University of Alberta

Comparison Theorem and its applications to Finance

by

Vladislav Krasin

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Mathematical finance

Department of Mathematical and Statistical Sciences

Fall 2010

Edmonton, Alberta

Examining Committee

B. Schmuland, Department of Mathematical and Statistical Sciences, University of Alberta

A. Melnikov, Department of Mathematical and Statistical Sciences, University of Alberta

A. Cadenillas, Department of Mathematical and Statistical Sciences, University of Alberta

T. Choulli, Department of Mathematical and Statistical Sciences, University of Alberta

C. Szepesvári, Department of Computing Science, University of Alberta

A. Swishchuk, Department of Mathematics and Statistics, University of Calgary

Abstract

The current Thesis is devoted to comprehensive studies of comparison, or stochastic domination, theorems. It represents a combination of theoretical research and practical ideas formulated in several specific examples.

Previously known results and their place it the theory of stochastic processes and stochastic differential equations is reviewed. This part of the work yielded three new theoretical results, formulated as theorems. Two of them are extensions of commonly used methods to more sophisticated processes and conditions. The third theorem is proven using previously not exploited technique. The place of all three results in the global theory is demonstrated by examining interconnections and possible distinctions between old and new theorems.

Second and equally important part of the work focuses on more practical issues. Its main goal is to demonstrate where and how various theoretical findings can be applied to typical financial problems, such as option pricing, hedging, risk management and others. The example chapter summarizes the best of the obtained results in this direction.

Acknowledgements

I would like to express deepest gratitude to my supervisor Dr. Alexander Melnikov. This project would never have been completed without his guidance, advice and help.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Stochastic processes in mathematical finance

The idea of describing various asset prices with the help of the theory of stochastic processes turned out to be very useful and yielded many quality applications. It can be argued that one of the main achievements of the modern financial mathematics are mechanisms for pricing and hedging of options. The celebrated Black Scholes model [6] provided a relatively simple and intuitive way of calculating prices of various derivative securities as well as a theoretical replicating strategy that could be used for risk management purposes.

Many researchers, however, agree that the above mentioned Black Scholes model is not accurate enough for the modern financial world. A log-normally distributed process is just too simple to fit into the data and is therefore not reliable enough when dealing with applications. Fischer Black in [5] summarized many undesirable features of the Black Scholes model when used in practical applications.

Various steps can be undertaken to improve quality of financial market models. One can consider Brownian motion-driven processes with non-constant volatility, be that a deterministic function of the stock price or a random process, possibly modelled via a stochastic differential equation (SDE). Probably the best known model of the first type was proposed by Cox and Ross in [9]. Merton in [38] suggested adding Poissonian jumps to the log-normal process, one of the first attempts to explain observable differences in European options' implied volatilities. Hull and White [23] as well as Heston [21] advocate a stochastic volatility approach - that is modelling stock price volatility with the help of a separate process with new Brownian motion as its source or randomness. Hobson and Rogers in [22] use yet another approach, which is based on path-dependent volatility parameter. Their work is interesting because, while demonstrating many features of stochastic volatility models, the proposed market remains complete.

All above mentioned works can be placed in a category of continuous time financial market models with asset price dynamics described with the help of stochastic differential equations. An unfortunate consequence of all those modifications is that the model will no longer admit explicit solution. Several problems arise from that fact. First, one has to be sure that the equation in question has a solution, preferably unique, before accepting the model. Existence and uniqueness of solution is one of the topics I study for that very reason.

Absence of explicit solutions results in other complications when dealing with stochastic processes in practice. Not only does it prevent one from performing exact calculations (which is one the main accomplishments of mathematical finance), but also reduces the amount of knowledge one has about many theoretical properties of the processes in question. While such simple things as sample path continuity are not difficult to establish other, more complex properties, might be of great interest in theory and practice and are not as obvious.

Similar questions can be traced back to the theory of (deterministic) differential equations: even simple-looking equations do not always admit explicit solutions. Keeping in mind close connections between two areas of mathematics, impossibility of obtaining explicit solutions of stochastic differential equations comes as no surprise.

However, theory of ordinary and partial differential equations has methods of establishing certain properties of solutions by examining the coefficients only. Existence and uniqueness of solution is, of course, one of them. Another question that can be studied is monotonicity of solution with respect to coefficients: how does increasing (or decreasing) coefficients of differential equations affect their solutions? It is possible to compare the values of deterministic processes by comparing their derivatives.

More specifically, considering two functions

$$\begin{aligned} dx(t) &= f_1(t, x(t))dt \\ dy(t) &= f_2(t, y(t))dt, \end{aligned}$$

the conditions (provided both equations admit a solution)

$$f_1(t, z) \leq f_2(t, z)$$

and $x(0) \leq y(0)$ ensures that $x(t) \leq y(t)$ for all t. However, extensions of this and other similar results to stochastic differential equations are not obvious.

Comparison of stochastic processes is a second part of my theoretical work. The topic represents a theoretical challenge and requires a lot of attention, because there is more than one dimension to the problem. First, comparison of random variables can be performed in several ways: it could be value related inequalities, or comparisons in mean where expected values are studied. Second aspect that distinguishes stochastic and ordinary differential equations is presence of different type of integrals. The simplest SDEs have a Brownian motion component, not to mention more general martingales with discontinuous parts. There are strong connections between the two areas of my interest, as both types of theorems use similar conditions and methods. All that acts as a further motivation to study both areas together.

At the same time the main motivation for studying comparison theorems is possible usages of these type of results in practical applications, namely mathematical finance. Better knowledge of theoretical properties of market models should be an asset when performing calculations and making other conclusions. Rather surprisingly, this idea does not seem to be used in literature, the fact that acts as further motivation behind my research.

1.2 Necessary notations

All considerations in this work are performed on a standard probability basis $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$, as defined, for example, in [26]. Filtration $\mathcal{F}_t$ is right-continuous and completed with all $\mathbf{P}$-null sets. All processes studied take values in $\mathbf{R}^d$ and are assumed to be càdlàg and adapted to the filtration $\mathbf{F}$.

Several specific classes of stochastic processes are defined. Let $\mathcal{A}^+$ be the set of non-decreasing processes with integrable variation, $\mathcal{M}^2$ be the set of square integrable martingales. The localized classes are denoted by $\mathcal{A}^+_{loc}$ and $\mathcal{M}^2_{loc}$ respectively.

