Ecole Doctorale de Paris Centre

Thèse de doctorat

pour l'obtention du titre de

Docteur de l'Université Pierre et Marie Curie (PARIS VI)

Spécialité : Mathématiques

présentée par

David BAKER

Titre

Martingales with specified marginals

Sous la direction de Marc YOR et Catherine DONATI-MARTIN

Soutenue le 18 Decembre 2012, devant le jury composé de:

M. Rama CONT	Université PARIS VI	Examinateur
Mme Catherine DONATI-MARTIN	Université de VERSAILLES	Co-directrice de these
M. David HOBSON	University of Warwick	Rapporteur
M. Svetlozar RACHEV	SUNY Stony Brook	Rapporteur
M. Peter TANKOV	Université PARIS VII	Examinateur
M. Marc YOR	Université PARIS VI	Directeur de thèse

Acknowledgements

I would like to thank my parents, Mark and Sylviane Baker for their support and encouragement throughout this thesis work.
I am extremely grateful for the guidance received from my advisors Marc Yor and Catherine Donati-Martin whose expertise, intellectual power, and patience were of immeasurable help.
I would like to thank David Hobson and Svetlozar Rachev for accepting to review the manuscript and for making helpful comments. Thanks also to Rama Cont, Aaron Kim and Peter Tankov for accepting to be on the jury.
Finally, I would like to thank Corentin Lacombe from the école doctorale whose professionalism was of great help in navigating through the administrative process.

Martingales with specified marginals

David Baker

December 18, 2012

Contents

Statement of results

This work contains results in two areas: the construction of martingales with specified marginals and the Skorokhod embedding problem. The contributions are the folllowing:

- A new solution to the Skorokhod embedding problem (published in Statistics and Probability Letters, see [Bak12])
- A Brownian sheet based construction of a martingale with the same marginals as the average of geometric Brownian motion. This provides a new proof that in the Black Scholes framework the the price of arithmetic Asian options are increasing in duration (joint work with Marc Yor, published in Electronic Journal of Probability [BY09])
- A sequence of Albin type continuous martingales with Brownian marginals and scaling (joint work with C. Donati-Martin and M. Yor, published in Seminaire de Probabilites, see [BDMY11])
- On Martingales with Given Marginals and the Scaling Property (joint work with M. Yor, published in Seminaire de Probabilites, see [BY11])
- A proof that the L2 quantization does not have the property of preserving the convex order (preprint submitted to Statistics and Probability Letters)
- A quantization method which we called $\mathcal{U}-$quantization and a proof that it has the property of preserving the convex order. Using this quantization we give new methods for constructing martingale transitions with specified marginals (preprint submitted to Electronic Journal of Probability)

Part I

Constructing martingale transitions through quantization of measures

Chapter 1

Introduction

The framework under consideration is the following: we are given two probability measures on $\mathbb{R}$, which we denote μ and ν and we wish to construct a martingale transition from μ to ν. It is known that a necessary and sufficient condition for the existence of a martingale transition from μ to ν is that μ and ν be ordered in the convex order, which is denoted $\mu \leq_{cx} \nu$ and defined as:

$$\mu \leq_{cx} \nu \iff \int_{\mathbb{R}} f(x)d\mu(x) \leq \int_{\mathbb{R}} f(x)d\nu(x) \text{ for every convex function } f$$

The method which we propose is to approximate μ by a sequence of discrete measures $(\hat{\mu}_n)_{n\in\mathbb{N}}$ which converges in law to μ. Similarly, we construct a sequence of discrete measures $(\hat{\nu}_n)_{n\in\mathbb{N}}$ which converges to ν. Then we provide methods which will construct, for each n, a martingale transition from $\hat{\mu}_n$ to $\hat{\nu}_n$.

Approximating a probability measure by a discrete measure is refered to as quantizing that measure. The method which is generally used to quantize probability measures is the $L2$ quantization. We will show that the $L2$ quantization cannot be used in this situation. Indeed, we will prove that the $L2$ quantization does not have the property of preserving the convex order. The consequence of this is that when $\mu \leq_{cx} \nu$, we may well have $\hat{\mu}_n \not\leq_{cx} \hat{\nu}_n$ for some n, in which case there exists no martingale transition from $\hat{\mu}_n$ to $\hat{\nu}_n$. It is necessary that the quantization method which we employ has the property of preserving the convex order. We define a quantization which has this property of preserving the convex order. This quantization method will be called $\mathcal{U}$-quantization. In theorem 2.4.11, we prove that $\mathcal{U}$-quantization preserves the convex order. This ensures that if there exists a martingale transition from μ to ν, then there also exists a martingale transition from $\hat{\mu}_n$ to $\hat{\nu}_n$.
The problem of the appropriate quantization method being settled, we show how, for each n, martingale transitions can be obtained from $\hat{\mu}_n$ to $\hat{\nu}_n$. We give 3 different methods of constructing such martingale transitions. The first method is straightforward but worth mentioning; it is linear programming and its solution is obtained by the simplex method. The second method is interesting because it relates the the-

ory of symmetric matrices with specified diagonal and spectrum to the theory of martingale transitions. Indeed, we show how by constructing a symmetric matrix with properly chosen eigenvalues and diagonal elements, we can produce a martingale transition from $\hat{\mu}_n$ to $\hat{\nu}_n$. The third method is the use of potential theory and an algorithm by Chacon and Walsh. This algorithm can be used here because $\mathcal{U}$-quantization has the property of preserving the convex order.

1.1 Relevance to risk management

1.1.1 Relevance to modeling of financial risks

In addition to being of theoretical interest, the problem of constructing martingales with specified marginals has important applications to financial risk, which we now briefly discuss. The observed market prices of European calls and puts on an instrument, provide the marginal laws of its process. There is of course some imprecision coming from the fact that not all strikes and maturities are traded. If all strikes and maturities were traded, then every marginal law of the process could be fully extracted from the observed prices. This imprecision and the need to interpolate are not addressed here. The framework under consideration is that of an observer having all marginal laws of a stochastic process, and wanting to infer additional information about this stochastic process. He or she may want the probability that the process will cross a threshold during a certain time interval. Or the quantity of interest may be the probability that the realized volatility will be greater than a certain value.
A common approach is to first suppose that the stochastic process belongs to a particular family (α-stable processes, variance-gamma processes, etc..). The next step, “model calibration”, would be to choose the member of this family which provides the closest fit to the observed marginals. The problem with this method is the model risk which it introduces. There is indeed no theoretical justification for the process belonging to some particular family. Postulating this will exclude from consideration processes which have no reason to be excluded. Any conclusions obtained by this method are subject to a potentially large and unquantified amount of additional model risk. Model free approaches, while technically more difficult, are increasingly becoming an active area of research. See for example D. Hobson’s lecture notes on the Skorokhod embedding problem and Model independent bounds for option prices ([Hob11]). In a model free approach, one does not assume that the underlying belongs to a particular class of processes. The only assumption made is that the underlying has the martingale property (after a change to the risk neutral measure). This is theoretically justified by an absence of arbitrage argument.

The idea then is to study the set of admissible martingales, which are the martingales having the required marginal laws. The set of admissible martingales is incredibly large and complex, and much more research is needed for it to become better understood. In the mean time, any method of constructing elements of this set (such as the methods presented here) improves our understanding of this set of admissible martingales.

1.1.2 Inferring marginal laws from option prices

In this work, we will take the marginal laws of the underlying as given. The problem of finding marginal laws compatible with observed option prices is an area of research in its own. Indeed it constitutes an inverse problem which in each case has many solutions. If there were a continuum of observed option prices, one for each strike, then the problem of recovering the marginal law would have a unique solution. This solution could be obtained by the Breeden and Litzenberger [BL78] formula.

Using the Breeden and Litzenberger formula one can extract the marginal laws from the option prices. We denote by ϕ the risk neutral density of the final spot S_T. As the call price is given by

$$C(S_0,K,T) = \int_{\mathbb{R}} (S_T - K)^+ \phi(S_T,T,S_0) dS_T$$

this can be differentiated twice with respect to the strike K to extract the density ϕ of the marginal law of S at time T.

$$\phi(K_T,S_0) = \frac{\partial^2 C(S_0,K,T)}{\partial K^2}$$

Since option prices are not available for the entire spectrum of strike values, interpolating the available values is necessary.

The state price density is often called the risk neutral density, in our framework it will be called the marginal density. In [JR96] a prior parametric density is postulated as the state price density. In [ASL98], kernel smoothing for this purpose is discussed.

Chapter 2

Quantization and preservation of the convex order

2.1 The convex order and the existence of martingale transitions between specified marginal laws

The convex order ($\leq_{cx}$) is a partial order on $\mathcal{P}(\mathbb{R})$, the space of probability measures on $\mathbb{R}$. It compares probability measures in terms of their dispersion.

Definition 2.1.1. Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$. We say that μ is dominated by ν in the convex order and write $\mu \leq_{cx} \nu$ if, for every convex function $\phi(x)$,

$$\int_{\mathbb{R}} \phi(x)\ d\mu(x) \leq \int_{\mathbb{R}} \phi(x)\ d\nu(x)$$

2.1.1 Characterizations of the convex order

The convex order can be characterized in several ways. In particular it can be characterized in terms of:

- potential functions
- distribution functions
- survival functions
- quantile functions
- put and call functions
- martingale transitions

These characterizations will be used throughout this work. Proofs of these characterizations can be found in the book by Shaked and Shanthikumar [SS06].

Characterization in terms of potential functions:

Definition 2.1.2. The potential function of a measure μ is given by

$$U\mu(t) = -\int_{\mathbb{R}} |t-x| d\mu(x)$$

Criterion 1. $\mu \leq_{cx} \nu$ iff $U\mu(t) \geq U\nu(t)$ for all t

Characterization in terms of distribution functions:

Definition 2.1.3. The distribution of a measure μ is the function $F(t) = \int_{-\infty}^{t} d\mu(x)$.

Criterion 2. Let μ be a measure with distribution function F and ν be a measure with distribution function G. Then

$$\mu \leq_{cx} \nu \iff \begin{cases} \mu \text{ and } \nu \text{ have equal means.} \\ \int_{-\infty}^{x} F(t)dt \leq \int_{-\infty}^{x} G(t)dt \quad \text{for every } x \in \mathbb{R} \end{cases}$$

Characterization in terms of survival functions:

Definition 2.1.4. The survival function of a measure μ is the function $\bar{F}(t) = \int_{t}^{\infty} d\mu(x)$.

Criterion 3. Let μ be a measure with survival function $\bar{F}$ and ν be a measure with survival function $\bar{G}$. Then

$$\mu \leq_{cx} \nu \iff \begin{cases} \mu \text{ and } \nu \text{ have equal means.} \\ \int_{x}^{\infty} \bar{F}(t)dt \leq \int_{x}^{\infty} \bar{G}(t)dt \quad \text{for every } x \in \mathbb{R} \end{cases}$$

Characterization in terms of quantile functions:

Definition 2.1.5. The quantile function of a probability measure with distribution function $F(x)$ is:

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : p \leq F(x)\}$$

Criterion 4.

$$\mu \leq_{cx} \nu \iff \begin{cases} \mu \text{ and } \nu \text{ have equal means.} \\ \int_{0}^{p} F^{-1}(u)du \geq \int_{0}^{p} G^{-1}(u)du \quad \text{for every } p \in [0,1] \end{cases}$$

Characterization in terms of call functions:

Definition 2.1.6. The following collection of functions, indexed by $K \in \mathbb{R}$, will be referred to as call functions and defined as:

$$C_K(x) = (x-K)^+ = \max(x-K, 0)$$

Criterion 5.

$$\mu \leq_{cx} \nu \iff \begin{cases} \mu \text{ and } \nu \text{ have equal means.} \\ \int_{\mathbb{R}} C_K(x) d\mu(x) \leq \int_{\mathbb{R}} C_K(x) d\nu(x) \quad \text{for every } K \in \mathbb{R} \end{cases}$$

Characterization in terms of put functions:

Definition 2.1.7. The following collection of functions, indexed by $K \in \mathbb{R}$, will be referred to as put functions and defined as:

$$P_K(x) = (K-x)^+ = \max(K-x, 0)$$

Criterion 6.

$$\mu \leq_{cx} \nu \iff \begin{cases} \mu \text{ and } \nu \text{ have equal means.} \\ \int_{\mathbb{R}} P_K(x) d\mu(x) \geq \int_{\mathbb{R}} P_K(x) d\nu(x) \quad \text{for every } K \in \mathbb{R} \end{cases}$$

Characterization in terms of martingale transitions:

Criterion 7. (Kellerer [Kel72]) $\mu \leq_{cx} \nu$ if and only if there exist random variables X and Y such that:

$$\begin{cases} X \sim \mu \\ Y \sim \nu \\ \mathbb{E}[Y|X] = X \end{cases}$$

2.1.2 Properties of the convex order

Equal means

Lemma 2.1.8. $\mu \leq_{cx} \nu$ implies that μ and ν have equal means.

The proof of this is straightforward:

Proof. $\phi_1(x) = x$ and $\phi_2(x) = -x$ are both convex functions.
Therefore $\mu \leq_{cx} \nu$ implies that $\int x\, d\mu \leq \int x\, d\nu$ and that $-\int x\, d\mu \leq -\int x\, d\nu$
Hence $\int x\, d\mu = \int x\, d\nu$ ◻

Relationship to variance

$\mu \leq_{cx} \nu$ implies that the variance of μ is at most as large as the variance of ν. This is straightforward as $f(x) = x^2$ is a convex function. The converse however is not true. In other words, μ can have a smaller variance than ν yet ν may not dominate μ in the convex order. An example of this is given in Rotschild and Stiglitz [RS70].

2.2 Quantization of measures on $\mathbb{R}$

To quantize a measure is to approximate it by a measure which is supported on a finite number of points. Quantizations of measures on $\mathbb{R}$ will play an important role in this work. We will use quantizations for two different purposes. We will use them to construct a new solution to the Skorokhod embedding problem. We will also use quantizations in order to build martingale transitions between specified marginals. When constructing martingales between specified marginals, we will be interested in quantizations which preserve the convex order. The commonly used quantization method in probability is the $L2$ quantization. We will prove that it does not have the property of preserving the convex order, and we will define a quantization which does have the property of preserving the convex order. Before we do all this we will devote this section to discussing the theory of quantization. In particular we will discuss the commonly used $L2$ quantization.

2.2.1 Voronoi style quantizations

Let μ be the probability measure on $\mathbb{R}$ which we wish to quantize. If we choose a vector of n points $(x_1,..,x_n)$ then a natural way to quantize μ is as follows: For each of x_i, construct an interval A_i, as follows:

$$\begin{cases} \text{if } i=1, & \text{then } A_i = \left(-\infty\,,\ \dfrac{x_1+x_2}{2}\right] \\ \text{if } 2 \leqslant i \leqslant n-1, & \text{then } A_i = \left[\dfrac{x_{i-1}+x_i}{2}\,,\ \dfrac{x_i+x_{i+1}}{2}\right] \\ \text{if } i=n, & \text{then } A_i = \left[\dfrac{x_{n-1}+x_n}{2}\,,\ +\infty\right) \end{cases}$$

Then a quantization of μ can be obtained as follows: For each i, place an atom of mass $\mu(A_i)$ at the position x_i. In other words,

$$\hat{\mu} = \sum_{i=1}^{n} \mu(A_i)\delta_{x_i}$$

This quantization is called the Voronoi quantization of μ, because the intervals A_i are the Voronoi cells corresponding to the points x_i.

Instead of choosing the points $(x_1,..,x_n)$, we could have chosen a partition of $\mathbb{R}$ as n intervals $(A_1, ..,A_n)$. A natural quantization of μ would then be: For each A_i, place an atom of mass $\mu(A_i)$ at the position $\frac{1}{\mu(A_i)}\int_{A_i} x d\mu(x)$. This is pretty much the same type of quantization as the Voronoi quantization. Indeed in the Voronoi quantization, the segments are obtained from the points and here the segments are given directly. We now prove that for these two types of quantizations, the original measure μ dominates its quantization $\hat{\mu}$ in the convex order, i.e. $\hat{\mu} \leq_{cx} \mu$.

Lemma 2.2.1. Let $\mathcal{J}$ be a partition of $\mathbb{R}$. Let $\hat{\mu}$ be the probability measure which is constructed from μ in the following way: for each $J \in \mathcal{J}$, an atom of mass $\mu(J)$ is placed at position $\frac{\int_J x \, d\mu(x)}{\mu(J)}$. Then,

$$\int_{\mathbb{R}} \phi(x) \, d\hat{\mu}(x) \leq \int_{\mathbb{R}} \phi(x) \, d\mu(x) \quad \text{for every convex function } \phi.$$

Proof. Let J be an arbitrary element of $\mathcal{J}$. By construction, $\hat{\mu}(J) = \mu(J)$ and $\int_J x \, d\hat{\mu}(x) = \int_J x \, d\mu(x)$. The measure $\frac{\mu(dx)}{\mu(J)}$ is a probability measure on J. Its expectation is $\frac{\int_J x \, d\mu(x)}{\mu(J)}$. Therefore, by Jensen's inequality, for every convex function ϕ,

$$\int_J \phi(x) \, \frac{d\mu(x)}{\mu(J)} \geq \phi\left(\frac{\int_J x \, d\mu(x)}{\mu(J)}\right)$$

As $\hat{\mu}(J) = \mu(J)$, the measure $\frac{\hat{\mu}(dx)}{\mu(J)}$ is a probability measure on J. It consists of a single Dirac point mass at the position $\frac{\int_J x \, d\mu(x)}{\mu(J)}$. Therefore,

$$\int_J \phi(x) \, \frac{d\hat{\mu}(x)}{\mu(J)} = \phi\left(\frac{\int_J x \, d\mu(x)}{\mu(J)}\right)$$

Combining the two above equations,

$$\int_J \phi(x) \, \frac{d\mu(x)}{\mu(J)} \geq \int_J \phi(x) \, \frac{d\hat{\mu}(x)}{\mu(J)}$$

which is equivalent to:

$$\int_J \phi(x) \, d\mu(x) \geq \int_J \phi(x) \, d\hat{\mu}(x)$$

As the above holds for each $J \in \mathcal{J}$, and together they constitute a partition of $\mathbb{R}$, it follows that

$$\int_{\mathbb{R}} \phi(x) \, d\mu(x) \geq \int_{\mathbb{R}} \phi(x) \, d\hat{\mu}(x)$$

□

2.2.2 L2 quantization

To quantize a random variable X is to approximate it by a random variable $\hat{X}$ which has a support consisting of n points. The resulting quadratic error is given by:

$$\mathbb{E}|X - \hat{X}|^2$$

The $L2$ quantization of X is the random variable $\hat{X}$, supported on n points which minimizes the quadratic error.

2.2.3 Lloyd's fixed point algorithm for performing $L2$ quantizations

Algorithm 1 Lloyd's fixed point algorithm for performing $L2$ quantizations

Let μ be a probability measure on $\mathbb{R}$. The $L2$ quantization of μ can be computed using Lloyd's [Llo82] algorithm as follows.

Initial step: *Seeding the algorithm.*
In order to seed the algorithm, pick n arbitray real numbers, $x_1,..,x_n$.

Step 1: *From points to intervals.*
Suppose that the x_i's are sorted in increasing order. For each of x_i, construct an interval A_i, as follows:

$$\begin{cases} \text{if } i=1, & \text{then } A_i = \left(-\infty\ ,\ \dfrac{x_1+x_2}{2}\right] \\ \text{if } 2 \leqslant i \leqslant n-1, & \text{then } A_i = \left[\dfrac{x_{i-1}+x_i}{2}\ ,\ \dfrac{x_i+x_{i+1}}{2}\right] \\ \text{if } i=n, & \text{then } A_i = \left[\dfrac{x_{n-1}+x_n}{2}\ ,\ +\infty\right) \end{cases}$$

Step 2: *From intervals to points.*
For each interval A_i, compute:

$$x_i^{'} = \frac{1}{\mu(A_i)} \int_{A_i} x d\mu(x)$$

update x_i to this new value.

Step 3: repeat steps 1 and 2 until convergence of the x_i's

Result The $L2$ quantization of μ is obtained as follows. For each i, place an atom of mass $\mu(A_i)$ at the position x_i. In other words,

$$\hat{\mu} = \sum_{i=1}^{n} \mu(A_i)\delta_{x_i}$$

Remark. In step 1, A_i consists of all points in $\mathbb{R}$ which are closer to x_i than to any of the other x_j. The intervals A_i correspond to Voronoi cells.

2.3 The $L2$ quantization does not preserve the convex order

2.3.1 Quantization of probability measures and the $L2$-quantization method.

A quantization of order n of a measure μ is a measure $\hat{\mu}$ which has a support consisting of at most n points. The measure $\hat{\mu}$ should also be a reasonably good approximation of μ.

Definition 2.3.1. Let μ be a probability measure on $\mathbb{R}$. Given a vector $(x_1,..,x_n) \in \mathbb{R}^n$, the *Voronoi quantization* of μ is defined as:

$$\hat{\mu} = \sum_{i=1}^{n} \mu(A_i)\, \delta_{x_i}$$

where A_i is the Voronoi cell of x_i defined as $A_i = \left\{x \in \mathbb{R} : |x - x_i| \leq |x - x_j| \text{ for all } 1 \leq j \leq n\right\}$ and δ_{x_i} denotes the Dirac point mass at x_i.

Remark. The vector of points $(x_1,..,x_n)$ is called the quantization grid. Note how the quantization grid together with μ uniquely defines $\hat{\mu}$.

Definition 2.3.2. The *quadratic error* of the Voronoi quantization defined above is given by:

$$\sum_{i=1}^{n} \int_{A_i} |x_i - u|^2 \, d\mu(u)$$

Definition 2.3.3. The Voronoi quantization which minimizes the quadratic error is called the *$L2$-quantization*.

2.3.2 The $L2$-quantization method does not preserve the convex order

In this section we show that the $L2$-quantization does not preserve the convex order. There are several characterizations of the convex order (see [SS06]). We will make use of two of these characterizations. The first one is in terms of potential functions, the second in terms of martingale transitions. These are given in Lemma 2.3.5 below.

Definition 2.3.4. The potential of a probability measure ρ is the function:

$$U\rho(t) = -\int_{\mathbb{R}} |\, x - t|\, d\rho(x)$$

Lemma 2.3.5. Let μ and ν be two probability measures on $\mathbb{R}$. The following are equivalent:

(i) $\mu \leq_{cx} \nu$

(ii) $U\mu(t) \geq U\nu(t)$ for all $t \in \mathbb{R}$

(iii) There exists random variables X and Y satisfying $X \sim \mu$, $Y \sim \nu$ and $\mathbb{E}[Y|X] = X$

Proof. see [SS06]. $\square$

Theorem 2.3.6. *The L2-quantization method does not preserve the convex order.*

Proof. The proof is based on exhibiting a counterexample. Consider the following two measures:

$$\mu = \frac{1}{2}(\delta_{\frac{1}{6}} + \delta_{\frac{5}{6}})$$
$$\nu = \frac{1}{3}(\delta_0 + \delta_{\frac{1}{2}} + \delta_1)$$

The proof proceeds in three steps. *i*) We first prove that $\mu \leq_{cx} \nu$. *ii*) Next we perform $L2$-quantization of μ and ν. *iii*) Finally, we show that the quantized measures are not ordered in the convex order.

Showing that the two original measures are ordered in the convex order.

To show that $\mu \leq_{cx} \nu$, it suffices by lemma 2.3.5 to exhibit two random variables X and Y which satisfy: $X \sim \mu$, $Y \sim \nu$ and $\mathbb{E}[Y|X] = X$. Let $X \sim \mu$ and $Y \sim \nu$ and define a transition as follows:

$$\begin{cases} \mathbb{P}(Y = 0 \mid X = \frac{1}{6}) = \frac{2}{3} \\ \mathbb{P}(Y = \frac{1}{2} \mid X = \frac{1}{6}) = \frac{1}{3} \\ \mathbb{P}(Y = \frac{1}{2} \mid X = \frac{5}{6}) = \frac{1}{3} \\ \mathbb{P}(Y = 1 \mid X = \frac{5}{6}) = \frac{2}{3} \end{cases}$$

We now check that this transition has the martingale property:

$$\mathbb{E}\left[Y \mid X = \frac{1}{6}\right] = 0 \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$
$$\mathbb{E}\left[Y \mid X = \frac{5}{6}\right] = \frac{1}{2} \cdot \frac{1}{3} + \frac{2}{3} \cdot 1 = \frac{5}{6}$$

Therefore $\mathbb{E}[Y \mid X] = X$ which by the criterion 7 (Kellerer) of Chapter 2, implies that $\mu \leq_{cx} \nu$.

Performing the $L2$-quantization of the two original measures.

Let $\hat{\mu}$ and $\hat{\nu}$ respectively denote the $L2$-quantization of order 2 of the measures μ and ν. The support of the measure μ consists of two points, it follows that $\hat{\mu}$ is equal to μ. Indeed taking the support of μ as the quantization grid leads to a quadratic error of zero.
Computing $\hat{\nu}$ amounts to performing the $L2$-quantization of order 2 of the measure $\frac{1}{3}(\delta_0 + \delta_{\frac{1}{2}} + \delta_1)$. This is a textbook example which can be found in the lecture notes of H. Pham (see [Pha12]) . For the purpose of completeness, we reproduce

and expand the calculations here.

