(s)ds$. Moreover, $a(t) \in C_0 := \{\omega : [0,1] \to \mathbb{R}^{l\times l}$ s.t $w(0) = \mathcal{O}_{l\times l}\}$.

Remark 7. Let B_t be the standard Brownian motion. Let $G_t \in \mathbb{R}^{l\times l}$ be a deterministic symmetric and positive definite matrix. Defining $\bar{B}_t := \int_0^1 G_s^{1/2} dB_s$, i.e. for $i = 1, \dots, l$ the coordinates are defined as $\bar{B}_t^i := \int_0^t \sum_{k=1}^l (G_s^{1/2})_{ik} dB_s^k$. Moreover, we have:

$$
\mathbb{E}[\bar{B}_t^i \bar{B}_t^j] = \mathbb{E}\Big[\sum_u \sum_v \int_0^t (G_s^{1/2})_{iu} dB_s^u \int_0^t (G_s^{1/2})_{jv} dB_s^v\Big].
$$

Note that when $u \neq v$ we have independence, thus:

$$
\begin{aligned}
&= \sum_u \mathbb{E}\Big[\int_0^t (G_s^{1/2})_{iu} dB_s^u \int_0^t (G_s^{1/2})_{ju} dB_s^u\Big] \\
&= \int_0^t \sum_u (G_s^{1/2})_{iu} (G_s^{1/2})_{ju} ds \\
&= \int_0^t (G_s^{1/2} (G_s^{1/2})^T)_{ij} ds \\
&= \int_0^t (G_s)_{ij} ds.
\end{aligned}
$$

Thus, for every $a \in C_0$ as above, we can define $\mathbb{Q}^a := \mathcal{L}(\bar{B}_t) = \mathcal{L}(\int_0^1 G_s^{1/2} dB_s) \in \mathcal{M}_l^2$ in such a way, by the previous remark, $\mathbb{Q}^a$ is l-Gaussian $\mathcal{N}_l(0_l, a(.))$ and it has independent increments. Note that strictly speaking, $\forall i = 1, \dots, l$, we want that $\int_0^1 G_{ii}(s)ds < \infty$. Finally, we denote with $\mathbb{B}$ the law of the standard Brownian motion B_t. Now we can see how Theorem 9 in [1] becomes with this multivariate setup.

Theorem 2.2. *In the previous environment we have:*

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

Moreover as $n \to \infty$,

$$\frac{1}{n}\log(\frac{d\mathbb{Q}^a}{d\mathbb{B}})|_{\mathcal{F}^n} \xrightarrow{L^1(\mathbb{Q}^a)} \frac{1}{2}\int_0^1 \big(tr(G(s)) - l - \log(\det G(s))\big)ds.$$

Before starting with the proof we need the following remark.

Remark 8. The function $F_l(X) := \frac{1}{2}(tr(X) - l - \log(\det X))$, with $l \in \mathbb{N}$ and $X \in \mathbb{R}^{l\times l}$ positive definite matrix, is convex. This means that given $\lambda \in \mathbb{R}$ and $X, Y \in \mathbb{R}^{l\times l}$, we have $\lambda F_l(X) + (1-\lambda)F_l(Y) \geq F_l(\lambda X + (1-\lambda)Y)$. Indeed:

$$\begin{aligned} &\lambda F_l(X) + (1-\lambda)F_l(Y) = \\ &= \lambda\frac{1}{2}\big(tr(X) - l - \log(\det X)\big) + (1-\lambda)\frac{1}{2}\big(tr(Y) - l - \log(\det Y)\big) \\ &= \frac{1}{2}\big(\lambda tr(X) + (1-\lambda)tr(Y) - l - \lambda\log(\det X) - (1-\lambda)\log(\det Y)\big). \end{aligned}$$

And on the other side:

$$\begin{aligned} &F_l(\lambda X + (1-\lambda)Y) = \\ &= \frac{1}{2}\big(tr(\lambda X + (1-\lambda)Y) - l - \log(\det(\lambda X + (1-\lambda)Y))\big) \\ &= \frac{1}{2}\big(\lambda tr(X) + (1-\lambda)tr(Y) - l - \log(\det \lambda X + (1-\lambda)Y)\big). \end{aligned}$$

So we want to show that:

$$-\log(\det(\lambda X + (1-\lambda)Y)) \leq -\lambda\log(\det X) - (1-\lambda)\log(\det Y).$$

That is:

$$\log(\det(\lambda X + (1-\lambda)Y)) \geq \lambda\log(\det X) + (1-\lambda)\log(\det Y).$$

Therefore, we want to show the concavity of the function $\log\det(.)$, i.e.:

$$\log\big(\det(\lambda X + (1-\lambda)Y)\big) \geq \lambda\log(\det X) + (1-\lambda)\log\det Y.$$

So passing to the exponentials:

$$\det(\lambda X + (1-\lambda)Y) \geq (\det X)^{\lambda}(\det Y)^{1-\lambda}.$$

Since, without loss of generality, X is positive definite such that $X^{1/2}X^{1/2} = X$ the left hand side can be written as:

$$\det(X^{1/2}(\lambda\mathcal{I} + (1-\lambda)X^{-1/2}YX^{-1/2})X^{1/2}) = \det X \det(\lambda\mathcal{I} + (1-\lambda)Z)$$

for $Z := X^{-1/2}YX^{-1/2}$. While the right hand side:

$$\det X^{1/2}(\det X^{-1/2}YX^{-1/2})^{1-\lambda} \det X^{1/2} = \det X^{1/2}(\det Z)^{1-\lambda} \det X^{1/2}.$$

So without loss of generality we want to show that

$$\det(\lambda\mathcal{I} + (1-\lambda)Z) \geq (\det Z)^{1-\lambda}.$$

Further, we notice that Z is clearly symmetric and also positive definite. Indeed, the smallest eigenvalue of Z is given by

$$\begin{aligned}\min_{|\mu|=1} \mu^T(X^{-1/2}YX^{-1/2})\mu &= \min_{|\mu|=1} (X^{-1/2}\mu)^T Y (X^{-1/2}\mu) \\ &= \min_{|\mu|=1} |X^{-1/2}\mu|^2 \frac{(X^{-1/2}\mu)^T}{|X^{-1/2}\mu|} Y \frac{(X^{-1/2}\mu)}{|X^{-1/2}\mu|}\end{aligned}$$

which by the same argument we know to be greater of the smallest eigenvalue of Y which is positive since Y is positive definite. Now considering the eigenvalues of Z, $(\gamma_i)_{i=1,\dots,l}$ it remains to show:

$$\prod_{i=1}^{l}(\lambda + (1-\lambda)\gamma_i) \geq \prod_{i=1}^{l} \gamma_i^{1-\lambda}$$

and passing to the logarithms:

$$\sum_{i=1}^{l} \log(\lambda + (1-\lambda)\gamma_i) \geq (1-\lambda)\sum_{i=1}^{l} \log \gamma_i$$

which holds by the concavity of the logarithm term by term. The concavity of $F_l(.)$ allows us to use the Jensen inequality.

Proof of Theorem 2.2. We proceed in parallel with the proof of Theorem 9 in [1]. Under $\mathbb{Q}^a$ the process X is l-Gaussian with independent increments and for all $s < t$,

$$\begin{aligned}&X_t - X_s \sim_{\mathbb{Q}^a} \mathcal{N}_l(0_l, a(t) - a(s)) \\ &B_{\frac{k}{n}} - B_{\frac{k-1}{n}} \sim \mathcal{N}_l(0_l, \frac{1}{n}\mathcal{I}).\end{aligned}$$

Recall that $H(\mathcal{N}_l(0_l, \Sigma_1)|\mathcal{N}_l(0_l, \Sigma_2)) = F_l(\Sigma_2^{-1}\Sigma_1)$ where $F_l(X) = \frac{1}{2}(tr(X) - l - \log\det(X)$. By Lemma 2.1 and independence of increments under $\mathbb{Q}^a$ as well as under $\mathbb{B}$, we can write:

$$\begin{aligned} H(\mathbb{Q}^a|\mathbb{B})|_{\mathcal{F}^n} &= \sum_{k=1}^{n} H(\mathcal{L}_{\mathbb{Q}^a}(\Delta_k^n X)|\mathcal{L}_B(\Delta_k^n X)) \\ &= \sum_{k=1}^{n} H(\mathcal{N}_l(0_l, \Delta_k^n a(.))|\mathcal{N}_l(0_l, \frac{1}{n}\mathcal{I})) \\ &= \sum_{k=1}^{n} F_l\big((\frac{1}{n}\mathcal{I})^{-1}\Delta_k^n a(.)\big). \end{aligned} \tag{3}$$

Now let $\mathcal{B}_n := \sigma\big((\frac{k-1}{n}, \frac{k}{n}] : k = 1, \ldots, n\big)$. Since we have $a(t) = \int_0^t G(s)ds$, where G is symmetric and positive definite,

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

with λ the Lebesgue density measure on $[0, 1]$. Moreover, we get in (3):

$$\begin{aligned} \sum_{k=1}^{n} F_l(n\Delta_k^n a(.)) &= n \int_0^1 F_l(n\Delta_k^n a(.)) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)ds \\ &= n \int_0^1 F_l\big(\mathbb{E}_\lambda[G|\mathcal{B}_n](s)\big) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)ds \end{aligned}$$

and dividing by n, we obtain,

$$\frac{1}{n} H(\mathbb{Q}^a|\mathbb{B})|_{\mathcal{F}^n} = \int_0^1 F_l(\mathbb{E}_\lambda[G|\mathcal{B}_n](s)) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(s)ds.$$

We now use that the function F is convex, Remark 8, so by Jensen's conditional inequality and the tower property of conditional expectation we obtain an upper bound for all $n \in \mathbb{N}$,

$$\begin{aligned} \frac{1}{n} H(\mathbb{Q}^a|\mathbb{B})|_{\mathcal{F}^n} &= \int_0^1 F_l(\mathbb{E}_\lambda[G|\mathcal{B}_n](s))ds \\ &\leq \int_0^1 \mathbb{E}_\lambda[F(G)|\mathcal{B}_n](s))ds \\ &= \int_0^1 F_l(G(s))ds. \end{aligned}$$

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

$$\begin{aligned}\int_0^1 F_l(G(s))ds &= \int_0^1 \liminf_{n\to\infty} F_l(\mathbb{E}_\lambda[G|\mathcal{B}_n](s))ds\\ &\leq \liminf_{n\to\infty} \int_0^1 F_l(\mathbb{E}_\lambda[G|\mathcal{B}_n](s))ds\\ &\leq \limsup_{n\to\infty} \int_0^1 F_l(\mathbb{E}_\lambda[G|\mathcal{B}_n](s))ds\\ &\leq \int_0^1 F_l(G(s))ds,\end{aligned}$$

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

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

which proves the first part of the theorem.
Now we assume that $\int_0^1 tr(G(s)) - l - \det G(s) ds < \infty$ and in particular $\Delta_k^n a(.) > 0$ for all $n \in \mathbb{N}$ and $k = 1, \ldots, n$. To show L^1 convergence, note that applying Lemma 2.1 plus some linear algebra:

$$\begin{aligned}&\frac{1}{n}\log\frac{d\mathbb{Q}^a}{d\mathbb{B}}|_{\mathcal{F}^n}(X) = \frac{1}{n}\log\frac{d\mathcal{N}_l(0_l, \Delta_k^n a(.))}{d\mathcal{N}_l(0_l, \frac{1}{n}\mathcal{I})}(\Delta_k^n X)\\ &= \frac{1}{2n}\sum_{k=1}^n \big(-\log(\det \Delta_k^n a(.)) - (\Delta_k^n X - 0_l)^T (\Delta_k^n a(.))^{-1}(\Delta_k^n X - 0_l) +\\ &+ \log(\det \frac{1}{n}\mathcal{I}) + (\Delta_k^n X - 0_l)^T (\frac{1}{n}\mathcal{I})^{-1}(\Delta_k^n X - 0_l)^T\big) =\\ &= \frac{1}{2n}\sum_{k=1}^n tr(n\mathcal{I}\Delta_k^n X(\Delta_k^n X)^T) - tr((\Delta_k^n a(.))^{-1}\Delta_k^n X(\Delta_k^n X)^T) - log(n^l \det \Delta_k^n a(.)) =\\ &= tr(\frac{1}{2n}\sum_{k=1}^n n\mathcal{I}\Delta_k^n X(\Delta_k^n X)^T) - tr(\frac{1}{2n}\sum_{k=1}^n (\Delta_k^n a(.))^{-1}\Delta_k^n X(\Delta_k^n X)^T) +\\ &- \frac{1}{2n}\sum_{k=1}^n log(n^l \det \Delta_k^n a(.)).\end{aligned} \tag{4}$$

Looking at the one dimensional case in [1] we get
$\sum_{k=1}^{n}(\Delta_k^n X)^2 \xrightarrow{L^2(\mathbb{Q}^a)} a(1)$ as $n\to\infty$, while $\frac{1}{n}\sum_{k=1}^{n}\frac{(\Delta_k^n X)^2}{\Delta_k^n a} \xrightarrow{L^2(\mathbb{Q}^a)} 1$ as $n\to\infty$. In the same way we have for the first term of (4):

$$\frac{1}{n}\sum_{k=1}^{n} n\mathcal{I}\Delta_k^n X(\Delta_k^n X)^T \xrightarrow{n\to\infty} a(1) = \int_0^1 G(s)ds.$$

While for the second term:

$$\frac{1}{n}\sum_{k=1}^{n}(\Delta_k^n a(.))^{-1}\Delta_k^n X(\Delta_k^n X)^T \xrightarrow{n\to\infty} \mathcal{I}.$$

For the last term, we note that from the first part of the proof we have:

$$\lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}^a}\Big[\frac{1}{2n}\sum_{k=1}^{n} tr(n\mathcal{I}\Delta_k^n X(\Delta_k^n X)^T) - tr((\Delta_k^n a(.))^{-1}\Delta_k^n X(\Delta_k^n X)^T)+$$
$$-\log(n^l \det\Delta_k^n a(.))\Big] = \int_0^1 F_l(G(s))ds = \frac{1}{2}\int_0^1 tr(G(s)) - l - \log(\det G(s))ds.$$

Therefore, we have obtained the deterministic term:

$$-log(n^l \det\Delta_k^n a(.)) \xrightarrow{n\to\infty} \int_0^1 -\log(\det G(s))ds.$$

In conclusion, we have that (4) as $n\to\infty$ becomes:

$$\frac{1}{2}\big(tr(a(1)) - tr(\mathcal{I} - \int_0^1 \log\det G(s)ds)\big) =$$
$$= \frac{1}{2}\int_0^1 \big(tr(G(s)) - l - \log(\det G(s)\big)ds =$$
$$= \int_0^1 F(G(s))ds.$$

□

For the next theorem we recall the following definition:

Definition 2.5. Let $T>0$, $t\in[0,T]$ and $l\in\mathbb{N}$. The quadratic variation for a multivariate process $X_t^l = (X_t^1,\ldots,X_t^l)$ is the matrix $\Sigma := \langle X\rangle$ such that $(\Sigma)_{ij} := \langle X^i, X^j\rangle_T$ with $i,j=1,\ldots,l$ and

$$\langle X^i, X^j\rangle_T = \lim_{|\Pi|\to 0}\sum_{k=1}(X_{t_k}^i - X_{t_{k-1}}^i)(X_{t_k}^j - X_{t_{k-1}}^j)$$

where $\Pi = \{0=t_0<t_1<\cdots<t_n=T\}$ is a partition of $[0,T]$ and $|\Pi|$ is the mesh size of the partition.

Remark 9. In the previous theorem we had $\langle X \rangle_t$ under $\mathbb{Q}^a$ is $a(t) = \int_0^t G(s)ds$.

Now we proceed with extending Theorem 10 in [1], where we consider more general martingale measures and we obtain a lower bound for the specific relative entropy along the dyadic sub-sequence.