Everything in this thesis, be that its theoretical part or ideas for possible applications, is done for stochastic processes, represented as solutions of stochastic differential equations. Those equations will be built with the help of a continuous non-decreasing process $a \in \mathcal{A}^+_{loc}$, a continuous local martingale $M \in \mathcal{M}^2_{loc}$ and a jump measure μ with a continuous compensator ν. The measure ν is assumed to be almost surely finite for every t, meaning that

$$\nu(t, \mathbf{R}^d \backslash \{0^d\}) < \infty \text{ (a.s.)}$$

A standard notation $< M, N >$ will be used for quadratic characteristic.

A typical representation of an SDE to be used in this work is given by

$$\begin{aligned} dX_t &= f(t, X_{t-})da_t + g(t, X_{t-})dM_t + I_{\{|u|\le 1\}}h(t, X_{t-}, u)d(\mu - \nu)_{t,u} \quad (1) \\ X_0 &= x_0 \end{aligned}$$

where x_0 is an $\mathcal{F}_0$ measurable random variable with finite expectation and measurable functions $f = f(t, x, \omega)$, $g = g(t, x, \omega)$, and $h = h(t, x, u, \omega)$ are continuous in (t, x). A (strong) solution of SDE (1) is an adapted process X_t such that

$$X_t = X_0 + \int_0^t f(s, X_{s-})da_s +$$
$$+ \int_0^t g(s, X_{s-})dM_s + \int_0^t \int_{|u|\le 1} h(s, X_{s-}, u)d(\mu - \nu)_{s,u}.$$

Since weak solutions are not studied in my work the word strong will be omitted henceforth.

To make sure the main point is not lost behind many technicalities, several steps will be undertaken to simplify proceedings. First, the time argument will be avoided whenever no confusion is possible and coefficients will be expresses as $f(x)$ (the same applies to functions g and h). Second step is to avoid writing double integrals when dealing with the jump component. Strictly speaking, integrals with respect to the jump measure μ and its compensator ν are two-dimensional integrals with time t and jump amplitude u as variables. In his two works Gal'chuk demonstrates how big jumps can be dealt with (lemma 3 of [13] for existence and uniqueness framework and lemma 2 of [14] for comparison framework). For this very reason equation (1) does not have a second jump component for magnitudes $|u| > 1$, and I will use single integral notations when dealing with random measures μ and ν.

The last adjustment concerns differences between martingales and local martingales. Strictly speaking, integrals with respect to local martingale M and compensated jump component $\mu - \nu$ are local martingales, but not necessarily true martingales. One can, however, cite a standard in the theory of stochastic processes localization procedure, described in [26] among others. It allows to assume, without any additional restrictions, that the above mentioned integral are indeed martingales. One can then safely take expectations of those integrals and use their martingale properties. This assumption is very common in literature and will be made throughout the thesis.

1.3 Contributions of this thesis

The organization of the thesis is as follows. Chapter 2 deals with existence and uniqueness of solutions. It starts with a review of previously established results in this area. Chapter 2.2 is devoted to the proof of the existence and uniqueness theorem for general semimartingale-driven processes. Another contribution of this theorem is to relax the condition on the drift coefficient f. It is followed by studies of relationships between this result and other works in the given framework.

Chapter 3 is devoted to comparison theorems and is organized in similar manner. It starts with a review of existing results, which is followed by two new theorems. One, presented in chapter 3.2 is based on a new approach to establishing path-wise comparisons of stochastic processes. The new method not only simplifies the proof, but also produces a more general result, as will be demonstrated in chapter 3.4. Another result, presented in chapter 3.3 is for comparisons of multidimensional processes and was published in [30].

The final part of my work is devoted to financial applications of comparison theorems to mathematical finance. Several numerical and theoretical examples are developed, some of which were also presented in [30]. They are designed to demonstrate the full range of usages of comparison theorems, as well as to try and establish general techniques that could be used in multiple settings.

2 Existence and uniqueness of solution[1]

2.1 Results overview

When considering an equation of the type (1), the main question of this chapter is: under what conditions on functions f, g and h does it have a unique solution.

One of the first efforts to establish existence and uniqueness of solutions of stochastic differential equations was performed by Kazamaki [28]. He showed that equation

$$dX_t = X_0 + \int_0^t f(X_{s-})da_s + \int_0^t g(X_{s-})dM_s,$$

where M_t is not necessarily continuous, admits a unique solution if $f(x)$ and $g(x)$ are differentiable with bounded derivatives.

The differentiability condition, though, can be relaxed. A standard way to ensure existence and uniqueness of solution is to require that all coefficients are Lipschitz-continuous. The work [12] is conducted under Lipschitz conditions:

$$|f(x) - f(y)| + |g(x) - g(y)| \leq C|x - y|,$$

for not necessarily continuous M and a.

The approach is to define an operator

$$QX_t = X_0 + \int_0^t f(X_{s-})da_s + \int_0^t g(X_{s-})dM_s,$$

and considering a simple iteration procedure

[1]A version of this chapter has been submitted for publication in [32]

$$X_t^{(n)} = QX_t^{(n-1)}$$

perform several steps to verify that the sequence X^n will converge.

Similar result under the Lipschitz condition for continuous driving local martingale case is studied by Protter in [42]. He also studies the question of exploding solutions: an explosion time R is defined as the moment when the process in question hits infinity. Protter then provides certain conditions under which $R = \infty$ (a.s.).

Finally, perhaps the most general result for Lipschitz continuous coefficients is provided by Gal'chuk in [13]. He considers processes with general driving semimartingale

$$\begin{aligned} X_t = X_0 + \int_0^t f(s, X_{s-}) da_s + \int_0^t g(s, X_{s-}) dM_s + \\ + \int_0^t \int_{|u|\le 1} h(s, X_{s-}, u) d(\mu - \nu)_{u,s} + \int_0^t \int_{|u|>1} k(s, X_{s-}, u) d\mu_{u,s} \end{aligned} \tag{2}$$

and proves that

$$||QX|| \le \alpha ||X|| \tag{3}$$

for some $\alpha < 1$ and an appropriately defined norm. Fixed point theorem then yields the desired result. By isolating all discontinuities with the help of jump measure μ, the conditions on big jump coefficient k is relaxed, making it only measurable.

Several authors have attempted to weaken the Lipschitz condition in different ways. One approach was used by Jacod [25] and Gyöngy and Krylov [17]. The same iteration procedure (3) is used, but its convergence is proven by studying the squared difference

$$|X_t^{(n+k)} - X_t^{(n)}|^2$$

with the help of Ito's formula. The main condition of those two works can be expressed (for the one dimensional continuous diffusion case) as

$$2(x-y)(f(x)-f(y)) + (g(x)-g(y))^2 \leq K(x-y)^2,$$

where K is a non-negative constant, and will be referred to as *the monotonity condition.* A growth restriction of the type

$$2xf(x) + |g(x)|^2 \leq K_2(1+x^2)$$

is also imposed.