Because $L2$-quantization is a Voronoi style quantization, $\hat{\nu}$ is determined by its support through:

$$\hat{\nu} = \sum_{i=1}^{n} \nu(A_i)\ \delta_{x_i}$$

where $(x_1,..,x_n)$ is the support of $\hat{\nu}$ and A_i is the Voronoi cell of x_i. In fact since $\hat{\nu}$ is a quantization of order 2, its support consists of at most 2 points. Let us denote these two support points by $a = x_1$ and $b = x_2$ and without loss of generality let $a \leq b$. Note that $\hat{\nu}$ will be supported by a single point if and only if $a = b$. We must determine a and b by minimizing the quadratic error function. It turns out that the quadratic error function has a different expression in each of the two following possible cases:

$$\begin{cases} \text{case } (i): & \mid a - \frac{1}{2} \mid < \mid b - \frac{1}{2} \mid \\ \text{case } (ii): & \mid a - \frac{1}{2} \mid > \mid b - \frac{1}{2} \mid \end{cases}$$

The point $\frac{1}{2}$ belongs to the Voronoi cell of a in case (i) and to the Voronoi cell of b in case (ii). Each of these two cases leads to a different quantization of ν. Let us determine the quantization resulting from the case (i) where $\mid a - \frac{1}{2} \mid < \mid b - \frac{1}{2} \mid$. The result which we establish holds true in case (ii) as well.
Since the point $\frac{1}{2}$ belongs to the Voronoi cell of a, the quadratic error function is given by:

$$\begin{aligned} E(a,b) &= \frac{1}{3}[(a-0)^2 + (a-\frac{1}{2})^2 + (b-1)^2] \\ &= \frac{1}{3}[a^2 + a^2 - a + \frac{1}{4} + (b-1)^2] \\ &= \frac{1}{3}[2(a-\frac{1}{4})^2 + \frac{1}{8} + (b-1)^2] \end{aligned}$$

This function is minimized when $a = \frac{1}{4}$ and $b = 1$. It follows that the support of $\hat{\nu}$ is $\left\{\frac{1}{4}, 1\right\}$. The resulting Voronoi cells are $A_1 = (-\infty, \frac{5}{8}]$ and $A_2 = (\frac{5}{8}, \infty)$. And so $\hat{\nu}$ is given by

$$\begin{aligned} \hat{\nu} &= \sum_{i=1}^{n} \nu(A_i)\ \delta_{x_i} \\ &= \nu(\ (-\infty, \frac{5}{8}]\ )\ \delta_{\frac{1}{4}} + \nu(\ (\frac{5}{8}, \infty)\ )\ \delta_1 \\ &= \frac{2}{3}\delta_{\frac{1}{4}} + \frac{1}{3}\delta_1 \end{aligned}$$

Showing that the two quantized measures are not ordered in the convex order.

By lemma 2.3.5, a necessary and sufficient condition for $\hat{\mu} \leq_{cx} \hat{\nu}$ is that $U\hat{\mu}(t) \geq U\hat{\nu}(t)$ holds for every $t \in \mathbb{R}$. It suffices therefore to exhibit a $t^* \in \mathbb{R}$ such that

$U\hat{\mu}(t^*) < U\hat{\nu}(t^*)$. This is the case when $t^* = \frac{1}{4}$ as we now show by evaluating the potential functions of $\hat{\mu}$ and $\hat{\nu}$.

$$
\begin{aligned}
U\hat{\mu}(\frac{1}{4}) &= -\int_{\mathbb{R}} \left| x - \frac{1}{4} \right| \, d\hat{\mu}(x) \\
&= -\frac{1}{2}\left|\frac{1}{6} - \frac{1}{4}\right| - \frac{1}{2}\left|\frac{5}{6} - \frac{1}{4}\right| \qquad\qquad \text{since } \hat{\mu} = \frac{1}{2}\delta_{\frac{1}{6}} + \frac{1}{2}\delta_{\frac{5}{6}} \\
&= -\frac{1}{3} \\
U\hat{\nu}(\frac{1}{4}) &= -\int_{\mathbb{R}} \left| x - \frac{1}{4} \right| \, d\hat{\nu}(x) \\
&= -\frac{1}{3}\left|\frac{1}{4} - 1\right| \qquad\qquad \text{since } \hat{\nu} = \frac{2}{3}\delta_{\frac{1}{4}} + \frac{1}{3}\delta_1 \\
&= -\frac{1}{4}
\end{aligned}
$$

□

Corollary 2.3.7. *Let μ and ν be a pair of measures which admits a martingale transition. Let $\hat{\mu}$ and $\hat{\nu}$ be their respective L2-quantizations. A martingale transition from $\hat{\mu}$ to $\hat{\nu}$ does not necessarily exist (because we do not necessarily have $\hat{\mu} \leq_{cx} \hat{\nu}$).*

Proof. This follows from Theorem 2.3.6 and the characterization of the convex order in terms of martingale transitions given in Lemma 2.3.5.

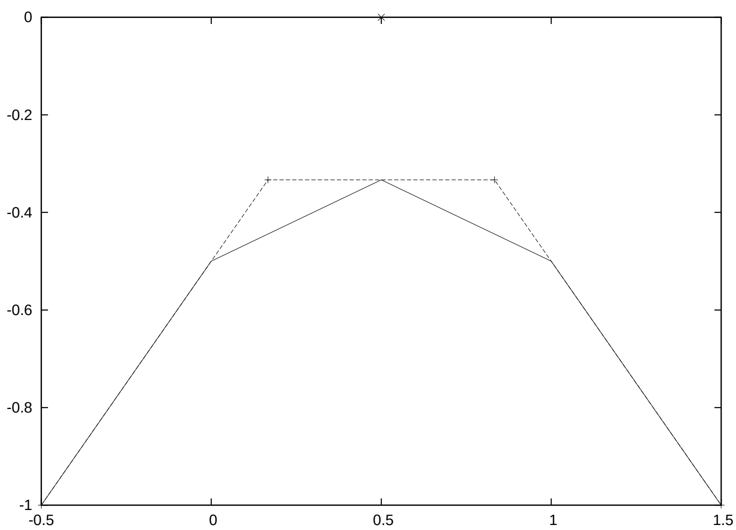


Figure 2.1: The potentials of the two measures before L2 quantization

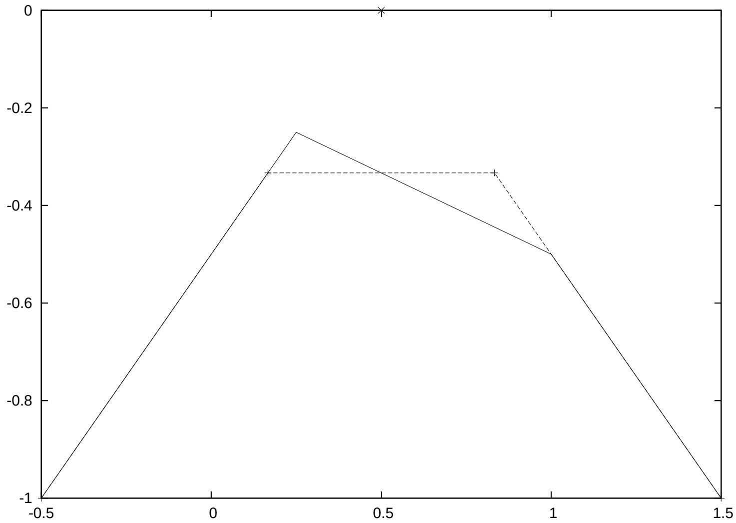


Figure 2.2: The potentials of the two measures after L2 quantization (note that by Criterion 1 of section 2, neither of the quantized measures dominates the other in the convex order).

□

2.4 A quantization which preserves the convex order

We have just seen that the $L2$ quantization, which is the commonly used method to quantize probability measures, does not preserve the convex order. In this section we provide a quantization which does have the property of preserving the convex order. This quantization will be called $\mathcal{U}$-quantization because it produces a quantization which is uniformly distributed on a finite number of support points.

2.4.1 Definition of $\mathcal{U}$-quantization

The $\mathcal{U}$-quantization of a measure is defined in terms of the quantile function of that measure. The quantile function of a measure is defined as follows:

Definition 2.4.1. The quantile function of a probability measure with distribution function $F(x)$ is:

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : p \leq F(x)\}$$

Definition 2.4.2. Choose an integer n. Let $\mu \in \mathcal{P}(\mathbb{R})$ with distribution function $F(u) = \int_{-\infty}^{u} d\mu(x)$.

$$\text{The } \mathcal{U}\text{-quantization of } \mu \text{ is} \begin{cases} U(a_1,..,a_n) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \delta_{a_i} \\ \text{where } a_i = n\displaystyle\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du \end{cases}$$

2.4.2 Numerical illustration of $\mathcal{U}$-quantization

Example 2.4.3. Let μ be the standard (mean 0, variance 1) Gaussian law and let ν be a (mean 0 and variance 2) Gaussian law. Let $U(a_1,..,a_{10})$ and $U(b_1,..,b_{10})$ be the respective quantizations of μ and ν (we chose n=10). Using numerical integration we can compute the vectors $(a_1,..,a_{10})$ and $(b_1,..,b_{10})$:

$$(a_1,..,a_{10})^T = \begin{bmatrix} -1.75498 \\ -1.04464 \\ -0.67731 \\ -0.38650 \\ -0.12600 \\ 0.12600 \\ 0.38650 \\ 0.67731 \\ 1.04464 \\ 1.75498 \end{bmatrix} \qquad (b_1,..,b_{10})^T = \begin{bmatrix} -2.48192 \\ -1.47734 \\ -0.95786 \\ -0.54659 \\ -0.17819 \\ 0.17819 \\ 0.54659 \\ 0.95786 \\ 1.47734 \\ 2.48192 \end{bmatrix}$$

Lemma 2.4.4. ($\mathcal{U}$-quantization preserves the mean of a measure) Let μ be a probability measure with distribution function F, and $U(a_1,..a_n)$ be its $\mathcal{U}$-quantization. Then μ and $U(a_1,..a_n)$ have the same mean.

Proof.

$$\begin{aligned}
\text{The mean of } U(a_1,..a_n) &= \frac{1}{n}\sum_{i=1}^{n} a_i \\
&= \frac{1}{n}\sum_{i=1}^{n} n\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du \\
&= \int_0^1 F^{-1}(u)du \\
&= \text{the mean of } \mu
\end{aligned}$$

□

2.4.3 $\mathcal{U}$-quantization preserves the convex order

To show that $\mathcal{U}$-quantization preserves the convex order we will need the notion of majorization which is a partial order which compares vectors of same length and equal mean in terms of the relative dispersion of their coordinates.

Definition 2.4.5. Let $(a_1,..,a_n)$ and $(b_1,..,b_n)$ be two vectors whose entries have been sorted in increasing order.
($a_1 \leq .. \leq a_n$ and $b_1 \leq .. \leq b_n$)
We say that $(a_1,..a_n)$ is majorized by $(b_1,..,b_n)$, and write $(a_1,..a_n) \prec (b_1,..,b_n)$ if:

$$\begin{cases} & (i)\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i \\ \text{and} & (ii)\sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} a_i \qquad \text{for } k = 1,..,n-1 \end{cases}$$

Examples 2.4.6. $(1,2,3) \prec (0,2,4)$ and $(1,1,1,1) \prec (0,0,0,4)$
The vectors $(1,6,6,9)$ and $(2,3,8,9)$ illustrate the fact that $(\prec)$ is a partial order, and not a total order. Both vectors have the same mean, but neither vector majorizes the other. ($1 < 2$ but $1+6 > 2+3$)

Remark. $(a_1,..,a_n) \prec (b_1,..,b_n)$ means that $(b_1,..,b_n)$ is more dispersed than $(a_1,..,a_n)$. In the literature there is no consensus as to the direction of the ordering. In economics $(\prec)$ is called the Lorenz order and compares income inequalities. There, $(a_1,..,a_n) \prec (b_1,..,b_n)$ if $(b_1,..,b_n)$ is less dispersed than $(a_1,..,a_n)$. In this paper, we are using majorization $(\prec)$ alongside the convex order $(\leq_{cx})$. For the convex order, $\mu \leq_{cx} \nu$ means that ν is more dispersed than μ. Therefore, it makes sense for us to choose the definition of $(\prec)$ which we have chosen.

The following lemma gives a characterization of the convex order in terms of the quantile function. We will use this frequently.

Lemma 2.4.7. Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ with distribution functions F and G, then :

$$\mu \leq_{cx} \nu \iff \begin{cases} \mu \text{ and } \nu \text{ have equal means.} \\ \displaystyle\int_0^p F^{-1}(u)du \geq \int_0^p G^{-1}(u)du \quad \text{for every } p \in [0,1] \end{cases}$$

Proof. See [SST94], page 112, Theorem 3.A.5. □

Definition 2.4.8. $U(a_1,..,a_n)$ will denote the law corresponding to $\frac{1}{n}\sum_{i=1}^{n}\delta_{a_i}$ where δ_x is the Dirac point mass at x.

In the following lemma we establish a relationship between the convex order and majorization.

Lemma 2.4.9. $(a_1,..,a_n) \prec (b_1,..,b_n) \iff U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$

Proof. Let us first determine the quantile functions of $U(a_1,..,a_n)$ and $U(b_1,..,b_n)$. Since $U(a_1,..,a_n)$ corresponds to $\frac{1}{n}\sum_{i=1}^{n}\delta_{a_i}$, its distribution function, F, is a piecewise constant function.

$$F(x) = \begin{cases} 0 & \text{for } x \leq a_1 \\ \frac{i}{n} & \text{for } x \in [a_i, a_{i+1}) \\ 1 & \text{for } x \geq a_n \end{cases}$$

The quantile function of $U(a_1,..,a_n)$ is by definition:

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : p \leq F(x)\}$$

It follows that F^{-1} is a piecewise constant function from $(0,1]$ to $\mathbb{R}$ which is given by:

$$F^{-1}(p) = a_i \text{ if } p \in \left(\frac{i-1}{n}, \frac{i}{n}\right]$$

Integrating a piecewise constant function is easy:

$$\int_0^{\frac{k}{n}} F^{-1}(p)\, dp = \frac{1}{n}\sum_{i=1}^{k} a_i$$

In the same way, G^{-1}, the quantile function of $U(b_1,..,b_n)$ satisfies:

$$\int_0^{\frac{k}{n}} G^{-1}(p)\, dp = \frac{1}{n}\sum_{i=1}^{k} b_i$$

Let us first show that $U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n) \Rightarrow (a_1,..,a_n) \prec (b_1,..,b_n)$.
By lemma 2.4.7, $U(a_1,..,a_n) \leq_{cx} U(a_1,..,a_n)$ implies that for each $t \in (0,1)$,

$$\begin{aligned} &\int_0^t F^{-1}(p)\,dp \geq \int_0^t G^{-1}(p)\,dp \\ &\quad \Rightarrow \sum_{i=1}^k a_i \geq \sum_{i=1}^k b_i \text{ for each } k \qquad (*) \end{aligned}$$

$U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$ implies by Lemma 2.4.9 and definition 2.4.5, that they have the same mean, and so:

$$\begin{aligned} &\int_0^1 F^{-1}(p)\,dp = \int_0^1 G^{-1}(p)\,dp \\ &\quad \Rightarrow \sum_{i=1}^n a_i = \sum_{i=1}^n b_i \qquad (**) \end{aligned}$$

Finally, $(*)$ together with $(**)$ imply that $(a_1,..,a_n) \prec (b_1,..,b_n)$.

Let us now show that $(a_1,..,a_n) \prec (b_1,..,b_n) \Rightarrow U(a_1,..,a_n) \leq_{cx} U(a_1,..,a_n)$.
We have seen that F^{-1} is a piecewise constant function which is constant on each of the intervals $\left[\frac{i}{n}, \frac{i+1}{n}\right]$.
It follows that $p \to \int_0^p F^{-1}(t)\,dt$ is a piecewise affine function, which is affine on these same intervals.
The same is true of the function $p \to \int_0^p G^{-1}(t)\,dt$. Therefore to show that

$$\int_0^p G^{-1}(t)dt \leq \int_0^p F^{-1}(t)\,dt \ \text{ for all } p \in (0,1)$$

it suffices to show that

$$\int_0^{i/n} G^{-1}(t)dt \leq \int_0^{i/n} F^{-1}(t)\,dt \ \text{ for each } i \in \{1,2,..,n\}$$

which by what we have shown at the beginning of the proof is equivalent to:

$$\frac{1}{n}\sum_{i=1}^k b_i \leq \frac{1}{n}\sum_{i=1}^k a_i \ \text{ for each } k \in \{1,2,..,n\}$$

which follows from the initial assumption that $(a_1,..,a_n) \prec (b_1,..,b_n)$. Therefore

$$\int_0^p G^{-1}(t)dt \leq \int_0^p F^{-1}(t)dt \ \text{ for each } p \text{ in } (0,1)$$

which gives, by lemma 2.4.7, that $U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$. □

$\mathcal{U}$-quantization is a bridge between the convex order ($\leq_{cx}$) and majorization ($\prec$). When measures are ordered in the convex order, the coordinates of their $\mathcal{U}$-quantizations are ordered in the majorization order.

Theorem 2.4.10. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$, with $\mathcal{U}$-quantizations $U(a_1,..,a_n)$ and $U(b_1,..,b_n)$. If $\mu \leq_{cx} \nu$ then $(a_1,..,a_n) \prec (b_1,..,b_n)$.*

Proof. Suppose that $\mu \leq_{cx} \nu$. By the definition of majorization ($\prec$), it suffices to show:

$$\begin{cases} (a_1,..,a_n) \text{ and } (b_1,..,b_n) \text{ have the same mean } & (i) \\ \sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} a_i \quad \text{for each } k \in \{1,..,n\} & (ii) \end{cases}$$

(i) Let us show that the vectors $(a_1,..,a_n)$ and $(b_1,..,b_n)$ have the same mean. Since $\mu \leq_{cx} \nu$, it follows that μ and ν have the same mean (see section on the convex order). Since $\mathcal{U}$-quantization preserves the mean of a probability measure, it follows that the measures $U(a_1,..,a_n)$ and $U(b_1,..,b_n)$ have the same mean. This implies that $\frac{1}{n}\sum_{i=1}^{n} a_i = \frac{1}{n}\sum_{i=1}^{n} b_i$, hence the vectors $(a_1,..,a_n)$ and $(b_1,..,b_n)$ have the same mean.

(ii) Letting F (resp. G) denote the distribution function of μ (resp. ν), we have:

$$a_i = n\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du$$
$$\sum_{i=1}^{k} a_i = \sum_{i=1}^{k} n\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du$$
$$\sum_{i=1}^{k} a_i = n\int_{0}^{\frac{k}{n}} F^{-1}(u)du$$

In the same way, one also obtains,

$$\sum_{i=1}^{k} b_i = n\int_{0}^{\frac{k}{n}} G^{-1}(u)du$$

As, $\mu \leq_{cx} \nu$ it follows by the characterization of the convex order using quantile functions that:

$$\int_0^p G^{-1}(u)du \leq \int_0^p F^{-1}(u)du \text{ for every } p \in [0,1]$$
$$\text{Hence } \sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} a_i \quad \text{for each } k \in \{1,..,n\}$$

□

Remark. It can be shown that $\mu \geq_{cx} U(a_1,..a_n)$ (see Section 4, lemma 3.6.5). Although this is common for several quantization methods, what is more remarkable is that this quantization preserves the convex order, as we now show:

Theorem 2.4.11. *($\mathcal{U}$-quantization preserves the convex order) Let $\mu,\nu \in \mathcal{P}(\mathbb{R})$ with quantizations $U(a_1,..,a_n)$ and $U(b_1,..,b_n)$. If $\mu \leq_{cx} \nu$ then $U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$.*

Proof. Suppose that $\mu \leq_{cx} \nu$. By Theorem 2.4.10 this implies that $(a_1,..,a_n) \prec (b_1,..,b_n)$. By Lemma 2.4.9 it follows that $U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$. $\square$

The quantization defined above would not be of much use if $U(a_1,..,a_n)$ did not converge to μ. Thankfully this is the case as the following theorem shows.

Theorem 2.4.12. *Let $\mu \in \mathcal{P}(\mathbb{R})$ with quantization $U(a_1,..,a_n)$. Then as n goes to infinity, $U(a_1,..,a_n)$ converges weakly to μ.*

Proof. Recall that

$$U(a_1,..,a_n) = \frac{1}{n}\sum_{i=1}^{n}\delta_{a_i} \qquad \text{where } a_i = n\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du$$

The cumulative distribution function F is a non-decreasing function, hence it follows that its inverse, F^{-1} is also a non-decreasing function. As the integrand is a non-decreasing function, the above integral may be bounded as follows:

$$\begin{aligned}
n(\frac{i}{n}-\frac{i-1}{n})F^{-1}(\frac{i-1}{n}) \leq \ & n\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du & \leq n\,(\frac{i}{n}-\frac{i-1}{n})\,F^{-1}(\frac{i}{n})\\
F^{-1}(\frac{i-1}{n}) \leq \ & n\int_{\frac{i-1}{n}}^{\frac{i}{n}} F^{-1}(u)du & \leq F^{-1}(\frac{i}{n})\\
F^{-1}(\frac{i-1}{n}) \leq \ & \qquad a_i & \leq F^{-1}(\frac{i}{n})
\end{aligned}$$

Let F_n denote the distribution function of $U(a_1,..,a_n)$. By the definition of $U(a_1,..,a_n)$,

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{a_i \leq t\}$$

where $\mathbf{1}$ denotes the indicator function.
Let us now examine $F(t)$ and $F_n(t)$ when $t \in [F^{-1}(\frac{i-1}{n})\ ,\ F^{-1}(\frac{i}{n})]$
Since F is a monotone increasing function, applying F to each term of $F^{-1}(\frac{i-1}{n}) \leq t \leq F^{-1}(\frac{i}{n})$, we obtain:

$$\frac{i-1}{n} \leq F(t) \leq \frac{i}{n} \qquad\qquad (*)$$

Again when $t \in [F^{-1}(\frac{i-1}{n})\ ,\ F^{-1}(\frac{i}{n})]$, we bound $F_n(t)$, the distribution function of $U(a_1,..,a_n)$, in the following way:
We have seen that:

$$a_{i-1} \leq F^{-1}(\frac{i-1}{n}) \leq a_i \leq F^{-1}(\frac{i}{n}) \leq a_{i+1}$$

It follows that when $F^{-1}(\frac{i-1}{n}) \leq t \leq F^{-1}(\frac{i}{n})$, one must have either $a_{i-1} \leq t \leq a_i$ or $a_i \leq t \leq a_{i+1}$. Therefore when $F^{-1}(\frac{i-1}{n}) \leq t \leq F^{-1}(\frac{i}{n})$, the distribution function F_n which is $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{a_i \leq t\}$ must be equal to one of the 3 following values: $\frac{i-1}{n}$ or $\frac{i}{n}$ or $\frac{i+1}{n}$.
It follows that when $F^{-1}(\frac{i-1}{n}) \leq t \leq F^{-1}(\frac{i}{n})$ the following must hold:

$$\frac{i-1}{n} \leq F_n(t) \leq \frac{i+1}{n} \qquad (**)$$

By (*) and (**), it follows that when $F^{-1}(\frac{i-1}{n}) \leq t \leq F^{-1}(\frac{i}{n})$, we have:

$$|F(t) - F_n(t)| \leq \frac{2}{n}$$

Now the collection of intervals $\{\big(F^{-1}(\frac{i-1}{n})\ ,\ F^{-1}(\frac{i}{n})\big]\ :\ 1 \leq i \leq n\}$ generate the support of μ, and therefore that of $U(a_1,..,a_n)$. From this we conclude that $\mid F_n(t) - F(t) \mid \leq \frac{2}{n}$ for all $t \in \mathbb{R}$. Hence as $n \to \infty$, the distribution function F_n converges pointwise to F. This means that as $n \to \infty$, the quantization $U(a_1,..,a_n)$ converges weakly to μ $\square$

Chapter 3

Construction of martingale transition between quantized measures

3.1 Martingale transitions though linear programming

3.1.1 Linear programming

Linear programming consists in optimizing a linear function subject to a set of linear constraints. The linear function to be optimized is called the objective function. Let the vector of variables be denoted $x = (x_1, .., x_n)$. A linear function in these variables is of the form:

$$\sum_{i=1}^{n} c_i x_i$$

where $c = (c_1, .., c_n)$ is the vector of coefficients of the objective function. Linear constraints can be of the following forms:

$$\sum_{i=1}^{n} a_i x_i \leq b \quad \text{(an upper bound constraint)}$$
$$\sum_{i=1}^{n} a_i x_i \geq b \quad \text{(a lower bound constraint)}$$
$$\sum_{i=1}^{n} a_i x_i = b \quad \text{(an equality constraint)}$$

where $a = (a_1, .., a_n)$ is a vector of constraint coefficients.
A collection of k linear constraints can be represented by:
a matrix A which has k rows and n columns,
a vector $b = (b_1..b_k)^T$,
in the following way:

$$Ax = b$$

(for k equality constraints)

3.1.2 The Linear programming problem in standard form

The linear programming problem can be stated as:

$$\begin{aligned}&\max c^T x\\&\text{subject to } Ax = b\\&\text{and } x \geq 0\end{aligned}$$

x is the vector of variables to be determined.
c is the vector of coefficients of the objective function.
A is the matrix of constraint coefficients.
b is the vector of right hand side values of the constraints.

The set of constraints:

$$\begin{aligned}&Ax = b\\&\text{and } x \geq 0\end{aligned}$$

specify a convex polytope over which the objective function $c^T x$ is to be optimized.

3.1.3 Solutions to linear programs

The simplex algorithm developed by Danzig solves a linear program when it has a solution (see [WD49], [Dan49] and [Dan98]).

3.1.4 Martingale transitions as solutions to linear programming problems

Given two specified marginal laws μ and ν, we have seen how $\mathcal{U}$-quantization provides us with two quantized measures $\hat{\mu}$ and $\hat{\nu}$. Both $\hat{\mu}$ and $\hat{\nu}$ are uniform laws on n support points.

$$\begin{aligned}\hat{\mu} &\sim U(a_1, .., a_n)\\ \hat{\nu} &\sim U(b_1, .., b_n)\end{aligned}$$

Let us now look at how linear programming provides us with martingale transitions from $\hat{\mu}$ to $\hat{\nu}$. A martingale transition from $\hat{\mu}$ to $\hat{\nu}$ can be expressed as a matrix $M = (m_{i,j})_{1 \leq i,j \leq n}$. The matrix M describes the transition probabilities through:

$$m_{i,j} = P(Y = b_j \mid X = a_i)$$

For each row i, we must have $\sum_{j=1}^{n} m_{i,j} = 1$.
For each column j, we must have $\sum_{i=1}^{n} m_{i,j} = 1$.

As there are n rows and n columns, together these row sums and column sums conditions impose $2n$ linear constraints on the entries of the matrix M. The martingale property of the transition matrix M translates to:

$$\text{for each row } i, \ \sum_{j=1}^{n} m_{i,j} b_j = a_i$$

As there are n rows, the martingale condition translates into n more linear constraints on the entries of the matrix M. We are in a situation with n^2 variables (the entries of the matrix M) subject to $3n$ linear constraints. Each of the linear constraints is an equality constraint. In order to have a linear programming problem we must specify a linear objective function which is to be maximized or minimized. Any vector of n^2 real numbers can be used as coefficients for the linear objective function. A vector c with n^2 entries defines a linear objective function through:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \left(m_{i,j} \ c_{i+n(j-1)} \right)$$

By choosing different vectors of objective function coefficients (the vector c), we can specify different linear programming problems and thus obtain different martingale transitions from $\hat{\mu}$ to $\hat{\nu}$.