Theorem 2.3. *Let B_t the standard Brownian motion and $\mathbb{B}$ its law. Let $\mathbb{Q} \in \mathcal{M}_l^2$ be such that $X_0 = 0$ a.s. and $\langle X \rangle_t = \int_0^1 G(s,x)ds$, with $G \in \mathbb{R}^{l\times l}$ symmetric positive definite and predictable matrix. Then we have:*

$$\liminf_{n\to\infty} \frac{1}{2^n} H(\mathbb{Q}|\mathbb{B})|_{\mathcal{F}^{2^n}} \geq \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\Big[\int_0^1 tr(G(s,X)) - l - \log(\det G(s,X))ds\Big]. \quad (5)$$

Proof. We proceed in parallel with the proof of Theorem 10 in [1]. Let $n \in \mathbb{N}$. We start by defining a new measure on $(C, \mathcal{F}^n)$. Let $\mathbb{Q}^n$ be such that for $k = 1, \ldots, n$:

$$\begin{aligned} X_{\frac{k}{n}} - X_{\frac{k-1}{n}} | X_0, \ldots, X_{\frac{k-1}{n}} &\sim_{\mathbb{Q}^n} \mathcal{N}_l\big(0_l, Var_{\mathbb{Q}}[X]\big) \\ &\sim_{\mathbb{Q}^n} \mathcal{N}_l\big(0_l, \mathbb{E}_{\mathbb{Q}}[\Delta_k^n X(\Delta_k^n X)^T | X_0, \ldots, X_{\frac{k-1}{n}}], \end{aligned}$$

i.e. the new measure is defined such that the increments of X have the same conditional variances as under $\mathbb{Q}$.
Assume that $H(\mathbb{Q}|\mathbb{B})|_{\mathcal{F}^n} < \infty$. Then it follows that, $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{B}|_{\mathcal{F}^n}$, which implies, $\mathbb{Q}|_{\mathcal{F}^n} \ll \mathbb{Q}^n$ and $\mathbb{Q}^n \ll \mathbb{B}|_{\mathcal{F}^n}$, and thus we have,

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

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

$$\begin{aligned} H(\mathbb{Q}|\mathbb{B})|_{\mathcal{F}^n} &= \int_C \log \frac{d\mathbb{Q}}{d\mathbb{Q}^n}|_{\mathcal{F}^n} d\mathbb{Q} + \int_C \log \frac{d\mathbb{Q}^n}{d\mathbb{B}} d\mathbb{Q} \\ &= H(\mathbb{Q}|\mathbb{Q}^n)|_{\mathcal{F}^n} + \int_C \frac{d\mathbb{Q}^n}{d\mathbb{B}}|_{\mathcal{F}^n} d\mathbb{Q}. \end{aligned} \quad (6)$$

In the remainder we will show that,

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

converges to the right-hand side of (5).
The statement then follows, since $H(\mathbb{Q}|\mathbb{Q}^n)|_{\mathcal{F}^n} \geq 0$ for all $n \in \mathbb{N}$. We will assume $H(\mathbb{Q}|B)|_{\mathcal{F}^{2^n}} < \infty$ for all $n \in \mathbb{N}$ and thus (6) holds for all $n \in \mathbb{N}$. The assumption is justified, since otherwise the left hand side in (5) is infinite and the inequality trivially true.
Note that for this argument we use that the dyadic numbers are nested. Let $\mathcal{P}_n := \sigma(A \times (s,t] \,|\, s < t \in \{0, \frac{1}{n}, \ldots, 1\}, A \in \sigma(X_0, X_{\frac{1}{n}}, \ldots, X_s))$. Since G is $\lambda \otimes \mathbb{Q}$-integrable, we can define,

$$G_n := \mathbb{E}_{\lambda\otimes\mathbb{Q}}[G|\mathcal{P}_n].$$

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

$$G_n(t) = n \sum_{k=1}^{n} \mathbb{E}_{\mathbb{Q}}\Big[\int_{\frac{k-1}{n}}^{\frac{k}{n}} G(s,X)ds | X_0, \dots, X_{\frac{k-1}{n}} \Big](X) \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n}]}(t).$$

Note that this means,

$$X_{\frac{k}{n}} - X_{\frac{k-1}{n}} | X_0, \dots, X_{\frac{k-1}{n}} \sim_{\mathbb{Q}^n} \mathcal{N}_l\Big(0_l, \frac{1}{n} G_n\big(\frac{k}{n}, X\big)\Big).$$

Indeed, this follows easily from,

$$\begin{aligned}
\frac{1}{n} G_n\big(\frac{k}{n}, X\big) &= \mathbb{E}_{\mathbb{Q}}\Big[\int_{\frac{k-1}{n}}^{\frac{k}{n}} G(s,X)ds | X_0, \dots, X_{\frac{k-1}{n}} \Big] \\
&= \mathbb{E}_{\mathbb{Q}}\Big[\mathbb{E}_{\mathbb{Q}}\Big[\int_{\frac{k-1}{n}}^{\frac{k}{n}} G(s,X)ds | \mathcal{F}_{\frac{k-1}{n}} \Big] | X_0, \dots, X_{\frac{k-1}{n}} \Big] \\
&= \mathbb{E}_{\mathbb{Q}}\Big[\mathbb{E}_{\mathbb{Q}}\Big[\Delta_k^n X (\Delta_k^n X)^T | \mathcal{F}_{\frac{k-1}{n}} \Big] | X_0, \dots, X_{\frac{k-1}{n}} \Big] \\
&= \mathbb{E}_{\mathbb{Q}}\big[\Delta_k^n X (\Delta_k^n X)^T | X_0, \dots, X_{\frac{k-1}{n}} \big]
\end{aligned}$$

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

$$\begin{aligned}
\frac{d\mathbb{Q}^n}{d\mathbb{B}}|_{\mathcal{F}^n}(X) &= \prod_{k=1}^{n} \frac{d\mathcal{L}_{\mathbb{Q}^n}(\Delta_k^n X | X_0, \dots, X_{\frac{k-1}{n}})}{d\mathcal{L}_B(\Delta_k^n X | X_0, \dots, X_{\frac{k-1}{n}})}(\Delta_k^n X) \\
&= \prod_{k=1}^{n} \frac{d\mathcal{N}_l(0_l, \frac{1}{n} G_n(\frac{k}{n}, X))}{d\mathcal{N}_l(0_l, \frac{1}{n}\mathcal{I})}(\Delta_k^n X) \\
&= \prod_{k=1}^{n} \sqrt{\frac{\det(\frac{1}{n}\mathcal{I})}{\det(\frac{1}{n}G_n)}} \frac{\exp(-\frac{1}{2}(\Delta_k^n X)^T (\frac{1}{n}G_n)^{-1} \Delta_k^n X)}{\exp(-\frac{1}{2}(\Delta_k^n X)^T (\frac{1}{n}\mathcal{I})^{-1} \Delta_k^n X)} \\
&= \prod_{k=1}^{n} \frac{1}{\sqrt{\det G_n}} \exp(-\frac{n}{2}(\Delta_k^n X)^T G_n^{-1} \Delta_k^n X) + \frac{n}{2}(\Delta_k^n X)^T \mathcal{I}^{-1} \Delta_k^n X)).
\end{aligned}$$

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

$$\begin{aligned}
&\frac{1}{n} \int_C \log \frac{d\mathbb{Q}^n}{d\mathbb{B}}|_{\mathcal{F}^n} d\mathbb{Q} = \\
&= \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_{\mathbb{Q}}\Big[\frac{n}{2}(\Delta_k^n X)^T (\mathcal{I})^{-1}) \Delta_k^n X - \frac{n}{2}(\Delta_k^n X)^T (G_n)^{-1}) \Delta_k^n X - \frac{1}{2}\log(\det G_n) \Big] \\
&= \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_{\mathbb{Q}}\Big[\frac{1}{2} tr(n\mathcal{I}^{-1} \Delta_k^n X (\Delta_k^n X)^T) - tr(nG_n^{-1} \Delta_k^n X (\Delta_k^n X)^T) - \frac{1}{2}\log(\det G_n) \Big].
\end{aligned} \tag{7}$$

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

$$\frac{1}{n} G_n\big(\frac{k}{n}, X\big) = \mathbb{E}_{\mathbb{Q}}[(\Delta_k^n X)^T \Delta_k^n X | X_0, \dots, X_{\frac{k-1}{n}}]$$

and therefore with the tower property of conditional expectation,

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\Big[n(G_n\big(\frac{k}{n}, X\big)^{-1} (\Delta_k^n X)^T \Delta_k^n X\Big] &= \mathbb{E}_{\mathbb{Q}}\Big[\mathbb{E}_{\mathbb{Q}}\Big[n(G_n\big(\frac{k}{n}, X\big)^{-1} (\Delta_k^n X)^T \Delta_k^n X | X_0, \dots, X_{\frac{k-1}{n}}\Big]\Big] \\
&= \mathbb{E}_{\mathbb{Q}}\Big[n(G_n\big(\frac{k}{n}, X\big)^{-1} \mathbb{E}_{\mathbb{Q}}\Big[(\Delta_k^n X)^T \Delta_k^n X | X_0, \dots, X_{\frac{k-1}{n}}\Big]\Big] \\
&= \mathcal{I}.
\end{aligned}$$

We continue in (7) with $F_l(X) := \frac{1}{2}(tr(X) - l - \log(\det X))$ as before,

$$\begin{aligned}
\frac{1}{n} \int_C \log \frac{d\mathbb{Q}^n}{d\mathbb{B}} |_{\mathcal{F}^n} d\mathbb{Q} &= \frac{1}{n} \sum_{k=1}^n \left(\frac{1}{2} tr\Big(G_n\big(\frac{k}{n}, X\big)\Big) - \frac{l}{2} - \log\Big(\det G_n\big(\frac{k}{n}, X\big)\Big) \right) \\
&= \frac{1}{n} \sum_{k=1}^n E_{\mathbb{Q}}\left[F_l\Big(G_n\big(\frac{k}{n}, X\big)\Big)\right] \\
&= E_{\mathbb{Q}}\left[\int_0^1 F_l\Big(G_n\big(\frac{k}{n}, X\big)\Big) ds\right]
\end{aligned}$$

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

$$\begin{aligned}
E_{\mathbb{Q}}\big[\int_0^1 F_l(G_n(\frac{k}{n}, X) ds\big] &= E_{\mathbb{Q}}\big[\int_0^1 F_l(\mathbb{E}_{\lambda \otimes \mathbb{Q}}[G|\mathcal{P}_n](s, X)) ds\big] \\
&\leq E_{\mathbb{Q}}\big[\int_0^1 \mathbb{E}_{\lambda \otimes \mathbb{Q}}[F_l(G)|\mathcal{P}_n](s, X)) ds\big] \\
&= E_{\mathbb{Q}}\big[\int_0^1 F_l(G) ds\big],
\end{aligned}$$

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

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

$$G_{2^n}(t, X) \xrightarrow{n \to \infty} G(t, X).$$

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

$$\begin{aligned}
E_{\mathbb{Q}}\big[\int_0^1 F_l(G(s,X)ds\big] &= E_{\mathbb{Q}}\big[\int_0^1 \liminf_{n\to\infty} F_l(G_{2^n}(s,X)ds\big] \\
&\leq \liminf_{n\to\infty} E_{\mathbb{Q}}\big[\int_0^1 F_l(G_{2^n}(s,X)ds\big] \\
&\leq \limsup_{n\to\infty} E_{\mathbb{Q}}\big[\int_0^1 F_l(G_{2^n}(s,X)ds\big] \\
&\leq E_{\mathbb{Q}}\big[\int_0^1 F_l(G(s,X)ds\big].
\end{aligned}$$

We conclude,

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

□

Remark 10. [1]. Note that all arguments in the proof of Theorem 2.3 still hold, if the sequence $(2^n)_{n\in\mathbb{N}}$ is replaced by $(p^n)_{n\in\mathbb{N}}$ for any $p \in \mathbb{N}$ and $p \geq 2$.

3 Applications and Properties

In this chapter we see three related examples of computing the specific relative entropy. In particular, we first compare the geometric Brownian motion with a Brownian motion, then with itself, and then extend this to the multivariate case.

The following two lemmas tell us how the multidimensional and the one-dimensional cases are connected, in relation with the independence of the coordinates.

Lemma 3.1. *Given two multivariate continuous martingales* $M = \begin{pmatrix} M^1 \\ \vdots \\ M^l \end{pmatrix}$ *and* $N = \begin{pmatrix} N^1 \\ \vdots \\ N^l \end{pmatrix}$ *with independent coordinates, then the specific relative entropy between M and N is equal to the sum of the specific relative entropy of the coordinates:*

$$h(\mathcal{L}(M)|\mathcal{L}(N)) = \sum_{i=1}^{l} h(\mathcal{L}(M^i)|\mathcal{L}(N^i)).$$

Proof. By definition $h(\mathcal{L}(M)|\mathcal{L}(N)) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{L}(M)|\mathcal{L}(N))|_{\mathcal{F}^n}$, and we want to show that it is equal to:

$$\sum_{i=1}^{l} h(\mathcal{L}(M^i)|\mathcal{L}(N^i)) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{l} H(\mathcal{L}(M^i)|\mathcal{L}(N^i))|_{\mathcal{F}^n}.$$

In particular, we want to show that

$$H(\mathcal{L}(M)|\mathcal{L}(N))|_{\mathcal{F}^n} = \sum_{i=1}^{l} H(\mathcal{L}(M^i)|\mathcal{L}(N^i))|_{\mathcal{F}^n}.$$

By the independence of the coordinates and Lemma 2.1 we have:

$$\begin{aligned}\sum_{i=1}^{l} H(\mathcal{L}(M^i)|\mathcal{L}(N^i))|_{\mathcal{F}^n} &= \sum_{i=1}^{l}\sum_{k=1}^{n} H(\mathcal{L}(\Delta_k^n M^i)|\mathcal{L}(\Delta_k^n N^i)) \\ &= \sum_{i=1}^{l}\sum_{k=1}^{n} \mathbb{E}_{\Delta_k^n M^i}\Big[\log\big(\frac{d(\Delta_k^n M^i)}{d(\Delta_k^n N^i)}\big)\Big]\end{aligned}$$

while on the other side we have:

$$\begin{aligned}H(\mathcal{L}(M)|\mathcal{L}(N))|_{\mathcal{F}^n} &= \sum_{k=1}^{n} H(\mathcal{L}(\Delta_k^n M)|\mathcal{L}(\Delta_k^n N)) \\ &= \sum_{k=1}^{n} \mathbb{E}_{\Delta_k^n M}\Big[\log\big(\frac{d(\Delta_k^n M)}{d(\Delta_k^n N)}\big)\Big] \\ &= \sum_{k=1}^{n} \mathbb{E}_{\Delta_k^n M}\Big[\log\prod_{i=1}^{l}\frac{d(\Delta_k^n M^i)}{d(\Delta_k^n N^i)}\Big] \\ &= \sum_{k=1}^{n} \mathbb{E}_{\Delta_k^n M}\Big[\sum_{i=1}^{l}\log\big(\frac{d(\Delta_k^n M^i)}{d(\Delta_k^n N)^i}\big)\Big] \\ &= \sum_{i=1}^{l}\sum_{k=1}^{n} \mathbb{E}_{\Delta_k^n M^i}\big[\log\big(\frac{d(\Delta_k^n M^i)}{d(\Delta_k^n N)^i}\big)\big].\end{aligned}$$

$\square$

Lemma 3.2. *Given two multivariate continuous martingales* $M = \begin{pmatrix} M^1 \\ \vdots \\ M^l \end{pmatrix}$ *and* $N = \begin{pmatrix} N^1 \\ \vdots \\ N^l \end{pmatrix}$*, where only* N *has independent coordinates, then the specific relative entropy between* M *and* N *is at least equal to the sum of the specific relative entropy of the coordinates:*

$$h(\mathcal{L}(M)|\mathcal{L}(N)) \geq \sum_{i=1}^{l} h(\mathcal{L}(M^i)|\mathcal{L}(N^i)).$$

Proof. Arguing as in the previous lemma we want to show that:

$$H(\mathcal{L}(\Delta_k^n M)|\mathcal{L}(\Delta_k^n N)) \geq \sum_{i=1}^{l} H(\mathcal{L}(\Delta_k^n M^i)|\mathcal{L}(\Delta_k^n N^i)).$$