As can be seen, the monotonity condition connects all coefficients in one inequality limiting growth of a certain expression, rather than dealing with each coefficient separately, and is another improvement on the standard Lipschitz setting.

The paper [49] uses the following conditions to prove path-wise uniqueness of solution in the continuous diffusion setting: let there exist non-negative increasing functions $k(x)$ and $\rho(x)$ such that $k(x)$ is concave,

$$\int_{0+} k^{-1}(u)du = \infty$$
$$\int_{0+} \rho^{-2}(u)du = \infty$$

and

$$|f(x) - f(y)| \leq k(|x - y|)$$
$$|g(x) - g(y)| \leq \rho(|x - y|).$$

The difference in conditions follows from changing of the approach. This time a sequence of approximations is constructed using a Euler scheme. It starts by considering piece-wise constant (with respect to time) coefficients. Then letting the partition size go to zero one can obtain a limit which solves the original SDE. This technique will be used to prove an existence and uniqueness theorem below.

2.2 Existence and uniqueness theorem

This part of the thesis presents a proof for the existence and uniqueness theorem. It studies the case of one-dimensional processes with general driving semimartingale represented by three components: a continuous non-decreasing process, a continuous local martingale and a compensated jump component:

$$\begin{aligned} dX_t &= f(t, X_{t-})da_t + g(t, X_{t-})dM_t + h(t, X_{t-}, u)d(\mu - \nu)_{t,u} \\ X_0 &= x_0. \end{aligned} \tag{4}$$

The theorem represents an improvement on the previously established results in several areas, as will be demonstrated in chapter 2.3.

Before stating the main theorem of this part it is necessary to introduce an auxiliary sequence of functions $\varphi_m(x)$ approximating $|x|$. Idea of such a sequence is used in [14, 50] among others an is fairly standard when dealing with stochastic differential equations.

Consider a continuous non-decreasing function ρ with $\rho(0) = 0$ such that for any $\varepsilon > 0$

$$\int_0^{\varepsilon} \rho^{-2}(s)ds = \infty. \tag{5}$$

Define $\{b_m\}$ as a decreasing sequence of positive numbers with $b_0 = 1$,

$$\int_{b_{m+1}}^{b_m} \rho^{-2}(x)dx = m+1, m = 0, 1, 2...$$

and let $\psi_m(x)$ be a sequence of non-negative continuous functions such that $\psi_m(x) = 0$ for $x \notin (b_m, b_{m-1})$ and

$$\psi_m(x) \leq \frac{2}{m}\rho^{-2}(x)$$
$$\int_{b_m}^{b_{m-1}} \psi_m(s)ds = 1.$$

The function $\varphi_m(x)$ is then defined as

$$\varphi_m(x) = \int_0^{|x|} \int_0^s \psi_m(z)dzds.$$

It is easy to see that $|x| - \varphi_m(x) \leq b_{m-1}$ and $\varphi_m(x) \uparrow |x|$ as $m \to \infty$.

One more condition is necessary when dealing with discontinuous processes. Following [14] assume that there exists a sequence of positive numbers $\{\varepsilon_m\}$ such that $b_m \leq b_{m-1} - \varepsilon_m$, the function ψ_m attains its maximum at $b_{m-1} - \varepsilon_m$ and

$$m^{-1}\rho^2(b_{m-1})\rho^{-2}(b_{m-1} - \varepsilon_m) \to 0 \tag{6}$$

as $m \to \infty$.

This specific condition will come into play in the form of the following lemma.

Lemma 1. Let $\varphi_m(x)$ be defined by (5) and (6) and $h(x)$ be a non-decreasing continuous function such that

$$(h(x) - h(y))^2 \leq \rho^2(|x - y|).$$

Then for any m, $0 \leq \theta \leq 1$ and any x, y

$$\varphi_m''(\theta(h(x) - h(y)) + x - y)(h(x) - h(y))^2 \leq \frac{2}{m}\rho^{-2}(b_{m-1} - \varepsilon_m)\rho^2(b_{m-1}).$$

Proof. By construction φ_m is symmetric, therefore it is sufficient to present the proof for $x \geq y$.

Then two options are possible:

$$x - y + \theta(h(x) - h(y)) > b_{m-1}$$

or

$$x - y + \theta(h(x) - h(y)) \leq b_{m-1}.$$

In the first case

$$\varphi_m''(\theta(h(x) - h(y)) + x - y) = 0$$

by the definition of φ_m and the proof is complete.

In the second case, note that since $h(\cdot)$ is a non-decreasing function then $h(x) - h(y) \geq 0$, therefore

$$x - y \leq b_{m-1}$$

which in turn means that

$$(h(x)-h(y))^2 \leq \rho^2(x-y) \leq \rho^2(b_{m-1}).$$

At the same time by the definition of φ_m

$$\begin{aligned} \varphi_m''(\theta(h(x)-h(y))+x-y) \leq \\ \leq \varphi_m''(b_{m-1}-\varepsilon_m) \leq \frac{2}{m}\rho^{-2}(b_{m-1}-\varepsilon_m) \quad . \end{aligned}$$

Altogether, the combined estimate is

$$\begin{aligned} \varphi_m''(x-y+\theta(h(x)-h(y)))(h(x)-h(y))^2 \leq \\ \leq \frac{2}{m}\rho^{-2}(b_{m-1}-\varepsilon_m)\rho^2(b_{m-1}). \end{aligned}$$

$\square$

Now it is possible to state the main theorem of this part.

Theorem 1. Suppose there exist non-negative predictable processes C, G such that for any $x \geq y$

$$f(t,x)-f(t,y) \leq C_t(x-y) \quad (7)$$

$$|g(t,x)-g(t,y)| \leq G_t\rho(x-y) \quad (8)$$

$$|h(t,x)-h(t,y)| \leq G_t\rho(x-y) \quad (9)$$

with h - non-decreasing in x and $\rho(x)$ satisfying (5) and (6) and

$$\mathbf{E}\int_0^T C_t e^{-\int_0^t C_s da_s} da_t < \infty,$$

$$\mathbf{E}\int_0^T G_t^2 d<M>_t < \infty$$

as well as

$$\mathbf{E}\int_0^T G_t^2 d\nu_t < \infty.$$

for any T.