We have seen that when a linear programming problem is given in standard form, the linear constraints are provided as a matrix of constraint coefficients together with a vector of right hand side values for the constraints. We will now provide an algorithm which constructs this matrix of constraint coefficients as well as the vector of right hand side values. This algorithm works under the assumptions outlined above (i.e. construction of a martingale transition from $\hat{\mu} = U(a_1, .., a_n)$ to $\hat{\nu} = U(b_1, .., b_n)$). This algorithm takes as arguments the two vectors of support points $(a_1, .., a_n)$ and $(b_1, .., b_n)$. It produces a $3n$ by n^2 matrix of constraint coefficients as well a vector of right hand side values which has length $3n$. These can then be used as inputs in a linear programming solver.

3.1.5 Algorithm to build the matrix of constraint coefficients

The following two algorithms build the matrix of constraint coefficients and the vector of right hand side values for the constraints.

Algorithm 2 Algorithm which constructs the matrix of constraint coefficients for the linear programming solver

---

for $i = 1 \to n$ **do** $\triangleright$ Linear constraints from the row sums.
  for $j = 1 \to n$ **do**
    $M(i,(i-1)n+j) \leftarrow 1$
  end for
end for
for $i = 1 \to n$ **do** $\triangleright$ Linear constraints from the column sums.
  for $j = 1 \to n$ **do**
    $M(n+i,(j-1)n+i) \leftarrow 1$
  end for
end for
for $i = 1 \to n$ **do** $\triangleright$ Linear constraints from the martingale property.
  for $j = 1 \to n$ **do**
    $M(2n+i,(i-1)n+j) \leftarrow b(j)$
  end for
end for

Algorithm 3 Algorithm which constructs the vector of right hand side constraints for the linear programming solver

---

for $i = 1 \to 2n$ **do** $\triangleright$ Because the matrix must be bistochastic.
  $R(i,1) \leftarrow 1$
end for
for $i = 1 \to n$ **do** $\triangleright$ For the martingale property.
  $R(2n+i,1) \leftarrow a(i)$
end for

Algorithm 4 Algorithm to turn the output of a linear programming solver from vector form into matrix form

---

for $i = 1 \to n$ **do**
  for $j = 1 \to n$ **do**
    $N(i,j) \leftarrow O((i-1)n+j)$
  end for
end for

Remark. GNU Octave provides a linear programming solver, the glpk routine.

3.1.6 Numerical Example

The following example illustrates the use of linear programming as a means of constructing martingale transitions between specified marginal laws. Consider two marginal laws, each one of which is a uniform distribution on the following vectors of support points:

$$\begin{pmatrix} -1.64683 \\ -0.89538 \\ -0.49135 \\ -0.15798 \\ 0.15798 \\ 0.49135 \\ 0.89538 \\ 1.64683 \end{pmatrix} \qquad \begin{pmatrix} -3.29366 \\ -1.79077 \\ -0.98270 \\ -0.31595 \\ 0.31595 \\ 0.98270 \\ 1.79077 \\ 3.29366 \end{pmatrix}$$

These vectors where obtained by performing a $\mathcal{U}$-quantization of order 8 of the following Gaussian laws: the first with parameters (mean 0, variance 1) and the second Gaussian law with parameters (mean 0, variance 2), (see Section 2.4 on $\mathcal{U}$-quantization). Let us take as a vector of objective coefficients, a vector of length $n^2 = 64$ with every entry equal to 1. We obtain the following martingale transition matrix:

$$\begin{pmatrix} 0.00000 & 0.97169 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.02831 \\ 0.00000 & 0.00000 & 0.92986 & 0.05890 & 0.00000 & 0.00000 & 0.00000 & 0.01123 \\ 0.05890 & 0.00000 & 0.00000 & 0.94110 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.13130 & 0.00000 & 0.00000 & 0.00000 & 0.86870 & 0.00000 & 0.00000 & 0.00000 \\ 0.17239 & 0.00000 & 0.00000 & 0.00000 & 0.13130 & 0.69632 & 0.00000 & 0.00000 \\ 0.20730 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.30368 & 0.48901 & 0.00000 \\ 0.20196 & 0.00000 & 0.07014 & 0.00000 & 0.00000 & 0.00000 & 0.51099 & 0.21691 \\ 0.22815 & 0.02831 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.74354 \end{pmatrix}$$

3.2 Martingale transitions obtained from symmetric matrices

Given two specified marginal laws μ and ν, we have seen how $\mathcal{U}$-quantization provides us with two quantized measures $\hat{\mu}$ and $\hat{\nu}$. Both $\hat{\mu}$ and $\hat{\nu}$ are uniform laws on n support points.

$$\hat{\mu} \sim U(a_1,..,a_n)$$
$$\hat{\nu} \sim U(b_1,..,b_n)$$

Any martingale transition from $\hat{\mu}$ to $\hat{\nu}$ can be expressed as a matrix $M = (m_{i,j})_{1\leq i,j\leq n}$ which describes the transition probabilities through:

$$m_{i,j} = P(Y = b_j \mid X = a_i)$$

A square matrix M of size n provides a martingale transition from $\hat{\mu} \sim U(a_1,..,a_n)$ to $\hat{\nu} \sim U(b_1,..,b_n)$ if and only if the following $3n$ conditions are verified:

(a) For each row i, $\sum_{j=1}^{n} m_{i,j} = 1$

(b) For each column j, $\sum_{i=1}^{n} m_{i,j} = 1$

(c) For each row i, $\sum_{j=1}^{n} m_{i,j} b_j = a_i$

3.2.1 Constructing a martingale transition from a symmetric matrix

Now suppose that we have a symmetric matrix S which has spectrum $(b_1,..,b_n)$ and diagonal elements $(a_1,..,a_n)$. We now describe how this matrix S can be used to construct a matrix M which provides a martingale transition from $\hat{\mu} \sim U(a_1,..,a_n)$ to $\hat{\nu} \sim U(b_1,..,b_n)$

By the spectral theorem for symmetric matrices there exits a real orthogonal matrix Q such that

$$S = Q^T \Lambda Q$$

The matrix Λ is the diagonal matrix with entries $(b_1,..,b_n)$.

Let us define M to be the matrix obtained by squaring the entries of Q^T. In other words, the (i,j) entry of M is given by $m_{i,j} = q_{i,j}^2$ where $q_{i,j}$ is the (i,j) entry of Q^T.

We now show that M satisfies the conditions (a), (b) and (c) above, which means that M provides a martingale transition from $\hat{\mu} \sim U(a_1,..,a_n)$ to $\nu \sim U(\hat{b}_1,..,b_n)$.

The rows of an orthogonal matrix form an orthonormal basis and similarly the columns of an orthogonal matrix also form an orthonormal basis. It follows that conditions (a) and (b) are verified.

Let us now verify that the transition described by the matrix M possesses the martingale property. This amounts to verifying condition (c):

$$\text{For each row } i,\ \sum_{j=1}^{n} m_{i,j} b_j = a_i$$

By hypothesis the diagonal elements of S are $(a_1,..,a_n)$, so the (i,i) entry of S is equal to a_i. We have seen that $S = Q^T \Lambda Q$. Let us perform these two matrix multiplications in order to calculate this (i,i) entry of S.

$$Q^T\Lambda = \begin{pmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{pmatrix} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_n \end{pmatrix}$$

From this we see that the matrix $Q^T\Lambda$ has (i,j) entry given by $q_{i,j}b_j$. The row i of the matrix $Q^T\Lambda$ is given by:

$$\begin{pmatrix} q_{i,1}b_1 & q_{i,2}b_2 & .. & q_{i,n}b_n \end{pmatrix}$$

The (i,i) entry of the matrix $Q^T\Lambda Q$ is the inner product of the row i of the matrix $Q^T\Lambda$ with the column j of the matrix Q.

$$\begin{pmatrix} q_{i,1}b_1 & q_{i,2}b_2 & .. & q_{i,n}b_n \end{pmatrix} \begin{pmatrix} q_{i,1} \\ q_{i,2} \\ .. \\ q_{i,n} \end{pmatrix} = \sum_{i=1}^{n} q_{i,j}^2 b_j$$

As $m_{i,j}$ was defined to be $q_{i,j}^2$ and the matrix S has the property that its (i,i) entry is a_i, the above line can be written as:

$$a_i = \sum_{j=1}^{n} m_{i,j} b_j$$

This completes the proof that the matrix M, obtained by squaring the entries of Q^T, provides a martingale transition from $\hat{\mu} \sim U(a_1,..,a_n)$ to $\hat{\mathrm{v}} \sim U(b_1,..,b_n)$.

3.2.2 Existence of symmetric matrices with given diagonal and spectrum

Now that we have seen how a symmetric matrix with properly chosen diagonal and spectrum can be used to produce a martingale transition, a natural question is: when does there exist a symmetric matrix with a given diagonal and spectrum? The answer is provided by the following theorem:

Theorem 3.2.1. *(Horn-Schur [Hor54], [Sch23]) There exists a symmetric matrix with diagonal $(a_1,..,a_n)$ and spectrum $(b_1,..,b_n)$ if and only if $(a_1,..,a_n) \prec (b_1,..,b_n)$*

The symbol $(\prec)$ denotes the partial ordering called majorization which is defined as follows:

Definition 3.2.2. Let $(a_1,..,a_n)$ and $(b_1,..,b_n)$ be two vectors whose entries have been sorted in increasing order ($a_1 \leq .. \leq a_n$ and $b_1 \leq .. \leq b_n$).
We say that $(a_1,..a_n)$ is majorized by $(b_1,..,b_n)$, and write $(a_1,..a_n) \prec (b_1,..,b_n)$ if:

$$(i) \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$$

$$(ii) \sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} a_i \qquad \text{for } k = 1,..,n-1$$

Examples 3.2.3. $(1,2,3) \prec (0,2,4)$ and $(1,1,1,1) \prec (0,0,0,4)$
The vectors $(1,6,6,9)$ and $(2,3,8,9)$ illustrate the fact that $(\prec)$ is a partial order, and not a total order. Both vectors have the same mean, but neither vector majorizes the other. ($1 < 2$ but $1+6 > 2+3$)

Remark. $(a_1,..,a_n) \prec (b_1,..,b_n)$ means that $(b_1,..,b_n)$ is more dispersed than $(a_1,..,a_n)$.

3.2.3 Algorithm for constructing matrices with specified diagonal and spectrum

Algorithm 5 Chan Li Algorithm to construct a symmetric matrix with specified diagonal and spectrum

(Chan-Li [CL83]) $(a_1,\ldots,a_n)$ and $(b_1,\ldots,b_n)$ are given vectors which satisfy $(a_1,\ldots,a_n) \prec (b_1,\ldots,b_n)$, this algorithm constructs a symmetric matrix with diagonal elements $(a_1,\ldots,a_n)$ and eigenvalues $(b_1,\ldots,b_n)$.

Proof. In the case where $n=2$, there is an explicit solution:
Suppose (b_1,b_2) and (a_1,a_2) are two vectors which satisfy $(a_1,a_2) \prec (b_1,b_2)$.
Define the following orthogonal matrix Q as

$$Q = \frac{1}{\sqrt{b_2-b_1}} \begin{bmatrix} \sqrt{b_2-a_1} & -\sqrt{a_1-b_1} \\ \sqrt{a_1-b_1} & \sqrt{b_2-a_1} \end{bmatrix}$$

Now,

$$Q^T \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix} Q = \begin{bmatrix} a_1 & * \\ * & a_2 \end{bmatrix}$$

Eigenvalues are left unchanged by conjugation with an orthogonal matrix. So the matrix on the right hand side is the desired matrix with spectrum (b_1, b_2) and diagonal (a_1, a_2).
In the case where $n > 2$, the algorithm proceeds in a recursive fashion. The main step of the algorithm reduces a problem of size k to a problem of size $k-1$. This main step is applied $n-2$ times, thus reducing a problem of dimension n down to a problem of dimension 2 which has the immediate solution given above. The algorithm starts with the diagonal matrix with entries $(b_1, .., b_n)$. This diagonal matrix is then conjugated $n-1$ times by properly chosen orthogonal matrices. At the end of this process, the diagonal entries are $(a_1, .., a_n)$ and the spectrum, left unchanged through conjugation by orthogonal matrices, is still $(b_1, .., b_n)$. The recursive step of the algorithm works as follows. You start with a square matrix of dimension n whose diagonal elements are $(b_1 \dots b_n)$ in any order. You conjugate it by a permutation matrix so that its (1,1) element is b_1 and its $(2,2)$ element is b_j. Then you conjugate it by an orthogonal matrix in the following way:

$$\begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & I_{n-2} \end{bmatrix} \begin{bmatrix} b_1 & \dots & \dots \\ \dots & b_j & \dots \\ \dots & \dots & \ddots \end{bmatrix} \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & I_{n-2} \end{bmatrix} = \begin{bmatrix} a_1 & \dots & \dots \\ \dots & b_1 + b_j - a_1 & \dots \\ \dots & \dots & \ddots \end{bmatrix}$$

I_{n-2} denotes the identity matrix of dimension $n-2$. The values of s and c are computed in the same way as in the $(n = 2)$ case. The main step of the algorithm is then recursively applied to the submatrix obtained by removing the first row and the first column of the right hand side matrix above. This is possible by lemma 3.2.4 □

The following lemma makes the recursive step in the algorithm possible.

Lemma 3.2.4. (Chan-Li [CL83]) Suppose $\vec{a} = (a_1, .., a_n) \prec (b_1, .., b_n) = \vec{b}$ are two given vectors whose entries have been sorted in increasing order. Denote by b_j the smallest element of $\vec{b}$ which is greater than or equal to a_1 (i.e. $b_{j-1} \leq a_1 \leq b_j$). Define two new vectors $\vec{a}^{new}$ and $\vec{b}^{new}$ as follows: $\vec{a}^{new}$ is obtained by removing a_1 from $\vec{a}$, and $\vec{b}^{new}$ is obtained by removing both b_1 and b_j from $\vec{b}$ and inserting the value $(b_1 + b_j - a_1)$. Then the following holds: $\vec{a}^{new} \prec \vec{b}^{new}$

Proof. We will use the following notation: for a vector $\vec{v}$, $\text{sum}(\vec{v})$ denotes the sum of the entries of $\vec{v}$.
Let us start by verifying that $\text{sum}(\vec{a}^{new}) = \text{sum}(\vec{b}^{new})$.

$$\begin{aligned} \text{sum}(\vec{b}^{new}) &= \text{sum}(\vec{b}) - b_1 - b_j + (b_1 + b_j - a_1) \\ &= \text{sum}(\vec{b}) - a_1 \\ &= \text{sum}(\vec{a}) - a_1 \quad (\text{ indeed } \vec{a} \prec \vec{b} \Rightarrow \text{sum}(\vec{a}) = \text{sum}(\vec{b}) \text{) } \\ &= \text{sum}(\vec{a}^{new}) \end{aligned}$$

Now $\vec{a}^{new}$ and $\vec{b}^{new}$ are both vectors of length $(n-1)$ with the same mean. Let us denote by $\mathrm{sum}(\vec{a}^{new},1,k)$ the sum of the k smallest elements of $\vec{a}$. As we have showed that $\mathrm{sum}(\vec{a}^{new}) = \mathrm{sum}(\vec{b}^{new})$, to show that $\vec{a}^{new} \prec \vec{b}^{new}$ it suffices to show that $\mathrm{sum}(\vec{b}^{new},1,k) \leq \mathrm{sum}(\vec{a}^{new},1,k)$ for each k from 1 to $k-1$. In order to prove this, let us first examine the relative position of the elements of $\vec{a}^{new}$ and $\vec{b}^{new}$.
One of the following statements must hold:

$$\text{either } b_1 \leq .. \leq b_{j-1} \leq a_1 \leq (b_1+b_j-a_1) \leq b_j$$

$$\text{or } b_1 \leq .. \leq b_{j-1} \leq (b_1+b_j-a_1) \leq a_1 \leq b_j$$

$$\text{or } b_1 \leq .. \leq (b_1+b_j-a_1) \leq b_{j-1} \leq a_1 \leq b_j$$

Indeed this is a consequence of the two following two observations:
(i) b_j was chosen so that $b_{j-1} \leq a_1 \leq b_j$.
(ii) (b_1,b_j) and $(a_1,(b_1+b_j-a_1))$ have the same mean and $b_1 \leq a_1$ (as $\vec{a} \leq \vec{b}$). Therefore either $b_1 \leq a_1 \leq (b_1+b_j-a_1) \leq b_j$ or $b_1 \leq (b_1+b_j-a_1) \leq a_1 \leq b_j$

Case $k \leq j-2$:
b_{j-1} is the $(j-2)^{th}$ smallest element of $\vec{b}^{new}$. Let us look at the sum of the k smallest elements of $\vec{b}^{new}$ when $k \leq j-2$. As b_j was chosen so that $b_{j-1} \leq a_1 \leq b_j$, it follows that when $k \leq j-2$, each element in $\mathrm{sum}(\vec{b}^{new},1,k)$ is less than a_1. Hence $\mathrm{sum}(\vec{b}^{new},1,k) \leq k\, a_1 \leq \mathrm{sum}(\vec{a}^{new},1,k)$.

Case $k = j-1$:

$$\begin{cases} \mathrm{sum}(\vec{b}^{new},1,k) & = \left(\sum_{i=2}^{j-1} b_i\right) + (b_1+b_j-a_1) \\ \mathrm{sum}(\vec{a}^{new},1,k) & = \sum_{i=2}^{j} a_i \end{cases}$$

$$\begin{aligned} & \text{Therefore } \mathrm{sum}(\vec{b}^{new},1,k) \leq \mathrm{sum}(\vec{a}^{new},1,k) \\ \iff & \left(\sum_{i=2}^{j-1} b_i\right) + (b_1+b_j-a_1) \ \leq \ \sum_{i=2}^{j} a_i \\ \iff & \sum_{i=1}^{j} b_i \leq \sum_{i=1}^{j} a_i \ \text{, which is true because } \vec{b}^{new} \succ \vec{a}^{new} \end{aligned}$$

Case $k \geq j$:
Note that b_{j+1} is the j^{th} smallest element of b^{new}

$$\begin{cases} \mathrm{sum}(\vec{b}^{new},1,k) & = \left(\sum_{i=2}^{j-1} b_i\right) + (b_1+b_j-a_1) + \sum_{i=j+1}^{k+1} b_i \\ \mathrm{sum}(\vec{a}^{new},1,k) & = \sum_{i=2}^{k+1} a_i \end{cases}$$

Therefore $\text{sum}(\vec{b}^{new}, 1, k) \leq \text{sum}(\vec{a}^{new}, 1, k)$

$$\iff \left(\sum_{i=2}^{j-1} b_i\right) + (b_1 + b_j - a_1) + \sum_{i=j+1}^{k+1} b_i \;\leq\; \sum_{i=2}^{j} a_i$$

$$\iff \sum_{i=1}^{k+1} b_i \leq \sum_{i=1}^{k+1} a_i \quad \text{which is true because } \vec{b}^{new} \succ \vec{a}^{new}$$

$\square$

3.2.4 Numerical example

The following example illustrates the construction of a martingale transition through the construction of a symmetric matrix with specified diagonal and spectrum.

$$\begin{pmatrix} -1.64683 \\ -0.89538 \\ -0.49135 \\ -0.15798 \\ 0.15798 \\ 0.49135 \\ 0.89538 \\ 1.64683 \end{pmatrix} \qquad \begin{pmatrix} -3.29366 \\ -1.79077 \\ -0.98270 \\ -0.31595 \\ 0.31595 \\ 0.98270 \\ 1.79077 \\ 3.29366 \end{pmatrix}$$

These vectors were obtained by performing a $\mathcal{U}$-quantization of order 8 of the following Gaussian laws: the first with parameters (mean 0, variance 1) and the second with parameters (mean 0, variance 2), (see Section 2.4 on $\mathcal{U}$-quantization). Using the Chan Li algorithm we can construct a symmetric matrix which has the first vector as diagonal and the second vector as spectrum. We then use the method described above to construct the martingale transition matrix.

We obtain the following martingale transition matrix:

$$\begin{pmatrix}
0.28738 & 0.00000 & 0.71262 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.41128 & 0.00000 & 0.00000 & 0.00000 & 0.58872 & 0.00000 & 0.00000 \\
0.17847 & 0.00000 & 0.07197 & 0.74955 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.18225 & 0.00000 & 0.07350 & 0.08545 & 0.65880 & 0.00000 & 0.00000 & 0.00000 \\
0.18940 & 0.00000 & 0.07638 & 0.08881 & 0.18365 & 0.00000 & 0.46176 & 0.00000 \\
0.00000 & 0.41833 & 0.00000 & 0.00000 & 0.00000 & 0.29224 & 0.00000 & 0.28943 \\
0.11589 & 0.04887 & 0.04674 & 0.05434 & 0.11237 & 0.03414 & 0.38388 & 0.20378 \\
0.04660 & 0.12153 & 0.01879 & 0.02185 & 0.04518 & 0.08490 & 0.15436 & 0.50679
\end{pmatrix}$$

3.3 Martingale transitions obtained by clipping potentials

We have established that $\mathcal{U}$-quantization has the property of preserving the convex order. That is, if $\mu \leq_{cx} \nu$, their $\mathcal{U}$-quantizations are also ordered in the convex order, i.e. $U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$. In this section we see that this property enables the use of an algorithm by Chacon and Walsh. In doing so, we can construct martingale transitions from $U(a_1,..,a_n)$ to $U(b_1,..,b_n)$.

Definition 3.3.1. The potential of a measure μ is defined to be the function:

$$t \to -\int_{-\infty}^{\infty} |x-t|d\mu(x)$$

The next lemma relates potentials of measures to the convex order. A proof of this lemma can be found in [SST94], on page 111.

Lemma 3.3.2. Let μ and ν be two probability measures on $\mathbb{R}$. Let f (resp. g) be the potential of μ (resp. ν).

$$\mu \leq_{cx} \nu \iff f \geq g$$

We now detail the main step of the Chacon-Walsh algorithm which Chacon and Walsh introduced to give a new solution to the Skorokhod embedding problem [CW76]. Let f be the potential function of a probability measure μ. Choose a line L which intersects the graph of f in two points. Denote these two points and their coordinates by $A = (A_x, A_y)$ and $B = (B_x, B_y)$. Without loss of generality, let $A_x < B_x$.

Define the function g by:

$$g(x) = \begin{cases} f(x) & \text{if } x \in (-\infty, A_x) \cup (B_x, \infty) \\ A_y + (x-A_x)\dfrac{(B_y - A_y)}{(B_x - A_x)} & \text{if } x \in [A_x, B_x] \end{cases}$$

The function g is also the potential function of a probability measure. Let B be a Brownian motion with initial law $B_0 \sim \mu$. Let T be the following stopping time for the Brownian motion B:

$$\begin{cases} \text{if } B_0 \in (-\infty, A_x) \cup (B_x, \infty) & \text{then } T = 0 \\ \text{if } B_0 \in [A_x, B_x] & \text{then } T = \inf\{t \geq 0 : B_t = A_x \text{ or } B_t = B_x\} \end{cases}$$

T is defined so that the law of B_T has potential function g.
Brownian motion is a martingale, and since $\mathbb{E}[T] < \infty$, it follows that $\mathbb{E}[B_T|B_0] = B_0$. Therefore the transition (B_0, B_T) has the martingale property. From the definition of T, the transition (B_0, B_T) is seen to be:

$$\begin{cases} \text{If } x \in (\infty, A_x) \cup (B_x, \infty) \text{ then } x \text{ transits to } x \text{ with probability } 1. \\ \text{If } x \in [A_x, B_x] \text{ then } \begin{cases} x \text{ transits to } A_x \text{ with probability } (B_x - x)/(B_x - Ax) \\ x \text{ transits to } B_x \text{ with probability } (x - A_x)/(B_x - Ax) \end{cases} \end{cases}$$

This step of the algorithm illustrates how to explicitly obtain the unique martingale transition between the probability measure with potential f and the probability measure with potential g.

Remark. We will call this procedure **clipping**. The potential f was clipped using L to produce g. Clipping using a segment will mean clipping using the line which contains that segment.

This brings us to the following corollary of Theorem 2.4.11

Corollary 3.3.3. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$, satisfying $\mu \leq_{cx} \nu$, and denote their quantizations by $U(a_1,..,a_n)$ and $U(b_1,..,b_n)$. A martingale transition from $U(a_1,..,a_n)$ to $U(b_1,..,b_n)$ can be generated by the Chacon-Walsh algorithm in (n+1) steps.*

Proof. Suppose $\mu \leq_{cx} \nu$. Since $\mathcal{U}$-quantization preserves the convex order (Theorem 2.4.11), it follows that $U(a_1,..,a_n) \leq_{cx} U(b_1,..,b_n)$. Denote by $f(t)$ the potential of $U(a_1,\dots a_n)$ and by $g(t)$ the potential of $U(b_1,\dots b_n)$. Lemma 3.3.2 implies that $f(t) \geq g(t)$ for every $t \in \mathbb{R}$. Note that f and g are both piecewise affine functions. The graph of g is composed of $(n-1)$ segments and 2 half-lines. We may clip f by each of these segments and half lines. By doing this we obtain after $n+1$ clippings a martingale transition from $U(a_1,\dots a_n)$ to $U(b_1,\dots b_n)$.