Define $\mu \in \mathcal{P}(\mathbb{R}^l)$ and $\gamma \in \mathcal{P}(\mathbb{R})$ where $\mu := \mathcal{L}(\Delta_k^n M_t)$ and by independence $\gamma \otimes \overset{l-times}{\cdots} \otimes \gamma = \mathcal{L}(\Delta_k^n N_t)$. Moreover, μ_i indicates the i-th marginale with density f_i, hence $\frac{d\mu(x_1,\dots,x_l)}{dx_1\dots dx_l} = f_1(x_1) f_{x_1,x_2}(x_2)\dots f_{x_1,\dots,x_l}(x_l)$.
For instance, $\frac{d\mu_1}{dx} = f_1$, $\frac{d \text{ cond. distr. of } x_2 \text{ given } x_1 \text{ under } \mu}{dx_2} = f_{x_1}(x_2)$ and $f_1(x_1) = \int_{x_2,\dots,x_l} f(x_1,\dots,x_l)dx_2\dots dx_l$. Moreover, in the following computations we assume the existence of dentisities, but this is not necessary really. Thus, we have:

$$
\begin{aligned}
&H(\mathcal{L}(\Delta_k^n M)|\mathcal{L}(\Delta_k^n N)) = H(\mu|\gamma_1 \otimes \overset{l-times}{\cdots} \otimes \gamma_l) = \\
&= \int_{x_1,\dots,x_l} \log\big(\frac{d\mu(x_1\dots x_l)}{d\gamma_1(x_1)\dots d\gamma_l(x_l)}\big) d\mu(x_1\dots x_l) = \\
&= \int_{x_1,\dots,x_l} \log\big(\frac{f_1(x_1)f_{x_1}(x_2)\dots f_{x_1,\dots,x_{l-1}}(x_l)}{\gamma_1(x_1)\dots\gamma_l(x_l)}\big) d\mu(x_1\dots x_l) = \\
&= \int_{x_1,\dots,x_l} \log\big(\frac{f_1(x_1)}{\gamma(x_1)}\big) d\mu(x_1\dots x_l) + \int_{x_1,\dots,x_l} \log\big(\frac{f_{x_1}(x_2)}{\gamma(x_2)}\big) d\mu(x_1\dots x_l) + \\
&+ \dots + \int_{x_1,\dots,x_l} \log\big(\frac{f_{x_1,\dots x_{l-1}}(x_l)}{\gamma(x_l)}\big) d\mu(x_1\dots x_l) = \\
&= \int_{x_1} \log\big(\frac{f_1(x_1)}{\gamma(x_1)}\big) df_1 f_{x_1},\dots,f_{x_1,\dots x_{l-1}}(x_1) + \\
&+ \int_{x_1,\dots,x_l} \log\big(\frac{f_{x_1}(x_2)}{\gamma(x_2)}\big) f_1(x_1)f_{x_1}(x_2)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l + \dots + \\
&+ \int_{x_1,\dots,x_l} \log\big(\frac{f_{x_1\dots x_{l-1}}(x_l)}{\gamma(x_l)}\big) f_1(x_1)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l = \\
&= \int_{x_1} \log\big(f_1(x_1)\big) df_1 f_{x_1},\dots,f_{x_1,\dots x_{l-1}}(x_1) + \\
&+ \int_{x_1,\dots,x_l} \log\big(f_{x_1}(x_2)\big) f_1(x_1)f_{x_1}(x_2)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l + \dots + \\
&+ \int_{x_1,\dots,x_l} \log\big(f_{x_1\dots x_{l-1}}(x_l)\big) f_1(x_1)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l + \\
&- \int_{x_1} \log\big(\gamma_1(x_1)\big) df_1 f_{x_1},\dots,f_{x_1,\dots x_{l-1}}(x_1) + \\
&- \int_{x_1,\dots,x_l} \log\big(\gamma_2(x_2)\big) f_1(x_1)f_{x_1}(x_2)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l - \dots \\
&- \int_{x_1,\dots,x_l} \log\big(\gamma_l(x_l)\big) f_1(x_1)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l
\end{aligned}
$$

In general we have that for any $j = 1,\dots,l$:

$$\int_{x_j} f_{x_1,\dots,x_j} dx_j = 1 \tag{8}$$

Moreover, if we define the convex function $g(x) = x\log x$, we have by the Jensen inequality and by (8):

$$
\begin{aligned}
&\int_{x_1,x_3\dots,x_l} g(f_{x_1}(x_2))f_1(x_2)f_{x_1,x_2}(x_3)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1dx_3\dots dx_l \geq \\
&\geq g\Big(\int_{x_1,x_3\dots,x_l} f_{x_1}(x_2)f_1(x_2)f_{x_1,x_2}(x_3)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1dx_3\dots dx_l\Big) = \\
&= g(f_2(x_2)) = f_2(x_2)\log f_2(x_2)
\end{aligned}
$$

And if we apply the Jensen inequality to any $j = 2, \dots, l$, by (8), we obtain:

$$
\begin{aligned}
&H(\mathcal{L}(\Delta_k^n M)|\mathcal{L}(\Delta_k^n N)) \geq \\
&\int_{x_1} \log\big(f_1(x_1)\big)f_1(x_1)dx_1 + \int_{x_2} \log f_2(x_2)f_2(x_2)dx_2 + \dots + \\
&+ \int_{x_l} \log f_l(x_l)f_l(x_l)dx_l - \int_{x_1} \log\big(\gamma_1(x_1)\big)df_1 f_{x_1}, \dots, f_{x_1,\dots x_{l-1}}(x_1) + \\
&- \int_{x_1,\dots,x_l} \log\big(\gamma_2(x_2)\big)f_1(x_1)f_{x_1,x_2}(x_2)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l + \\
&- \dots - \int_{x_1,\dots,x_l} \log\big(\gamma_l(x_l)\big)f_1(x_1)\dots f_{x_1,\dots,x_{l-1}}(x_l)dx_1\dots dx_l = \\
&= \int_{x_1} \log\big(\frac{f_1(x_1)}{\gamma(x_1)}\big)f_1(x_1)dx_1 + \int_{x_2} \log\big(\frac{f_2(x_2)}{\gamma(x_2)}\big)f_1(x_1)f_1(x_1)f_{x_1}(x_2)dx_2 + \\
&+ \dots + \int_{x_l} \log\big(\frac{f_l(x_l)}{\gamma_l(x_l)}\big)f_1(x_1)f_{x_1,x_2}(x_2)\dots f_{x_1,\dots,x_{l-1}}dx_l = \\
&= H(\mu_1|\gamma_1) + H(\mu_2|\gamma_2) + \dots + H(\mu_l|\gamma_l).
\end{aligned}
$$

□

Let's proceed with some applications.

3.1 Specific Relative Entropy between the standard geometric Brownian motion and the Brownian motion.

Lemma 3.3. *The specific relative entropy between the standard geometric Brownian motion M_t and the Brownian motion B_t, both starting at 1, is:*

$$
h(M|B) = \frac{e}{2} - \frac{3}{4}
$$

.

Proof. Let $B_t \sim \mathcal{N}_1(1,t)$ be the Brownian motion with $B_0 = 1$. Let $M_t = e^{B_t - 1 - \frac{t}{2}} \sim lognormal(-\frac{t}{2}, t)$ be the standard geometric Brownian motion with $M_0 = 1$. We are first interested to compute:

$$
H(M|B)|_{F^n} = H\big(\mathcal{L}(M_0, M_{\frac{1}{n}}, \dots, M_{\frac{n-1}{n}}, M_1)\,|\,\mathcal{L}(B_0, B_{\frac{1}{n}}, \dots, B_{\frac{n-1}{n}}, B_1)\big)
$$

Now, by Lemma 1.3:

$$= H(\mathcal{L}(M_0)|\mathcal{L}(B_0)) + H(\mathcal{L}(M_{\frac{1}{n}}|M_0)|\mathcal{L}(B_{\frac{1}{n}}|B_0))+$$
$$+ \int_y H(\mathcal{L}(M_{\frac{2}{n}}|M_0, M_{\frac{1}{n}} = y)|\mathcal{L}(B_{\frac{2}{n}}|B_0, B_{\frac{1}{n}} = y))d\mathcal{L}(M_{\frac{1}{n}})(y)+$$
$$+ \int_{y_1,y_2} H(\mathcal{L}(M_{\frac{3}{n}}|M_0, M_{\frac{1}{n}} = y_1, M_{\frac{2}{n}} = y_2)|\mathcal{L}(B_{\frac{3}{n}}|B_0, B_{\frac{3}{n}} = y_1, B_{\frac{2}{n}} = y_2))$$
$$d\mathcal{L}(M_{\frac{1}{n}}, M_{\frac{2}{n}})(y_1, y_2) + \dots$$

Thanks to the Markov property of geometric Brownian motions, in the conditioning we can just consider the M_t with the greatest time t. Moreover, since $M_0 = 1$ and $B_0 = 1$:

$$= 0 + H(\mathcal{L}(M_{\frac{1}{n}})|\mathcal{L}(B_{\frac{1}{n}})) + \int_y H(\mathcal{L}(M_{\frac{2}{n}}|M_{\frac{1}{n}} = y)|\mathcal{L}(B_{\frac{2}{n}}|B_{\frac{1}{n}} = y))d\mathcal{L}(M_{\frac{1}{n}})(y)+$$
$$+ \int_y H(\mathcal{L}(M_{\frac{3}{n}}|M_{\frac{2}{n}} = y)|\mathcal{L}(B_{\frac{3}{n}}|B_{\frac{2}{n}} = y))d\mathcal{L}(M_{\frac{2}{n}})(y) + \dots =$$

So, we want to compute:

$$H(\mathcal{L}(M_{\frac{1}{n}})|\mathcal{L}(B_{\frac{1}{n}})) + \sum_{k=2}^{n} \int_y H(\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)|\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}} = y))d\mathcal{L}(M_{\frac{k-1}{n}})(y).$$

The first term is given by Example 1.3 with $t = \frac{1}{n}$:

$$H(\mathcal{L}(M_{\frac{1}{n}})|\mathcal{L}(B_{\frac{1}{n}})) = \frac{1}{2}(\frac{1}{n} - 1 + n(e^{\frac{1}{n}} - 1)). \tag{9}$$

For the rest, we first want to compute $H(\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)|\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}} = y))$. While it is immediate to see $\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}} = y) \sim \mathcal{N}_1(y, \frac{1}{n})$, for $\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)$ we first need to notice that the relative increments

$$\begin{aligned} RI(t_i, t_{i+1}) &:= \frac{e^{B_{t_{i+1}}-1-\frac{1}{2}t_{i+1}} - e^{B_{t_i}-1-\frac{1}{2}t_i}}{e^{B_{t_i}-1-\frac{1}{2}t_i}} \\ &= e^{(B_{t_{i+1}}-B_{t_i})-\frac{1}{2}(t_{i+1}-t_i)} - 1 \\ &= g(B_{t_{i+1}} - B_{t_i}) \end{aligned}$$

are independent to each other:

$$RI(0, \frac{1}{n}) \perp\!\!\!\perp RI(\frac{1}{n}, \frac{2}{n}) \perp\!\!\!\perp \dots \perp\!\!\!\perp RI(\frac{n-1}{n}, \frac{n}{n}).$$

In this way, we obtain:

$$\begin{aligned}\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y) &= \mathcal{L}\Big(M_{\frac{k-1}{n}}\big(\frac{M_{\frac{k}{n}}-M_{\frac{k-1}{n}}}{M_{\frac{k-1}{n}}}+1\big)|M_{\frac{k-1}{n}}=y\Big)\\ &\sim y\exp\big(B_{\frac{k}{n}}-1-\frac{k}{2n}-(B_{\frac{k-1}{n}}-1-\frac{k-1}{2n})\big)\\ &\sim ye^{\mathcal{N}_1(-\frac{1}{2n},\frac{1}{n})}\\ &\sim y\, lognormal(-\frac{1}{2n},\frac{1}{n})\\ &\sim lognormal(-\frac{1}{2n}+\log y,\frac{1}{n}).\end{aligned}$$

Now we can proceed looking at the logarithms of the conditioned distributions:

$$\begin{aligned}&\log(g(x)) = \log(\frac{1}{\sqrt{2\pi\frac{1}{n}x^2}}\exp(-(\log x+\frac{1}{2n}-\log y)^2)\frac{1}{2\frac{1}{n}})\\ &= -\frac{1}{2}\log(2\pi\frac{1}{n})-\log x-(\log^2 x+\frac{1}{4n^2}+\log^2 y+\frac{1}{n}\log x-2\log x\log y-\frac{1}{n}\log y)\frac{n}{2}\end{aligned}$$

where $g(x)$ is the distribution of $lognormal(-\frac{1}{2n}+\log y,\frac{1}{n})$, and,

$$\begin{aligned}\log(n(x)) &= \log(\frac{1}{\sqrt{2\pi\frac{1}{n}}}\exp(-\frac{(x-y)^2}{2\frac{1}{n}}))\\ &= -\frac{1}{2}\log(2\pi\frac{1}{n})-(x^2+y^2-2xy)\frac{n}{2}\end{aligned}$$

where $n(x)$ is the distribution of $\mathcal{N}_1(y,\frac{1}{n})$. Moreover, for the expectations we need by Remark 3:

$$\begin{aligned}&\mathbb{E}_{M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y}[x] = e^{(-\frac{1}{2n}+\log y)+\frac{1}{2n}} = y\\ &Var_{M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y}[x] = (e^{\frac{1}{n}}-1)e^{2(-\frac{1}{2n}+\log y)+\frac{1}{n}} = (e^{\frac{1}{n}}-1)y^2\\ &\mathbb{E}_{M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y}[x^2] = (e^{\frac{1}{n}}-1)y^2+y^2 = e^{\frac{1}{n}}y^2.\end{aligned}$$

Hence we get:

$$\begin{aligned}&\mathbb{E}_{\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y)}\big[\log(g(x))\big] =\\ &= -\frac{1}{2}\log(2\pi\frac{1}{n})+\frac{1}{2n}-\log y-\big((\frac{1}{4n^2}-\frac{1}{n}\log y+\log^2 y)+\frac{1}{n}+\frac{1}{4n^2}+\\ &+\log^2 y+\frac{1}{n}(-\frac{1}{2n}+\log y)-2(\log y-\frac{1}{2n})\log y-\frac{1}{n}\log y\big)\frac{n}{2} =\\ &= -\frac{1}{2}\log(2\pi\frac{1}{n})+\frac{1}{2n}-\log y-\big(\frac{1}{4n^2}-\frac{1}{n}\log y+\log^2 y+\\ &+\frac{1}{n}+\frac{1}{4n^2}+\log^2 y-\frac{1}{2n^2}+\frac{1}{n}\log y-2\log^2 y+\frac{1}{n}\log y-\frac{1}{n}\log y\big)\frac{n}{2} =\\ &= -\frac{1}{2}\log(2\pi\frac{1}{n})-\log y-\frac{1}{2}+\frac{1}{4n}\end{aligned}$$

and,

$$\mathbb{E}_{\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y)}\big[\log(n(x))\big] = -\frac{1}{2}\log(2\pi\frac{1}{n}) - (e^{\frac{1}{n}}y^2 + y^2 - 2y^2)\frac{n}{2}$$
$$= -\frac{1}{2}\log(2\pi\frac{1}{n}) - \frac{ny^2}{2}(e^{\frac{1}{n}} - 1).$$

Thus, the relative entropy becomes:

$$H(\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)|\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}=y})) =$$
$$= \mathbb{E}_{\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}}=y)}\big[\log(g(x)) - \log(n(x))\big]$$
$$= -\frac{1}{2}\log(2\pi\frac{1}{n}) - \log y - \frac{1}{2} + \frac{1}{4n} - (-\frac{1}{2}\log(2\pi\frac{1}{n}) - \frac{ny^2}{2}(e^{\frac{1}{n}} - 1))$$
$$= -\log y - \frac{1}{2} + \frac{1}{2n} + \frac{ny^2}{2}(e^{\frac{1}{n}} - 1).$$