Assume there exist a $b \in \mathcal{A}_{loc}^+$ and predictable processes β and L such that $d\nu = \eta db$, $d < M >= \gamma db$ and

$$\gamma(g^2(x) - \beta x) + \eta h^2(x) \le L. \tag{10}$$

Furthermore, define a process $Y_t = \mathcal{E}(-\int_0^t 2(C_s + 1)da_s - \frac{1}{2}\int_0^t \beta_s dM_s)$, where $\mathcal{E}()$ denotes stochastic exponential, and assume that

$$(\int_0^t Y_{s-} f^2(s,0)da_s + \int_0^t Y_{s-} L_s db_s) \in \mathcal{A}_{loc}^+.$$

Then equation (4) admits a unique strong solution.

For the sake of simplicity an additional assumption will be made:

$$\mathbf{E}(\int_0^\infty Y_{s-} f^2(s,0)da_s + \int_0^\infty Y_{s-} L_s db_s) < \infty. \tag{11}$$

It allows to conduct the proof on time interval $(0,\infty)$, as opposed to considering finite intervals and extending the constructed solution to any t.

The proof of Theorem 1 will be conducted in several steps represented by the following lemmas.

Lemma 2. Assume solution of (4) exists. Then it is unique.

Proof. Let X_t and Y_t be solutions of (4). Then by Ito's formula

$$\mathbf{E}\varphi_m(X_t - Y_t)e^{-\int_0^t C_z da_z} = \mathbf{E}(I_1 + 1/2 I_2 + I_3),$$

where

$$I_1 = \int_0^t e^{-\int_0^s C_z da_z}(\varphi_m'(X_{s-} - Y_{s-})(f(X_{s-}) - f(Y_{s-})) -$$
$$-C_s\varphi_m(X_{s-} - Y_{s-}))da_s$$
$$I_2 = \int_0^t e^{-\int_0^s C_z da_z}\varphi_m''(X_{s-} - Y_{s-})(g(X_{s-}) - g(Y_{s-}))^2 < M >_s$$

and

$$I_3 = \int_0^t e^{-\int_0^s C_z da_z}(\varphi_m(X_{s-} - Y_{s-} + \bar{h}_s) -$$
$$-\varphi_m(X_{s-} - Y_{s-}) - \varphi_m'(X_{s-} - Y_{s-})\bar{h}_s)d\nu_s,$$

where

$$\bar{h}_s = h(X_{s-}) - h(Y_{s-}).$$

Since by (8) and definition of φ_m

$$\varphi_m''(X_{s-} - Y_{s-})(g(X_{s-}) - g(Y_{s-}))^2 \leq \max[\psi_m(x)\rho^2(x)]G_s^2 \leq \frac{2}{m}G_s^2$$

then as $m \to \infty$

$$\mathbf{E}I_2 \leq \frac{2}{m}\mathbf{E}\int_0^t e^{-\int_0^s C_z da_z}G_s^2 d < M >_s \to 0.$$

For I_3 using Taylor's decomposition for some $0 \leq \theta_s \leq 1$

$$I_3 = \int_0^t e^{-\int_0^s C_z da_z}\varphi_m''(X_{s-} - Y_{s-} + \theta_s\bar{h}_s)\bar{h}_s^2 d\nu_s.$$

To deal with the above it is necessary to estimate the following expression:

$$\varphi_m''(x-y+\theta(h(x)-h(y)))(h(x)-h(y))^2.$$

Here lemma 1 along with (9) can be used to produce an estimate

$$\varphi_m''(x-y+\theta(h(x)-h(y)))(h(x)-h(y))^2 \le$$
$$\le \frac{2}{m}\rho^{-2}(b_{m-1}-\varepsilon_m)\rho^2(b_{m-1})G^2.$$

Thus, from (6)

$$\mathbf{E}I_3 \le \frac{2}{m}\rho^{-2}(b_{m-1}-\varepsilon_m)\rho^2(b_{m-1})\mathbf{E}\int_0^t e^{-\int_0^s C_z da_z} G_s^2 d\nu_s \to 0$$

as $m \to \infty$.

The first part of the integrand for I_1 can be expressed as

$$\varphi_m'(x-y)(f(x)-f(y)),$$

which is estimated using the following steps:

if $x \ge y$, then $0 \le \varphi_m'(x-y) \le 1$ and from (7)

$$\varphi_m'(x-y)(f(x)-f(y)) \le C_s\varphi_m'(x-y)(x-y) \le C_s(x-y).$$

If $x < y$ then $-1 \le \varphi_m'(x-y) \le 0$. From (7)

$$(f(x)-f(y)) \ge C_s(x-y),$$

therefore

$$
\begin{aligned}
\varphi_m'(x-y)(f(x)-f(y)) &\leq C_s\varphi_m'(x-y)(x-y) \leq \\
\leq -C_s(x-y) &= C_s|x-y|.
\end{aligned}
$$

Combining both cases one gets:

$$
\varphi_m'(x-y)(f(x)-f(y)) \leq C_s|x-y|.
$$

Altogether:

$$
\begin{aligned}
&\mathbf{E}\varphi_m(X_t-Y_t)e^{-\int_0^t C_z da_z} \leq \\
&\leq \mathbf{E}(I_2+I_3)+\mathbf{E}\int_0^t e^{-\int_0^s C_z da_z} C_s(|X_{s-}-Y_{s-}|-\varphi_m(X_{s-}-Y_{s-}))da_s \leq \\
&\leq \mathbf{E}(I_2+I_3)+b_{m-1}\mathbf{E}\int_0^t e^{-\int_0^s C_z da_z} C_s da_s.
\end{aligned}
$$

Letting m go to infinity yields $\mathbf{E}|X_t-Y_t|e^{-\int_0^t C_z da_z}=0$, therefore $X_t=Y_t$ (a.s.) for any t. □

The next step is to build a sequence of approximating solutions and prove its convergence. Divide the time interval $[0,T]$ into n equal parts and denote partition points by t_i, $i=0,1,..n-1$. Let $t^n(t)$ be

$$
t^n(t) = t_i \text{ for } t_i \leq t < t_{i+1}. \tag{12}
$$

Then define a process:

$$
\begin{aligned}
X_t^n &= X_0 + \\
+ & \int_0^t f(s^n, X^n_{s^n-})da_s + \int_0^t g(s^n, X^n_{s^n-})dM_s + \int_0^t h(s^n, X^n_{s^n-})d(\mu-\nu)_s
\end{aligned} \tag{13}
$$

where $s^n = t^n(s)$.