□

3.3.1 Implementation of the Chacon Walsh algorithm for $\mathcal{U}$-quantization

To implement the algorithm, one needs the coordinate of the intersection point of two lines (see [Wei12])

Let L_1 and L_2 be two lines in the plane, with L_1 going through the points (x_1, y_1) and (x_2, y_2) and L_2 going through the points (x_3, y_3) and (x_4, y_4). Then the intersection point has the following x and y coordinates: The x coordinate of the intersection point is given by:

$$x = \frac{\begin{vmatrix} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} & \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} \\ \begin{vmatrix} x_3 & y_3 \\ x_4 & y_4 \end{vmatrix} & \begin{vmatrix} x_3 & 1 \\ x_4 & 1 \end{vmatrix} \end{vmatrix}}{\begin{vmatrix} \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} & \begin{vmatrix} y_1 & 1 \\ y_2 & 1 \end{vmatrix} \\ \begin{vmatrix} x_3 & 1 \\ x_4 & 1 \end{vmatrix} & \begin{vmatrix} y_3 & 1 \\ y_4 & 1 \end{vmatrix} \end{vmatrix}} = \frac{\begin{vmatrix} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} & x_1 - x_2 \\ \begin{vmatrix} x_3 & y_3 \\ x_4 & y_4 \end{vmatrix} & x_3 - x_4 \end{vmatrix}}{\begin{vmatrix} x_1 - x_2 & y_1 - y_2 \\ x_3 - x_4 & y_3 - y_4 \end{vmatrix}}$$

Similarly, the y coordinate of the intersection point is given by

$$y = \frac{\begin{vmatrix} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} & \begin{vmatrix} y_1 & 1 \\ y_2 & 1 \end{vmatrix} \\ \begin{vmatrix} x_3 & y_3 \\ x_4 & y_4 \end{vmatrix} & \begin{vmatrix} y_3 & 1 \\ y_4 & 1 \end{vmatrix} \end{vmatrix}}{\begin{vmatrix} \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} & \begin{vmatrix} y_1 & 1 \\ y_2 & 1 \end{vmatrix} \\ \begin{vmatrix} x_3 & 1 \\ x_4 & 1 \end{vmatrix} & \begin{vmatrix} y_3 & 1 \\ y_4 & 1 \end{vmatrix} \end{vmatrix}} = \frac{\begin{vmatrix} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} & y_1 - y_2 \\ \begin{vmatrix} x_3 & y_3 \\ x_4 & y_4 \end{vmatrix} & y_3 - y_4 \end{vmatrix}}{\begin{vmatrix} x_1 - x_2 & y_1 - y_2 \\ x_3 - x_4 & y_3 - y_4 \end{vmatrix}}$$

Here

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

denotes the determinant of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

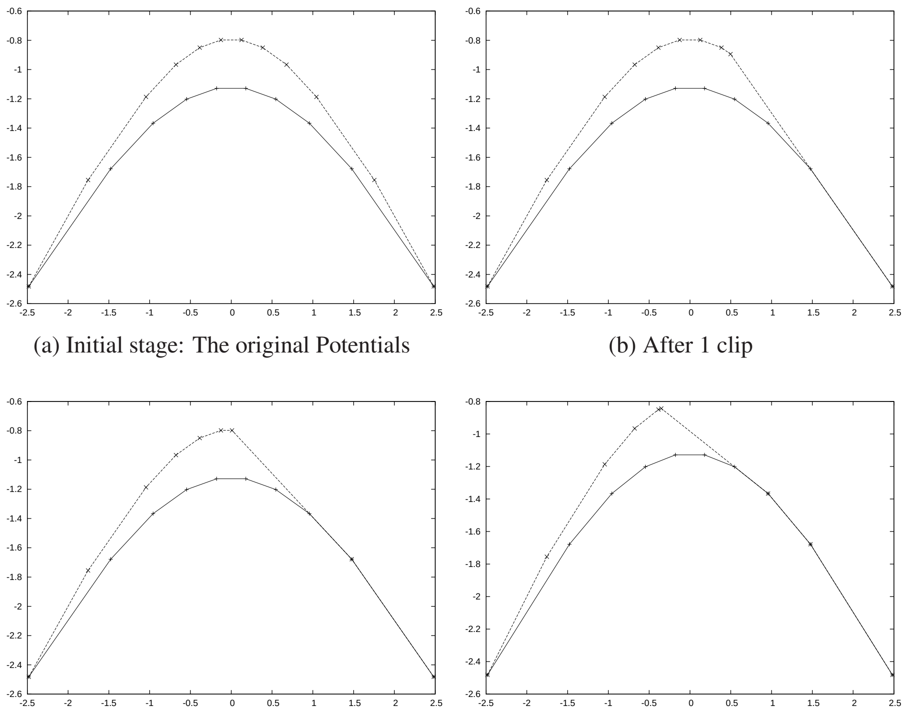
Example 3.3.4. We have obtained the $\mathcal{U}$-quantizations $U(a_1,..,a_8)$ and $U(b_1,..,b_8)$ for the Gaussian laws $N(0,1)$ and $N(0,2)$.

$$\begin{pmatrix} -1.64683 \\ -0.89538 \\ -0.49135 \\ -0.15798 \\ 0.15798 \\ 0.49135 \\ 0.89538 \\ 1.64683 \end{pmatrix} \qquad \begin{pmatrix} -3.29366 \\ -1.79077 \\ -0.98270 \\ -0.31595 \\ 0.31595 \\ 0.98270 \\ 1.79077 \\ 3.29366 \end{pmatrix}$$

We can now use the method described in this section to compute a martingale transition matrix from $U(a_1,..,a_8)$ to $U(b_1,..,b_8)$. This martingale transition is represented by a matrix $M = \{m_{ij}\}$ which provides the transition probabilities $m_{ij} = \mathbb{P}(Y = b_j \mid X = a_i)$.

There are 4 canonical ways to clip:
1: clockwise
2: counter clockwise
3: extremities to center
4: center to extremities

The following diagrams illustrate *counterclockwise* clipping of the potential function.



(a) Initial stage: The original Potentials

(b) After 1 clip

(c) After 2 clips

(d) After 3 clips

Figure 3.1: Counter clockwise clipping of the potential: the first 3 steps

3.3.2 Clipping from extremities to center

Clipping *from extremities to center* produces the following martingale transition:

$$\begin{pmatrix} 0.505115 & 0.122688 & 0.083704 & 0.072123 & 0.056251 & 0.053373 & 0.053373 & 0.053373 \\ 0.279299 & 0.178670 & 0.121898 & 0.105033 & 0.081918 & 0.077727 & 0.077727 & 0.077727 \\ 0.157884 & 0.208771 & 0.142434 & 0.122728 & 0.095718 & 0.090822 & 0.090822 & 0.090822 \\ 0.057702 & 0.233607 & 0.159378 & 0.137328 & 0.107105 & 0.101626 & 0.101626 & 0.101626 \\ 0.000000 & 0.196832 & 0.180625 & 0.155636 & 0.121384 & 0.115175 & 0.115175 & 0.115175 \\ 0.000000 & 0.059433 & 0.211525 & 0.182261 & 0.142149 & 0.134878 & 0.134878 & 0.134878 \\ 0.000000 & 0.000000 & 0.100437 & 0.224891 & 0.175397 & 0.166425 & 0.166425 & 0.166425 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.220078 & 0.259974 & 0.259974 & 0.259974 \end{pmatrix}$$

The expected variance of a martingale transition from $\hat{\mu}$ to $\hat{\nu}$ is entirely determined by those marginals ($\hat{\mu}$ and $\hat{\nu}$). In fact it is equal to the area between the potentials of $\hat{\mu}$ and $\hat{\nu}$. We now show this. For this reason it will be of interest to study the conditional variance of each martingale transition from $\hat{\mu}$ to $\hat{\nu}$ as these are different.

Lemma 3.3.5. Let (X,Y) be a 2 step martingale. The variance of the martingale increment $Y-X$ is uniquely determined by the variance of X and that of Y. More precisely, $\mathbb{V}ar[Y-X] = \mathbb{V}ar[Y] - \mathbb{V}ar[X]$

Proof. We show it when $E[X] = E[Y] = 0$, the proof can easily be extended to the general case.

$Var(Y-X) = E(Y-X)^2 = E(Y^2 - 2XY + X^2) = E(Y^2) - 2E(XY) + E(X^2)$
$= E(Y^2) - 2E(X.(X + (Y-X))) + E(X^2)$
$= E(Y^2) - E(X^2) - 2E[X(Y-X])$
then you condition on X and integrate with respect to the law of X.
$Var[Y] - Var[X] - E[2E[X(Y-X)|X]]$
$Var[Y] - Var[X]$

$\square$

Lemma 3.3.6. If μ and ν are two centered measures with finite support such that $\mu \leq_{cx} \nu$, then the area between their potential functions is equal to the second moment of ν minus the second moment of μ.

Proof. Let K_1 and K_2 be such that $\mu((K_1, K_2)) = 1$ and $\nu((K_1, K_2)) = 1$.

$$
\begin{aligned}
\mathcal{A} &= \int_{\mathbb{R}} U\mu(t) - U\nu(t)dt \\
&= \int_{K_1}^{K_2} (\int_{K_1}^{K_2} -|x-t|d\mu(x) - \int_{K_1}^{K_2} -|x-t|d\nu(x))dt \\
&= \int_{K_1}^{K_2} (\int_{K_1}^{K_2} |x-t|d\nu(x) - \int_{K_1}^{K_2} |x-t|d\mu(x))dt \\
&= \int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|d\nu(x)dt - \int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|d\mu(x)dt \\
&= \int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|\ dt\ d\nu(x) - \int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|\ dt\ d\mu(x) \text{ by Fubini}
\end{aligned}
$$

Now,

$$
\begin{aligned}
\int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|\ dt\ d\mu(x) &= \int_{K_1}^{K_2} \int_{K_1}^{x} |x-t|dt\ d\mu(x) + \int_{K_1}^{K_2} \int_{x}^{K_2} |x-t|dt\ d\mu(x) \\
&= \int_{K_1}^{K_2} \int_{K_1}^{x} x-t\ dt\ d\mu(x) + \int_{K_1}^{K_2} \int_{x}^{K_2} t-x\ dt\ d\mu(x) \\
&= \int_{K_1}^{K_2} \frac{1}{2}x^2 - xK_1 - \frac{1}{2}K_1^2 d\mu(x) + \int_{K_1}^{K_2} \frac{1}{2}x^2 + \frac{1}{2}K_2^2 - xK_2 d\mu(x) \\
&= \int_{K_1}^{K_2} x^2 d\mu(x) - \int_{K_1}^{K_2} xK_1 - \frac{1}{2}K_1^2 + \frac{1}{2}K_2^2 - xK_2 d\mu(x)
\end{aligned}
$$

As $\int_{K_1}^{K_2} d\mu(x) = 1$ and $\int_{K_1}^{K_2} xd\mu(x) = 0$, we get

$$
\int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|\ dt\ d\nu(x) = \int_{K_1}^{K_2} x^2 d\mu(x) + \frac{1}{2}K_1^2 - \frac{1}{2}K_2^2
$$

As $\int_{K_1}^{K_2} \int_{K_1}^{K_2} |x-t|\ dt\ d\mu(x)$ is of the same form with ν replaced by μ, it follows that

$$
\mathcal{A} = \int_{\mathbb{R}} x^2 d\nu(x) - \int_{\mathbb{R}} x^2 d\mu(x)
$$

□

As we have just seen that every martingale transition from $\hat{\mu}$ to $\hat{\nu}$ has the same variance, it is interesting to look at the variance of the martingale increment conditioned on the value before the transition. To examine this, we define the conditional variance function:

Definition 3.3.7. (Conditional variance function)
Given two random variables X and Y, we define the function

$$
x \to Var[Y - X \mid X = x]
$$

which we call the conditional variance function.

For a two step martingale (X,Y) we can plot the variance of Y conditioned on $X = x$. We will call this the conditional variance function: $x \to \mathbb{V}ar[Y|X = x]$.
There are several interesting cases:
• conditional variance function can be a constant (for ex. for a Brownian transition law this is the case)
• the graph of the conditional variance function can be convex or smile shaped.
This means that conditioned on a big movement the expectation of the magnitude of the next movement is larger than if the initial movement had been small.
(This is likely to be the dynamics of a stock price)
• The graph of the conditional variance function can be unimodal, meaning that the middle diffuses more than the extremities.

We can plot the conditional variance function of this martingale function. For clarity we produce the plot with quantizations of order larger than 8.

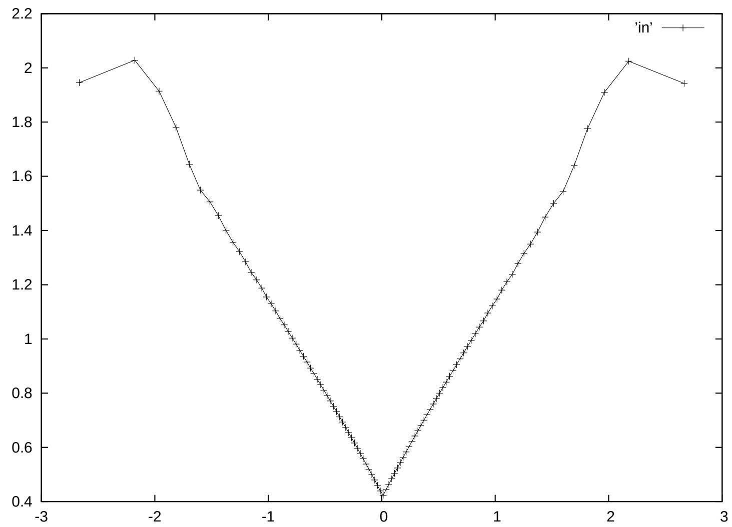


Figure 3.2: Conditional variance: clipping tails first

Remark. This martingale transition exhibits a phenomenon called persistence of volatility. We are dealing with a 2 step martingale. The initial law $(t = 0)$ is the Dirac at 0. The next law (at $t = 1$) is the $\mathcal{U}$-quantization of the $N(0,1)$ law. The third law (at $t = 2$) is the $\mathcal{U}$-quantization of the $N(0,2)$ law. One sees that if the first martingale increment ($t = 0$ to $t = 1$) is large in absolute value then the second increment ($t = 1$ to $t = 2$) can also be expected to be large in absolute value. This is a phenomenon of persistence of volatility.

3.3.3 Clipping from center to extremities

Clipping the potential function of $\hat{v}$ from *center to extremities* produces the following martingale transition:

$$\begin{pmatrix}
0.505115 & 0.081628 & 0.081628 & 0.081628 & 0.081628 & 0.081628 & 0.081628 & 0.005115 \\
0.279299 & 0.118875 & 0.118875 & 0.118875 & 0.118875 & 0.118875 & 0.118875 & 0.007448 \\
0.157884 & 0.138902 & 0.138902 & 0.138902 & 0.138902 & 0.138902 & 0.138902 & 0.008703 \\
0.057702 & 0.155427 & 0.155427 & 0.155427 & 0.155427 & 0.155427 & 0.155427 & 0.009738 \\
0.000000 & 0.158673 & 0.158673 & 0.158673 & 0.158673 & 0.158673 & 0.158673 & 0.047964 \\
0.000000 & 0.141803 & 0.141803 & 0.141803 & 0.141803 & 0.141803 & 0.141803 & 0.149181 \\
0.000000 & 0.121358 & 0.121358 & 0.121358 & 0.121358 & 0.121358 & 0.121358 & 0.271851 \\
0.000000 & 0.083333 & 0.083333 & 0.083333 & 0.083333 & 0.083333 & 0.083333 & 0.500000
\end{pmatrix}$$

We now look at the corresponding conditional variance function:

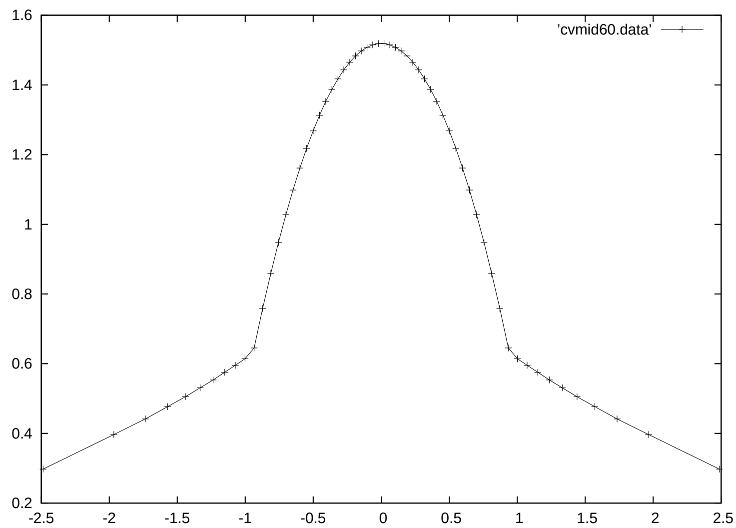


Figure 3.3: Conditional variance: clipping from center to extremities

Remark. This martingale transition exhibits the opposite phenomenon to that of persistance of volatility. If $|M_0 - M_1|$ is small, then the expected variance of $M_2 - M_1$ is large.

3.3.4 Clockwise clipping

Clipping the potential function of $\hat{v}$ *clockwise* produces the following martingale transition:

$$\begin{pmatrix}
0.259974 & 0.259974 & 0.259974 & 0.220078 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\
0.166425 & 0.166425 & 0.166425 & 0.175397 & 0.224891 & 0.100437 & 0.000000 & 0.000000 \\
0.134878 & 0.134878 & 0.134878 & 0.142149 & 0.182261 & 0.211525 & 0.059433 & 0.000000 \\
0.115175 & 0.115175 & 0.115175 & 0.121384 & 0.155636 & 0.180625 & 0.196832 & 0.000000 \\
0.101626 & 0.101626 & 0.101626 & 0.107105 & 0.137328 & 0.159378 & 0.233607 & 0.057702 \\
0.090822 & 0.090822 & 0.090822 & 0.095718 & 0.122728 & 0.142434 & 0.208771 & 0.157884 \\
0.077727 & 0.077727 & 0.077727 & 0.081918 & 0.105033 & 0.121898 & 0.178670 & 0.279299 \\
0.053373 & 0.053373 & 0.053373 & 0.056251 & 0.072123 & 0.083704 & 0.122688 & 0.505115
\end{pmatrix}$$

3.3.5 Counter clockwise clipping

Clipping the potential function of $\hat{v}$ *counter clockwise* produces the following martingale transition:

$$\begin{pmatrix}
0.259974 & 0.259974 & 0.259974 & 0.220078 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\
0.193288 & 0.193288 & 0.193288 & 0.203709 & 0.056265 & 0.053387 & 0.053387 & 0.053387 \\
0.161724 & 0.161724 & 0.161724 & 0.170442 & 0.089532 & 0.084952 & 0.084952 & 0.084952 \\
0.135679 & 0.135679 & 0.135679 & 0.142994 & 0.116980 & 0.110996 & 0.110996 & 0.110996 \\
0.110996 & 0.110996 & 0.110996 & 0.116980 & 0.142994 & 0.135679 & 0.135679 & 0.135679 \\
0.084952 & 0.084952 & 0.084952 & 0.089532 & 0.170442 & 0.161724 & 0.161724 & 0.161724 \\
0.053387 & 0.053387 & 0.053387 & 0.056265 & 0.203709 & 0.193288 & 0.193288 & 0.193288 \\
0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.220078 & 0.259974 & 0.259974 & 0.259974
\end{pmatrix}$$

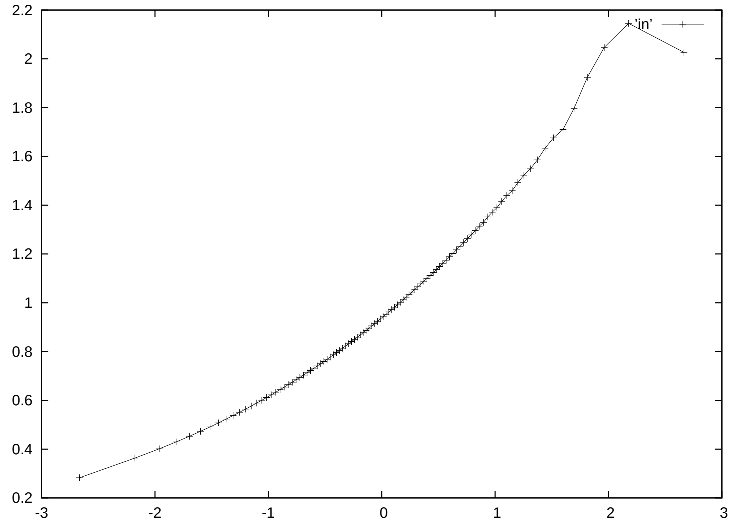


Figure 3.4: Conditional variance: clipping counterclockwise

Part II

The Skorokhod embedding problem and constructions of martingales with specified marginals

3.4 Introduction

We will give a new solution to the Skorokhod embedding problem (SEP) that was published in [Bak12]. We will also discuss the use of solutions to the SEP as means of constructing martingales with specified marginals, as well as the limitations of this approach.

3.4.1 Martingales as time changed Brownian motion

Before we discuss the Skorokhod embedding problem and its use as a way of constructing martingales with specified marginals, it will be usefull to recall some theory about how martingales can be represented as time changed Brownian motions. By Dambis Dubins-Schwarz (see [RY94]), if M is a continuous martingale starting from 0 (i.e. $M_0 = 0$) with $< M, M >_\infty = \infty$, then

$$M_t = B_{<M,M>_t}$$

for some Brownian Motion (B_u).

3.4.2 The Skorokhod embedding problem

The Skorokhod embedding problem (SEP), see [Sko65], is the problem of embedding a probability measure into Brownian motion by means of a stopping time. Formally, Skorokhod's original definition of the Skorokhod embedding problem is the following:

Definition 3.4.1. Given a Brownian motion W and a probability measure μ on $\mathbb{R}$ which satisfies $\int_{\mathbb{R}} x \, d\mu(x) = 0$ and $\int_{\mathbb{R}} x^2 \, d\mu(x) < \infty$, a solution to the Skorokhod embedding problem is a stopping time T for W, such that:

$$W_T \text{ has law } \mu \text{ and } \mathbb{E}[T] < \infty$$

It turns out that the requirement that μ have a finite second moment is not necessary. A generalized definition of the Skorokhod embedding problem is as follows:

Definition 3.4.2. Given a Brownian motion W and a probability measure μ on $\mathbb{R}$ which satisfies $\int_{\mathbb{R}} x d\mu = 0$ and $\int_{\mathbb{R}} |x| d\mu < \infty$, a solution to the Skorokhod embedding problem is a stopping time T for W, such that:
W_T has law μ and $W_{t \wedge T}$ is uniformly integrable.

In [Sko65], Skorokhod gives a solution to the SEP. The solution given by Skorokhod however requires an additional random variable which is independent of the Brownian motion. The solution given by Dubins, see [Dub68], is the first solution which does not require an additional independent random variable. Since then, a variety of other solutions have been given. For an extensive survey of existing solutions, see [Obł04].

It turns out that every stopping time which is a solution to the Skorokhod embedding problem for a measure μ has the same expectation. The value of its expectation is the variance of the measure μ. Indeed, let T be a solution to the Skorokhod embedding problem for μ. By Ito's formula, $M_t = B_t^2 - t$ is a martingale. By the optional stopping theorem, $\mathbb{E}[T] < \infty$ implies that

$$\begin{aligned}
&\mathbb{E}[M_0] = \mathbb{E}[M_T] \\
&\Rightarrow 0 = \mathbb{E}[B_T^2 - T] \\
&\Rightarrow \mathbb{E}[T] = \mathbb{E}[B_T^2]
\end{aligned}$$

Now since T is a solution to the Skorokhod embedding problem for μ, it follows that $B_T \sim \mu$ and so $\mathbb{E}[B_T^2]$ is equal to the second moment of μ.

3.4.3 Using the Skorokhod embedding problem to construct martingales with specified marginals

This past decade has seen renewed interest in solutions to the SEP. This interest is due to the use of such solutions to construct martingales with specified marginals (see, e.g., [HP02], [Hob98], [CH07], [MY02], and [HPRY11]). New solutions to the SEP can in turn lead to new constructions of martingales with specified marginals. Model-free methods for pricing financial instruments rely on constructions of martingales with given marginals (see, e.g. [Hob11]). For these reasons, new solutions to the SEP can ultimately lead to improved bonds on model-free prices of financial instruments.

• In the two marginal setting: Let μ and ν be given measures which satisfy $\mu \leq_{cx} \nu$. In order to construct a martingale transition from μ to ν it suffices to construct stopping times τ_1 and τ_2 which satisfy:

$$\begin{cases}
\tau_1 \text{ is a solution to the SEP for } \mu. \\
\tau_2 \text{ is a solution to the SEP for } \nu. \\
\tau_1 \leq \tau_2 \text{ a.s.}
\end{cases}$$

Indeed, the bivariate law (W_{τ_1}, W_{τ_2}) by construction has the required marginals, $W_{\tau_1} \sim \mu$ and $W_{\tau_2} \sim \nu$, as well as the martingale property $\mathbb{E}[W_{\tau_2} | W_{\tau_1}] = W_{\tau_1}$

• In the continuous time setting: Let $(\mu_t)_{t \in \mathbb{R}^+}$ be a time indexed collection of marginals which satisfy $\mu_s \leq_{cx} \mu_t$ whenever $s \leq t$. Then in order to construct a martingale M which satisfies $M_t \sim \mu_t$ for each $t \in \mathbb{R}^+$, it suffices to obtain a collection of solutions to the SEP: τ_t being a solution to the SEP for μ_t and $\tau_s \leq \tau_t$ a.s. when $s \leq t$. Then $(W_{\tau_t})_{t \in \mathbb{R}^+}$ is a martingale which satisfies $M_t \sim \mu_t$ for each $t \in \mathbb{R}^+$. Indeed $\mathbb{E}[W_{\tau_t} | W_{\tau_s}] = W_{\tau_s}$ for all s, t satisfying $s \leq t$.

Definition 3.4.3. (the barycenter function of a measure) The barycenter function of a probability measure μ is defined as follows:

$$\phi(x) = \frac{\int_x^\infty y\mu(dx)}{\int_x^\infty \mu(dx)}$$

The Azema Yor solution to the Skorokhod embedding problem is defined as

Definition 3.4.4. (The Azema-Yor solution to the Skorokhod embedding problem see [AY79]) Let M(t) denote the maximum value to date of the Brownian motion B.

$$M(t) = \sup_{0 \le s \le t} B(s)$$

The Azema-Yor solution to the Skorokhod embedding problem is the stopping time τ defined as

$$\tau = \inf\{s | M(s) \ge \phi(B(s))\}$$

Let $(\mu_t)_{t \in \mathbb{R}^+}$ be a collection of probability measures such that the function

$$\phi(x,t) = \frac{\int_x^\infty y g(y,t) dy}{\int_x^\infty g(y,t) dy}$$

is increasing in t for each x. Here $g(y,t)$ denotes the density of the measure μ_t. Under this condition, Madan and Yor (see [MY02]) use the Azema-Yor solution to the Skorokhod embedding problem to construct a martingale M with

$$M_t \sim \mu_t$$

The martingale M is defined as $M_t = B_{\tau_t}$, with τ_t being the Azema-Yor solution to the Skorokhod embedding problem for μ_t. The recent book by Hirsch, Profeta, Roynette, and Yor (see [HPRY11]) contains numerous other constructions.