Now we want to integrate this quantity:

$$\int_y H(\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)|\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}} = y))\mathcal{L}(M_{\frac{k-1}{n}})dy =$$
$$= \int_y(-\log y - \frac{1}{2} + \frac{1}{4n} + \frac{ny^2}{2}(e^{\frac{1}{n}} - 1))dy =$$
$$= -\mathbb{E}\big[\log(M_{\frac{k-1}{n}})\big] - \frac{1}{2} + \frac{1}{4n} + \frac{n}{2}(e^{\frac{1}{n}} - 1)\mathbb{E}\big[M^2_{\frac{k-1}{n}}\big].$$

And since by Remark 3:

$$\mathbb{E}\big[M_{\frac{k-1}{n}}\big] = e^{-\frac{k-1}{2n}+\frac{k-1}{2n}} = 1$$
$$Var\big[M_{\frac{k-1}{n}}\big] = (e^{\frac{k-1}{n}} - 1)e^{-\frac{k-1}{2n}2+\frac{k-1}{n}} = e^{\frac{k-1}{n}} - 1$$
$$\mathbb{E}\big[M^2_{\frac{k-1}{n}}\big] = e^{\frac{k-1}{n}} - 1 + 1^2 = e^{\frac{k-1}{n}}$$
$$\log(M_t) \sim \mathcal{N}_1(-\frac{t}{2}, t).$$

We get:

$$\begin{aligned}&\int_y H(\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)|\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}} = y))\mathcal{L}(M_{\frac{k-1}{n}})dy =\\ &= \frac{k}{2n} - \frac{1}{4n} - \frac{1}{2} + \frac{n}{2}(e^{\frac{1}{n}} - 1)e^{\frac{k-1}{n}}.\end{aligned} \tag{10}$$

Summing up we have that:

$$H(M|B)|_{F^n} = H(\mathcal{L}(M_{\frac{1}{n}}))|\mathcal{L}(B_{\frac{1}{n}})+$$
$$+\sum_{k=2}^{n}\int H(\mathcal{L}(M_{\frac{k}{n}}|M_{\frac{k-1}{n}} = y)|\mathcal{L}(B_{\frac{k}{n}}|B_{\frac{k-1}{n}} = y))\mathcal{L}(M_{\frac{k-1}{n}})dy$$

Thus by equations (9) and (10),

$$= \frac{1}{2}(\frac{1}{n} - 1 + n(e^{\frac{1}{n}} - 1)) + \sum_{k=2}^{n} \Big(\frac{k}{2n} - \frac{1}{4n} - \frac{1}{2} + \frac{n}{2}(e^{\frac{1}{n}} - 1)e^{\frac{k-1}{n}} \Big).$$

For the specific relative entropy we look at the asymptotic behaviour of each term multiplied by $\frac{1}{n}$. Regarding the first term we have:

$$\frac{1}{2n}\big(\frac{1}{n} - 1 + n(e^{\frac{1}{n}} - 1)\big) \xrightarrow{n\to\infty} 0$$

while for the other terms:

$$\begin{aligned}
&\frac{1}{n}(\sum_{k=2}^{n} \frac{k}{2n} - \frac{1}{4n} - \frac{1}{2} + \frac{n}{2}(e^{\frac{1}{n}} - 1)e^{\frac{k-1}{n}}) = \\
&= \frac{1}{2n^2}(\sum_{k=2}^{n} k) - \frac{n-1}{4n^2} - \frac{n-1}{2n} + \frac{1}{2}(e^{\frac{1}{n}} - 1)\sum_{k=2}^{n} e^{\frac{k-1}{n}} = \\
&= \frac{1}{2n^2}(\frac{n(n-1)}{2} - 1) - \frac{n-1}{4n^2} - \frac{n-1}{2n} + \frac{1}{2}(e^{\frac{1}{n}} - 1)(\frac{e - e^{\frac{1}{n}}}{e^{\frac{1}{n}} - 1}) \\
&\xrightarrow{n\to\infty} \frac{1}{4} - \frac{1}{2} + \frac{e-1}{2} = -\frac{3}{4} + \frac{e}{2}.
\end{aligned}$$

In conclusion we get the specific relative entropy:

$$\begin{aligned}
h(M|B) &= \lim_{n\to\infty} \frac{1}{n} H(M_t|B_t)|_{F^n} \\
&= 0 - \frac{3}{4} + \frac{e}{2} = \frac{e}{2} - \frac{3}{4}.
\end{aligned}$$

□

Remark 11. From [1] we can conjecture that if M_t is such that $dM_t = \sigma_t dB_t$ then $h(M|B) = \frac{1}{2}\mathbb{E}\Big[\int_0^1 (\sigma_t^2 - \log\sigma_t^2 - 1)dt\Big]$. In our case, $\sigma_t := M_t$, so it translates in:

$$h(M|B) = \frac{1}{2}\mathbb{E}\big[\int_0^1 (M_t^2 - \log(M_t^2) - 1)dt\big]. \tag{11}$$

To check if it holds we need to recall Remark 3:

$$\begin{aligned}
&M_t = e^{B_t - 1 - \frac{t}{2}} \sim \mathit{lognormal}(-\frac{t}{2}, t) \\
&\mathbb{E}\big[M_t\big] = e^{-\frac{t}{2}+\frac{t}{2}} = 1 \\
&Var\big[M_t\big] = (e^t - 1)e^{-\frac{t}{2}2+t} = e^t - 1 \\
&\mathbb{E}\big[M_t^2\big] = e^t - 1 + 1^2 = e^t \\
&\mathbb{E}\big[\log M_t\big] = \mathbb{E}\big[\mathcal{N}_1(-\frac{t}{2}, t)\big] = -\frac{t}{2}
\end{aligned}$$

So, applying Fubini to the right hand side of (11) we get:

$$
\begin{aligned}
h(M|B) &= \frac{1}{2}\int_0^1 (\mathbb{E}\big[M_t^2\big] - \mathbb{E}\big[\log M_t^2\big] - 1)dt \\
&= \frac{1}{2}\int_0^1 (e^t + t - 1)dt \\
&= \frac{1}{2}(e^t + \frac{t^2}{2} - t|_0^1 \\
&= \frac{1}{2}(e + \frac{1}{2} - 1 - 1) \\
&= \frac{e}{2} - \frac{3}{4}.
\end{aligned}
$$

As we expected from Lemma 3.3.

3.2 Specific Relative Entropy between two geometric Brownian motion.

Definition 3.1. We say that M_t^σ with $\sigma \in \mathbb{R}$, is a geometric Brownian motion with respect to σ if $dM_t^\sigma := \sigma M_t dB_t$, i.e. if $M_t^\sigma := e^{\sigma B_t - \frac{\sigma^2}{2}t}$.
In particular, $M_t^\sigma \sim lognormal(-\frac{\sigma^2}{2}t, \sigma^2 t)$.

Lemma 3.4. *The specific relative entropy between two geometric Brownian motion M^{σ_i} such that $\sigma_i \in \mathbb{R}$ for $i = 1, 2$ is:*

$$
h(M^{\sigma_1}|M^{\sigma_2}) = \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1).
$$

Proof. By Lemma 1.2:

$$
H(M^{\sigma_1}|M^{\sigma_2})|_{\mathcal{F}^n} = H\Big(\mathcal{L}\big(M^{\sigma_1}_{\frac{1}{n}}, \dots, M^{\sigma_1}_{n-1/n}, M^{\sigma_1}_1\big) \,|\, \mathcal{L}\big(M^{\sigma_2}_{\frac{1}{n}}, \dots, M^{\sigma_2}_{n-1/n}, M^{\sigma_2}_1\big)\Big)
$$

$$
= H\Big(\mathcal{L}\big(M^{\sigma_1}_{\frac{1}{n}}, M^{\sigma_1}_{\frac{2}{n}}/M^{\sigma_1}_{\frac{1}{n}}, \dots, M^{\sigma_1}_1/M^{\sigma_1}_{\frac{n-1}{n}}\big) \,|\, \mathcal{L}\big(M^{\sigma_2}_{\frac{1}{n}}, M^{\sigma_2}_{\frac{2}{n}}/M^{\sigma_2}_{\frac{1}{n}}, \dots, M^{\sigma_2}_1/M^{\sigma_2}_{\frac{n-1}{n}}\big)\Big)
$$

and by Lemma 1.3:

$$
= H(\mathcal{L}(M^{\sigma_1}_{\frac{1}{n}})|\mathcal{L}(M^{\sigma_2}_{\frac{1}{n}})) + \int_y H(\mathcal{L}(\frac{M^{\sigma_1}_{\frac{2}{n}}}{M^{\sigma_1}_{\frac{1}{n}}}|M^{\sigma_1}_{\frac{1}{n}} = y)|\mathcal{L}(\frac{M^{\sigma_2}_{\frac{2}{n}}}{M^{\sigma_2}_{\frac{1}{n}}}|M^{\sigma_2}_{\frac{1}{n}} = y))\mathcal{L}(M^{\sigma_1}_{\frac{1}{n}})dy +
$$

$$
+ \int_{y,z} H(\mathcal{L}(\frac{M^{\sigma_1}_{\frac{3}{n}}}{M^{\sigma_1}_{\frac{2}{n}}}|M^{\sigma_1}_{\frac{1}{n}} = y, \frac{M^{\sigma_1}_{\frac{2}{n}}}{M^{\sigma_1}_{\frac{1}{n}}} = z)|\mathcal{L}(\frac{M^{\sigma_2}_{\frac{3}{n}}}{M^{\sigma_2}_{\frac{2}{n}}}|M^{\sigma_2}_{\frac{1}{n}} = y, \frac{M^{\sigma_2}_{\frac{2}{n}}}{M^{\sigma_2}_{\frac{1}{n}}} = z))
$$

$$
\mathcal{L}(M^{\sigma_2}_{\frac{1}{n}})dy\mathcal{L}(\frac{M^{\sigma_1}_{\frac{2}{n}}}{M^{\sigma_1}_{\frac{1}{n}}})dz + \dots + \int_{y_1,\dots,y_n} H(\mathcal{L}(\frac{M^{\sigma_1}_1}{M^{\sigma_1}_{\frac{n-1}{n}}}|M^{\sigma_1}_{\frac{1}{n}} = y_1, \dots, \frac{M^{\sigma_1}_{\frac{n-1}{n}}}{M^{\sigma_1}_{\frac{n-2}{n}}} = y_{n-1})|
$$

$$
\mathcal{L}(\frac{M^{\sigma_2}_1}{M^{\sigma_2}_{\frac{n-1}{n}}}|M^{\sigma_2}_{\frac{1}{n}} = y_1, \dots, \frac{M^{\sigma_2}_{\frac{n-1}{n}}}{M^{\sigma_2}_{\frac{n-2}{n}}} = y_{n-1}))\mathcal{L}(M^{\sigma_1}_{\frac{1}{n}})dy_1 \dots \mathcal{L}(\frac{M^{\sigma_2}_{\frac{n-1}{n}}}{M^{\sigma_2}_{\frac{n-2}{n}}})dy_{n-1}.
$$

Thanks to the Markov property, in the conditioning we can just consider the $M_t^{\sigma_i}$ with the greatest time t:

$$= H(\mathcal{L}(M_{\frac{1}{n}}^{\sigma_1})|\mathcal{L}(M_{\frac{1}{n}}^{\sigma_2})) + \int_y H(\mathcal{L}(\frac{M_{\frac{2}{n}}^{\sigma_1}}{M_{\frac{1}{n}}^{\sigma_1}}|M_{\frac{1}{n}}^{\sigma_1} = y)|\mathcal{L}(\frac{M_{\frac{2}{n}}^{\sigma_2}}{M_{\frac{1}{n}}^{\sigma_2}}|M_{\frac{1}{n}}^{\sigma_2} = y))\mathcal{L}(M_{\frac{1}{n}}^{\sigma_1})dy +$$
$$+ \int_y H(\mathcal{L}(\frac{M_{\frac{3}{n}}^{\sigma_1}}{M_{\frac{2}{n}}^{\sigma_1}}|M_{\frac{2}{n}}^{\sigma_1} = y)|\mathcal{L}(\frac{M_{\frac{3}{n}}^{\sigma_2}}{M_{\frac{2}{n}}^{\sigma_2}}|M_{\frac{2}{n}}^{\sigma_2} = y))\mathcal{L}(M_{\frac{2}{n}}^{\sigma_1})dy + \cdots +$$
$$+ \int_y H(\mathcal{L}(\frac{M_1^{\sigma_1}}{M_{\frac{n-1}{n}}^{\sigma_1}}|M_{\frac{n-1}{n}}^{\sigma_1} = y)|\mathcal{L}(\frac{M_1^{\sigma_2}}{M_{\frac{n-1}{n}}^{\sigma_2}}|M_{\frac{n-1}{n}}^{\sigma_2} = y))\mathcal{L}(M_{\frac{n-1}{n}}^{\sigma_1})dy.$$

Now recall that for $i = 1, 2$:

$$M^{\sigma_i} = \exp(\sigma_i B_t - \frac{\sigma^2}{2}t) \sim lognormal(-\frac{\sigma_i^2}{2}t, \sigma_i^2 t).$$

Then for $t > s$ we have:

$$\mathcal{L}(\frac{e^{\sigma B_t - \frac{\sigma^2}{2}t}}{e^{\sigma B_s - \frac{\sigma^2}{2}s}}|e^{\sigma B_s - \frac{\sigma^2}{2}s} = y) = \mathcal{L}(e^{\sigma(B_t - B_s) - \frac{\sigma^2}{2}(t-s)}|e^{\sigma B_s - \frac{\sigma^2}{2}s} = y)$$

and, since $B_t - B_s$ is independent to B_s:

$$\sim lognormal(-\frac{\sigma^2}{2}(t-s), \sigma^2(t-s)).$$

Therefore,

$$\mathcal{L}(\frac{M_{\frac{k}{n}}}{M_{\frac{k-1}{n}}}|M_{\frac{k-1}{n}}^{\sigma} = y) \sim lognormal(-\frac{\sigma^2}{2}\frac{1}{n}, \sigma^2\frac{1}{n})$$

Now we have that:

$$H\big(\mathcal{L}(\frac{M_{\frac{k}{n}}^{\sigma_1}}{M_{\frac{k-1}{n}}^{\sigma_1}}|M_{\frac{k-1}{n}}^{\sigma_1} = y)|\mathcal{L}(\frac{M_{\frac{k}{n}}^{\sigma_2}}{M_{\frac{k-1}{n}}^{\sigma_2}}|M_{\frac{k-1}{n}}^{\sigma_2} = y)) =$$
$$= H(lognormal(-\frac{\sigma_1^2}{2}\frac{1}{n}, \sigma_1^2\frac{1}{n})|lognormal(-\frac{\sigma_2^2}{2}\frac{1}{n}, \sigma_2^2\frac{1}{n})\big).$$

But, in Example 1.4 we've already showed that the relative entropy between two lognormal-distributions coincides withe the relative entropy between the corresponding normal distributions, therefore:

$$= H(\mathcal{N}_1(-\frac{\sigma_1^2}{2}\frac{1}{n}, \sigma_1^2\frac{1}{n})|\mathcal{N}_1(-\frac{\sigma_2^2}{2}\frac{1}{n}, \sigma_2^2\frac{1}{n})) =$$
$$= \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} + (\sigma_1^2 - \sigma_2^2)\frac{1}{4n\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1).$$

Since we don't have any dependence on the variable "y", we have:

$$\int_y H(\mathcal{L}(\frac{M^{\sigma_1}_{\frac{k}{n}}}{M^{\sigma_1}_{\frac{k-1}{n}}}|M^{\sigma_1}_{\frac{k-1}{n}} = y)|\mathcal{L}(\frac{M^{\sigma_2}_{\frac{k}{n}}}{M^{\sigma_2}_{\frac{k-1}{n}}}|M^{\sigma_2}_{\frac{k-1}{n}} = y))\mathcal{L}(M^{\sigma_1}_{\frac{k-1}{n}})dy =$$
$$= \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} + (\sigma_1^2 - \sigma_2^2)\frac{1}{4n\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1).$$