Two following lemmas will be proven under one additional condition.

Assumption 1. For every t on $[0,t]$ the functions f, g and h are uniformly continuous and bounded in both variables. In this case $||f||_t = \sup_{s\le t,x}|f(s,x)| < \infty$ for every t and the same holds for g and h.

Lemma 3. As $n\to\infty$ the following holds:

$$\begin{aligned}\mathbf{E}\int_0^t |f(s,X^n_{s-}) - f(s^n, X^n_{s^n-})|da_s &\to 0\\ \mathbf{E}\int_0^t (g(s,X^n_{s-}) - g(s^n, X^n_{s^n-}))^2 d<M>_s &\to 0\\ \mathbf{E}\int_0^t (h(s,X^n_{s-}) - h(s^n, X^n_{s^n-}))^2 d\nu_s &\to 0.\end{aligned}$$

Proof. The proof will be presented for the second integral. Every other claim can be proven in the same manner.

Fix $\varepsilon > 0$ and $t > 0$. From assumption 1 on the time interval $[0,t]$ there exists a positive δ, such that

$$(g(s_1,x) - g(s_2,y))^2 \le \varepsilon$$

if $|s_1 - s_2| \le \delta$ and $|x-y|\le\delta$.

Let n be such that $\sup_s |s - s^n| \le \delta$. Then

$$\begin{aligned}&\mathbf{E}\int_0^t (g(s,X^n_{s-}) - g(s^n, X^n_{s^n-}))^2 d<M>_s =\\ &= \mathbf{E}\Big(\int_0^t (g(s,X^n_{s-}) - g(s^n, X^n_{s^n-}))^2 \mathbf{I}_{\{|X^n_{s-} - X^n_{s^n-}|\le\delta\}} d<M>_s +\\ &+\int_0^t (g(s,X^n_{s-}) - g(s^n, X^n_{s^n-}))^2 \mathbf{I}_{\{|X^n_{s-} - X^n_{s^n-}|>\delta\}} d<M>_s\Big) \le\\ &\le (\varepsilon + 4||g||_t^2 \mathbf{P}(\sup_{s\le t}|X^n_{s-} - X^n_{s^n-}| > \delta))\mathbf{E}<M>_t .\end{aligned}$$

It remains to show that $\mathbf{P}(\sup_{s\le t}|X^n_{s-} - X^n_{s^n-}| > \delta)$ converges to zero as $n\to\infty$ to complete the proof of the lemma.

To do so note that by construction

$$\sup_{s\le t}|X^n_{s-}-X^n_{s^n-}|=$$
$$=\sup_{s\le t}|f(s^n,X^n_{s^n-})(a_s-a_{s^n})+g(s^n,X^n_{s^n-})(M_s-M_{s^n})+$$
$$+h(s^n,X^n_{s^n-})\int_s^{s^n}1d(\mu-\nu)_z| \quad ,$$

therefore

$$\mathbf{P}(\sup_{s\le t}|X^n_{s-}-X^n_{s^n-}|>\delta)\le \mathbf{P}(\sup_{s\le t}|a_s-a_{s^n}|>\frac{\delta}{||f||_t})+$$
$$+\mathbf{P}(\sup_{s\le t}|M_s-M_{s^n}|>\frac{\delta}{||g||_t})+\mathbf{P}(\sup_{s\le t}|\int_s^{s^n}1d(\mu-\nu)_z|>\frac{\delta}{||h||_t}).$$

Focusing on the middle term define a stopping time

$$\tau=\inf\{s:|M_s-M_{s^n}|>\frac{\delta}{||g||_t}\}.$$

Then

$$\mathbf{P}(\sup_{s\le t}|M_s-M_{s^n}|>\frac{\delta}{||g||_t})\le \mathbf{P}(|M_\tau-M_{\tau^n}|>\frac{\delta}{||g||_t})\le$$
$$\le\frac{||g||_t^2}{\delta^2}\mathbf{E}|M_\tau-M_{\tau^n}|^2=\frac{||g||_t^2}{\delta^2}\mathbf{E}(<M>_\tau-<M>_{\tau^n}).$$

After similar considerations for the first and third terms, assertion of the lemma follows from continuity of processes a, $<M>$ and ν. $\square$

Lemma 4. Under conditions of theorem 1 and assumption 1 equation (4) admits a solution.

Proof. Denote $R^n_s=X^{n+k}_{s-}-X^n_{s-}$ and $\bar{h}^n_s=h(s^{n+k},X^{n+k}_{s^{n+k}-})-h(s^n,X^n_{s^n-})$ then by Ito's formula

$$\varphi_m(X_t^{n+k} - X_t^n)e^{-\int_0^t C_s da_s} = \text{local martingale} + I_1 + 1/2I_2 + I_3, \tag{14}$$

where

$$\begin{aligned}
I_1 &= \int_0^t e^{-\int_0^s C_z da_z}(\varphi_m^{'}(R_s^n)(f(X_{s^{n+k}-}^{n+k}) - f(X_{s^n-}^n)) - C_s\varphi_m(R_s^n))da_s \\
I_2 &= \int_0^t \varphi_m^{''}(R_s^n)e^{-\int_0^s C_z da_z}(g(X_{s^{n+k}-}^{n+k}) - g(X_{s^n-}^n))^2 < M >_s \\
I_3 &= \int_0^t e^{-\int_0^s C_z da_z}(\varphi_m(R_s^n + \bar{h}_s^n) - \varphi_m(R_s^n) - \varphi_m^{'}(R_s^n)\bar{h}_s^n)d\nu_s.
\end{aligned}$$

Decompose the integrals in the following manner:

$$\begin{aligned}
\mathbf{E}I_1 \le \mathbf{E}\int_0^t e^{-\int_0^s C_z da_z}(\varphi_m^{\prime}(R_s^n)\{(f(X_{s^{n+k}-}^{n+k}) - f(X_{s-}^{n+k})) + \\
+(f(X_{s-}^n) - f(X_{s^n-}^n)) + (f(X_{s-}^{n+k}) - f(X_{s-}^n) - C_s\varphi_m(R_s^n))\})da_s = \\
= \mathbf{E}(I_1^{(1)} + I_1^{(2)} + I_1^{(3)}).
\end{aligned}$$