3.4.4 Limitations of the Skorokhod embedding problem as a means of constructing martingales with specified marginals

Here we discuss the reason why most solutions to the SEP are unable to construct martingale transition between every pair of measures which admits a martingale transition. This is a consequence of a Theorem by Meilijson given in [Mei82]. We now explain why this is the case.

In the following $(W_t)_{(t \in \mathbb{R}^+)}$ will be a standard Brownian motion.

Definition 3.4.5. A procedure which associates to each measure in the set

$$\{\, \mu \in \mathcal{P}(\mathbb{R}) \ : \ \int_{\mathbb{R}} |x| \, d\mu(x) < \infty\}$$

a unique stopping time τ which is a solution to the Skorokhod embedding problem for μ, i.e. :

$$\begin{cases} W_\tau \sim \mu \\ W_{t\wedge\tau} \text{ is a u.i. martingale} \end{cases}$$

will be called a *standard* solution to the SEP.

Remark. Most currently published solutions to the SEP are *standard* solutions. This includes the solutions given by Dubins [Dub68], Azema-Yor [AY79], Vallois [Val83], as well as the new one presented in this work. A solution to the SEP which is not *standard* can produce several different stopping times for a given measure μ.

Definition 3.4.6. Let ν be a probability measure on $\mathbb{R}$. A solution τ to the Skorokhod embedding problem for μ is said to be *ultimate* if:
For every measure μ with $\mu \leqslant_{cx} \nu$, there exists a stopping time $\tau^{'}$ satisfying

$$\begin{cases} \tau^{'} \leqslant \tau \text{ a.s.} \\ W_{\tau'} \sim \mu \end{cases}$$

Definition 3.4.7. The hitting time of the level a, for the Brownian motion W is denoted T_a and defined as:

$$T_a = \inf\{t \geqslant 0 : W_t = a\}$$

Theorem 3.4.8. *(Meilijson see [Mei82]) τ is ultimate if and only if $\tau = T_a \wedge T_b$ for some $a < 0 < b$.*

This leads to the following limitation of *standard* solution to the SEP as means of constructing martingale transitions between specified marginals:

Corollary 3.4.9. *Every* standard *solution to the SEP is unable to construct a martingale transition between certain pairs of measures which admit a martingale transition*

Proof. Let ν be a measure which is not supported on two points, i.e.

$$\nu \neq \alpha\delta_a + (1-\alpha)\delta_b \;\text{ with } a, b \in \mathbb{R}, \alpha \in [0,1]$$

Consider a standard solution to the SEP. Denote by $\tau(\nu)$ the stopping time which this solution associates to ν. As ν is not supported by two points it follows by Meilijson's theorem that $\tau(\nu)$ is not ultimate. Therefore there exists a measure μ with $\mu \leqslant_{cx} \nu$ for which there exists no stopping time $\tau(\mu)$ satisfying:

$$\begin{cases} \tau(\mu) \leqslant \tau(\nu) \\ W_{\tau(\mu)} \sim \mu \end{cases}$$

Therefore the standard solution under consideration is unable to construct a martingale transition from μ to ν. Of course since $\mu \leqslant_{cx} \nu$, a martingale transition from μ to ν necessarily exists. □

Remark. Hobson, Brown and Rogers (see [BHR01]) have modified the Azema-Yor solution in order that it no longer be *standard* and that it be able to construct martingale transitions between arbitrary pairs of marginals which admit a martingale transition.

3.5 The Dubins solution to the Skorokhod embedding problem

In this section, we describe the Dubins solution to the SEP given in [Dub68]. The presentation here differs from the original presentation because we wish to emphasize a framework which we will use in the next section to construct a new Dubins type solution to the SEP.

Let μ be the probability measure which is to be embedded in Brownian motion. A sequence of partitions of $\mathbb{R}$ is defined recursively. The initial partition, $\mathcal{P}artition(0)$ is $\{\mathbb{R}\}$. The following partitions are obtained recursively. $\mathcal{P}artition(n+1)$ is obtained by cutting each interval $[a,b] \in \mathcal{P}artition(n)$ into two, as follows:

$[a,b] \to [a,c]$ and $[c,b]$ where

$$c = \frac{1}{\mu([a,b])} \int_{[a,b]} x \, d\mu(x) \quad \text{(note that } a \leq c \leq b)$$

If a or b is $+\infty$ or $-\infty$, the value of c is calculated in the same way, and the cutting is also done in the same way. For each $n \in \mathbb{N}$, a measure μ_n is obtained from $\mathcal{P}artition(n)$ in the following way: for each interval $[a,b] \in \mathcal{P}artition(n)$, place an atom of mass $\mu([a,b])$ at position $\frac{1}{\mu([a,b])} \int_{[a,b]} x \, d\mu(x)$

An increasing sequence of stopping times is defined by

$$\tau_n = \inf\{t \geq \tau_{n-1} : W_t \in \text{ support of } \mu_n\}$$

and Dubins' solution to the SEP is the stopping time τ defined by

$$\tau := \sup\{\tau_n\}$$

3.6 A new solution to the Skorokhod embedding problem

Dubins in [Dub68] gave the first solution to the Skorokhod embedding problem (SEP) based solely on the underlying Brownian motion, and thus requiring no additional independent random variable. The Dubins solution to the SEP, can be expressed as $\tau := \sup\{\tau_n\}$ with $\tau_n = \inf\{t \geq \tau_{n-1} : W_t \in \text{ support of } \mu_n\}$. Since the measures μ_n are defined recursively, in order to compute μ_n, each of $\mu_0, .., \mu_{n-1}$ must first be computed. We now give a new solution to the SEP by showing how to construct a different sequence of measures $\{\mu_n\}_{n\in\mathbb{N}}$. The advantage of this solution is that for any given n, the measure μ_n can be constructed directly without prior computation of the measures $\mu_0, .., \mu_{n-1}$.

We will define a sequence of measures μ_n and a corresponding increasing sequence of stopping times τ_n. First, we will prove that μ_n converges to μ, then we will prove that $W_{\tau_n} \sim \mu_n$ for each $n \in \mathbb{N}$. Finally, defining τ to be $sup_n\{\tau_n\}$, we will obtain $W_\tau \sim \mu$.

Let F be the cumulative distribution function of μ. Its inverse, $F^{-1}(x)$ is called the quantile function of μ. Since F is a non decreasing function from $\mathbb{R}$ to $[0,1]$, its inverse, F^{-1} is a non decreasing function from $(0,1]$ to $\mathbb{R}$.

Definition 3.6.1. For $n \geq 0$, define μ_n to be the uniform measure on the following 2^n coordinates:

$$a_i = 2^n \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} F^{-1}(u)\, du \qquad\qquad \text{with } i \text{ ranging from } 0 \text{ to } 2^n - 1$$

Lemma 3.6.2. μ_n converges weakly to μ.

Proof. Let F be the cumulative distribution function of μ, and F_n be the cumulative distribution function of μ_n. Showing that μ_n converges weakly to μ amounts to showing that F_n converges pointwise to F. The collection of intervals $\{\big(F^{-1}(\frac{i-1}{2^n})\ ,\ F^{-1}(\frac{i}{2^n})\big] : 1 \leq i \leq 2^n\}$ generate the support of μ. We will proceed by establishing bounds for F and F_n when t belongs to such an interval, i.e. when $t \in [F^{-1}(\frac{i-1}{2^n})\ ,\ F^{-1}(\frac{i}{2^n})]$.
Bounding F is straightforward: since F is non decreasing,

$$F^{-1}(\frac{i-1}{2^n}) \leq t \leq F^{-1}(\frac{i}{2^n}) \Rightarrow F\left(F^{-1}(\frac{i-1}{2^n})\right) \leq F(t) \leq F\left(F^{-1}(\frac{i}{2^n})\right)$$
$$\Rightarrow \frac{i-1}{2^n} \leq F(t) \leq \frac{i}{2^n}$$

We now proceed to obtain bounds for F_n. Recall that μ_n is the uniform distribution on the following 2^n points:

$$a_i = 2^n \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} F^{-1}(u)\, du \qquad\qquad i \text{ ranging from } 0 \text{ to } 2^n - 1$$

Since F^{-1} is a non decreasing function, we obtain a bound for a_i by bounding the above integral:

$$2^n(\frac{i+1}{2^n} - \frac{i}{2^n})\, F^{-1}(\frac{i}{2^n}) \leq a_i \leq 2^n(\frac{i+1}{2^n} - \frac{i}{2^n})\, F^{-1}(\frac{i+1}{2^n}) \quad \Rightarrow \quad F^{-1}(\frac{i}{2^n}) \leq a_i \leq F^{-1}(\frac{i+1}{2^n})$$

From these bounds for a_i, we will obtain bounds for the cumulative distribution function F_n of μ. Letting $\mathbf{1}$ denote the indicator function, we have the following expression for the cumulative distribution function F_n of μ_n:

$$F_n(t) = \frac{1}{2^n} \sum_{i=0}^{2^n-1} \mathbf{1}\{a_i \leq t\}$$

Since we have seen that

$$a_{i-1} \leq F^{-1}(\frac{i}{2^n}) \leq a_i \leq F^{-1}(\frac{i+1}{2^n}) \leq a_{i+1}$$

therefore,

$$\begin{aligned} F^{-1}(\frac{i-1}{2^n}) \leq \ t \ \leq F^{-1}(\frac{i}{2^n}) \ &\Rightarrow \ a_{i-1} \leq \ t \ \leq a_{i+1} \\ &\Rightarrow \frac{i-1}{2^n} \leq \ \frac{1}{2^n} \sum_{i=0}^{2^n-1} \mathbf{1}\{a_i \leq t\} \ \leq \frac{i+1}{2^n} \\ &\Rightarrow \frac{i-1}{2^n} \leq \ F_n(t) \ \leq \frac{i+1}{2^n} \end{aligned}$$

The bounds for F, together with the bounds for F_n , give bounds for $\mid F(t) - F_n(t) \mid$:

$$\begin{aligned} t \in [F^{-1}(\frac{i-1}{2^n}) \ , \ F^{-1}(\frac{i}{2^n})] \ \Rightarrow \ \frac{i-1}{2^n} \leq F(t) \leq \frac{i}{2^n} \ \text{ and } \ \frac{i-1}{2^n} \leq F_n(t) \leq \frac{i+1}{2^n} \\ \Rightarrow \mid F(t) - F_n(t) \mid \leq \ \frac{2}{2^n} \end{aligned}$$

Since the collection of intervals $\{\left(F^{-1}(\frac{i-1}{2^n}) \ , \ F^{-1}(\frac{i}{2^n})\right] : \ 1 \leq i \leq n\}$ generates the support of μ, and therefore the support of μ_n, we obtain:

$$\mid F_n(t) - F(t) \mid \leq \ \frac{2}{2^n} \quad \forall \, t \in \mathbb{R}$$

Hence F_n converges pointwise to F, and this implies that μ_n converges weakly to μ. $\square$

Definition 3.6.3. Define the following collection of stopping times:

$$\tau_0 = 0 \quad \text{and for } n \geq 1, \quad \tau_n = \inf\{t \geq \tau_{n-1} : W_t \in \text{ support of } \mu_n\}$$

Theorem 3.6.4. *W_{τ_n} has law μ_n.*

Proof. We prove this by induction. We first verify that $W_{\tau_0} \sim \mu_0$. Now τ_0 is defined to be 0, so $W_{\tau_0} = W_0 = 0$. Also μ_0 is defined to be the Dirac at $\int_0^1 F^{-1}(u)\, du = \int_{\mathbb{R}} x \, d\mu = 0$. Hence $W_{\tau_0} \sim \mu_0$, and so the statement is true for $n = 0$.

Suppose that the statement is true for n, in other words, suppose that $W_{\tau_n} \sim \mu_n$. We will determine the law of $W_{\tau_{n+1}}$, conditioned on W_{τ_n} having law μ_n. To do this, we take an arbitrary point m in the support of μ_n and calculate the distribution of $W_{\tau_{n+1}}$ conditioned on $\{W_{\tau_n} = m\}$. Since m belongs to the support of μ_n, there exists an $i \in \{0,..,2^n - 1\}$ such that

$$m = 2^n \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} F^{-1}(u)\, du$$

.

Denote by u (for upper) the smallest element of the support of μ_{n+1} which is greater or equal to m. Also denote by l (for lower) the largest element of the support of μ_{n+1} which is less than or equal to m. Since F^{-1} is a non decreasing function, we can write these two elements u and l as:

$$\begin{aligned} u &= 2^{n+1} \int_{\frac{i}{2^n}+\frac{1}{2^{n+1}}}^{\frac{i+1}{2^n}} F^{-1}(u)\, du \\ l &= 2^{n+1} \int_{\frac{i}{2^n}}^{\frac{i}{2^n}+\frac{1}{2^{n+1}}} F^{-1}(u)\, du \end{aligned} \tag{3.1}$$

Notice that $u + l = 2m$ and so $m = \frac{l+u}{2}$. It follows that m is equidistant from l and u. And so, conditioned on $\{W_{\tau_n} = m\}$, the events $\{W_{\tau_{n+1}} = l\}$ and $\{W_{\tau_{n+1}} = u\}$ are equiprobable. In other words,

$$\begin{aligned} \mathbb{P}(W_{\tau_{n+1}} = l \mid W_{\tau_n} = m) &= \frac{1}{2} \\ \mathbb{P}(W_{\tau_{n+1}} = u \mid W_{\tau_n} = m) &= \frac{1}{2} \end{aligned} \tag{3.2}$$

By a straightforward iterative argument, it follows that

$$\begin{aligned} \mathbb{P}(W_{\tau_{n+1}} = l) &= \frac{1}{2}P(W_{\tau_n} = m) = \frac{1}{2} \times \frac{1}{2^n} = \frac{1}{2^{n+1}} \\ \mathbb{P}(W_{\tau_{n+1}} = u) &= \frac{1}{2}P(W_{\tau_n} = m) = \frac{1}{2} \times \frac{1}{2^n} = \frac{1}{2^{n+1}} \end{aligned} \tag{3.3}$$

Therefore $W_{\tau_{n+1}}$ follows a uniform distribution which has the same support as μ_{n+1}. Since μ_{n+1} is itself a discrete uniform law, it follows that $W_{\tau_{n+1}} \sim \mu_{n+1}$. $\square$

We now provide a lemma which shows that, for a large class of quantizations, the quantized measure is dominated in the convex order by the original measure (i.e. $\hat{\mu} \leq_{cx} \mu$).

Lemma 3.6.5. Let $\mathcal{J}$ be a partition of $\mathbb{R}$. Let $\hat{\mu}$ be the probability measure which is constructed from μ in the following way: for each $J \in \mathcal{J}$, an atom of mass $\mu(J)$ is placed at position $\dfrac{\int_J x\, d\mu(x)}{\mu(J)}$. Then,

$$\int_{\mathbb{R}} \phi(x)\, d\hat{\mu}(x) \leq \int_{\mathbb{R}} \phi(x)\, d\mu(x) \quad \text{for every convex function } \phi.$$

Proof. Let J be an arbitrary element of $\mathcal{J}$. By construction, $\hat{\mu}(J) = \mu(J)$ and $\int_J x \, d\hat{\mu}(x) = \int_J x \, d\mu(x)$. The measure $\frac{\mu(dx)}{\mu(J)}$ is a probability measure on J. Its expectation is $\frac{\int_J x \, d\mu(x)}{\mu(J)}$. Therefore, by Jensen's inequality, for every convex function ϕ,

$$\int_J \phi(x) \, \frac{d\mu(x)}{\mu(J)} \geq \phi\left(\frac{\int_J x \, d\mu(x)}{\mu(J)}\right)$$

As $\hat{\mu}(J) = \mu(J)$, the measure $\frac{\hat{\mu}(dx)}{\mu(J)}$ is a probability measure on J. It consists of a single Dirac point mass at the position $\frac{\int_J x \, d\mu(x)}{\mu(J)}$. Therefore,

$$\int_J \phi(x) \, \frac{d\hat{\mu}(x)}{\mu(J)} = \phi\left(\frac{\int_J x \, d\mu(x)}{\mu(J)}\right)$$

Combining the two above equations yields

$$\int_J \phi(x) \, \frac{d\mu(x)}{\mu(J)} \geq \int_J \phi(x) \, \frac{d\hat{\mu}(x)}{\mu(J)}$$

which is equivalent to

$$\int_J \phi(x) \, d\mu(x) \geq \int_J \phi(x) \, d\hat{\mu}(x)$$

As the above holds for each $J \in \mathcal{J}$, and together they constitute a partition of $\mathbb{R}$, we obtain

$$\int_{\mathbb{R}} \phi(x) \, d\mu(x) \geq \int_{\mathbb{R}} \phi(x) \, d\hat{\mu}(x)$$

$\square$

Theorem 3.6.6. $\tau := sup\{\tau_n\}$ *is a solution to the Skorokhod Embedding Problem (SEP).*

We have shown that μ_n converges weakly to μ and that $W_{\tau_n} \sim \mu_n$, which together imply that $W_\tau \sim \mu$. For τ to be a solution to the SEP, it remains to verify that $\mathbb{E}[\tau] < \infty$. In order to do this, we first check that τ_n is a solution to the SEP for μ_n. Now,

$$\mathbb{E}[\tau_n] = \mathbb{E}[\tau_n - \tau_{n-1}] + \mathbb{E}[\tau_{n-1} - \tau_{n-2}] + .. + \mathbb{E}[\tau_2 - \tau_1] + \mathbb{E}[\tau_1]$$

Each expectation on the right hand side is a weighted average of expected exit times from strips for the Brownian motion W, and is therefore finite. This together with Theorem 3.6.4 ($W_{\tau_n} \sim \mu_n$) implies that τ_n is a solution to the SEP for μ_n. It follows that,

$$\mathbb{E}[\tau_n] = \int_{\mathbb{R}} x^2 \, d\mu_n(x) \text{ for each } n.$$

Let a_n denote the sequence $n \to \mathbb{E}[\tau_n]$ which, by the above equality, is identical to the sequence $n \to \int_{\mathbb{R}} x^2 \, d\mu_n(x)$. We show that a_n converges by showing that it is increasing and bounded. It is increasing because $n \to \mathbb{E}[\tau_n]$ is increasing. It is bounded because by lemma 3.6.5, we have

$$\int_{\mathbb{R}} x^2 \, d\mu_n \leq \int x^2 \, d\mu \text{ for each } n.$$

Therefore a_n converges. In other words $\mathbb{E}[\tau_n]$ converges. This means that $\mathbb{E}[\tau] < \infty$.

3.7 Numerical illustration

Using the standard Gaussian law, $\mathcal{N}(0,1)$, we numerically show that the sequence of measures $\{\mu_n\}_{n\in\mathbb{N}}$ constructed by our solution to the SEP is different from the sequence of measures generated by the Dubins solution. The first 4 partitions of $\mathbb{R}$ produced by Dubins solution are:

```
Partition(0) = { [-inf,inf] }
Partition(1) = { [-inf,0] [0,inf] }
Partition(2) = { [-inf,-0.797885] [-0.797885,0] [0,0.797885] [0.797885,inf] }
Partition(3) = { [-inf,-1.36576] [-1.36576,-0.797885] [-0.797885,-0.378257] [-0.378257
                 [0,0.378257] [0.378257,0.797885] [0.797885,1.36576] [1.36576,inf] }
```

For each interval $[a,b]$, we compute $\mu([a,b])$:

```
Partition(0) : 1
Partition(1) : 0.5 0.5
Partition(2) : 0.212469 0.287531 0.287531 0.212469
Partition(3) : 0.086007 0.126462 0.140151 0.14738 0.14738 0.140151 0.126462 0.086007
```

Our solution to the SEP generates a sequence of measures $\{\mu_n\}_{n\in\mathbb{N}}$, such that each μ_n is a uniform law on 2^n support points. The Dubins construction for the $\mathcal{N}(0,1)$ law has produced laws μ_2 and μ_3 which are not uniformly distributed. It follows that our construction is different from that of Dubins.

Part III

Continuous time martingales with specified marginals: some constructions

Chapter 4

Overview of existence and uniqueness results

Definition 4.0.1. Let $(\mu_t)_{t\in\mathbb{R}^+}$ be a collecting of probability laws on $\mathbb{R}$. We say that this collection is *increasing in the convex order* if:

$$\text{for all } s \leqslant t, \quad \mu_s \leqslant_{cx} \mu_t$$

Definition 4.0.2. The following *call transform* associates a function of two variables $C(t,k)$ to a collection of marginal laws $(\mu_t)_{t\in\mathbb{R}^+}$:

$$C(t,k) = \int_{\mathbb{R}} (x-k)^+ d\mu_t(x)$$

Existence and uniqueness results have been obtained in the following 3 frameworks:

Framework A:

$$\begin{cases} (\mu_t)_{t\in\mathbb{R}^+} \text{ have constant means.} \\ C(t,k) \text{ is increasing in } t. \end{cases}$$

Framework B:

$$\begin{cases} (\mu_t)_{t\in\mathbb{R}^+} \text{ have constant means.} \\ C(t,k) \text{ is increasing and continuous in } t. \end{cases}$$

Framework C:

$$\begin{cases} (\mu_t)_{t\in\mathbb{R}^+} \text{ have constant means.} \\ C(t,k) \text{ is increasing and differentiable in } t. \end{cases}$$

Remark. Framework A is equivalent to $(\mu_t)_{t\in\mathbb{R}^+}$ being increasing in the convex order.

Remark. These 3 frameworks are from least to most restrictive

$$C \Rightarrow B \Rightarrow A$$

Theorem 4.0.3. *(Kellerer [Kel72], Lowther [Low08], Dupire [Dup94])*

- *Under framework A, there exists a martingale M with $M_t \sim \mu_t$. The martingale M is generally not unique.*
  (see Kellerer [Kel72], see also [HR12] for an alternate proof of this result)

- *Under framework B, uniqueness is obtained when restricting consideration to the class of martingales which are almost continuous diffusions. (see Lowther [Low08]). See below for the definition of an almost continuous diffusion.*

- *Under framework C, uniqueness is obtained when restricting consideration to the class of martingales which are continuous diffusions. (see Dupire [Dup94])*

Definition 4.0.4. A process X is an *almost continuous diffusion* if it is strong Markov with cadlag paths and given two independent processes X and Z distributed as X, the following holds:

$$\forall s,t \in \mathbb{R}^+ \text{with } s<t,\ \mathbb{P}\left(Y_s < Z_s, Y_t > Z_t \text{ and } Y_u \neq Z_u\ \forall u \in (s,t)\right) = 0$$

4.1 The Kellerer existence theorem

Kellerer's celebrated theorem for the existence of martingales with specified marginals is as follows.

Theorem 4.1.1. *(Kellerer [Kel72]) Let $(\mu_t)_{t\in\mathbb{R}^+}$ be a specified collection of marginals. If this collection is increasing in the convex order then there exists a martingale $(M_t)_{t\in\mathbb{R}^+}$ which has the Markov property and satisfies $\forall t \in \mathbb{R}^+,\ M_t \sim \mu_t$.*

The paper containing this result, ([Kel72]) is published in German, so we give an outline of the proof here. This theorem is in fact proved as the consequence of a more general existence theorem. This general theorem provides sufficient conditions for the existence of a Markov process which is compatible with a collection of marginals as well as with a collection of bivariate transition laws.

4.1.1 Kernels and disintegration of measures

We will need the definition of a transition kernel:

Definition 4.1.2. Let $\mathcal{B}$ denote the set of Borel sets of $\mathbb{R}$. A *Kernel* is a map from $(\mathbb{R},\mathcal{B})$ such that:
(i)$\forall x \in \mathbb{R}$, the function $B \to K(x,B)$ is a probability measure on $\mathbb{R}$.
(ii)$\forall B \in \mathcal{B}$, the function $x \to K(x,B)$ is measurable.

Remark. (Disintegration of measures) A measure $\rho \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$ can be represented as a measure $\mu \in \mathcal{P}(\mathbb{R})$ and a transition kernel $K(x,dy)$ as follows:

$$\rho(A\times B) = \int_A K(x;B)\, d\mu(x)$$

4.1.2 Framework of Kellerer's theorem

Kellerer's main theorem is based on the following setting:
• The marginal laws are specified:

$$\forall t \in \mathbb{R}^+,\ \ \mu_t \in \mathcal{P}(\mathbb{R}) \text{ is given}$$

• For each pair of times (s,t), a collection of bivariate laws denoted $\mathcal{L}_{s,t}$ is specified:

$$\forall (s,t) \in \mathbb{R}^+\times\mathbb{R}^+,\ \mathcal{L}_{s,t} \subseteq \mathcal{P}(\mathbb{R}\times\mathbb{R}) \text{ is given.}$$
$$\text{Each } \rho \in \mathcal{L}_{s,t} \text{ must have marginals } \mu_s \text{ and } \mu_t.$$

4.1.3 The Kellerer existence theorem

Theorem 4.1.3. *(Kellerer) A sufficient condition for the existence of a stochastic process $(X_t)_{t\in\mathbb{R}^+}$ satisfying:*

$$\begin{cases} \forall t \in \mathbb{R}^+, & X_t \sim \mu_t \\ \forall s,t \in \mathbb{R}^+, & \text{the law of } (X_s,X_t) \text{ belongs to } \mathcal{L}_{s,t} \end{cases}$$

is that for all r,s,t with $r<s<t$,

$$\begin{cases} \mathcal{L}_{s,t} \text{ be closed and non empty.} \\ \mathcal{L}_{s,t} \subseteq \mathcal{S} \text{ (the set } \mathcal{S} \text{ is a special class of bivariate laws} - \text{see below for its definition)} \\ \text{If } \rho_1 \in \mathcal{L}_{r,s} \text{ and } \rho_2 \in \mathcal{L}_{s,t}, \text{ then their composition must belong to } \mathcal{L}_{r,t} \end{cases}$$

Definition 4.1.4. The set $\mathcal{S}$ denotes the following special class of bivariate laws. A measure $\rho \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$ belongs to $\mathcal{S}$ if and only if there exists a disintegration of ρ as

$$\rho(A\times B) = \int_A K(x;B)\, d\mu(x)$$

with $\mu \in \mathcal{P}(\mathbb{R})$ and a kernel $K(x,dy)$ which satisfies:

$$||K(a;.) - K(b,.)|| \leq ||a-b|| \qquad \text{for all } a,b \in \text{support}(\mu)$$

Lemma 4.1.5. The set $\mathcal{S}$ is closed with respect to the weak topology.