It remains to compute the first term:

$$\begin{aligned} H(\mathcal{L}(M^{\sigma_1}_{\frac{1}{n}})|\mathcal{L}(M^{\sigma_2}_{\frac{1}{n}})) &= H(lognormal(-\frac{\sigma_1^2}{2}\frac{1}{n}, \sigma_1^2\frac{1}{n})|lognormal(-\frac{\sigma_2^2}{2}\frac{1}{n}, \sigma_2^2\frac{1}{n})) \\ &= H(\mathcal{N}_1(-\frac{\sigma_1^2}{2}\frac{1}{n}, \sigma_1^2\frac{1}{n})|\mathcal{N}_1(-\frac{\sigma_2^2}{2}\frac{1}{n}, \sigma_2^2\frac{1}{n})) \\ &= \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} + (\sigma_1^2 - \sigma_2^2)^2\frac{1}{4n\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1). \end{aligned}$$

To sum up we have:

$$\begin{aligned} H(M^{\sigma_1}|M^{\sigma_2})|_{\mathcal{F}^n} &= \sum_{k=0}^{n-1}\frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} + (\sigma_1^2 - \sigma_2^2)\frac{1}{4n\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1) \\ &= \frac{n}{2}(\frac{\sigma_1^2}{\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1 + (\sigma_1^2 - \sigma_2^2)^2\frac{1}{4\sigma_2^2}. \end{aligned}$$

In conclusion we obtain the specific relative entropy:

$$\begin{aligned} h(M^{\sigma_1}|M^{\sigma_2}) &= \lim_{n\to\infty}\frac{1}{n}H(M^{\sigma_1}|M^{\sigma_2})|_{\mathcal{F}^n} \\ &= \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1). \end{aligned}$$

□

Remark 12. Let $i = 1, 2$. From [1] we expect that if $M^{\Sigma_i}_t$ is such that $dM^{\Sigma_i}_t = \Sigma_i(M_t)dB_t$ then almost always:

$$h(M^{\Sigma_1}|M^{\Sigma_2}) = \frac{1}{2}\mathbb{E}_{M^{\Sigma_1}}\left[\int_0^1(\frac{\Sigma_1(M)^2}{\Sigma_2(M)^2} - \log(\frac{\Sigma_1(M)^2}{\Sigma_2(M)^2}) - 1)dt\right].$$

In our case $\Sigma_i = \sigma_i M$ and so it translates in:

$$\frac{1}{2}\mathbb{E}\left[\int_0^1(\frac{\sigma_1^2(M^{\sigma_1})^2}{\sigma_2^2(M^{\sigma_1})^2} - \log(\frac{\sigma_1^2(M^{\sigma_1})^2}{\sigma_2^2(M^{\sigma_1})^2}) - 1)dt\right] = \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} - \log(\frac{\sigma_1^2}{\sigma_2^2}) - 1).$$

As we expected from Lemma 3.4.

3.3 Specific Relative Entropy between two multivariate geometric Brownian motions

Lemma 3.5. *Defined an l-dimensional geometric Brownian motion $dM_t^\Gamma := diag(M_t^\Gamma)\Gamma dB_t$, where $dM_t^i := M_t^i(\sum_{k=1}^l \Gamma_{ik}dB_t^k)$ and $l \in \mathbb{N}$, then the specific relative entropy between two l-dimensional geometric Brownian motion with parameters $\Gamma_1, \Gamma_2 \in \mathbb{R}^{l\times l}$ is:*

$$h(M^{\Gamma_1}|M^{\Gamma_2}) = \frac{1}{2}(tr(\Sigma_2^{-1}\Sigma_1) - l - \log\frac{\det\Sigma_1}{\det\Sigma_2}) \tag{12}$$

where for $u = 1, 2$, $(\Sigma_u)_{ij} = \frac{1}{n}\Sigma_{k=1}^l(\Gamma_u)_{ik}(\Gamma_u)_{jk} = \frac{1}{n}(\Gamma_u\Gamma_u^T)_{ij}$.

Proof. We have:

$$H(M^{\Sigma_1}|M^{\Sigma_2})|_{\mathcal{F}^n} = H\Big(\mathcal{L}\big(M_{1/n}^{\Gamma_1}, M_{\frac{2}{n}}^{\Gamma_1}, \cdots, M_1^{\Gamma_1}\big)|\mathcal{L}\big(M_{1/n}^{\Gamma_2}, M_{\frac{2}{n}}^{\Gamma_2}, \cdots, M_1^{\Gamma_2}\big)\Big)$$

explicitly,

$$= H\Big(\mathcal{L}\big(M_{\frac{1}{n}}^{1\,\Gamma_1}, \ldots, M_{1/n}^{l\,\Gamma_1}, M_{\frac{2}{n}}^{1\,\Gamma_1}, \ldots, M_{\frac{2}{n}}^{l\,\Gamma_1}, \ldots\big)|\mathcal{L}\big(M_{\frac{1}{n}}^{1\,\Gamma_2}, \ldots, M_{1/n}^{l\,\Gamma_2}, M_{\frac{2}{n}}^{1\,\Gamma_2}, \ldots, M_{\frac{2}{n}}^{l\,\Gamma_2}, \ldots\big)\Big)$$

applying lemma 1.2,

$$= H\left(\mathcal{L}\Big(M_{\frac{1}{n}}^{1,\Gamma_1}, \ldots, M_{\frac{1}{n}}^{l,\Gamma_1}, \frac{M_{2/n}^{1,\Gamma_1}}{M_{1/n}^{1\,\Gamma_1}}, \ldots, \frac{M_{\frac{2}{n}}^{2,\Gamma_1}}{M_{\frac{1}{n}}^{2\,\Gamma_1}}\Big)|\mathcal{L}\Big(M_{\frac{1}{n}}^{1,\Gamma_2}, \ldots, M_{\frac{1}{n}}^{l,\Gamma_2}, \frac{M_{2/n}^{1,\Gamma_2}}{M_{1/n}^{1,\Gamma_2}}, \ldots, \frac{M_{\frac{2}{n}}^{2,\Gamma_2}}{M_{\frac{1}{n}}^{2,\Gamma_2}}\Big)\right)$$

and by Lemma 1.3,

$$\begin{aligned}
&H(M^{\Sigma_1}|M^{\Sigma_2})|_{\mathcal{F}^n} = H\Big(\mathcal{L}(M_{\frac{1}{n}}^{1,\Gamma_1}, \ldots, M_{\frac{1}{n}}^{l,\Gamma_1})\Big|\mathcal{L}\Big(M_{\frac{1}{n}}^{1,\Gamma_2}, \ldots, M_{\frac{1}{n}}^{l,\Gamma_2}\Big)+\\
&+\int_{y_1,\ldots,y_l} H\left(\mathcal{L}\Big(\frac{M_{\frac{2}{n}}^{1,\Gamma_1}}{M_{\frac{1}{n}}^{1\,\Gamma_1}}, \ldots, \frac{M_{\frac{2}{n}}^{l,\Gamma_1}}{M_{\frac{1}{n}}^{l,\Gamma_1}}\Big|M_{\frac{1}{n}}^{1,\Gamma_1} = y_1, \ldots, M_{\frac{1}{n}}^{l,\Gamma_1} = y_l\Big)\right|\\
&\left.\mathcal{L}\Big(\frac{M_{\frac{2}{n}}^{1,\Gamma_2}}{M_{\frac{1}{n}}^{1,\Gamma_2}}, \ldots, \frac{M_{\frac{2}{n}}^{l,\Gamma_2}}{M_{\frac{1}{n}}^{l,\Gamma_2}}\Big|M_{\frac{1}{n}}^{1,\Gamma_2} = y_1, \ldots, M_{\frac{1}{n}}^{l,\Gamma_2} = y_l\Big)\right)d\mathcal{L}(M_{\frac{1}{n}}^{1,\Gamma_1}, \ldots, M_{\frac{1}{n}}^{l,\Gamma_1})(y_1, \ldots, y_l)+\\
&+\int_{y_1,\ldots,y_l} H\left(\mathcal{L}\Big(\frac{M_{\frac{3}{n}}^{1,\Gamma_1}}{M_{\frac{2}{n}}^{1,\Gamma_1}}, \ldots, \frac{M_{\frac{3}{n}}^{l,\Gamma_1}}{M_{\frac{2}{n}}^{l,\Gamma_1}}\Big|M_{\frac{1}{n}}^{1,\Gamma_1} = y_1, \ldots, M_{\frac{1}{n}}^{l,\Gamma_1} = y_l, M_{\frac{2}{n}}^{1,\Gamma_1} = z_1, \ldots, M_{\frac{2}{n}}^{l,\Gamma_1} = z_l\Big)\right.\\
&\left.\Big|\mathcal{L}\Big(\frac{M_{\frac{3}{n}}^{1,\Gamma_2}}{M_{\frac{2}{n}}^{1,\Gamma_2}}, \ldots, \frac{M_{\frac{3}{n}}^{l,\Gamma_2}}{M_{\frac{2}{n}}^{l,\Gamma_2}}\Big|M_{\frac{1}{n}}^{1,\Gamma_2} = y_1, \ldots, M_{\frac{1}{n}}^{l,\Gamma_2} = y_l, M_{\frac{2}{n}}^{1,\Gamma_2} = z_1, \ldots, M_{\frac{2}{n}}^{l,\Gamma_2} = z_l\Big)\right)\\
&d\mathcal{L}(M_{\frac{1}{n}}^{1,\Gamma_1}, \ldots, M_{\frac{1}{n}}^{l\,\Gamma_1})(y_1, \ldots, y_l)d\mathcal{L}\Big(M_{\frac{2}{n}}^{1,\Gamma_1}, \ldots, M_{\frac{2}{n}}^{l,\Gamma_1}\Big)(z_1, \ldots, z_l) + \ldots
\end{aligned}$$

Note that, by the Markov property and the independence of the increments of the Brownian motion, the rations are all independents and so we can ignore all the conditionings. Therefore we want to compute the first term and for $m = 2, \dots, n$:

$$H\Bigg(\mathcal{L}\Big(\frac{M^{1,\Gamma_1}_{\frac{m}{n}}}{M^{1,\Gamma_1}_{\frac{m-1}{n}}}, \dots, \frac{M^{l,\Gamma_1}_{\frac{m}{n}}}{M^{l,\Gamma_1}_{\frac{m-1}{n}}}\Big)\Big|\mathcal{L}\Big(\frac{M^{1,\Gamma_2}_{\frac{1}{n}}}{M^{1,\Gamma_2}_{\frac{m}{n}}}, \dots, \frac{M^{l,\Gamma_2}_{\frac{1}{n}}}{M^{l,\Gamma_2}_{\frac{m-1}{n}}}\Big)\Bigg). \tag{13}$$

Now, dropping the subscript "$1, 2$" for a moment, we look at the distributions of each coordinate $i = 1, \dots l$:

$$\begin{aligned}\frac{M^{i}_{\frac{m}{n}}}{M^{i}_{\frac{m-1}{n}}} &= e^{\sum_{k=1}^{l}\Gamma_{ik}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}\Gamma^2_{ik}} \\ &\sim lognormal\Big(\frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{ik}, \frac{1}{n}\sum_{k=1}^{l}\Gamma^2_{ik}\Big)\end{aligned}$$

So equation (13) becomes:

$$\begin{aligned}&= H\Bigg(\mathcal{L}\Big(e^{\sum_{k=1}^{l}\Gamma_{1,\,1k}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{1,\,1k}}, \dots, e^{\sum_{k=1}^{l}\Gamma_{1,\,lk}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{1,\,lk}}\Big)\Big| \\ &\mathcal{L}\Big(e^{\sum_{k=1}^{l}\Gamma_{2,\,1k}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{2,\,1k}}, \dots, e^{\sum_{k=1}^{l}\Gamma_{2,\,lk}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{2,\,lk}}\Big)\Bigg).\end{aligned}$$

and again by Lemma 1.2:

$$\begin{aligned}&= H\Bigg(\mathcal{L}\Big(\sum_{k=1}^{l}\Gamma_{1,\,1k}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{1,\,1k}, \dots, \sum_{k=1}^{l}\Gamma_{1,\,lk}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{1,\,lk}\Big)\Big| \\ &\mathcal{L}\Big(\sum_{k=1}^{l}\Gamma_{2,\,1k}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{2,\,1k}, \dots, \sum_{k=1}^{l}\Gamma_{2,\,lk}B^k_{1/n} - \frac{1}{2n}\sum_{k=1}^{l}\Gamma^2_{2,\,lk}\Big)\Bigg).\end{aligned} \tag{14}$$

Now in any coordinate there is a normal distributions so we want to compute the relative entropy between two multivariate Gaussians: $H\big(\mathcal{N}_l(\mu_1, \Sigma_1)|\mathcal{N}_l(\mu_2, \Sigma_2)\big)$ where $\mu_{1,2}$ and $\Sigma_{1,2}$ are respectively the mean vector and the covariance matrix given by the coordinates of (14): $(\Sigma_{1,2})_{ij} = \frac{1}{n}\Sigma_{k=1}^{l}(\Gamma_{1,2})_{ik}(\Gamma_{1,2})_{jk}$ and $\mu_{1,2} = -\frac{1}{2n}\big(\sum_{k=1}^{l}(\Gamma_{1,2,})^2_{1k}, \dots, \sum_{k=1}^{l}(\Gamma_{1,2})^2_{lk}\big)^T$.
To show how to get the covariance matrix, we first define $Z_i := \sum_{k=1}^{l}\Gamma_i B^k_{1/n} - \frac{1}{2n}\sum_{k=1}\Gamma_i^2$ and $\bar{Z}_i := \sum_{k=1}^{l}\Gamma_i B^k_{1/n}$. Ignoring the subscripts for simplicity we have:

$$\begin{aligned}(\Sigma)_{ij} = Cov[(Z_i, Z_j)] &= \mathbb{E}\big[(Z_i - \mathbb{E}[Z_i])(Z_j - \mathbb{E}[Z_j])\big] \\ &= \mathbb{E}\big[\bar{Z}_i\bar{Z}_j\big] = \mathbb{E}\big[\sum_{k=1}^{l}\Gamma_{ik}B^k_{1/n}\sum_{k=1}^{l}\Gamma_{jk}B^k_{1/n}\big] \\ &= \sum_{k=1}^{l}\Gamma_i\Gamma_j\mathbb{E}\big[(B^k_{1/n})^2\big] = \frac{1}{n}\sum_{k=1}^{l}\Gamma_i\Gamma_j.\end{aligned}$$

After noticing that we can argue in the same way for the first term since $M_{\frac{1}{n}}^{i,\Gamma_{1,2}} \sim lognormal(\frac{1}{2n}\sum_{k=1}^{l}(\Gamma_{1,2})_{ik}^2, \frac{1}{n}\sum_{k=1}^{l}(\Gamma_{1,2})_{ik}^2)$ then by Lemma 2.1 and using Example 1.1 we have:

$$\begin{aligned}
H\big(M^{\Gamma_1}|M^{\Gamma_2}\big)|_{\mathcal{F}^n} &= \sum_{m=1}^{n} H(\mathcal{N}_l(\mu_1,\Sigma_1)|\mathcal{N}_l(\mu_2,\Sigma_2)) \\
&= nH(\mathcal{N}_l(\mu_1,\Sigma_1)|\mathcal{N}_l(\mu_2,\Sigma_2)) \\
&= \frac{n}{2}(tr(\Sigma_2^{-1}\Sigma_1) + (\mu_2-\mu_1)^T\Sigma_2^{-1}(\mu_2-\mu_1) - l - \log\frac{\det\Sigma_2}{\det\Sigma_2}).
\end{aligned}$$

Looking at the asymptotic behaviour of the terms, we expect that:

$$tr(\Sigma_2^{-1}\Sigma_1) \text{ and } \log(\frac{\det\Sigma_2}{\det\Sigma_2}) \sim \frac{n}{n}$$

while the quadratic term:

$$(\mu_2-\mu_1)^T\Sigma_2^{-1}(\mu_2-\mu_1) \sim \frac{1}{n^2}\, n.$$

In this way, we get the specific relative entropy desired:

$$\begin{aligned}
h(M_1^{\Sigma}|M_2^{\Sigma}) &= \lim_{n\to\infty}\frac{1}{n}H(M_1^{\Sigma}|M_2^{\Sigma})|_{\mathcal{F}^n} \\
&= \frac{1}{2}(tr(\Sigma_2^{-1}\Sigma_1) - l - \log\frac{\det\Sigma_1}{\det\Sigma_2}).
\end{aligned}$$

□

Remark 13. In Lemma 3.5 we have that $\Sigma_u = \frac{1}{n}\Gamma_u\Gamma_u^T$ with $u=1,2$. Indeed, $(\frac{1}{n}\Gamma_u\Gamma_u^T)_{ij} = \frac{1}{n}\sum_{k=1}^{l}(\Gamma_u)_{ik}(\Gamma_u)_{kj}^T = \frac{1}{n}\sum_{k=1}^{l}(\Gamma_u)_{ik}(\Gamma_u)_{jk} = (\Sigma_u)_{ij}$.