Similarly:

$$\begin{aligned}
\mathbf{E}I_2 \le 3\mathbf{E}\int_0^t \varphi_m^{\prime\prime}(R_s^n)e^{-\int_0^s C_z da_z}\{(g(X_{s^{n+k}-}^{n+k}) - g(X_{s-}^{n+k}))^2 + \\
+(g(X_{s-}^n) - g(X_{s^n-}^n))^2 + (g(X_{s-}^{n+k}) - g(X_{s-}^n))^2\}d < M >_s = \\
= 3\mathbf{E}(I_2^{(1)} + I_2^{(2)} + I_2^{(3)})
\end{aligned}$$

and for some $0 \le \theta_s \le 1$:

$$\begin{aligned}
\mathbf{E}I_3 = \int_0^t e^{-\int_0^s C_z da_z}\varphi_m^{\prime\prime}(R_s^n + \theta_s\bar{h}_s^n)(h(X_{s^{n+k}-}^{n+k}) - h(X_{s^n-}^n))^2 d\nu_s \le \\
\le 3\mathbf{E}\int_0^t e^{-\int_0^s C_z da_z}\varphi_m^{\prime\prime}(R_s^n + \theta_s\bar{h}_s^n)\{(h(X_{s^{n+k}-}^{n+k}) - h(X_{s-}^{n+k}))^2 + \\
+(h(X_{s-}^n) - h(X_{s^n-}^n))^2 + (h(X_{s-}^{n+k}) - h(X_{s-}^n))^2\}d\nu_s = \\
= 3\mathbf{E}(I_3^{(1)} + I_3^{(2)} + I_3^{(3)}).
\end{aligned}$$

Integrals $I_1^{(3)}$, $I_2^{(3)}$ and $I_3^{(3)}$ are dealt with in the same way as in lemma 2. Fix $\varepsilon > 0$ and choose m such that $b_{m-1} \leq \varepsilon$ and

$$\mathbf{E}(I_2^{(3)} + I_3^{(3)}) \leq \varepsilon. \tag{15}$$

.

With m fixed $\varphi_m''(R)$ is bounded and by lemma 3 for large enough n:

$$\mathbf{E}(I_1^{(1)} + I_1^{(2)} + I_2^{(1)} + I_2^{(2)} + I_3^{(1)} + I_3^{(2)}) \leq \varepsilon. \tag{16}$$

Taking expectation of (14) and using (15) and (16) results in:

$$\begin{aligned}
\mathbf{E}\varphi_m(R_t^n)e^{-\int_0^t C_s da_s} &\leq 2\varepsilon + \mathbf{E}\int_0^t e^{-\int_0^s C_z da_z} C_s(|R_s^n| - \varphi_m(R_s^n))da_s \leq \\
&\leq 2\varepsilon + b_{m-1}\mathbf{E}\int_0^t e^{-\int_0^s C_z da_z} C_s da_s \leq K\varepsilon.
\end{aligned}$$

Thus there exists a constant K' such that

$$\mathbf{E}|R_t^n|e^{-\int_0^t C_z da_z} \leq \mathbf{E}(\varphi_m(R_t^n) + b_{m-1})e^{-\int_0^t C_z da_z} \leq K'\varepsilon,$$

therefore as $n \to \infty$

$$\mathbf{E}|X_t^{n+k} - X_t^n|e^{-\int_0^t C_s da_s} \to 0. \tag{17}$$

Fix T and define a stopping time

$$\tau = \inf\{t : |X_t^{n+k} - X_t^n|e^{-\int_0^t C_s da_s} \geq \delta\} \wedge T.$$

Since τ is bounded (17) will hold with t replaced by τ as well. Then

$$
\begin{aligned}
&P(\sup_{t\le T}|X_t^{n+k}-X_t^n|e^{-\int_0^t C_s da_s}\ge\delta)\le\\
&\le P(|X_\tau^{n+k}-X_\tau^n|e^{-\int_0^\tau C_s da_s}\ge\delta)\le\ \ \tfrac{1}{\delta}\mathbf{E}|X_\tau^{n+k}-X_\tau^n|e^{-\int_0^\tau C_s da_s}
\end{aligned}
$$

and $P(\sup_{t\le T}|X_t^{n+k}-X_t^n|e^{-\int_0^t C_s da_s}\ge\delta)\to 0$ as $n\to\infty$. Therefore X_t^n converges in probability and this convergence in uniform in t on every bounded interval. One can then select a subsequence converging with probability 1 and denote its limit as X_t. It remains to show that X_t indeed satisfies the desired equation (1). By continuity of f, g and h and Lebesgue dominated convergence theorem

$$
\begin{aligned}
\int_0^t f(s^n,X_{s-}^n)da_s &\to \int_0^t f(s,X_{s-})da_s\\
\int_0^t g(s^n,X_{s-}^n)dM_s &\to \int_0^t g(s,X_{s-})dM_s\\
\int_0^t h(s^n,X_{s-}^n)d\nu_s &\to \int_0^t h(s,X_{s-})d\nu_s.
\end{aligned}
$$

□

Proof of Theorem 1. Uniqueness was proved in lemma 2, while lemma 4 provides existence in the case of bounded coefficients. In the general case, let

$$
f_k(t,x)=\begin{cases} f(t,x), & \text{for } |x|\le k,\\ f(t,k), & \text{for } x\ge k,\\ f(t,-k), & \text{for } x\le -k,\end{cases}
$$

for any positive integer k.

Define g_k and h_k in the same way and let X_t^k be a solution of

$$
\begin{aligned}
dX_t^k &= f_k(t,X_{t-}^k)da_t+g_k(t,X_{t-}^k)dM_t+h_k(t,X_{t-}^k)d(\mu-\nu)_t\\
X_0^k &= x_0.
\end{aligned}
$$

Define stopping times $\tau_k = \inf\{t : |X_t^k| \geq k\}$. For $k_1 < k_2$ on the time interval $[0, \tau_{k_1})$ uniqueness of solution implies $X_t^{k_1} = X_t^{k_2}$, therefore τ_k is a non-decreasing sequence of stopping times. It remains to show that

$$\lim_{k \to \infty} \tau_k = \infty \text{ (a.s.)}.$$

First, note that from (7) for positive x and $y = 0$

$$f(t, x) \leq f(t, 0) + C_t x.$$

Letting $x = 0$ and $y < 0$:

$$f(t, y) \geq f(t, 0) + C_t y.$$

Altogether for any x:

$$x f(t, x) \leq f(t, 0) x + C_t x^2.$$

Therefore

$$x f(t, x) - C_t x^2 - x^2 \leq f(t, 0) x - x^2 \leq f^2(t, 0)/2,$$

where a simple inequality $ax - x^2 \leq a^2/2$ was used.