The proof of Kellerer's theorem relies on the following lemmas:

Lemma 4.1.6. Let $(\mu_t)_{t\in\mathbb{R}^+}$ be any specified collection of marginals.
Consider the set of real valued processes which have these marginals:

$$A = \{(X_t)_{t\in\mathbb{R}^+} \ : \ X_t \sim \mu_t \ \forall t \in \mathbb{R}^+\}$$

Then the set of measures on $\mathbb{R}^{[0,\infty)}$ corresponding to the above set of processes is compact with respect to the weak topology (see [Kel72]).

Lemma 4.1.7. Let $t_1,..,t_n$ be an increasing collection of times.
Define the set $A_{t_1,..,t_n}$ as:

$$A_{t_1,..,t_n} = \{(X_t)_{t\in\mathbb{R}^+} \in A : (X_{t_i}, X_{t_{i+1}}) \in \mathcal{L}_{t_i,t_{i+1}} \text{ for each } i \text{ from } 1 \text{ to } n-1\}$$

Then $A_{t_1,..,t_n}$ is closed and non empty.

Lemma 4.1.8. if $\{s_1,..,s_k\} \subseteq \{t_1,..,t_n\}$ then,

$$A_{t_1,..,t_n} \subseteq A_{s_1,..,s_k}$$

Proof. (of Theorem 4.1.3) The intersection of the sets $A_{t_1,..,t_n}$ over all finite collections of times $t_1,..,t_n$ is non empty. □

Chapter 5

A Brownian sheet martingale with the same marginals as the arithmetic average of geometric Brownian motion.

This section is based on a joint article with M. Yor see [BY09]

Definition 5.0.9. Brownian sheet is the two parameter centered a Gaussian process with covariance function

$$\mathbb{E}[B(s,t)B(s^{'},t^{'})] = \min(s,s^{'}) \times \min(t,t^{'})$$

Note that this implies that $\mathbb{V}ar[B(s,t)] = st$.

5.1 Introduction and Main Result

We construct a martingale which has the same marginals as the arithmetic average of geometric Brownian motion. This provides a short proof of the recent result due to P. Carr et al [CEX08] that the arithmetic average of geometric Brownian motion is increasing in the convex order. The Brownian sheet plays an essential role in the construction. Our method may also be applied when the Brownian motion is replaced by a stable subordinator.

To $(B_t, t \geqslant 0)$ a 1-dimensional Brownian motion, starting from 0, we associate the geometric Brownian motion:

$$\mathscr{E}_t = \exp(B_t - \frac{t}{2}), \qquad t \geqslant 0$$

and its arithmetic average:

$$\frac{1}{t}A_t = \frac{1}{t}\int_0^t ds\, \mathscr{E}_s, \qquad t \geqslant 0$$

A recent striking result by P. Carr et al [CEX08] is the following:

Theorem 5.1.1. *i) The process* $(\frac{1}{t}A_t, t \geqslant 0)$ *is increasing in the convex order, that is: for every convex function* $g : \mathbb{R}_+ \to \mathbb{R}$*, such that* $\mathbb{E}\left[\left|g\left(\frac{1}{t}A_t\right)\right|\right] < \infty$ *for every* $t > 0$*, the function:*

$$t \to \mathbb{E}\left[g\left(\frac{1}{t}A_t\right)\right] \text{ is increasing}$$

ii) In particular, for any $K \geqslant 0$*, the call and put prices of the Asian option which we define as:*

$$C^+(t,K) = \mathbb{E}\left[\left(\frac{1}{t}A_t - K\right)^+\right] \text{ and } C^-(t,K) = \mathbb{E}\left[\left(K - \frac{1}{t}A_t\right)^+\right]$$

are increasing functions of $t \geqslant 0$.

Comments on Theorem 5.1.1

a) One of the difficulties inherent to the proof of ii), say, is that the law of A_t for fixed t, is complicated, as can be seen from the literature on Asian options.

b) A common belief among practitioners is that any "decent" option price should be increasing with maturity. But examples involving "strict local martingales" show that this need not be the case. See e.g. Pal-Protter [PP08], Delbaen-Schachermayer [DS95]. On the other hand Theorem 5.1.1 offers a proof of the increase in maturity for Asian options.

The proof of Theorem 5.1.1 as given in [CEX08] (see also [BY08] for a slight variation) is not particularly easy, as it involves the use of either a maximum principle argument (in [CEX08]) or a supermartingale argument (in [BY08]). We note that the proofs given in [CEX08] and [BY08] show that for any individual convex function g, the associated function $G(t) = \mathbb{E}[g(\frac{1}{t}A_t)]$ is increasing. In contrast, in the present paper we obtain directly the result of Theorem 5.1.1 as a consequence of Jensen's inequality, thanks to the following

Theorem 5.1.2. *i)There exists a filtered probability space* $(\Omega, \mathscr{G}, \mathscr{G}_t, \mathbb{Q})$ *and a continuous martingale* $(M_t, t \geqslant 0)$ *on this space such that:*
for every fixed $t \geqslant 0$*,* $\frac{1}{t}A_t \overset{(law)}{=} M_t$
ii)More precisely, if $(W_{u,t}, u \geqslant 0, t \geqslant 0)$ *denotes the standard Brownian sheet and* $\mathscr{F}_{u,t} = \sigma\{W_{v,s}, v \leqslant u, s \leqslant t\}$ *its natural increasing family of* σ*-fields, one may choose:*

$$M_t = \int_0^1 du \exp(W_{u,t} - \frac{ut}{2}) \quad , \qquad t \geq 0$$

which is a continuous martingale with respect to $(\mathscr{F}_{\infty,t}, t \geqslant 0)$

We note that in [MY02] several methods have been developed to construct martingales with given marginals, an important problem considered by Strassen, Doob, Kellerer among others. See, e.g., references in [MY02]. Theorem 5.1.2 may also be considered in this light, providing a martingale whose one-dimensional marginals are those of $(\frac{1}{t}A_t, t \geqslant 0)$. In Section 2, we give our (very simple!) proof of Theorem 5.1.2, and we comment on how we arrived gradually at the formulation of Theorem 5.1.2. We also obtain a variant of Theorem 5.1.2 when $(\exp(B_t - \frac{t}{2}), t \geqslant 0)$ is replaced by $(\exp(B_t - at), t \geqslant 0)$ for any $a \in \mathbb{R}$.

In Section 3, we study various possible extensions of Theorem 5.1.2, i.e. : when the original Brownian motion $(B_t, t \geqslant 0)$ is replaced by certain Lévy processes, in particular stable subordinators and self-decomposable Lévy processes. In Section 4, we study some consequences of Theorem 5.1.1.

5.2 Proof of Theorem 5.1.2, and Comments

(2.1) We first make the change of variables: $u = vt$, in the integral

$$A_t = \int_0^t du \exp(B_u - \frac{u}{2})$$

We get: $\frac{1}{t}A_t = \int_0^1 dv \exp(B_{vt} - \frac{vt}{2})$
It is now immediate that since, for fixed t,

$$(B_{vt}, v \geqslant 0) \overset{(law)}{=} (W_{v,t}, v \geqslant 0), \text{ then:}$$

$$\text{for fixed } t, \qquad \frac{1}{t}A_t \overset{(law)}{=} \int_0^1 dv \exp(W_{v,t} - \frac{vt}{2})$$

Denoting by (M_t) the right-hand side, it remains to prove that it is a $(\mathscr{F}_{\infty,t}, t \geqslant 0)$ martingale. However, let $s < t$, then:

$$\mathbb{E}\left[M_t \middle| \mathscr{F}_{\infty,s}\right] = \int_0^1 dv\, \mathbb{E}\left[\exp(W_{v,t} - \frac{vt}{2}) \middle| \mathscr{F}_{\infty,s}\right].$$

Since $(W_{v,t} - W_{v,s})$ is independent from $\mathscr{F}_{\infty,s}$, we get:

$$\mathbb{E}\left[\exp(W_{v,t} - \frac{vt}{2}) \middle| \mathscr{F}_{\infty,s}\right] = \exp(W_{v,s} - \frac{vs}{2})$$

so that, finally: $\mathbb{E}\left[M_t \middle| \mathscr{F}_{\infty,s}\right] = M_s$.
This ends the proof of Theorem 5.1.2.

Remark: The same argument of independence allows to show more generally that, if $f : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is space-time harmonic, i.e. $(f(B_t, t), t \geqslant 0)$ is a martingale, then:

$$M_t^{(f)} \overset{def}{=} \int_0^1 du\, f(W_{u,t}, ut)$$

is a $(\mathscr{F}_{\infty,t}, t \geqslant 0)$ martingale. Thus in particular, for any $n \in \mathbb{N}$, one gets:

$$\text{for fixed } t, \qquad \frac{1}{t}\int_0^t du\, H_n(B_u, u) \overset{(law)}{=} M_t^{(n)}$$

$$\text{where: } M_t^{(n)} = \int_0^1 du\, H_n(W_{u,t}, ut)$$

and $H_n(x,t) = t^{n/2} h_n(\frac{x}{\sqrt{t}})$ denotes the n^{th} Hermite polynomial in the two variables $(x,t) \in \mathbb{R} \times \mathbb{R}_+$

Consequently, in that generality,

$$\left(\frac{1}{t}\int_0^t du\, f(B_u, u), t \geqslant 0\right)$$

is increasing in the convex order sense.

(2.2) At this point, we feel that a few words of comments on how we arrived gradually at the statement of Theorem 5.1.2 may not be useless.
(2.2.1) We first recall the basic result of Rothschild and Stiglitz [RS70]. The notation $\leqslant_{cv}$ means domination in the convex order sense; see [SS94], [SS06].

Proposition 5.2.1. *Two variables X and Y on a probability space satisfy:*
$X \leqslant_{cv} Y$ if and only if on some (other) probability space, there exists $\hat{X}$ and $\hat{Y}$ such that:

$$(i)\; X \overset{(law)}{=} \hat{X} \qquad (ii)\; Y \overset{(law)}{=} \hat{Y} \qquad (iii)\; \mathbb{E}\left[\hat{Y} \middle| \hat{X}\right] = \hat{X}$$

For discussions, variants, amplifications of the RS result, we refer the reader to the books of Shaked-Shantikumar ([SS94], [SS06]). Thus in order to show that a process $(H_t, t \geqslant 0)$ is increasing in the convex order sense, one is led naturally to look for a martingale $(M_t^H, t \geqslant 0)$ such that:

$$\text{for fixed } t, \qquad H_t \overset{(law)}{=} M_t^H$$

In fact the papers of Strassen, Doob and Kellerer, refered in [MY02], show that there exists such a martingale $(M_t^H, t \geq 0)$.
(2.2.2) The following variants of Proposition 1 shall lead us to consider properties of the process:

$$\frac{1}{t}A_t^{(a)} \equiv \frac{1}{t}\int_0^t ds \exp(B_s - as)$$

for any $a \in \mathbb{R}$.
The notation [icv], resp. [dcv] used below indicates the notion of "increasing convex ", resp. "decreasing convex" order. (See e.g. [SS94], [SS06] for details; in particular, Theorem 2.A.3 in [SS94] and Theorem 3.A.4 in [SS06])

Proposition 5.2.2. *Two variables X and Y on a probability space satisfy:*
$X \leqslant_{[icv]} Y$ if and only if there exists on some (other) probability space, a pair $(\hat{X}, \hat{Y})$ such that:

$$(i)\; X \overset{(law)}{=} \hat{X} \qquad (ii)\; Y \overset{(law)}{=} \hat{Y} \qquad (iii)_{\uparrow}\quad \hat{X} \leqslant \mathbb{E}\left[\hat{Y} \middle| \hat{X}\right]$$

Proposition 5.2.3. *Same as Proposition 5.2.2, but where [icv] is replaced by [dcv], and* $(iii)_\uparrow$ *by:* $(iii)_\downarrow \quad \hat{X} \geqslant \mathbb{E}\left[\hat{Y}\middle|\hat{X}\right]$

We now apply Propositions 5.2.2 and 5.2.3 to the process $(\frac{1}{t}A_t^{(a)}, t \geqslant 0)$

Theorem 5.2.4. *1) Let* $a \leqslant \frac{1}{2}$. *Then the process* $(\frac{1}{t}A_t^{(a)}, t \geqslant 0)$ *increases in the [icv] sense*
2) Let $a \geqslant \frac{1}{2}$. *Then, the process* $(\frac{1}{t}A_t^{(a)}, t \geqslant 0)$ *increases in the [dcv] sense.*

We leave the details of the proof of Theorem 5.2.4 to the reader as it is extremely similar to that of Theorem 5.1.2.

(2.2.3) The following statement is presented here in order to help with our explanation of how we arrived gradually at the statement of Theorem 5.1.2.

Proposition 5.2.5. *Let* (Z_u) *and* $(Z_u^{'})$ *denote two processes. Then under obvious adequate integrability assumptions, we have:*

$$\int_0^1 du\, Z_u\, \mathbb{E}\left[Z_u^{'}\middle|Z\right] \leqslant_{cv} \int_0^1 du\, Z_u\, Z_u^{'}$$

Again, the proof is an immediate application of Jensen's inequality.
We now explain how we arrived at Theorem 5.1.2:
we first showed that, for $0 < \sigma^{'} < \sigma$, there is the inequality:

$$I_{\sigma'} \equiv \int_0^1 du \exp(\sigma^{'} B_u - \frac{\sigma^{'2}}{2}u) \leqslant_{cv} \int_0^1 du \exp(\sigma B_u - \frac{\sigma^2 u}{2}) \equiv I_\sigma \qquad (2)$$

Indeed, to obtain (2) as a consequence of Proposition 5.2.5, it suffices to write: $(\sigma B_u, u \geqslant 0) \overset{(law)}{=} (\sigma^{'} B_u + \gamma\beta_u, u \geqslant 0)$ where $(\beta_u, u \geqslant 0)$, is a BM independent from $(B_u, u \geqslant 0)$
and $\sigma^2 = (\sigma^{'})^2 + \gamma^2$, i.e. $\gamma = \sqrt{\sigma^2 - (\sigma^{'})^2}$
Once we had made this remark, it seemed natural to look for a "process" argument (with respect to the parameter σ), and this is how the Brownian sheet comes naturally into the picture.

5.3 Variants involving stable subordinators and self-decomposable Lévy processes

(3.1) Here is an analogue of Theorem 5.1.1 when we replace Brownian motion by a (α)-stable subordinator (T_t), for $0 < \alpha < 1$, whose law is characterized by:

$$\mathbb{E}\left[\exp(-\lambda T_t)\right] = \exp(-t\lambda^\alpha) \quad , \quad t \geqslant 0,\ \lambda \geqslant 0$$

Theorem 5.3.1. *The process* $\frac{1}{t}A_t^{(\alpha)} \overset{def}{=} \frac{1}{t}\int_0^t ds \exp(-\lambda T_s + s\lambda^\alpha)$ *is increasing for the convex order.*

We prove Theorem 5.3.1 quite similarly to the way we proved Theorem 5.1.1, namely: there exists a α-stable sheet $(T_{s,t}, s \geqslant 0, t \geqslant 0)$ which may be described as follows:

$(T(A), A \in \mathcal{B}(\mathbb{R}_+^2), |A| < \infty)$ is a random measure such that:
i) for all $A_1, \ldots, A_k$ disjoint Borel sets with $|A_i| < \infty$,
$T(A_1), .., T(A_k)$ are independent random variables,
ii) $\mathbb{E}[\exp(-\lambda T(A_i))] = \exp(-|A_i|\lambda^\alpha), \lambda \geq 0$.
($T(A_i)$ is an α-stable random variable)

Then we denote $T_{s,t} = T(R_{s,t})$, with $R_{s,t} \equiv [0,s] \times [0,t]$
See, e.g., [ST94] for the existence of such measures. The result of Theorem 5.3.1 is a consequence of:

Theorem 5.3.2. *The process* $M_t^{(\alpha)} = \int_0^1 du \exp(-\lambda T_{u,t} + ut\lambda^\alpha)$ *is a* $\mathscr{F}_{\infty,t}^{(\alpha)} \equiv \sigma\{T_{h,k}, h \geqslant 0, k \leqslant t\}$ *martingale, and for fixed t:*

$$\frac{1}{t} A_t^{(\alpha)} \overset{(law)}{=} M_t^{(\alpha)}$$

(3.2) We now consider a self-decomposable Lévy process.
(See e.g., Jeanblanc-Pitman-Yor [JPY02] for a number of properties of these processes.)
Assuming that: $\forall \alpha > 0$, $\mathbb{E}[\exp(\alpha X_u)] < \infty$, then:

$$\mathbb{E}[\exp(\alpha X_u)] = \exp(u\varphi(\alpha)), \text{ for some function } \varphi.$$

In this framework, we show the following.

Theorem 5.3.3. *The process* $(I_\alpha = \int_0^1 du \exp(\alpha X_u - u\varphi(\alpha)),\ \alpha \geqslant 0)$ *is increasing in the convex order.*

Proof. Since $(X_u, u \geqslant 0)$ is self-decomposable, there exists, for any $c \in (0,1)$, another Lévy process $(\eta_u^{(c)}, u \geqslant 0)$ such that:
$(X_u, u \geqslant 0) \overset{(law)}{=} (cX_u + \eta_u^{(c)}, u \geqslant 0)$, with independence of X and $\eta^{(c)}$. Consequently, we obtain, for any $(\alpha, c) \in (0,\infty) \times (0,1)$

$$I_\alpha \overset{(law)}{=} \int_0^1 du \exp(\alpha c X_u - u\varphi(\alpha c)) \exp(\alpha \eta_u^{(c)} - u\varphi_c(\alpha)) \tag{3}$$

where on the RHS of (3), X and $\eta^{(c)}$ are assumed to be independent.
Denote by I'_α the RHS of (3), then :

$$\mathbb{E}\left[I'_\alpha \middle| X\right] = \int_0^1 du \exp(\alpha c X_u - u\varphi(\alpha c)) = I_{\alpha c}$$

which implies, from Jensen's inequality: for every convex function g,

$$\mathbb{E}\left[g(I_{\alpha c})\right] \leqslant \mathbb{E}\left[g(I_{\alpha})\right]$$

$\square$

However we have not found, in this case, a martingale $(\mu_\alpha,\ \alpha \geqslant 0)$ such that:

$$\text{for every fixed } \alpha, \qquad I_\alpha \overset{(law)}{=} \mu_\alpha$$

Remark: We note that the above argument is a particular case of the argument presented in Proposition 5.2.5, which involves two processes Z and Z'.

5.4 Some consequences

Since the process $(\frac{1}{t}A_t, t \geqslant 0)$ is increasing in the convex order, we find, by differentiating the increasing function of t: $\mathbb{E}[(K - \frac{1}{t}A_t)^+]$

$$\text{for every } K \geqslant 0 \text{ and } t \geqslant 0, \qquad \mathbb{E}\left[\mathbf{1}(\frac{1}{t}A_t < K)\ \ (\mathscr{E}_t - \frac{1}{t}A_t)\right] \geqslant 0,$$

although, it is not true that: $\mathbb{E}\left[\mathscr{E}_t \mid \frac{1}{t}A_t\right]$ is greater than or equal to $\frac{1}{t}A_t$, since this would imply that: $\frac{1}{t}A_t = \mathcal{E}_t$, as the common expectation of both quantities is 1.

(4.1) More generally, the following proposition presents a remarkable consequence of the increasing property of the process $(\frac{1}{t}A_t, t \geq 0)$ in the convex order sense.

Proposition 5.4.1. *For every increasing Borel function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ there is the inequality:*

$$\mathbb{E}\left[\varphi\left(\frac{1}{t}A_t\right)\left(\frac{1}{t}A_t\right)\right] \leq \mathbb{E}\left[\varphi\left(\frac{1}{t}A_t\right)\mathcal{E}_t\right]. \qquad (\star)$$

Equivalently,

$$\mathbb{E}\left[\varphi\left(\frac{1}{t}A_t\right)\left(\frac{1}{t}A_t\right)\right] \leq \mathbb{E}\left[\varphi\left(\frac{1}{t}\tilde{A}_t\right)\right], \qquad (\star\star)$$

where $\tilde{A}_t = \int_0^t du \exp(B_u + \frac{u}{2})$

Proof. We may assume φ bounded. Then, $g(x) = \int_0^x dy \varphi(y)$ is convex (its derivative is increasing), and formula $(\star)$ follows by differentiating the increasing function:

$$t \to \mathbb{E}\left[g\left(\frac{1}{t}A_t\right)\right].$$

Formula $(\star\star)$ follows from $(\star)$ by using the Cameron-Martin relationship between $(B_u, u \leq t)$ and $(B_u + u, u \leq t)$

$\square$

(4.2) As a partial check on the previous result $(\star)$, we now prove directly that, for every integer $n \geq 1, t \to \mathbb{E}[(\frac{1}{t}A_t)^n]$ is increasing and that: $\mathbb{E}[(\frac{1}{t}A_t)^n] \leq \mathbb{E}[(\frac{1}{t}A_t)^{n-1}\mathcal{E}_t]$

Here are two explicit formulae for: $\alpha_n(t) = \mathbb{E}[(\frac{1}{t}A_t)^n]$, and $\beta_n(t) = \mathbb{E}[(\frac{1}{t}A_t)^{n-1}\mathcal{E}_t]$.

$$\begin{aligned}\alpha_n(t) &= \frac{n!}{t^n}\mathbb{E}[\int_0^t ds_1 \int_{s_1}^t ds_2 ... \int_{s_{n-1}}^t ds_n \exp((B_{s_1} + ... + B_{s_n}) - \frac{1}{2}(s_1 + .. + s_n))] \\ &= \frac{n!}{t^n}\int_0^t ds_1 \int_{s_1}^t ds_2 ... \int_{s_{n-1}}^t ds_n \exp(\frac{1}{2}C(s_1, .., s_n))\end{aligned}$$

$$\begin{aligned}\text{where } C(s_1, .., s_n) &= \mathbb{E}[(B_{s_1} + B_{s_2} + .. + B_{s_n})^2] - (s_1 + ... + s_n) \\ &= 2\sum_{1 \leq i \leq n} s_i(n-i) \qquad (\geqslant 0)\end{aligned}$$

Consequently:

$$\alpha_n(t) = n!\int_0^1 du_1 ... \int_{u_{n-1}}^1 du_n \exp(\frac{t}{2}C(u_1, ..., u_n)) \qquad (3)$$

from which it follows that $\alpha_n(t)$ is increasing in t.

$$\begin{aligned}&\text{Now } \beta_n(t) = \frac{(n-1)!}{t^{n-1}} \times \\ &\int_0^t ds_1 ... \int_{s_{n-1}}^t ds_{n-1}\mathbb{E}\left[\exp((B_{s_1} + ... + B_{s_{n-1}} + B_t) - \frac{1}{2}(s_1 + ... + s_{n-1} + t))\right] \\ &= (n-1)!\int_0^1 du_1 ... \int_{u_{n-2}}^1 du_{n-1} \exp(\frac{t}{2}C(u_1, ..., u_{n-1}, 1)) \qquad (4)\end{aligned}$$

We have already seen from formula (3), that $\alpha_n(t)$ is increasing in t; consequently: $\alpha_n'(t) \geq 0$ and by definition of α_n:

$$\begin{aligned}\alpha_n'(t) &= n\mathbb{E}\left[\left(\frac{1}{t}A_t\right)^{n-1}\left(-\frac{1}{t^2}A_t + \frac{\mathcal{E}_t}{t}\right)\right] \\ &= \frac{n}{t}\{\beta_n(t) - \alpha_n(t)\}\end{aligned}$$

Hence: $\beta_n(t) \geq \alpha_n(t)$.

(4.3) To conclude this paper, let us connect the properties of increase of the functions α_n and β_n with our method of proving Theorem 5.1.1 using the Wiener sheet, as performed in Theorem 5.1.2.

Indeed, the same argument as in Theorem 5.1.2 shows that for any positive measure $\mu(du_1,...,du_n)$ on $[0,1]^n$ the process:

$$\int \mu(du_1,..,du_n)\prod_{i=1}^{n}\mathcal{E}_{(u_it)} \qquad (5)$$

admits the same one-dimensional marginals as the $(\mathcal{W}_t')$ submartingale

$$\int \mu(du_1,..,du_n)\prod_{i=1}^{n}\mathcal{E}_t^{(u_i)}(W) \qquad (6)$$

where $\mathcal{E}_t^{(u)}(W) = \exp(W_{u,t} - \frac{ut}{2})$.
Hence, the common expectation of (5) and (6) increases with t; $\alpha_n(t)$ and $\beta_n(t)$ constitute particular examples of this.
A final Note: Pushing further the use of the Brownian sheet and a variation from the construction of the Ornstein-Uhlenbeck process on the canonical path-space $C([0,1];\mathbb{R})$ in terms of that sheet, Hirsch-Yor [HY09] obtain a large class of processes, adapted to the brownian filtration, which admit the one-dimensional marginals of a martingale.

Chapter 6

When the greeks of Asian options are positive supermartingales

6.1 Introduction

In [CEX08] Carr, Ewald and Xiao prove that under the assumptions of the Black-Scholes Model a convex payoff arithmetic Asian option's value is a monotonically increasing function of the volatility. In this paper we present a supermartingale argument which is used to obtain this monotonicity result for all diffusions with affine coefficients. This includes the geometric Brownian motion of the Black-Scholes model as well as processes such as the mean reversing Ornstein-Uhlenbeck of the Vasicek model. This is of practical importance because due to their averaging feature Asian options are often written on exchange rates, interest rates or commodities which do not follow the dynamics of the Black-Scholes model but can be modeled by the Vasicek model.

By showing the vega of an Asian call is a positive supermartingale in addition to the monotonicity implications in σ this yields additional information on an investor's exposure to volatility through this instrument. Not only does this instrument make him long on volatility, but in addition his expected future exposure to volatility through this instrument is less than his current exposure. This clearly provides useful insights for risk management.