Corollary 3.6. *If* $\Gamma_1 = diag(\gamma_i)_{i=1,\dots,l}$ *and* $\Gamma_2 = diag(\bar{\gamma}_i)_{i=1,\dots,l}$ *then* $h(M^{\Gamma_1}|M^{\Gamma_2}) = \frac{1}{2}\big(\sum_{i=1}^{l}\frac{\gamma_i^2}{\bar{\gamma}_i^2} - l - \sum_{i=1}^{l}\log\frac{\gamma_i^2}{\bar{\gamma}_i^2}\big)$.

Proof. We can show it in two ways. In the first one we follow the steps of the previous Lemma and we obtain:

$$(\Gamma_1)_{ij} = diag(\frac{1}{n}\gamma_i^2) \text{ and } (\Gamma_2)_{ij} = diag(\frac{1}{n}\bar{\gamma}_i^2)$$

So,

$$\Sigma_2^{-1}\Sigma_1 = diag(\frac{n}{\bar{\gamma}_i^2}\frac{1}{n}\gamma_i^2) = diag(\frac{\gamma_i^2}{\bar{\gamma}_i^2}), \text{ hence, } tr(\Sigma_2^{-1}\Sigma_1) = \sum_{i=1}^{l}\frac{\gamma_i^2}{\bar{\gamma}_i^2}$$

while the logarithm term becomes:

$$\begin{aligned}
\log\frac{\det\Sigma_1}{\det\Sigma_2} &= \log(\det(\Sigma_2^{-1}\Sigma_1)) \\
&= \log\prod_{i=1}^{l}\frac{\gamma_i^2}{\bar{\gamma}_i^2} = \sum_{i=1}^{l}\log\frac{\gamma_i^2}{\bar{\gamma}_i^2}
\end{aligned}$$

and since the mean vectors are $\mu_1 = -\frac{1}{2n}(\gamma_1^2, \ldots, \gamma_l^2)$ and $\mu_2 = -\frac{1}{2n}(\bar{\gamma}_1^2, \ldots, \bar{\gamma}_l^2)$ we can compute also the quadratic term:

$$(\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1) = \frac{1}{4n}(\frac{1}{\bar{\gamma}_i^2} \sum_{i=1}^{l} (\gamma_i^2 - \bar{\gamma}_i^2)).$$

Now, as in (3.3) we have:

$$\begin{aligned} H(M^{\Gamma_1}|M^{\Gamma_2})|_{\mathcal{F}^n} &= \sum_{m=1}^{n} \frac{1}{2}\big(tr(\Sigma_2^{-1}\Sigma_1) - l + (\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1) - \log \frac{\det \Sigma_1}{\det \Sigma_2}\big) \\ &= \frac{n}{2}\big(tr(\Sigma_2^{-1}\Sigma_1) - l + (\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1) - \log \frac{\det \Sigma_1}{\det \Sigma_2}\big). \end{aligned}$$

Finally, the desired specific relative entropy is:

$$\begin{aligned} h(M^{\Sigma_1}|M^{\Sigma_2}) &= \lim_{n \to \infty} \frac{1}{n} H(M^{\Sigma_1}|M^{\Sigma_2})|_{\mathcal{F}^n} \\ &= \frac{1}{2}(\sum_{i=1}^{l} \frac{\gamma_i^2}{\bar{\gamma}_i^2} - l - \sum_{i=1}^{l} \log \frac{\gamma_i^2}{\bar{\gamma}_i^2}). \end{aligned}$$

The second way to show this is noticing that we have independence among the coordinates, so we can apply Lemma 3.1 and for $m = 2, \ldots, n$ we get:

$$H\Big(\mathcal{L}\big(M^{1,\Gamma_1}_{\frac{m}{n}}/M^{1,\Gamma_1}_{\frac{m-1}{n}}, \ldots, M^{l,\Gamma_1}_{\frac{m}{n}}/M^{l,\Gamma_1}_{\frac{m-1}{n}}\big)|\mathcal{L}\big(M^{1,\Gamma_2}_{\frac{1}{n}}/M^{1,\Gamma_2}_{\frac{m}{n}}, \ldots, M^{l,\Gamma_2}_{\frac{1}{n}}/M^{l,\Gamma_2}_{\frac{m-1}{n}}\big)\Big) =$$
$$= \sum_{i=1}^{l} H(M^{i,\Gamma_1}_{\frac{m}{n}}/M^{i,\Gamma_1}_{\frac{m-1}{n}}|M^{i,\Gamma_2}_{\frac{m}{n}}/M^{i,\Gamma_2}_{\frac{m-1}{n}}).$$

Now, since each coordinate is a geometric Brownian motion we are in the case of 3.4, therefore we have:

$$H(M^{\Sigma_1}|M^{\Sigma_2})|_{\mathcal{F}^n} = \sum_{k=1}^{l} \frac{1}{2}\Big(\sum_{i=1}^{l} \frac{\gamma_i^2}{\bar{\gamma}_i^2} - l - \sum_{i=1}^{l} \log(\frac{\gamma_i^2}{\bar{\gamma}_i^2}) + \frac{1}{4n}\sum_{i=1}^{l}(\gamma_i^2 - \bar{\gamma}_i^2)\Big)$$

And finally,

$$h(M^{\Sigma_1}|M^{\Sigma_2}) = \frac{1}{2}(\sum_{i=1}^{l} \frac{\gamma_i^2}{\bar{\gamma}_i^2} - l - \sum_{i=1}^{l} \log \frac{\gamma_i}{\bar{\gamma}_i^2}).$$

$\square$

Corollary 3.7. *If Γ_1 and Γ_2 are triangular matrices then also in this case $h(M^{\Gamma_1}|M^{\Gamma_2}) = \frac{1}{2}\big(\sum_{i=1}^{l} \frac{\gamma_i^2}{\bar{\gamma}_i^2} - l - \sum_{i=1}^{l} \log \frac{\gamma_i^2}{\bar{\gamma}_i^2}\big)$ where $(\gamma_i)_{i=1,\ldots,l}$ are the diagonal elements of Γ_1 and $(\bar{\gamma}_i)_{i=1,\ldots,l}$ are the diagonal elements of Γ_2.*

Proof. We can suppose that Γ_1 and Γ_2 are upper triangular matrices, although the proof reveals that the two matrices can be also both lower triangular or of a different type of triangularity. By Remark 13 recall that $\Sigma_1 = \frac{1}{n}\Gamma_1\Gamma_1^T$ and $\Sigma_2 = \frac{1}{n}\Gamma_2\Gamma_2^T$. Thus, we have:

$$\begin{aligned} \det\Sigma_1 &= \det(\frac{1}{n}\Gamma_1\Gamma_1^T) \\ &= \frac{1}{n^l}\det(\Gamma_1)\det(\Gamma_1^T). \end{aligned}$$

And since the transpose of a upper (lower) triangular matrix is a lower (upper) triangular and since Γ_1 is a (upper) triangular matrix, we have

$$= \frac{1}{n^l}\prod_{i=1}^{l}\gamma_i^2$$

where $(\gamma_i)_{i=1,\dots,l}$ are the diagonal elements of Γ_1. On the other side,

$$\begin{aligned} \det\Sigma_2^{-1} &= \det\big(\frac{1}{n}\Gamma_2(\Gamma_2^T)\big)^{-1} \\ &= n^l\det(\Gamma_2^T)^{-1}\det(\Gamma_2)^{-1} \end{aligned}$$

Since the inverse of an upper (lower) triangular matrix is a lower (upper) triangular and has in the diagonal the inverse of the diagonal elements of the starting matrix, we have:

$$= n^l\frac{1}{\prod_{i=1}^{l}\bar{\gamma}_i^2}$$

where $(\bar{\gamma}_i)_{i=1,\dots,l}$ are the diagonal elements of Γ_2.
Therefore, we obtain the logarithm term of (12):

$$\begin{aligned} \log\frac{\det\Sigma_1}{\det\Sigma_2} &= \log(\det\Sigma_2^{-1}\det\Sigma_1) \\ &= \log\frac{\prod_{i=1}^{l}\gamma_i^2}{\prod_{i=1}^{l}\bar{\gamma}_i^2} \\ &= \sum_{i=1}^{l}\log\frac{\gamma_i^2}{\bar{\gamma}_i^2}. \end{aligned}$$

To compute the trace of $\Sigma_2^{-1}\Sigma_1$ we need a little more of work. Firstly, by the commutativity of the trace:

$$\begin{aligned} tr(\Sigma_2^{-1}\Sigma_1) &= tr\big((\frac{1}{n}\Gamma_2\Gamma_2^T)^{-1}(\frac{1}{n}\Gamma_1\Gamma_1^T)\big) \\ &= tr\big(((\Gamma_2^T)^{-1}\Gamma_1)\,((\Gamma_2)^{-1}\Gamma_1^T)\big). \end{aligned}$$

Since the product of two triangular matrix is also a triangular matrix of the same type we have that $(\Gamma_2^T)^{-1}\Gamma_1$ is an upper triangular matrix while $(\Gamma_2)^{-1}\Gamma_1^T$ is a lower triangular matrix. Now we use the following. If U, L are respectively upper and lower square l-matrix then:

$$\begin{aligned}
tr(UL) &= \sum_{i=1}^{l}(UL)_{ii} \\
&= \sum_{i=1}^{l}(\sum_{k}^{l} U_{ik}L_{ki}) \\
&= \sum_{i=1}^{l} U_{ii}L_{ii}.
\end{aligned}$$

Moreover, since the product of two upper (lower) triangular matrix has in the diagonal the product of the elements of the diagonal of the two matrices we have $((\Gamma_2^T)^{-1}\Gamma_1)_{ii} = ((\Gamma_2^T)^{-1})_{ii}(\Gamma_1)_{ii}$ and $((\Gamma_2)^{-1}\Gamma_1^T)_{ii} = ((\Gamma_2)^{-1})_{ii}(\Gamma_1^T)_{ii}$. Therefore we have obtained the first term in (12):

$$\begin{aligned}
tr(\Sigma_2^{-1}\Sigma_1) &= \sum_{i=1}^{l}((\Gamma_2^T)^{-1})_{ii}(\Gamma_1)_{ii}((\Gamma_2)^{-1})_{ii}(\Gamma_1^T)_{ii} \\
&= \sum_{i=1}^{l}\frac{\gamma_i^2}{\bar{\gamma}_i^2}
\end{aligned}$$

where $(\gamma_i)_{i=1,\dots,l}, (\bar{\gamma}_i)_{i=1,\dots,l}$ are respectively the diagonal elements of Γ_1 and Γ_2.

□

Remark 14. Let $l \in \mathbb{N}$. Let $\Sigma_t, \bar{\Sigma}_t \in \mathbb{R}^{l\times l}$ positive definite. Let $B_t \in \mathbb{R}^l$ multivariate Brownian motion and $M_t, \bar{M}_t \in \mathbb{R}$ such that $dM_t = \Sigma(M_t)dB_t$ and $d\bar{M}_t = \Sigma(\bar{M}_t)dB_t$ then from [1] we guess that almost always we can approximate $h(M|\bar{M})$ with:

$$\frac{1}{2}\mathbb{E}\Big[\int_0^1 \big(tr((\bar{\Sigma}\bar{\Sigma}^T)^{-1}\Sigma\Sigma^T(M_t)) - l - \log(\det(\bar{\Sigma}\bar{\Sigma}^T)^{-1}\Sigma\Sigma^T(M_t))\big)\Big].$$

In our case if we define $\Lambda_t := diag(M_t)\Gamma_1$ and $\bar{\Lambda}_t := diag(M_t)\Gamma_2$ both in $\mathbb{R}^{l\times l}$ then it translates in:

$$\mathbb{E}\Big[\int_0^1 \frac{1}{2}\big(tr((\bar{\Lambda}_t\bar{\Lambda}_t^T)^{-1}\Lambda_t\Lambda_t^T(M_t)) - l - \log\det((\bar{\Lambda}_t\bar{\Lambda}_t^T)^{-1}\Lambda_t\Lambda_t^T(M_t))\big)dt\Big]. \quad (15)$$

Since, $(\Lambda_t)_{ij} = m_i(\Gamma_1)_{ij}$, where $m_i := (M_t)_{ii}$ and $i, j = 1, \dots, l$, we have:

$$\begin{aligned}
&(\Lambda_t\Lambda_t^T)_{ij} = \sum_{k=1}^{l}\Lambda_{ik}\Lambda_{kj}^T = \sum_{k=1}^{l}\Lambda_{ik}\Lambda_{jk} = \sum_{k=1}^{l} m_i(\Gamma_1)_{ik}m_j(\Gamma_1)_{jk} = \\
&= m_i m_j\sum_{k=1}^{l}(\Gamma_1)_{ik}(\Gamma_1)_{jk} = m_i m_j(\Gamma_1\Gamma_1^T)_{ij}
\end{aligned}$$

and similarly for $\bar{\Lambda}$ we get: $(\bar{\Lambda}_t\bar{\Lambda}_t^T)_{ij}^{-1} = \frac{1}{m_i m_j}(\Gamma_2\Gamma_2^T)_{ij}^{-1}$. Finally, by Remark 13:

$$\left((\bar{\Lambda}_t\bar{\Lambda}_t^T)^{-1}\Lambda_t\Lambda_t^T(M_t)\right)_{ij} = \frac{1}{m_i m_j}(\Gamma_2\Gamma_2^T)_{ij}^{-1} m_i m_j (\Gamma_1\Gamma_1^T)_{ij} = (\Sigma_2)_{ij}^{-1}(\Sigma_1)_{ij}.$$

Thus, we have showed that (15) is equal to $\frac{1}{2}(tr(\Sigma_2^{-1}\Sigma_1) - l - \log\frac{\det\Sigma_1}{\det\Sigma_2})$ which is exactly the specific relative entropy computed in Lemma 3.5.

4 Simulations.

In this chapter we will define the "Heston-model" and the "SABR-Model", in order to study their specific relative entropy through simulations on Python.

4.1 Heston-model.

Definition 4.1. Let S_t and V_t be real valued stochastic processes representing respectively the price and the volatility of an asset. Let B_t the usual Browning motion, and let $k, \theta, \xi \in \mathbb{R}$ be positive parameters, then the Heston-model (with correlation 1) is defined as follow:

$$\begin{aligned} dS_t &= S_t \sqrt{V_t} dB_t \\ dV_t &= k(\theta - V_t)dt + \xi \sqrt{V_t} dB_t. \end{aligned}$$

Intuitively, θ indicates the long term variance of the price, k indicates how the volatility of the price reverts to θ and ξ the volatility of the volatility of the price. Moreover, we assume that $2k\theta > \xi^2$, which guarantees well-posedness of the above SDE. Using the Euler method for solving stochastic differential equations, we can generate simulations of the Heston-model (Figure 1).