Now let Y_t be a solution of

$$\begin{aligned} dY_t &= Y_{t-}(-2(C_t + 1) da_t - \frac{\beta_t}{2} dM_t) \\ Y_0 &= 1. \end{aligned}$$

Then using Ito's formula, (10) and (11) yields:

$$\mathbf{E}(X^k_{\tau_k})^2 Y_{\tau_k}\mathbf{I}_{\{\tau_k<\infty\}} \le \mathbf{E}x_0^2 +$$
$$+\mathbf{E}\int_0^{\tau_k} 2Y_{s-}(X^k_{s-}f(X^k_{s-}) - C_s(X^k_{s-})^2 - (X^k_{s-})^2)da_s +$$
$$+\mathbf{E}\int_0^{\tau_k} Y_{s-}(g^2(X^k_{s-}) - \beta_s X^k_{s-})d<M>_s +\mathbf{E}\int_0^{\tau_k} Y_{s-}h^2(s, X^k_{s-})d\nu_s \le$$
$$\le \mathbf{E}x_0^2 + \mathbf{E}\int_0^{\tau_k} Y_{s-}f^2(s,0)da_s + \mathbf{E}\int_0^{\tau_k} Y_{s-}L_s db_s < \infty.$$

But by the definition of the stopping time $|X^k_{\tau_k}| \ge k$ and therefore

$$k^2\mathbf{E}Y_{\tau_k}\mathbf{I}_{\{\tau_k<\infty\}} \le \mathbf{E}(X^k_{\tau_k})^2 Y_{\tau_k}\mathbf{I}_{\{\tau_k<\infty\}} < \infty$$

for all k, which implies that $\tau_k \to \infty$ as $k \to \infty$. ■

In theorem 1 the works [49] and [51] have been extended by considering a more general semimartingale and relaxing Lipschitz condition on the drift f. In the next part connections between the above result and another type of existence and uniqueness theorem are examined.

2.3 Relationships between results

The focus of this chapter is the monotonity condition, derived in [25] and later in [17]. Before comparing it with theorem 1 it is important to note several differences in the set-ups of the two theorems. First, the reference work [17] deals with multidimensional processes, while theorem 1 considers the one-dimensional case. Second distinction is in decomposition of semimartingales. It was assumed that all discontinuity is absorbed by the jump component μ and the other two (a and M) are continuous. Gyöngy and Krylov do not make that assumption. Moreover, in the same paper they show that a separate jump component is not required at all if M takes values in a Hilbert space. Nevertheless, they provide the monotonity condition for the separate jump component as well. The last and probably least important difference is that it is not necessary to assume absolute continuity of a and $<M>$ with respect to each other to prove theorem 1.

Differences between two results are best demonstrated in the simplest possible scenario. That is considering a continuous one-dimensional case with Brownian motion as the driving martingale and letting f and g be deterministic functions of space variable only.

In this terms the monotonity condition becomes

$$2(x-y)(f(x)-f(y))+(g(x)-g(y))^2 \leq C(x-y)^2 \tag{18}$$

for some positive constant C.

Due to symmetry of the above inequality it is sufficient to consider the case $x>y$ without loss of generality. It follows that

$$2(x-y)(f(x)-f(y)) \leq C(x-y)^2$$

and dividing both sides by $2(x-y)$ results in condition (7).

To derive condition (8) divide (18) by $(x-y)$:

$$2(f(x)-f(y))+\frac{(g(x)-g(y))^2}{x-y} \leq C(x-y)$$

and letting $y \to x-$ yields

$$\lim_{y \to x-} \frac{(g(x)-g(y))^2}{x-y} = 0.$$

Notice how in the proof of theorem 1 cut-off functions f_k and g_k were used to satisfy uniform continuity assumption. In similar fashion it is possible to limit considerations to bounded intervals $x, y \in [-K, K]$ and assume that convergence of $\frac{(g(x)-g(y))^2}{x-y}$ to zero is actually uniform in (x, y). Thus, for y

sufficiently close to x

$$(g(x) - g(y))^2 \leq (x - y).$$

As can be seen from the previous chapter, it only matters how the function ρ behaves close to zero. The function $\rho^2(x) = x$ clearly satisfies (5), while condition (6) is not necessary for the continuous case.

In conclusion, result proven in theorem 1 is more general than the monotonity condition for the continuous case.

3 Comparison of stochastic processes

3.1 Results overview

The name comparison theorems in literature usually refers to results establishing inequality-like relationships between stochastic processes. A typical comparison theorem would study stochastic differential equations of the type

$$\begin{aligned} dX_t &= f(t, X_{t-})da_t + g(t, X_{t-})dM_t + h(t, X_{t-}, u)d(\mu - \nu)_{t,u} \\ X_0 &= x_0 \end{aligned} \tag{19}$$

and place certain conditions on coefficients f, g and h as well as processes a and M and measures μ and ν to ensure one of the following:

1. path-wise almost surely comparison, i.e.

$$X_t \leq Y_t \qquad (a.s.), \tag{20}$$

2. mean comparisons, i.e.

$$\mathbf{E}k(X_t) \leq \mathbf{E}k(Y_t) \tag{21}$$

for a certain class of functions k. As was the case with existence and uniqueness of solution results, big jumps need not be considered in the body of main theorems and can be dealt with separately, as demonstrated in [14].

The first result of the first type belongs to Skorokhod [47], who proved a comparison theorem for processes with constant diffusion. It was later generalized by Yamada [50] and can be briefly stated as:

considering two diffusion processes

$$\begin{aligned} dX_t &= f(X_t)dt + g(X_t)dW_t \\ dY_t &= \tilde{f}(Y_t)dt + g(Y_t)dW_t, \end{aligned}$$

with continuous functions f, $\tilde{f}$ and g the following conditions

$$\begin{aligned} f(x) &< \tilde{f}(x) \text{ for all } x \qquad (22) \\ |g(x) - g(y)| &\leq \rho(|x-y|) \qquad (23) \end{aligned}$$

ensure that for all t

$$X_t \leq Y_t \qquad \text{(a.s.)}.$$

Here the function $\rho(x)$ satisfies to the standard condition (5). The above mentioned result will from now on be referred to as standard one-dimensional path-wise comparison theorem.

The idea behind Yamada's proof is to demonstrate that for all t the following inequality holds true:

$$\mathbf{E}|Y_t - X_t| \leq \mathbf{E}(Y_t - X_t)$$

The fact that $|x|$ is not a C^2 function (thus Ito's formula cannot be applied directly) presents a complication. To resolve it, a series of smooth approximations for absolute value is constructed (see φ_m used in chapter 2.2).