An option is a financial contract whose value depends on another economic variable called the underlying. The underlying could be for exemple a stock, an exchange rate, an interest rate, a commodity. It is not surprising that the properties of a particular Asian option are highly dependent on those of its underlying. In a given financial model the dynamics of the underlying will be specified by a stochastic process. Under the the Black-Scholes model assumption the underlying, which is a Stock, follows a geometric Brownian motion with drift. Under the Vasicek model the underlying which is the short term interest rate follows a mean reversing Ornstein-Uhlenbeck process. The short term interest rate can also be modeled by a Cox-Ingersoll-Ross (CIR) process. In the following pages we develop a methodology to study the impact of volatility of Asian options, which we apply to successively larger classes of underlying processes.

An Asian option is a path dependent option which means that its value is dependent on the entire trajectory of the underlying from the initial time t=0 to maturity (t=T). given a specified function g the holder of the asian option with maturity T on the underlying X receives at maturity the following payoff

$$g(\frac{1}{T}\int_0^T X_u du) = g(A_T) \text{ where } A_t = \frac{1}{t}\int_0^t X_u du$$

In the case of an Asian call $g(x) = (x-K)^+$ and in the case of an Asian put $g(x) = (x-K)^-$

By the risk neutral pricing formula (see for example [Shr00]), the value of the Asian option is

$$v(t,T,x,y,\sigma) = \mathbb{E}(g(\frac{1}{T}\int_0^T X_u du) \mid \mathcal{F}_t) = \mathbb{E}(g(\frac{1}{T}\int_0^T X_u du) \mid X_t = x, \int_0^t X_u du = y)$$

We see that the value of an Asian option is a functional of the running average of the underlying process. This paper describes an approach to evaluate how changes in volatility affect the value of Asian options on certain underlyings.

The content of the paper is distributed as follows: Section 2 contains a discussion of the problem and an outline of the supermartingale argument which will be employed in the following three sections. In Section 3 we work under the Black-Scholes model assumptions and use the supermartingale argument to get the results obtained in [CEX08]. In section 4, we show that this is true for all diffusions with affine coefficients and such that the volatility parameter does not appear in the drift term . In Section 5 we prove the result for an interesting case where the drift of the process is dependent on volatility. In Section 6 the result is examined under the expected utility framework. Finally in section 7 we use a result from the theory of expected utility to obtain a property of the running average process of geometric Brownian motion.

6.2 The supermartingale argument

We now give an outline of our supermartingale argument, The method we use is to focus on the time indexed process v_σ which describes the sensitivity of the option to volatility. This sensitivity is of course a function of the two state variables X_t and Y_t. Heuristically the two key phenomena and their mathematical formulations are

- The present vega (v_σ) is larger than the expected future vega (v_σ is a supermartingale)

- at maturity the option is no longer affected by volatility ($v_\sigma(t=T)=0$)

These two properties imply that vega cannot be negative, for if it was negative at a given time (call it t_0) then vega would be expected to increase (actually certain to increase) between t_0 and and T contradicting the supermartingale property. i.e. the expectation of the future value is less than the present value.

The state variables of an arithmetic asian option are

- The underlying (for example an exchange rate, an interest rate, a stock price) which is modeled as a diffusion process. $dX_t = b(X_t)dt + a(X_t)\sigma dW_t$ Note that this encompasses a fairly large class of diffusion processes containing Geometric Brownian motion, the CIR process ($dr_t = -\theta(r_t-\mu)dt + \sigma\sqrt{r_t}dW_t$) etc...
  Note However that the class of diffusion which we are considering excludes those for which the volatility parameter appears in the drift term
- the running integral of the underlying which by definition is $Y_t = \int_0^t X_s ds$

6.3 Black Scholes model

Lemma 6.3.1. let g be a stricly convex (resp. concave) function. Under the assumptions of the Black Scholes model [BS73], the Asian option with payoff $\mathbb{E}[g(\int_0^1 duX_u^{(\sigma)})]$ has a positive (resp. negative) gamma (v_{xx})

Proof.

$$v_{xx}(t,x,y,\sigma) = \frac{d^2}{dx^2}\mathbb{E}[g(y+x\int_0^{1-t} duX_u^{(\sigma)})] = \mathbb{E}[g''(y+x\int_0^{1-t} duX_u^{(\sigma)})(\int_0^{1-t} duX_u^{(\sigma)})^2] > 0$$

□

Theorem 6.3.2. *If $v_{xx} \geq 0$ then $v_\sigma(t, X_t^{(\sigma)}, \int_0^t duX_u^{(\sigma)}, \sigma)$ $(0 \leq t \leq T)$ is an $\mathcal{F}_t -$ supermartingale,which takes the value 0 at t=T*

Proof. By a regularity result for the solutions of PDEs, $v_\sigma(t,x,y)$ is C^1 in t and C^2 in x and y.

We may apply Ito's formula to obtain

$$\begin{aligned} dv_\sigma(t,X_t,Y_t) =& v_{\sigma t}dt + v_{\sigma x}b(X_t)dt + v_{\sigma x}\sigma a(X_t)dW_t + xv_{\sigma y}dt + \frac{1}{2}\sigma^2 a^2(X_t)v_{\sigma xx}dt \\ =& (v_{\sigma t} + v_{\sigma x}b(X_t) + xv_{\sigma y} + \frac{1}{2}\sigma^2 a^2(X_t)v_{\sigma xx})dt + v_{\sigma x}\sigma a(X_t)dW_t \end{aligned}$$

Now we focus on the finite variation process (the dt term) and proceed as in Carr, Ewald, Xiao [CEX08] in order to obtain an alternate representation for it.

We are placing ourselves under the martingale equivalent measure. It follows that $v(t,X_t,Y_t)$ is a martingale.
This implies that the dt term in its Ito development is 0.
in other words

$$v_t + v_x b(x_t) + x v_y + \frac{1}{2}\sigma^2 a^2(x) v_{xx} = 0$$

Differentiating with respect to σ which we can do by regularity results for the solutions to PDEs

$$v_{t\sigma} + v_{x\sigma} b(x) + x v_{y\sigma} + \sigma a^2(x) v_{xx} + \frac{1}{2}\sigma^2 a^2(x) v_{xx\sigma} = 0$$

$$\Rightarrow v_{t\sigma} + v_{x\sigma} b(x) + x v_{y\sigma} + \frac{1}{2}\sigma^2 a^2(x) v_{xx\sigma} = -\sigma a^2(x) v_{xx}$$

Returning to the semimartingale representation of v_σ we see that

$$dv_\sigma(t,X_t,Y_t) = -\sigma^2 a^2(X_t) v_{xx} dt + v_{\sigma x} \sigma a(X_t) dW_t$$

by the equation above, if $v_{xx} > 0$ the finite variation process in the semimartingale representation of v_σ is decreasing which implies that $v_\sigma(t,X_t,Y_t)$ is an $\mathcal{F}_t-$supermartingale.
Moreover $v(1,x,y,\sigma) = g(y)$ so $v_\sigma(1,x,y,\sigma) = 0$
this completes the proof. Now we show the non-negativity of the supermartingale

$$0 = \mathbb{E}[v_\sigma(1, X_1^{(\sigma)}, \int_0^1 X_u^{(\sigma)} du, \sigma) | \mathcal{F}_t] \le v_\sigma(t, X_t^{(\sigma)}, \int_0^t du X_u^{(\sigma)}, \sigma)$$

□

Corollary 6.3.1. for any $t \in [0,T)$ if g is a convex function (resp. concave) the option's value is increasing in volatility (resp. decreasing).

Proof. We have shown that for g convex (resp. concave) v_σ is a postive supermartingale (resp. negative submartingale) The monotonicity follows from the constant sign of the partial derivative. □

Corollary 6.3.2. for $t = 0$ If g is a convex function (resp. concave) the option's value is increasing (resp. decreasing) in time to maturity (T).

Proof. By scaling we reduce the problem to $T = 1$ and to a discussion in $\sigma(= \sqrt{T})$
Indeed

$$\mathbb{E}[g(\frac{1}{T}\int_0^T \exp(B_s - \frac{s}{2}) ds)] = \mathbb{E}[g(\int_0^1 \exp(\sigma B_u - \frac{\sigma^2 u}{2}) du)]$$

□

6.4 Diffusions with affine coefficients

This section is devoted to situations where the dynamics of the underlying are described by a diffusion process with affine coefficients. More precisely X_t is the unique solution to the stochastic differential equation $dX_t = b(X_t)dt + a(X_t)\sigma dW_t$ where a and b are affine functions. this class encompasses the following processes

- geometric Brownian motion which is the solution to the following SDE $dX_t = \sigma X_t dW_t$
- geometric Brownian motion with constant drift which is the solution to the following SDE $dX_t = \mu X_t dt + \sigma X_t dW_t$ This is the process chosen to model the stock price in the Black-Scholes model
- the mean-reversing Ornstein-Uhlenbeck process which is the solution to the following SDE: $dr_t = -\theta(r_t - \mu)dt + \sigma dW_t$ This is the process chosen to model interest rates in the Vasicek model for the short term interest rate.

Theorem 6.4.1. *If $v_{xx} \geq 0$ then $v_\sigma(t, X_t^{(\sigma)}, \int_0^t du X_u^{(\sigma)}, \sigma)$ $(0 \leq t \leq T)$ is an $\mathcal{F}_t-$ supermartingale,which takes the value 0 at t=T*

Proof. By a regularity result for the solutions of PDEs, $v_\sigma(t,x,y)$ is C^1 in t and C^2 in x and y.

We may apply Ito's formula to obtain

$$\begin{aligned} dv_\sigma(t, X_t, Y_t) =& v_{\sigma t}dt + v_{\sigma x}b(X_t)dt + v_{\sigma x}\sigma a(X_t)dW_t + xv_{\sigma y}dt + \frac{1}{2}\sigma^2 a^2(X_t)v_{\sigma xx}dt \\ =& (v_{\sigma t} + v_{\sigma x}b(X_t) + xv_{\sigma y} + \frac{1}{2}\sigma^2 a^2(X_t)v_{\sigma xx})dt + v_{\sigma x}\sigma a(X_t)dW_t \end{aligned}$$

Now we focus on the finite variation process (the dt term) and proceed as in Carr, Ewald, Xiao [CEX08] in order to obtain an alternate representation for it.

We are placing ourselves under the the martingale equivalent measure. It follows that $v(t, X_t, Y_t)$ is a martingale.
This implies that the dt term in its Ito development is 0.
in other words

$$v_t + v_x b(x_t) + xv_y + \frac{1}{2}\sigma^2 a^2(x)v_{xx} = 0$$

Differentiating with respect to σ which we can do by regularity results for the solutions to PDEs

$$v_{t\sigma} + v_{x\sigma}b(x) + xv_{y\sigma} + \sigma a^2(x)v_{xx} + \frac{1}{2}\sigma^2 a^2(x)v_{xx\sigma} = 0$$

$$\Rightarrow v_{t\sigma} + v_{x\sigma}b(x) + xv_{y\sigma} + \frac{1}{2}\sigma^2 a^2(x)v_{xx\sigma} = -\sigma a^2(x)v_{xx}$$

Returning to the semimartingale representation of v_σ we see that

$$dv_\sigma(t,X_t,Y_t) = -\sigma^2 a^2(X_t) v_{xx} dt + v_{\sigma x}\sigma a(X_t) dW_t$$

by the equation above, if $v_{xx} > 0$ the finite variation process in the semimartingale representation of v_σ is decreasing which implies that $v_\sigma(t,X_t,Y_t)$ is an $\mathcal{F}_t-$supermartingale. Moreover $v(1,x,y,\sigma) = g(y)$ so $v_\sigma(1,x,y,\sigma) = 0$
this completes the proof. Now we show the non-negativity of the supermartingale

$$0 = \mathbb{E}[v_\sigma(1,X_1^{(\sigma)},\int_0^1 X_u^{(\sigma)} du,\sigma)|\mathcal{F}_t] \leq v_\sigma(t,X_t^{(\sigma)},\int_0^t duX_u^{(\sigma)},\sigma)$$

$\square$

Implications We have reduced our study of v_σ to the study of v_{xx} This is called the gamma (greek letter Γ) of the option. We now obtain an expression for v_{xx} interms of the underlying process X_t

Recall that $v(t,y,x,\sigma) = \mathbb{E}[g(\int_0^1 X_s ds)|\mathcal{F}_t] = \mathbb{E}[g(y+\int_0^{1-t} duX_u^x)]$ with $y = \int_0^t dsX_s$ and $x = X_t$
Differentiating with respect to the initial condition,

$$v_x(t,y,x,\sigma) = \mathbb{E}[g^{'}(\int_0^{1-t} X_s^x ds)(\int_0^{1-t} X_s^{'x} ds)]$$

Differentiating again with respect to the initial condition,

$$v_{xx} = \mathbb{E}[g^{''}(y+\int_0^{1-t} X_s^x ds)(\int_0^{1-t} X_s^{'x})^2 ds + g^{'}(y+\int_0^{1-t} X_s^x ds)(\int_0^{1-t} X_s^{''x^2} ds)]$$

We must now examine the derivative of the process X_t with respect to the initial condition.

Recall $X_t^x = x + \int_0^t a(X_s^x)\sigma d\beta_s + \int_0^t b(X_s^x) ds$

$$\Rightarrow X_t^{'x} = 1 + \int_0^t [a^{'}(X_s^x)\sigma d\beta_s + b^{'}(X_s^x)ds]X_s^{'x}$$

$$= \exp(\int_0^t a^{'}(X_s^x)\sigma d\beta_s + b^{'}(X_s^x)ds) - \frac{1}{2}\int_0^t (a^{'}(X_s^x))^2\sigma^2 ds)$$

We see that if a and b are affine functions $\frac{d}{dx}X_t$ does not depend on x . Which means that $\frac{d^2}{dx^2}X_t$ is the identically zero process And so

$$v_{xx}(t,y,x,\sigma) = \mathbb{E}[g^{''}(y+\int_0^{1-t} X_s^x)(\int_0^{1-t} X_s^{'x} ds)^2]$$

Discussion : When the underlying is a diffusion with affine coefficients whose drift is unaffected by volatility, the following holds: (a) The value of an Asian call is increasing in volatility.

(b) The value of an Asian put is decreasing in volatility.
More generally:
When the payoff function g is convex, the vega (ν_σ) of an asian option is a positive supermartingale
When the payoff function g is concave, the vega (ν_σ) of an asian option is a negative submartingale.

- Black Scholes (the undelying follows a geometric Brownian motion)
- Vasicek (the underlying follows a mean reversing Ornstein-Uhlenbeck process)

Directions for further research The above results characterize the qualitative effect of volatility on an important class of underlying diffusions. A direction for further research would be to examine the effect of volatility on processes with non-affine coefficients such as the Cox-Ingersoll-Ross (CIR) process, which is the solution to the following stochastic differential equation.

$$dr_t = -\theta(r_t - \mu)dt + \sigma\sqrt{X_t}dW_t$$

It has the property of staying positive which is useful when modeling the short term interest rate.

6.5 An interesting case where σ appears in the drift

The case where the underlying follows a geometric Brownian motion with drift is an interesting one because it describes the dynamics of the underlying in the Black-Scholes model. In this section X_t is taken to be the solution to the following linear stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

Using Ito's formula we can check that the process X_t given by

$$X_t = X_0 \exp(\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t)$$

is the solution to the above stochastic differential equation starting from X_0 at time 0. μ constant was covered in section 4. This Section is devoted to the case where μ depends on σ. There are several reasons why this is of interest. One of these reasons is that in the Capital Asset Pricing Model (CAPM) expected return is correlated with volatility. The justification for this is that risk adverse investors require a larger rate of expected return to hold on to an asset when its risk increases. As a result a realistic stock model might incorporate a dependency in σ into the drift

term.
Another reason is that this process appears when studying the following functional.

$$T \to g(\frac{1}{T}\int_0^T \exp(W_s - as)ds)$$

which with the following change of variable $T = \sigma^2$ is equal in law to

$$g(\int_0^1 \exp(\sigma W_t - a\sigma^2 t)dt)$$

And monotonicity of the second expression in σ would lead to monotonicity of the first in T. Now Observe that:

$$[g(\int_0^1 dt \exp(\sigma B_t - a\sigma^2 t))|\mathcal{F}_t] = [g(\int_0^1 dt \exp(\sigma B_t - \frac{1}{2}\sigma^2 t(a - \frac{1}{2})\sigma^2 t))|\mathcal{F}_t]$$

It follows that we are indeed dealing with a geometric Brownian motion for which the drift coefficient is $\mu = (\frac{1}{2} - a)\sigma^2$
When $a = \frac{1}{2}$ we are in the case of geometric Brownian motion which was covered in Section 4. The process $\exp(\sigma W_t - a\sigma^2 t)$ is a diffusion which is the solution to the SDE: $dX_t = (\frac{1}{2} - a)\sigma^2 X_t dt + \sigma X_t dW_t$ And so its infinitesimal generator is

$$\mathcal{L}f = \sigma^2 \frac{x^2}{2}\frac{d^2 f}{dx^2} + \sigma^2(\frac{1}{2} - a)x\frac{df}{dx}$$

This is more complex than the framework used in the previous sections because here the volatility parameter σ appears not only in the diffusion term but also in the drift term.

Geometric Brownian motion with drift has the same scaling property as plain geometric Brownian motion.

$$\begin{aligned} v(t,x,y,\sigma) =& \mathbb{E}[g(\int_0^1 X_u du)|\mathcal{F}_t] \\ =& \mathbb{E}[g(y + x\int_0^{1-t} X_u du)] \text{ where } x = X_t, y = \int_0^t X_u du \end{aligned}$$

Under the risk neutral probability measure $\mathbb{P}$ all asset prices including this one are martingales. $v(t, S_t^{(\sigma)}, \int_0^t du S_u^{(\sigma)}, \sigma)$ being an $(\mathbb{P}, \mathcal{F}_t)$ martingale implies that its finite variation process is identically zero. By an application of Ito's formula this property translates into.

$$v_t + \frac{1}{2}\sigma^2 v_{xx} + \sigma^2(\frac{1}{2} - a)v_x + xv_y = 0$$

In order to obtain an alternate representation for the dt term of v_σ we differentiate the above equation with respect to σ

$$v_{t\sigma} + \frac{\sigma^2}{2} v_{xx\sigma} + \sigma^2(\frac{1}{2} - a) v_{x\sigma} + x v_{y\sigma} = -\sigma v_{xx} - 2\sigma(\frac{1}{2} - a) v_x$$
$$= -\sigma(v_{xx} + 2\sigma(\frac{1}{2} - a) v_x)$$

An application of Ito's formula to $v_\sigma(t, X_t, \int_0^t X_s ds, \sigma)$ shows that the above is the dt term of v_σ. We see that if $v_{xx} + 2(\frac{1}{2} - a)v_x \geq 0$ then this dt term is non-increasing which means that v_σ is a supermartingale This works if $\frac{1}{2} - a \geq 0$ Conclusion:
If $g' \geq 0$ and $\frac{1}{2} - a \geq 0$ the supermartingale argument holds
also if $g' \leq 0$ and $\frac{1}{2} \leq a$ the supermartingale argument holds again

As in the previous sections, the process v_σ attains the value 0 at $t = 1$ because at maturity the option price is independent of σ; indeed, $v(1, x, y, \sigma) = g(y)$.

6.6 Financial theory : Implications of the result

Given that this probability problem stems from the desire to hedge financial risk, it is tempting to look at the economic justifications for (or consequences of) this mathematical result. To do so we look at this problem from the expected utility viewpoint.

- $\int_0^1 \exp(\sigma B_s - \frac{\sigma^2}{2} s) ds$ is a random payoff
- g is the utility function of an investor. g is always assumed to be increasing and concave.
- $\mathbb{E}\left[g(\int_0^1 \exp(\sigma B_s - \frac{\sigma^2}{2} s) ds)\right]$ is the expected utility of the random payoff to an investor with utility curve g. $U(X) = \mathbb{E}[g(X)]$

In this framework the main result says that any increase in volatility (σ) results in a decrease of the expected utility of all risk adverse investors.

$$\sigma_1 > \sigma_2 \Rightarrow U\left(\int_0^1 \exp(\sigma_1 B_s - \frac{1}{2}\sigma_1^2 s) ds\right) < U\left(\int_0^1 \exp(\sigma_2 B_s - \frac{1}{2}\sigma_2^2 s) ds\right)$$

We now explain why this is not just a simple consequence of

$$\sigma_1 > \sigma_2 \Rightarrow Var\left(\int_0^1 \exp(\sigma_1 B_s - \frac{1}{2}\sigma_1^2 s) ds\right) > Var\left(\int_0^1 \exp\left(\sigma_2 B_s - \frac{1}{2}\sigma_2^2 s\right) ds\right)$$

Note that the above equation is indeed true; it follows from the characterization of the second moment of the payoff in [Yor92]

Within a location-scale family of probability measures, an increase in the variance does not always translate into a decrease in utility for all risk adverse investors. Variance is often used as a proxy for financial risk because it is easy to use but it is not always consistent with the economic notion of financial risk.

Rothschild and Stiglitz [RS70] give the following family of measures as an example where an increase in variance results in an increase in expected utility for some risk adverse investors. This family is indexed by $a, c > 0$ and given by the corresponding distribution function

$a, c > 0$

$$
\begin{aligned}
F_{a,c}(x) =& 0 && \text{for } x \leq 1 - 0.25/a \\
=& ax + 0.25 - a && \text{for } 1 - 0.25/a \leq x \leq 1 + (2c - 0.5)/(c - a) \\
=& cx + 0.75 - 3c && \text{for } 1 + (2c - 0.5)/(c - a) \leq x \leq 3 + 0.25/c \\
=& 1 && \text{for } x > 3 + 0.25/c
\end{aligned}
$$

and if we keep μ constant then $\frac{dT(y)}{d\sigma^2}$ changes sign where $T(y, \sigma^2, \mu) = \int_0^y F(x, \sigma^2, \mu)dx$ this implies that some investors with concave utility functions are better off with an increase in variance.

In our case every investor will experience a loss in expected utility when the volatility increases.

Haim Levy in [Lev92] considers utility curves with increase at a decreasing rate that is $u' \geq 0$ and $u'' \leq 0$. He shows that if X and Y are two random variables and $\mathbb{E}g(X) > \mathbb{E}g(Y)$ for all concave increasing functions g then this translates into a property of second order stochastic dominance

In our case we have shown that if $\sigma_1 < \sigma_2$
all risk adverse rational investors will prefer $\int_0^1 \exp(\sigma_1 B_s - \frac{1}{2}\sigma_1^2 s)ds$ to $\int_0^1 \exp(\sigma_2 B_s - \frac{1}{2}\sigma_2^2 s)ds$

6.7 Implications for the running average of geometric Brownian motion

Given two random variables X and Y Rothschild and Stiglitz in [RS70] show that the following are equivalent:

- $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for all concave increasing function u
- Y is equal in law to $X + Z$ with $\mathbb{E}(Z|X) = 0$

Denote the running average process by $A_t := \frac{1}{t}\int_0^t S_u du$ where S_u is a geometric Brownian Motion for $s < t$

We have shown that $\mathbb{E}u(A_s) \geq \mathbb{E}u(A_t)$ for any concave function u.

It then follows that A_t is equal in law to $A_s + Z$ with $\mathbb{E}[Z|A_s] = 0$

References also include [DMMY00], [CS04] and [Yor92]

Chapter 7

A sequence of Albin type continuous martingales with Brownian marginals and scaling

This chapter is based on a joint article with C. Donati-Martin and M. Yor (see [BDMY11]) Closely inspired by Albin's method which relies ultimately on the duplication formula for the Gamma function, we exploit Gauss' multiplication formula to construct a sequence of continuous martingales with Brownian marginals and scaling.

7.1 Motivation and main results

(1.1) Knowing the law of a "real world" random phenomena, i.e. random process, $(X_t, t \geq 0)$ is often extremely difficult and in most instances, one avails only of the knowledge of the 1-dimensional marginals of $(X_t, t \geq 0)$. However, there may be many different processes with the same given 1-dimensional marginals.

In the present paper, we make explicit a sequence of continuous martingales $(M_m(t), t \geq 0)$ indexed by $m \in \mathbb{N}$ such that for each m,

i) $(M_m(t), t \geq 0)$ enjoys the Brownian scaling property: for any $c > 0$,

$$(M_m(c^2 t), t \geq 0) \overset{(law)}{=} (cM_m(t), t \geq 0)$$

ii) $M_m(1)$ is standard Gaussian.

Note that, combining i) and ii), we get, for any $t > 0$

$$M_m(t) \overset{(law)}{=} B_t,$$

where $(B_t, t \geq 0)$ is a Brownian motion, i.e. M_m admits the same 1-dimensional marginals as Brownian motion.

(1.2) Our main result is the following extension of Albin's construction [Alb08] from $m = 1$ to any integer m.

Theorem 7.1.1. *Let $m \in \mathbb{N}$. Then, there exists a continuous martingale $(M_m(t), t \geq 0)$ which enjoys i) and ii) and is defined as follows:*

$$M_m(t) = X_t^{(1)} \dots X_t^{(m+1)} Z_m \tag{7.1}$$

where $(X_t^{(i)}, t \geq 0)$, for $i = 1, \dots, m+1$, are independent copies of the solution of the SDE

$$dX_t = \frac{1}{m+1} \frac{dB_t}{X_t^m}; \quad X_0 = 0 \tag{7.2}$$

and, furthermore, Z_m is independent from $(X^{(1)}, \dots, X^{(m+1)})$ and

$$Z_m \overset{(law)}{=} (m+1)^{1/2} \left(\prod_{j=0}^{m-1} \beta(\frac{1+2j}{2(m+1)}, \frac{m-j}{m+1}) \right)^{\frac{1}{2(m+1)}} \tag{7.3}$$

where $\beta(a,b)$ denotes a beta variable with parameter (a,b) with density

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} 1_{[0,1]}(x)$$

and the beta variables on the right-hand side of (7.3) are independent.

Remark: For $m = 1$, $Z_1 = \sqrt{2} \left(\beta(\frac{1}{4}, \frac{1}{2})\right)^{1/4}$ and we recover the distribution of $Y := Z_1$ given by (2) in [Alb08].