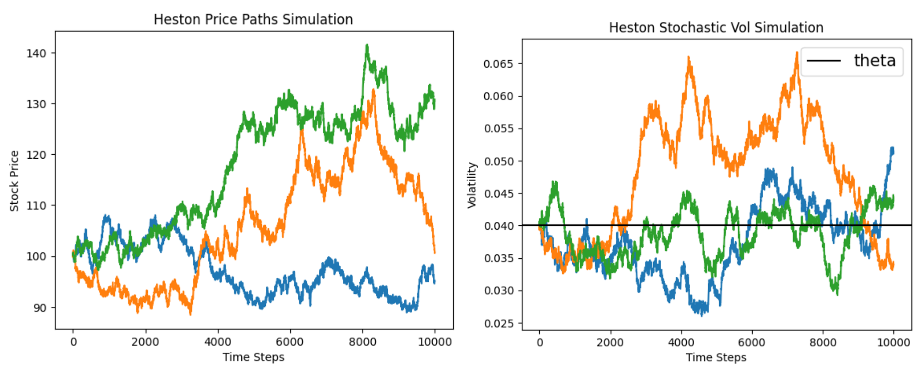
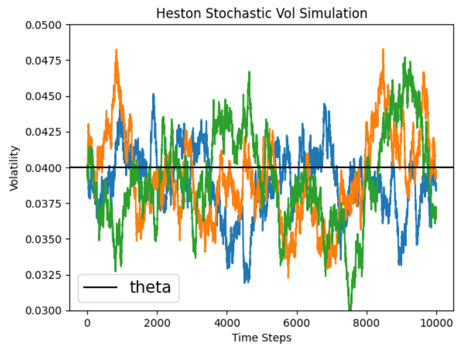
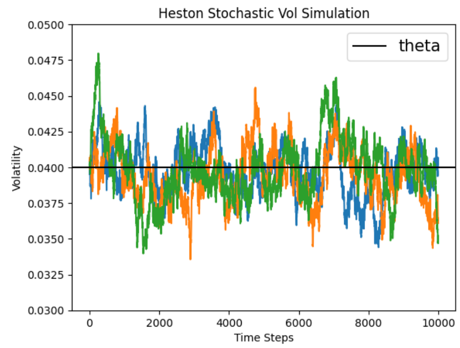


Figure 1: We have plotted three independent batches of the Heston-model with $k = 3, \theta = 0.04, \xi = 0.1$ (A.1).

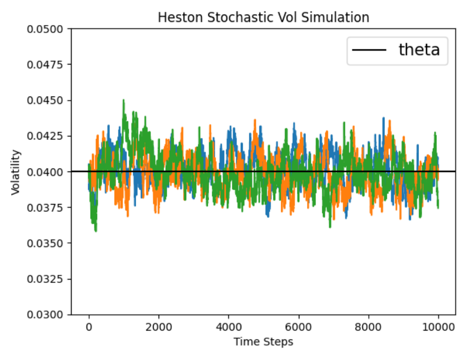
We can notice that all the paths in Figure 1 looks very random and we cannot extrapolate any information. However, if we make k increasing the volatility gathers around the value of θ (Figure 2).



(a) Volatility simulation $k = 25$.



(b) Volatility simulation $k = 50$.



(c) Volatility simulation $k = 100$.

Figure 2: We have plotted three new independent batches of the Heston-model changing the value of k with fixed $\theta = 0.04$ and $\xi = 0.1$ (A.1).

Indeed, we can obtain the mean of the volatility from:

$$\frac{d\mathbb{E}[V_t]}{dt} = k(\theta - \mathbb{E}[V_t]).$$

So defining $y_t := \mathbb{E}[V_t]$ we obtain the ODE:

$$y_t' = k(\theta - y_t)$$

and dividing by e^{kt} we have:

$$e^{kt}y_t' + ke^{kt}y_t = k\theta e^{kt}$$

which can be written as

$$\frac{d}{dt}(e^{kt} + y_t) = k\theta e^{kt}$$

So up to a constant C:

$$e^{kt}y_t = \theta e^{kt} + C$$

Hence:

$$y_t = \mathbb{E}[V_t] = \theta + Ce^{-kt}.$$

Accordingly, if we fix θ, for large k we expect that the mean of the volatility converges to θ, as we can see from the simulation (Figure 3).

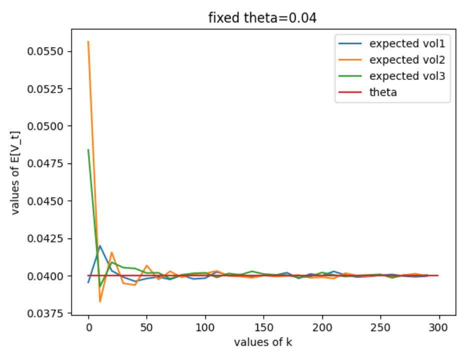


Figure 3: As k varies we have computed the volatility mean with respect to the time of three independent batches, hence we have plotted for each simulation the function $k \to \mathbb{E}[V_t^k]$ with fixed $\theta = 0.04$ and $\xi = 0.1$ (A.2).

To study the specific relative entropy of the Heston-model we could proceed as in [Chapter 3] by the discretization way. However, for the Heston-model we don't have any close density, so that way would be very challenging. Nevertheless, since we are interested in $h(S, M)$, where $dM_t := M_t dB_t$, we expect from [1] that to get the specific relative entropy it is enough to compute:

$$\frac{1}{2}\mathbb{E}\Big[\int_0^1 (V_t - \log V_t - 1)dt\Big]. \tag{16}$$

In this way we can plot a surface where we see how $h(S, M)$ depends on k and θ (Figure 4).

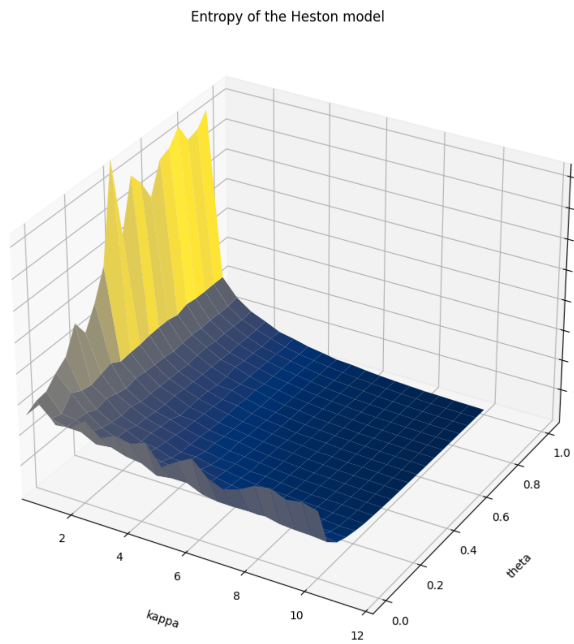


Figure 4: We have plotted a surface represented the entropy of a simulation of the Heston-model with respect to the parameter k and θ (A.3).

We notice that for large k it appears a convex function. This happens because, applying Fubini and the Jensen inequality, equation (16) can be approximate with $\frac{1}{2}\int_0^1(\mathbb{E}[V_t] - \log \mathbb{E}[V_t] - 1)dt$ but since $\mathbb{E}[V_t] \xrightarrow{k\to+\infty} \theta$ we have:

$$h(S|M) \approx \frac{1}{2}(\theta - \log\theta - 1)$$

which is a convex function of θ.

Moreover, we expect that the specific relative entropy of the Heston-model converges to the specific relative entropy of a geometric Brownian motion $M_t^{\sqrt{\theta}}$ with variance $\sqrt{\theta}$. Indeed, for large k and small ξ we can ignore the volatility of V_t and, as we discussed so far, approximate V_t with θ. In this way, the price becomes a geometric Brownian motion with variance $\sqrt{\theta}$:

$$dS_t \approx S_t\sqrt{\theta}dB_t \text{ becomes } dM_t^{\sqrt{\theta}} = \sqrt{\theta}M_t^{\sqrt{\theta}}dB_t.$$

In other words, we have that for $\xi \ll 1$, $h(S|M) \xrightarrow{k\to+\infty} h(M^{\sqrt{\theta}}|M)$. Finally, we can compute the latter thanks to Lemma 3.4, and obtain $h(M^{\sqrt{\theta}}|M) = \frac{1}{2}(\theta - \log\theta - 1)$. Note that this is consistent also with the approximation made above. We can show what has just been discussed graphically in Figure 5.

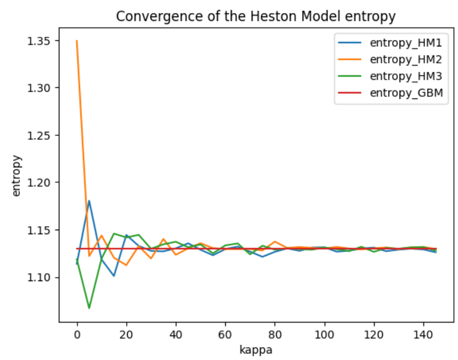


Figure 5: We have plotted the entropy of three independent batches of the Heston-model and the entropy of the geometric Brownian motion $M^{\sqrt{\theta}}$ with $\theta = 0.1$ (A.4) .

4.2 SABR-Model.

Definition 4.2. Let S_t and σ_t be real valued stochastic processes representing respectively the price and the volatility of an asset. Let B_t the usual Brownian motion. Let $\alpha > 0$ and $0 < \beta < 1$ be real valued parameters, then the SABR-model is defined as follow:

$$\begin{aligned} dS_t &= \sigma_t (S_t)^\beta dB_t \\ d\sigma_t &= \alpha \sigma_t dB_t \end{aligned} \tag{17}$$

Using the Euler method for solving stochastic differential equations, we can generate simulations of the SABR-model (Figure 6). Graphically, we see that although we can run the simulation with different values of the parameters we cannot extract any useful intuition.

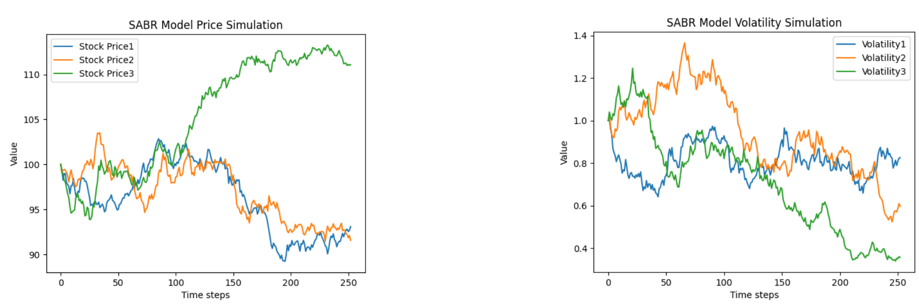


Figure 6: We have plotted three independent batches of the SABR-Model with $\beta = 0.5, \alpha = 0.5$ (B.1).

Regarding the specific relative entropy $h(S|M)$, since we don't have any closed density for the SABR-model we conjecture from [1] that:

$$h(S|M) = \frac{1}{2}\mathbb{E}\Big[\int_0^1 \big(\sigma_t^2 S_t^{2(\beta-1)} - 1 - \log(\sigma_t^2 S_t^{2(\beta-1)})\big)dt\Big] \tag{18}$$

where M_t is the standard geometric Brownian motion defined as $dM_t = M_t dB_t$.

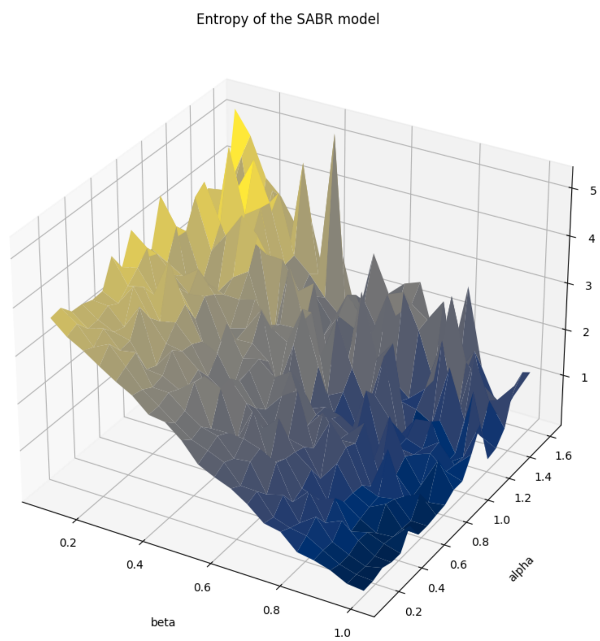


Figure 7: We have plotted a surface represented the entropy of a simulation of the SABR-model with respect to the parameter α and β (B.2).

Graphically (Figure 7), we can deduce that for a fixed α the specific relative entropy decreases in an almost linear way with respect to β. Actually, the slope with respect to β is given by $\frac{d}{d\beta}h(S|M)$, so applying Fubini to (18) we compute:

$$\begin{aligned}
&\int_0^1 \mathbb{E}\Big[\frac{d}{d\beta}\Big\{\frac{1}{2}\big(\sigma_t^2 S_t^{2(\beta-1)} - 1 - \log(\sigma_t^2 S_t^{2(\beta-1)})\big)\Big\}\Big]dy = \\
&= \int_0^1 \mathbb{E}\Big[\frac{1}{2}\big(2\sigma_t^2 S_t^{2(\beta-1)}\log(S_t) - 2\log(S_t)\big)\Big]dy = \\
&= \mathbb{E}\big[(\sigma_t^2 S_t^{2(\beta-1)} - 1)\log S_t\big] =: f_{SABR}(\beta).
\end{aligned}$$

Therefore, the inclination of the surface in Figure 7 is given by $f_{SABR}(\beta)$ which decreases roughly linearly, as we can see from Figure 8.

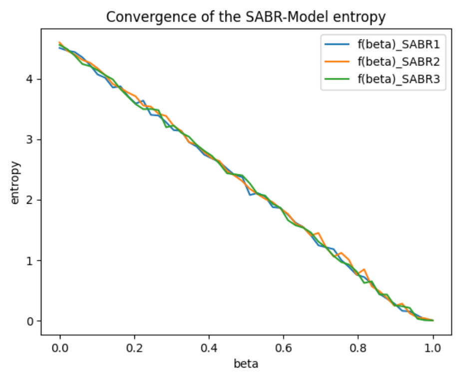


Figure 8: We have plotted the slope of three independent batches of the SABR-model, i.e the respective function $f_{SABR}(\beta)$ for a fixed $\alpha = 0.8$ (B.3).

Moreover, in Figure 7 we can notice that, for small α and for β near to one, the specific relative entropy is almost null. Indeed, if $\alpha \to 0$ then the volatility becomes a constant, hence we have $\sigma_t = \sigma_0 := 1$ and $dS_t \to (S_t)^{2(\beta-1)} dB_t$. Now, if $\beta \to 1$ then $dS_t \to S_t dB_t$, i.e. the stock price S_t of the SABR-model tends to the standard geometric Brownian motion M_t. Therefore, we have that, as $\alpha \to 0$ and $\beta \to 1$, $h(S|M) \to h(M|M) = 0$, so graphically we can see that the specific relative entropy of the SABR-model vanishes both when we fix $\alpha \ll 1$ and $\beta \to 1$, and when we fix $\beta \approx 1$ and $\alpha \to 0$ (Figure 9).

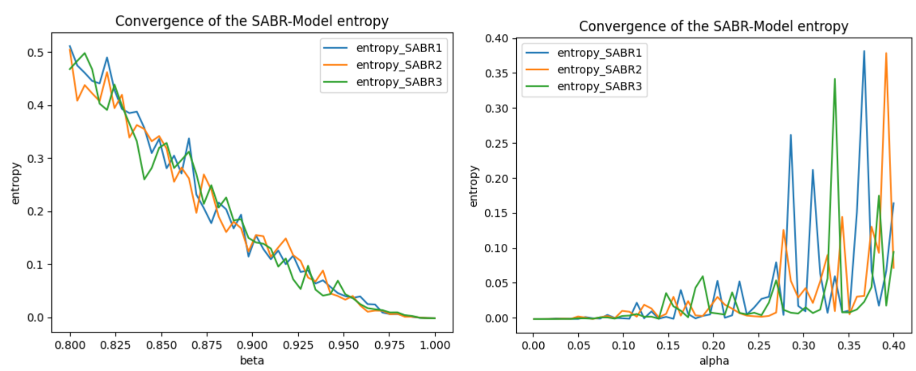


Figure 9: We have plotted the specific relative entropy of three independent batches of the SABR-model, on the left for a fixed $\alpha = 0.001$ while on the right for a fixed $\beta = 0.999$ (B.4).