(1.3) For the convenience of the reader, we also recall that, if one drops the continuity assumption when searching for martingales $(M(t); t \geq 0)$ satisfying i) and ii), then, the Madan-Yor construction [MY02] based on the "Azéma-Yor under scaling" method provides such a martingale.
Precisely, starting from a Brownian motion $(B_u, u \geq 0)$ and denoting $S_u = \sup_{s \leq u} B_s$, introduce the family of stopping times

$$\tau_t = \inf\{u, S_u \geq \psi_t(B_u)\}$$

where ψ_t denotes the Hardy-Littlewood function associated with the centered Gaussian distribution μ_t with variance t, i.e.

$$\begin{aligned} \psi_t(x) &= \frac{1}{\mu_t([x,\infty[)} \int_x^\infty y \exp(-\frac{y^2}{2t}) \frac{dy}{\sqrt{2\pi t}} \\ &= \sqrt{t} \exp(-\frac{x^2}{2t}) / \mathcal{N}(x/\sqrt{t}) \end{aligned}$$

where $\mathcal{N}(a) = \int_a^\infty \exp(-\frac{y^2}{2}) dy$. Then, $M_t = B_{\tau_t}$ is a martingale with Brownian marginals.
Another solution has been given by Hamza and Klebaner [HK07].

(1.4) In section 3, we prove that Theorem 7.1.1 is actually the best we can do in our generalisation of Albin's construction: we cannot generalize (7.1) by allowing the $X^{(i)}$'s to be solution of (7.2) associated to different m_i's.
Finally, we study the asymptotic behavior of $X_t^{(1)} \dots X_t^{(m+1)}$ as $m \longrightarrow \infty$.

7.2 Proof of the theorem

Step 1: For $m \in \mathbb{R}$ and $c \in \mathbb{R}$, we consider the stochastic equation:

$$dX_t = c\frac{dB_t}{X_t^m}, \quad X_0 = 0.$$

This equation has a unique weak solution which can be defined as a time-changed Brownian motion

$$(X_t) \overset{(law)}{=} W(\alpha^{(-1)}(t))$$

where W is a Brownian motion starting from 0 and $\alpha^{(-1)}$ is the (continuous) inverse of the increasing process

$$\alpha(t) = \frac{1}{c^2}\int_0^t W_u^{2m} du.$$

We look for $k \in \mathbb{N}$ and c such that $(X_t^{2k}, t \geq 0)$ is a squared Bessel process of some dimension d. It turns out, by application of Itô's formula, that we need to take $k = m+1$ and $c = \frac{1}{m+1}$. Thus, we find that $(X_t^{2(m+1)}, t \geq 0)$ is a squared Bessel process with dimension $d = k(2k-1)c^2 = \frac{2m+1}{m+1}$.
Note that the law of a BESQ(d) process at time 1 is well known to be that of $2\gamma_{d/2}$, where γ_a denotes a gamma variable with parameter a. Thus, we have:

$$|X_1| \overset{(law)}{=} \left(2\gamma_{\frac{2m+1}{2(m+1)}}\right)^{\frac{1}{2(m+1)}} \tag{7.4}$$

Step 2: We now discuss the scaling property of the solution of (7.2). From the scaling property of Brownian motion, it is easily shown that , for any $\lambda > 0$, we get:

$$(X_{\lambda t}, t \geq 0) \overset{(law)}{=} (\lambda^{\alpha} X_t, t \geq 0)$$

with $\alpha = \frac{1}{2(m+1)}$, that is, the process $(X_t, t \geq 0)$ enjoys the scaling property of order $\frac{1}{2(m+1)}$.

Step 3: Consequently, if we multiply $m+1$ independent copies of the process $(X_t, t \geq 0)$ solution of (7.2), we get a process

$$Y_t = X_t^{(1)} \dots X_t^{(m+1)}$$

which is a martingale and has the scaling property of order $\frac{1}{2}$.

Step 4: Finally, it suffices to find a random variable Z_m independent of the processes $X_t^{(1)}, \dots, X_t^{(m+1)}$ and which satisfies:

$$N \overset{(law)}{=} X_1^{(1)} \dots X_1^{(m+1)} Z_m \tag{7.5}$$

where N denotes a standard Gaussian variable. Note that the distribution of any of the $X_1^{(i)}$'s is symmetric. We shall take $Z_m \geq 0$; thus, the distribution of Z_m shall be determined by its Mellin transform $\mathcal{M}(s) = \mathbb{E}(Z_m^s)$. From (7.5), $\mathcal{M}(s)$ satisfies:

$$\mathbb{E}[(2\gamma_{1/2})^{s/2}] = \left(\mathbb{E}[(2\gamma_{d/2})^{s/2(m+1)}]\right)^{m+1} \mathcal{M}(s)$$

with $d = \frac{2m+1}{m+1}$, that is:

$$2^{s/2} \frac{\Gamma(\frac{1+s}{2})}{\Gamma(\frac{1}{2})} = 2^{s/2} \left(\frac{\Gamma(\frac{d}{2} + \frac{s}{2(m+1)})}{\Gamma(\frac{d}{2})} \right)^{m+1} \mathcal{M}(s)$$

that is precisely:

$$\frac{\Gamma(\frac{1+s}{2})}{\Gamma(\frac{1}{2})} = \left(\frac{\Gamma(\frac{2m+1+s}{2(m+1)})}{\Gamma(\frac{2m+1}{2(m+1)})} \right)^{m+1} \mathcal{M}(s). \tag{7.6}$$

Now, we recall Gauss multiplication formula ([AAR99], see also [CY03])

$$\Gamma(kz) = \frac{k^{kz-1/2}}{(2\pi)^{\frac{k-1}{2}}} \prod_{j=0}^{k-1} \Gamma(z + \frac{j}{k}) \tag{7.7}$$

which we apply with $k = m+1$ and $z = \frac{1+s}{2(m+1)}$. We then obtain, from (7.7)

$$\frac{\Gamma(\frac{1+s}{2})}{\sqrt{\pi}} = \frac{(m+1)^{s/2}}{(2\pi)^{m/2}} \frac{1}{\sqrt{\pi}} \prod_{j=0}^{m} \Gamma(\frac{1+s+2j}{2(m+1)}) \tag{7.8}$$

$$= (m+1)^{s/2} \prod_{j=0}^{m} \left(\frac{\Gamma(\frac{1+s+2j}{2(m+1)})}{\Gamma(\frac{1+2j}{2(m+1)})} \right) \tag{7.9}$$

since the two sides of (7.8) are equal to 1 for $s = 0$. We now plug (7.9) into (7.6) and obtain

$$(m+1)^{s/2} \prod_{j=0}^{m} \left(\frac{\Gamma(\frac{1+s+2j}{2(m+1)})}{\Gamma(\frac{1+2j}{2(m+1)})} \right) = \left(\frac{\Gamma(\frac{2m+1+s}{2(m+1)})}{\Gamma(\frac{2m+1}{2(m+1)})} \right)^{m+1} \mathcal{M}(s) \tag{7.10}$$

We note that for $j = m$, the same term appears on both sides of (7.10), thus (7.10) may be written as:

$$(m+1)^{s/2} \prod_{j=0}^{m-1} \left(\frac{\Gamma(\frac{1+s+2j}{2(m+1)})}{\Gamma(\frac{1+2j}{2(m+1)})} \right) = \left(\frac{\Gamma(\frac{2m+1+s}{2(m+1)})}{\Gamma(\frac{2m+1}{2(m+1)})} \right)^{m} \mathcal{M}(s) \tag{7.11}$$

In terms of independent gamma variables, the left-hand side of (7.11) equals:

$$(m+1)^{s/2}\mathbb{E}\left[\left(\prod_{j=0}^{m-1}\gamma^{(j)}_{\frac{1+2j}{2(m+1)}}\right)^{\frac{s}{2(m+1)}}\right] \tag{7.12}$$

whereas the right-hand side of (7.11) equals:

$$\mathbb{E}\left[\left(\prod_{j=0}^{m-1}\gamma^{(j)}_{\frac{1+2m}{2(m+1)}}\right)^{\frac{s}{2(m+1)}}\right]\mathcal{M}(s) \tag{7.13}$$

where the $\gamma^{(j)}_{a_j}$ denote independent gamma variables with respective parameters a_j.
Now, from the beta-gamma algebra, we deduce, for any $j \leq m-1$:

$$\gamma^{(j)}_{\frac{1+2j}{2(m+1)}} \overset{(law)}{=} \gamma^{(j)}_{\frac{1+2m}{2(m+1)}}\beta(\frac{1+2j}{2(m+1)},\frac{m-j}{m+1}).$$

Thus, we obtain, again by comparing (7.12) and (7.13):

$$\mathcal{M}(s) = (m+1)^{s/2}\,\mathbb{E}\left[\left(\prod_{j=0}^{m-1}\beta(\frac{1+2j}{2(m+1)},\frac{m-j}{m+1})\right)^{\frac{s}{2(m+1)}}\right]$$

which entails:

$$\mathbb{E}[Z_m^s] = (m+1)^{s/2}\,\mathbb{E}\left[\left(\prod_{j=0}^{m-1}\beta(\frac{1+2j}{2(m+1)},\frac{m-j}{m+1})\right)^{\frac{s}{2(m+1)}}\right]$$

that is, equivalently,

$$Z_m \overset{(law)}{=} (m+1)^{1/2}\left(\prod_{j=0}^{m-1}\beta(\frac{1+2j}{2(m+1)},\frac{m-j}{m+1})\right)^{\frac{1}{2(m+1)}}$$

7.3 Some remarks about Theorem 7.1.1

7.3.1 A further extension

We tried to extend Theorem 7.1.1 by taking a product of independent martingales $X^{(i)}$, solution of (7.2) with different m_i's. Here are the details of our attempt.
We are looking for the existence of a variable Z such that the martingale

$$M(t) = \left(\prod_{j=0}^{p-1}X_t^{(m_j)}\right)Z$$

satisfies the properties i) and ii). Here $p, (m_j)_{0\leq j\leq p-1}$ are integers and $X^{(m_j)}$ is the solution of the EDS (7.2) associated to m_j, the martingales being independent for j varying. In order that M enjoys the Brownian scaling property, we need the following relation

$$\sum_{j=0}^{p-1} \frac{1}{m_j+1} = 1. \tag{7.14}$$

Following the previous computations, see (7.6), the Mellin transform $\mathcal{M}(s)$ of Z should satisfy

$$\frac{\Gamma(\frac{1+s}{2})}{\Gamma(\frac{1}{2})} = \left(\prod_{j=0}^{p-1} \frac{\Gamma(\frac{2m_j+1+s}{2(m_j+1)})}{\Gamma(\frac{2m_j+1}{2(m_j+1)})}\right) \mathcal{M}(s). \tag{7.15}$$

We recall (see (7.9)) the Gauss multiplication formula

$$\frac{\Gamma(\frac{1+s}{2})}{\sqrt{\pi}} = p^{s/2} \prod_{j=0}^{p-1} \left(\frac{\Gamma(\frac{1+s+2j}{2p})}{\Gamma(\frac{1+2j}{2p})}\right) \tag{7.16}$$

To find $\mathcal{M}(s)$ from (7.15), (7.16), we give some probabilistic interpretation:

$$\frac{\Gamma(\frac{1+s+2j}{2p})}{\Gamma(\frac{1+2j}{2p})} = \mathbb{E}[\gamma_{(1+2j)/2p}^{s/2p}]$$

whereas

$$\frac{\Gamma(\frac{2m_j+1+s}{2(m_j+1)})}{\Gamma(\frac{2m_j+1}{2(m_j+1)})} = \mathbb{E}[\gamma_{(1+2m_j)/2(m_j+1)}^{s/2(m_j+1)}].$$

Thus, we would like to factorize

$$\gamma_{(1+2j)/2p}^{1/2p} \overset{(law)}{=} \gamma_{(1+2m_j)/2(m_j+1)}^{1/2(m_j+1)} z_{m_j,p}^{(j)} \tag{7.17}$$

for some variable $z_{m_j,p}^{(j)}$ to conclude that

$$Z = p^{1/2} \prod_{j=0}^{p-1} z_{m_j,p}^{(j)}.$$

It remains to find under which condition the identity (7.17) may be fulfilled. We write

$$\gamma_{(1+2j)/2p} \overset{(law)}{=} \gamma_{(1+2m_j)/2(m_j+1)}^{p/(m_j+1)} (z_{m_j,p}^{(j)})^{2p}. \tag{7.18}$$

Now, if $\frac{1+2j}{2p} < \frac{1+2m_j}{2(m_j+1)}$, we may apply the beta-gamma algebra to obtain

$$\gamma_{(1+2j)/2p} \overset{(law)}{=} \gamma_{(1+2m_j)/2(m_j+1)} \beta\left(\frac{1+2j}{2p}, \frac{1+2m_j}{2(m_j+1)} - \frac{1+2j}{2p}\right)$$

but in (7.18), we need to have on the right-hand side $\gamma_{(1+2m_j)/2(m_j+1)}^{p/(m_j+1)}$ instead of $\gamma_{(1+2m_j)/2(m_j+1)}$.
However, it is known that

$$\gamma_a \overset{(law)}{=} \gamma_a^c \gamma_{a,c}$$

for some variable $\gamma_{a,c}$ independent of γ_a for any $c \in (0,1]$. This follows from the self-decomposable character of $\ln(\gamma_a)$. Thus, we seem to need $\frac{p}{m_j+1} \leq 1$. But, this condition is not compatible with (7.14) unless $m_j = m = p-1$.

7.3.2 Asymptotic study

We study the behavior of the product $X_1^{(1)} \dots X_1^{(m+1)}$, resp. Z_m, appearing in the right-hand side of the equality in law (7.5), when $m \longrightarrow \infty$. Recall from (7.4) that

$$|X_1| \stackrel{(law)}{=} \left(2\gamma_{\frac{2m+1}{2(m+1)}}\right)^{\frac{1}{2(m+1)}}.$$

We are thus led to consider the product

$$\Theta_{a,b,c}^{(p)} = \left(\prod_{i=1}^{p} \gamma_{a-b/p}^{(i)}\right)^{c/p}$$

where in our set up of Theorem 7.1.1, $p = m+1$, $a = 1$, $b = c = 1/2$.

$$\begin{aligned}
\mathbb{E}[(\Theta_{a,b,c}^{(p)})^s] &= \prod_{i=1}^{p} \mathbb{E}[\left(\gamma_{a-b/p}^{(i)}\right)^{cs/p}] \\
&= \left(\frac{\Gamma(a-\frac{b}{p}+\frac{cs}{p})}{\Gamma(a-\frac{b}{p})}\right)^{p} \\
&= \exp[p(\ln(\Gamma(a+\frac{cs-b}{p})) - \ln(\Gamma(a-\frac{b}{p})))] \\
&\underset{p\to\infty}{\longrightarrow} \exp(\frac{\Gamma'(a)}{\Gamma(a)}cs).
\end{aligned}$$

Thus, it follows that

$$\Theta_{a,b,c}^{(p)} \xrightarrow[p\to\infty]{\mathbb{P}} \exp(\frac{\Gamma'(a)}{\Gamma(a)}c),$$

implying that

$$|X_1^{(1)} \dots X_1^{(m+1)}| \xrightarrow[m\to\infty]{\mathbb{P}} \exp(-\gamma/2) \tag{7.19}$$

and

$$\exp(-\gamma/2) Z_m \xrightarrow[m\to\infty]{(law)} |N|. \tag{7.20}$$

where $\gamma = -\Gamma'(1)$ is the Euler constant.

We now look for a central limit theorem for $\Theta_{a,b,c}^{(p)}$. We consider the limiting distribution of

$$\sqrt{p}\left\{\frac{c}{p}\sum_{i=1}^{p}\ln(\gamma_{a-b/p}^{(i)}) - c\frac{\Gamma'(a)}{\Gamma(a)}\right\}.$$

$$
\begin{aligned}
&\mathbb{E}\left(\exp\left[cs\sqrt{p}\left\{\frac{1}{p}\sum_{i=1}^{p}\ln(\gamma_{a-b/p}^{(i)})-\frac{\Gamma'(a)}{\Gamma(a)}\right\}\right]\right)\\
&\quad=\mathbb{E}\left[\prod_{i=1}^{p}\left(\gamma_{a-b/p}^{(i)}\right)^{cs/\sqrt{p}}\right]\exp(-cs\sqrt{p}\frac{\Gamma'(a)}{\Gamma(a)})\\
&\quad=\mathbb{E}\left[\left(\gamma_{a-b/p}^{(i)}\right)^{cs/\sqrt{p}}\right]^{p}\exp(-cs\sqrt{p}\frac{\Gamma'(a)}{\Gamma(a)})\\
&\quad=\left(\frac{\Gamma(a-\frac{b}{p}+\frac{cs}{\sqrt{p}})}{\Gamma(a-\frac{b}{p})}\right)^{p}\exp(-cs\sqrt{p}\frac{\Gamma'(a)}{\Gamma(a)})\\
&\quad=\exp[p(\ln(\Gamma(a-\frac{b}{p}+\frac{cs}{\sqrt{p}}))-\ln(\Gamma(a-\frac{b}{p})))-cs\sqrt{p}\frac{\Gamma'(a)}{\Gamma(a)}]\\
&\quad=\exp(\frac{c^2s^2}{2}(\ln(\Gamma))''(a)+O(m^{-1/2}))
\end{aligned}
$$

We thus obtain that

$$
\sqrt{p}\left\{\frac{c}{m}\sum_{i=1}^{m}\ln(\gamma_{a-b/m}^{(i)})-c\frac{\Gamma'(a)}{\Gamma(a)}\right\}\overset{(law)}{\longrightarrow}N(0,\sigma^2) \tag{7.21}
$$

where $N(0,\sigma^2)$ denotes a centered Gaussian variable with variance:

$$
\sigma^2=c^2(\ln(\Gamma))''(a)=c^2\left[\frac{\Gamma''(a)}{\Gamma(a)}-\left(\frac{\Gamma'(a)}{\Gamma(a)}\right)^2\right].
$$

or, equivalently

$$
\left(\Theta_{a,b,c}^{(p)}\exp(\frac{\Gamma'(a)}{\Gamma(a)}c)\right)^{\sqrt{p}}\underset{p\to\infty}{\overset{(law)}{\longrightarrow}}\exp(N(0,c^2(\ln(\Gamma))''(a))). \tag{7.22}
$$

List of Algorithms

List of Figures

Bibliography

[AAR99] G.E. Andrews, R. Askey, and R. Roy. Special Functions, volume 71 of Encyclopedia of Mathematics and its Applications, 1999.

[Alb08] JMP Albin. A continuous non-Brownian motion martingale with Brownian motion marginal distributions. *Statistics & Probability Letters*, 78(6):682–686, 2008.

[ASL98] Y. Aït-Sahalia and A.W. Lo. Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *The Journal of Finance*, 53(2):499–547, 1998.

[AY79] J. Azéma and M. Yor. Une solution simple au probleme de Skorokhod. *Séminaire de Probabilités XIII*, pages 90–115, 1979.

[Bak12] D. Baker. A Dubins type solution to the Skorokhod embedding problem. *Statistics & Probability Letters*, 82(6):1054–1058, 2012.

[BDMY11] D. Baker, C. Donati-Martin, and M. Yor. A sequence of Albin type continuous martingales with Brownian marginals and scaling. *Séminaire de Probabilités XLIII*, pages 441–449, 2011.

[BHR01] H. Brown, D. Hobson, and LCG Rogers. The maximum maximum of a martingale constrained by an intermediate law. *Probability Theory and Related Fields*, 119(4):558–578, 2001.

[BL78] D.T. Breeden and R.H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651, 1978.

[BS73] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political economy, Vol 81, No 3*, 1973.

[BY08] D. Baker and M. Yor. A proof of the increase in maturity of the expectation of a convex function of the arithmetic average of geometric Brownian motion. *Preprint, Sep. 2008*, 2008.

[BY09] D. Baker and M. Yor. A Brownian sheet martingale with the same marginals as the arithmetic average of geometric Brownian motion. *Electronic Journal of Probability*, 14:1532–1540, 2009.

[BY11] D. Baker and M. Yor. On Martingales with Given Marginals and the Scaling Property. *Séminaire de Probabilités XLIII*, pages 437–439, 2011.

[CEX08] P. Carr, C.-0. Ewald, and Y. Xiao. On the qualitative effect of volatility and duration on prices of Asian Options. *Finance Research Letters, Vol. 5, Issue 3, p. 162-171 ,Sep. 2008*, 2008.

[CH07] A. Cox and D. Hobson. A unifying class of Skorokhod embeddings: connecting the Azéma–Yor and Vallois embeddings. *Bernoulli*, 13(1):114–130, 2007.

[CL83] N.N. Chan and K.H. Li. Diagonal elements and eigenvalues of a real symmetric matrix. *Journal of Mathematical Analysis and Applications*, 91(2):562–566, 1983.

[CS04] P. Carr and M. Schroder. Bessel processes, the integral of geometric Brownian motion, and asian options. *Theory probab. appl., Vol 48, No 3*, 2004.

[CW76] R. Chacon and J. Walsh. One-dimensional potential embedding. *Séminaire de Probabilités X Université de Strasbourg*, pages 19–23, 1976.

[CY03] L. Chaumont and M. Yor. *Exercises in Probability: a guided tour from measure theory to random processes, via conditioning*. Cambridge Univ Pr, 2003.

[Dan49] G.B. Dantzig. Programming of interdependent activities: II mathematical model. *Econometrica, Journal of the Econometric Society*, pages 200–211, 1949.

[Dan98] G.B. Dantzig. *Linear programming and extensions*. Princeton Univ Pr, 1998.

[DMMY00] C. Donati-Martin, H. Matsumoto, and M. Yor. The law of Geometric Brownian motion and its integral revisited: Application to conditional moments. *Proceeding of the Bachelier Congress*, 2000.

[DS95] F. Delbaen and W. Schachermayer. Arbitrage possibilities in Bessel processes and their relation to local martingales. *Prob. Theory Relat. Fields 102, 357-366*, 1995.

[Dub68] L.E. Dubins. On a theorem of Skorohod. *The Annals of Mathematical Statistics*, 39(6):2094–2097, 1968.

[Dup94] B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[HK07] K. Hamza and F.C. Klebaner. A family of non-Gaussian martingales with Gaussian marginals. *Journal of Applied Mathematics and Stochastic Analysis*, 2007(1):92723, 2007.

[Hob98] D. Hobson. The maximum maximum of a martingale. *Séminaire de Probabilités XXXII*, pages 250–263, 1998.

[Hob11] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. *Paris-Princeton Lectures on Mathematical Finance 2010*, pages 267–318, 2011.

[Hor54] A. Horn. Doubly stochastic matrices and the diagonal of a rotation matrix. *American Journal of Mathematics*, pages 620–630, 1954.

[HP02] D.G. Hobson and J.L. Pedersen. The minimum maximum of a continuous martingale with given initial and terminal laws. *Annals of probability*, pages 978–999, 2002.

[HPRY11] F. Hirsch, C. Profeta, B. Roynette, and M. Yor. *Peacocks and associated martingales, with explicit constructions*, volume 3. Bocconi-Springer, 2011.

[HR12] F. Hirsch and B. Roynette. A new proof of Kellerer's theorem. *ESAIM: Probability and Statistics, to appear*, 2012.

[HY09] F. Hirsch and M. Yor. A construction of processes with one dimensional martingale marginals, based upon path-space Ornstein-Uhlenbeck processes and the Brownian sheet. *Kyoto Journal of Mathematics*, 49(2):389–417, 2009.

[JPY02] M. Jeanblanc, J. Pitman, and M. Yor. Self-similar processes with independent increments associated with Lévy and Bessel processes. *Stoch. Proc. and their App., Vol. 100, p. 223-232*, 2002.

[JR96] J.C. Jackwerth and M. Rubinstein. Recovering probability distributions from option prices. *Journal of Finance*, pages 1611–1631, 1996.

[Kel72] H.G. Kellerer. Markov-Komposition und eine Anwendung auf Martingale. *Mathematische Annalen*, 198(3):99–122, 1972.

[Lev92] H. Levy. Stochastic dominance and expected utility: survey and analysis. *Management Science, Vol 38, No 4*, 1992.

[Llo82] S. Lloyd. Least squares quantization in PCM. *Information Theory, IEEE Transactions on*, 28(2):129–137, 1982.

[Low08] G. Lowther. Fitting martingales to given marginals. *Arxiv preprint arXiv:0808.2319*, 2008.

[Mei82] I. Meilijson. There exists no ultimate solution to Skorokhod's problem. *Séminaire de Probabilités XVI 1980/81*, pages 392–399, 1982.

[MY02] D.B. Madan and M. Yor. Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536, 2002.

[Obł04] J. Obłój. The Skorokhod embedding problem and its offspring. *Probability Surveys*, 1:321–392, 2004.

[Pha12] H. Pham. Methodes de quantification optimales et Applications en Finance. http://felix.proba.jussieu.fr/pageperso/pham/coursquantif.pdf, 2012.

[PP08] S. Pal and P. Protter. Strict local martingales, bubbles, and no early exercise. *Preprint, Nov. 2007*, 2008.

[RS70] M. Rothschild and J.E. Stiglitz. Increasing risk: I. A definition. *Journal of Economic theory*, 2(3):225–243, 1970.

[RY94] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, 1994.

[Sch23] I. Schur. Uber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, 22:9–20, 1923.

[Shr00] S. Shreve. *Stochastic Calculus for Finance II*. Springer, 2000.

[Sko65] A.V. Skorokhod. *Studies in the Theory of Random Processes*. Addison-Wesley Educational Publishers Inc, 1965.

[SS94] M. Shaked and J. Shanthikumar. *Stochastic orders with applications*. Academic press, 1994.

[SS06] M. Shaked and J. Shanthikumar. *Stochastic orders*. Springer Statistics, 2006.

[SST94] M. Shaked, J. Shanthikumar, and Y. Tong. *Stochastic orders and their applications*. Academic press New York, 1994.

[ST94] G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman-Hall, 1994.

[Val83] P. Vallois. Le probleme de Skorokhod sur IR: une approche avec le temps local. *Séminaire de Probabilités XVII 1981/82*, pages 227–239, 1983.

[WD49] M.K. Wood and G.B. Dantzig. Programming of interdependent activities: I general discussion. *Econometrica, Journal of the Econometric Society*, pages 193–199, 1949.

[Wei12] E. Weisstein. Line-Line Intersection. http://mathworld.wolfram.com/Line-LineIntersection.html, 2012.

[Yor92] M. Yor. On some Exponential Functional of Brownian Motion. *Advances in applied probability, Vol 24, No 3*, 1992.