In conclusion, the specific relative entropy of the Heston-Model and the SABR-Model with respect to the standard geometric Brownian motion is an useful tool for describing the behaviour of the respective stock prices. Indeed, thanks to the study of the specific relative entropy, we saw that, under certain conditions, the Heston-model tends to "approach" to a geometric Brownian motion while the SABR-model to a standard geometric Brownian motion. Therefore, if we think about the specific relative entropy as a measure of "distance" then we have showed in which cases the two models are "near" to be geometric Brownian motions.

References

[1] Master's Thesis, *"On the Specific Relative Entropy between Martingale Measures"*, Clara Unterberger (MSc), Ass.-Prof. Dr. Julio Daniel Backhoff-Veraguas, University of Vienna, 2022.

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

Appendix

Below we provide the Python codes relating to the simulations of [Chapter 4].

```python
import numpy as np
import matplotlib.pyplot as plt
import math
from mpl_toolkits.mplot3d import Axes3D
```

Listing 1: Libraries that we need.

A Heston-Model.

A.1 Price and volatility simulation.

```python
# Function of Heston-Model simulation
def generate_heston_paths(S, T, mu, kappa, theta, v_0, rho, xi, steps, Npaths):
    dt = T/steps
    size = (Npaths, steps)
    prices = np.zeros(size)
    sigs = np.zeros(size)
    log_v = np.zeros(size)
    S_t = S
    v_t = v_0
    for t in range(steps):
        WT = np.random.multivariate_normal(np.array([0, 0]), cov=np.array([[1, rho], [rho, 1]]), size=Npaths) * np.sqrt(dt)

        S_t = S_t * (np.exp((mu - 0.5 * v_t) * dt + np.sqrt(v_t) * WT[:, 0]))
        v_t = np.abs(v_t + kappa * (theta - v_t) * dt + xi * np.sqrt(v_t) * WT[:, 1])
        prices[:, t] = S_t
        sigs[:, t] = v_t
        log_v[:, t] = np.log(v_t)

    return prices, sigs, log_v
```

Listing 2: Function of Heston-Model simulation

```python
# Set parameters
T = 1.    # Time
steps = 10000    # Number of steps
kappa = 100 # kappa parameter
theta = 0.04    # theta parameter
v_0 = 0.04     # Initial value of the volatility
xi = 0.1      # xi parameter
```

```python
mu = 0.0     # mu parameter
S = 100      # Initial value of the price
paths = 3    # Number of simulations desired
rho = 0  # rho parameter

# Simulate Heston-Model
prices, volatility, log_v = generate_heston_paths(S, T, mu,
    kappa, theta, v_0, rho, xi, steps, paths)

# Plot the results
plt.plot(prices.T)
plt.title('Heston Price Paths Simulation')
plt.xlabel('Time Steps')
plt.ylabel('Stock Price')
plt.show()

# Plot the results
plt.plot(volatility.T)
plt.axhline(theta, color='black', label=r'$\theta$')
plt.title('Heston Stochastic Vol Simulation')
plt.xlabel('Time Steps')
plt.ylabel('Volatility')
plt.ylim([0.03, 0.05])
plt.legend(fontsize=15)
plt.show()
```

Listing 3: Setting Parameters and Simulating Heston-Model

A.2 Volatility mean simulation

```python
# Set parameters
c = 300
k2 = np.arange(0, c, 10)
fixed_theta = 0.04
x = np.arange(0, c, 1)
y = np.zeros(c)
for i in range(c):
  y[i] = fixed_theta

# Computing volatility mean of three different simulations
E_volatility4 = np.zeros(len(k2))
E_volatility3 = np.zeros(len(k2))
E_volatility5 = np.zeros(len(k2))
for i in range(len(k2)):
  prices3, volatility3, log_v3 = generate_heston_paths(S, T,
     mu, k2[i], fixed_theta, v_0, rho, xi, steps, 1)
  E_volatility3[i] = np.mean(volatility3)

```

```python
  prices4, volatility4, log_v4 = generate_heston_paths(S, T, mu, k2[i], fixed_theta, v_0, rho, xi, steps, 1)
  E_volatility4[i] = np.mean(volatility4)

  prices5, volatility5, log_v5 = generate_heston_paths(S, T, mu, k2[i], fixed_theta, v_0, rho, xi, steps, 1)
  E_volatility5[i] = np.mean(volatility5)

# Plot the results
plt.plot(k2, E_volatility3)
plt.plot(k2, E_volatility4)
plt.plot(k2, E_volatility5)
plt.plot(x, y)
plt.xlabel("Values of k")
plt.ylabel("Values of E[V_t]")
plt.title(f"Fixed theta={fixed_theta}")
plt.legend(["Expected Volatility 1", "Expected Volatility 2", "Expected Volatility 3", "theta"], loc="upper right")
plt.show()
```

Listing 4: Setting Parameters and Computing Volatility Mean

A.3 Entropy simulation

```python
# Set parameters
c0 = 20
k = np.linspace(0, 12, c0)
th = np.linspace(0.00, 1, c0)
entropy = np.zeros((len(k), len(th)))

# Compute the entropy
for i in range(len(k)):
  for j in range(len(th)):
    prices1, volatility1, log_v1 = generate_heston_paths(S, T, mu, k[i], th[j], v_0, rho, xi, steps, paths)
    entropy[i, j] = 0.5 * (np.mean(volatility1) - np.mean(log_v1) - 1)
```

Listing 5: Setting Parameters and Computing Entropy

```python
# Plot the results
fig = plt.figure(figsize=(12, 10))
ax = plt.axes(projection='3d')

KAPPA, THETA = np.meshgrid(k, th)

surf = ax.plot_surface(THETA, KAPPA, entropy, cmap=plt.cm.cividis)

```

```
# Set axes label
ax.set_xlabel('theta', labelpad=20)
ax.set_ylabel('kappa', labelpad=20)
ax.set_zlabel('Entropy', labelpad=20)
ax.set_title('Entropy of the Heston model')
ax.set_xlim([0.0, 1.0])  # Adjust the limits as needed
ax.set_ylim([0.5, 12.0])  # Adjust the limits as needed

# Uncomment the line below if you want to use a log scale on
    the x-axis
# ax.set_xscale("log")

# Uncomment the line below if you want to add a colorbar
# fig.colorbar(surf, shrink=0.5, aspect=8)

plt.show()
```

Listing 6: Plotting the 3D Surface

A.4 Entropy convergence simulation

```
# Set parameters
th = 0.04
c4 = 80
k3 = np.arange(0, c4, 5)
entropy4 = np.zeros(len(k3))
entropy5 = np.zeros(len(k3))
entropy6 = np.zeros(len(k3))

# Computing entropy of GBM
h_GBM2 = np.zeros(len(k3))
f_h_GBM2 = 0.5 * (th - np.log(th) - 1)

for i in range(len(k3)):
  h_GBM2[i] = f_h_GBM2

# Computing entropy of H-M
for i in range(len(k3)):
  prices7, volatility7, log_v7 = generate_heston_paths(S, T,
     mu, k3[i], th, v_0, rho, xi, steps, 1)
  entropy4[i] = 0.5 * (np.mean(volatility7) - np.mean(log_v7
    ) - 1)

  prices8, volatility8, log_v8 = generate_heston_paths(S, T,
     mu, k3[i], th, v_0, rho, xi, steps, 1)
  entropy5[i] = 0.5 * (np.mean(volatility8) - np.mean(log_v8
    ) - 1)

```

```
    prices9, volatility9, log_v9 = generate_heston_paths(S, T,
        mu, k3[i], th, v_0, rho, xi, steps, 1)
    entropy6[i] = 0.5 * (np.mean(volatility9) - np.mean(log_v9
        ) - 1)

# Plot the results
plt.plot(k3, entropy4)
plt.plot(k3, entropy5)
plt.plot(k3, entropy6)
plt.plot(k3, h_GBM2)
plt.xlabel("kappa")
plt.ylabel("entropy")
plt.legend(["Entropy HM1", "Entropy HM2", "Entropy HM3", "
    Entropy GBM"])
plt.title("Convergence of the Heston Model Entropy")
plt.show()
```

Listing 7: Setting Parameters and Plotting Entropy Convergence

B SABR-Model

B.1 Price and volatility simulation

```
# Set parameters
T = 1.0               # Time to maturity
num_steps = 252       # Number of time steps
beta = 0.5            # Beta parameter
rho = 0               # Rho parameter
alpha = 0.5           # Alpha parameter
initial_price = 100.0        # Initial stock price
initial_volatility = 1       # Initial volatility

# Simulate SABR model
price_path, volatility_path, product, log_Ssigma =
    sabr_model_simulation(T, num_steps, beta, rho, alpha,
    initial_price, initial_volatility)
price_path2, volatility_path2, product2, log_Ssigma2 =
    sabr_model_simulation(T, num_steps, beta, rho, alpha,
    initial_price, initial_volatility)
price_path3, volatility_path3, product3, log_Ssigma3 =
    sabr_model_simulation(T, num_steps, beta, rho, alpha,
    initial_price, initial_volatility)

# Plot the results
time_steps = np.linspace(0, T, num_steps + 1)
plt.plot(price_path.T, label='Stock Price1')
plt.plot(price_path2.T, label='Stock Price2')
plt.plot(price_path3.T, label='Stock Price3')
```

```
plt.xlabel('Time steps')
plt.ylabel('Value')
plt.title('SABR Model Price Simulation')
plt.legend()
plt.show()

# Plot the results
time_steps = np.linspace(0, T, num_steps + 1)
plt.plot(volatility_path.T, label='Volatility1')
plt.plot(volatility_path2.T, label='Volatility2')
plt.plot(volatility_path3.T, label='Volatility3')
plt.xlabel('Time steps')
plt.ylabel('Value')
plt.title('SABR Model Volatility Simulation')
plt.legend()
plt.show()
```

Listing 8: Setting Parameters and Simulating SABR Model

B.2 Entropy Simulation

```
# Set parameters
c5 = 20
b = np.linspace(0.1, 0.99999, c5)
a = np.linspace(0.1, 1.6, c5)

h_sabr = np.zeros((len(b), len(a)))

for i in range(len(b)):
  for j in range(len(a)):
    price_pathA, volatility_pathA, productA, log_SsigmaA =
      sabr_model_simulation(T, num_steps, b[i], rho, a[j],
      initial_price, initial_volatility)
    h_sabr[j, i] = 0.5 * (np.mean(productA) - np.mean(
      log_SsigmaA) - 1)
```

Listing 9: Setting Parameters and Computing SABR Model Entropy

```
# Plot the results
fig = plt.figure(figsize=(12, 10))
ax = plt.axes(projection='3d')

BETA, ALPHA = np.meshgrid(b, a)
surf = ax.plot_surface(BETA, ALPHA, h_sabr, cmap=plt.cm.
    cividis)

# Set axes label
ax.set_xlabel('Beta', labelpad=20)
ax.set_ylabel('Alpha', labelpad=20)
```

```python
ax.set_zlabel('Entropy', labelpad=20)
ax.set_title('Entropy of the SABR model')

# Uncomment the lines below if you want to use a log scale
    on the axes
# ax.set_xscale("log")
# ax.set_yscale("log")
# ax.set_zscale("log")

# Uncomment the line below if you want to add a colorbar
# fig.colorbar(surf, shrink=0.5, aspect=8)

plt.show()
```

Listing 10: Plotting 3D Surface

B.3 Entropy slope simulation

```python
# Set the parameters
c6 = 50
alph = 0.08  # alpha parameter
bet = np.linspace(0, 0.99999, c6)  # beta parameter

y_valuesC = np.zeros(len(bet))
y_valuesC2 = np.zeros(len(bet))
y_valuesC3 = np.zeros(len(bet))

for i in range(len(bet)):
    price_pathC, volatility_pathC, productC, log_SsigmaC =
    sabr_model_simulation(T, num_steps, bet[i], rho, alph,
    initial_price, initial_volatility)
    y_valuesC[i] = 0.5 * np.mean((productC - 1) *
    log_SsigmaC)

    price_pathC2, volatility_pathC2, productC2, log_SsigmaC2
     = sabr_model_simulation(T, num_steps, bet[i], rho, alph,
     initial_price, initial_volatility)
    y_valuesC2[i] = 0.5 * np.mean((productC2 - 1) *
    log_SsigmaC2)

    price_pathC3, volatility_pathC3, productC3, log_SsigmaC3
     = sabr_model_simulation(T, num_steps, bet[i], rho, alph,
     initial_price, initial_volatility)
    y_valuesC3[i] = 0.5 * np.mean((productC3 - 1) *
    log_SsigmaC3)

# Plot the results
plt.plot(bet, y_valuesC)
plt.plot(bet, y_valuesC2)
```

```python
plt.plot(bet, y_valuesC3)
plt.xlabel("beta")
plt.ylabel("entropy")
plt.legend(["f(beta)_SABR1", "f(beta)_SABR2", "f(beta)_SABR3
    "])
plt.title("Convergence of the SABR-Model entropy")
plt.show()
```

Listing 11: Plotting $f_{SABR}(\beta)$

B.4 Entropy convergence simulation

```python
# Set parameters
c6 = 50
alph = 0.001  # alpha parameter
bet = np.linspace(0.8, 0.99999, c6)  # beta parameter

h_sabrB = np.zeros(len(bet))
h_sabrB2 = np.zeros(len(bet))
h_sabrB3 = np.zeros(len(bet))

for i in range(len(bet)):
  price_pathB, volatility_pathB, productB, log_SsigmaB =
    sabr_model_simulation(T, num_steps, bet[i], rho, alph,
    initial_price, initial_volatility)
  h_sabrB[i] = 0.5 * (np.mean(productB) - np.mean(
    log_SsigmaB) - 1)
  price_pathB2, volatility_pathB2, productB2, log_SsigmaB2 =
     sabr_model_simulation(T, num_steps, bet[i], rho, alph,
    initial_price, initial_volatility)
  h_sabrB2[i] = 0.5 * (np.mean(productB2) - np.mean(
    log_SsigmaB2) - 1)
  price_pathB3, volatility_pathB3, productB3, log_SsigmaB3 =
     sabr_model_simulation(T, num_steps, bet[i], rho, alph,
    initial_price, initial_volatility)
  h_sabrB3[i] = 0.5 * (np.mean(productB3) - np.mean(
    log_SsigmaB3) - 1)

# Plot the results
plt.plot(bet, h_sabrB)
plt.plot(bet, h_sabrB2)
plt.plot(bet, h_sabrB3)
plt.xlabel("Beta")
plt.ylabel("Entropy")
plt.legend(["Entropy SABR1", "Entropy SABR2", "Entropy SABR3
    "])
plt.title("Convergence of the SABR-Model Entropy")
plt.show()
```

Listing 12: Plotting Entropy Convergence with fixed alpha

```python
# Set parameters
c7 = 50
be = 0.999  # beta parameter
alp = np.linspace(0.001, 0.4, c7)  # alpha parameters

h_sabrC = np.zeros(len(alp))
h_sabrC2 = np.zeros(len(alp))
h_sabrC3 = np.zeros(len(alp))

for i in range(len(alp)):
  price_pathC, volatility_pathC, productC, log_SsigmaC = sabr_model_simulation(T, num_steps, be, rho, alp[i], initial_price, initial_volatility)
  h_sabrC[i] = 0.5 * (np.mean(productC) - np.mean(log_SsigmaC) - 1)
  price_pathC2, volatility_pathC2, productC2, log_SsigmaC2 = sabr_model_simulation(T, num_steps, be, rho, alp[i], initial_price, initial_volatility)
  h_sabrC2[i] = 0.5 * (np.mean(productC2) - np.mean(log_SsigmaC2) - 1)
  price_pathC3, volatility_pathC3, productC3, log_SsigmaC3 = sabr_model_simulation(T, num_steps, be, rho, alp[i], initial_price, initial_volatility)
  h_sabrC3[i] = 0.5 * (np.mean(productC3) - np.mean(log_SsigmaC3) - 1)

# Plot the results
plt.plot(alp, h_sabrC)
plt.plot(alp, h_sabrC2)
plt.plot(alp, h_sabrC3)
plt.xlabel("Alpha")
plt.ylabel("Entropy")
plt.legend(["Entropy SABR1", "Entropy SABR2", "Entropy SABR3"])
plt.title("Convergence of the SABR-Model Entropy")
plt.show()
```

Listing 13: Plotting Entropy Convergence for fixed beta