Similar method was used in [35, 14] to extend the one-dimensional path-wise comparison theorem to the general semimartingale case. Addition of jumps (in the form of an integral with respect to a jump measure) required one extra condition on the function $\rho(x)$, which was demonstrated by (6). The discontinuous case was also considered before in [2]. Another extension of path-wise comparison theorem came in the form of multidimensional continuous diffusion equations, studied in [36]. Many of these results are summarized in the book [46].

An alternative approach to overcoming the complication presented by non-smoothness of the absolute value function is presented by Ding and Wu in

[11]. The paper studies comparison of multidimensional processes represented as solutions of stochastic differential inequalities. It is based on application of extended Ito's formula, which uses the concept of local time to obtain a stochastic integral representation for $|Y_t - X_t|$. Despite difference in proofs their conditions, when limited to the one-dimensional case, are very similar to those of Yamada.

One of the biggest weaknesses of the path-wise theorems is that they only allow comparison of processes with identical diffusion coefficients. O'Brien [39] and later Gal'chuk and Davis [15] developed a comparison theorem for SDE's with different diffusions, which require more specific conditions on initial values. The idea of that method is to use transformations: for a processes X_t satisfying

$$dX_t = f(X_t)dt + g(X_t)dW_t$$

define $G(x) = \int_{x_0}^{x} g^{-1}(u)du$. Assuming necessary smoothness, dynamics of $Z_t = G(X_t)$ is then given by

$$dZ_t = (f(X_t)/g(X_t) - 0.5g'(X_t))dt + dW_t.$$

Modified in such way processes can be compared using standard approaches. As far as applications go the method itself appears to be rather complicated and not very useful. Nevertheless, similar transformations can become a powerful tool and will be used in a number of examples later.

Finally, Peng and Zhu in [40] proved that if initial conditions are not specified, then identical diffusion coefficient is a necessity for comparison of solutions of one-dimensional SDE's.

The first theorem of the second type was derived by Hajek [18]. Given two processes

$$X_t = x_0 + \int_0^t \sigma_s dW_s$$
$$Y_t = x_0 + \int_0^t \rho(Y_s) dW_s$$

condition $|\sigma_s| \leq \rho(x_s)$ implies that $\mathbf{E}g(X_t) \leq \mathbf{E}g(Y_t)$ for any convex function g. The proof is based on time-change and can be extended to the non-zero drift case.

A general approach to mean comparisons is demonstrated by Bergenthum and Ruschendorf in [4]. Let two processes be described by

$$\begin{aligned} dX_t &= \tilde{f}(X_t)dt + \tilde{g}(X_t)dW_t \\ dY_t &= f(Y_t)dt + g(Y_t)dW_t \\ X_0 &= Y_0. \end{aligned}$$

It is based on properties of the propagation operator, defined by

$$\mathcal{H}(t, x) = \mathbf{E}(h(X_T)|X_t = x).$$

By construction, $\mathcal{H}(t, X_t)$ is a martingale, therefore

$$\mathcal{H}_t(t, x) + \mathcal{H}_x(t, x)\tilde{f}(x) + \frac{1}{2}\mathcal{H}_{xx}(t, x)\tilde{g}^2(x) = 0.$$

Since by construction

$$\mathcal{H}(0, X_0) = \mathbf{E}h(X_T)$$

one can use Ito's formula to show that

$$\mathbf{E}h(Y_T) = \mathbf{E}h(X_T) + \mathbf{E}\int_0^T \mathcal{H}_x(f(Y_t) - \tilde{f}(Y_t)) + \frac{1}{2}\mathcal{H}_{xx}(g^2(Y_t) - \tilde{g}^2(Y_t))dt.$$

At the same time if $h(x)$ is monotone, convex or both, then $H(t, x)$ will be same (in space variable). Using this fact allows to find certain conditions under which $\mathcal{H}(t, Y_t)$ becomes super(sub)-martingale. The desired comparison of expectations of the type (21) for monotone and/or convex functions can therefore be obtained. The method, which can be seen as a union of two types of comparison theorems, does not improve the previously known conditions. Nevertheless, its idea is very clear and simple and will be used in one of the examples below.

The next part presents a proof to the path-wise comparison theorem.

3.2 Path-wise comparison theorem[2]

As can be seen from the previous part the basic idea behind path-wise comparisons is to prove that (under certain conditions)

$$\mathbf{E}(|R_t| - R_t) \leq 0. \tag{24}$$

where R_t is the difference of two processes one wishes to compare. The function $|x| - x$ is non-negative therefore the only way (24) is possible is if the expectation is actually equal to zero which means that $R_t \geq 0$ (a.s.).

The leading idea of this part is to replace $|x| - x$ with a smooth non-negative function $\gamma(x)$ which is equal to zero to one side of zero. The simplest example would be to take $\gamma(x) = x^2\mathbf{I}_{\{x>0\}}$. Strictly speaking, this function is not in C^2, but if considered as a limit of $x^{2+1/n}\mathbf{I}_{\{x>0\}}$ correctness of Ito's formula will follow. In any case, the proposed method reduces the amount of work to be done and (as will be shown below) provides a somewhat more general result.

To demonstrate the concept, consider the following proposition.

[2]A version of this chapter has been submitted for publication in [32]

Proposition 1. Let X_t and Y_t be two diffusion processes satisfying

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t$$

and

$$dY_t = \tilde{f}(t, Y_t)dt + g(t, Y_t)dW_t$$

where f, $\tilde{f}$ and g are deterministic functions and W_t is a Brownian motion.

If $X_0 \leq Y_0$ and there exists a constant C such that for any t and $x \geq y$

$$2(x-y)(f(t,x) - \tilde{f}(t,y)) + (g(x) - g(y))^2 \leq C(x-y)^2 \tag{25}$$

then for all t

$$X_t \leq Y_t \qquad (a.s.)$$

Proof. Define $R_t = X_t - Y_t$ and $\gamma_\varepsilon(t,x) = e^{-Ct}x^{2+\varepsilon}I_{\{x\geq 0\}}$. Applying Ito's formula to $\gamma_{t,\varepsilon}(R_t)$ yields:

$$\begin{aligned}\gamma_\varepsilon(t, R_t) = -\int_0^t C(X_s - Y_s)^{2+\varepsilon} I_{\{R_s \geq 0\}} e^{-Cs} ds + \\ + \int_0^t (2+\varepsilon)(X_s - Y_s)^{1+\varepsilon}(f(s, X_s) - \tilde{f}(s, Y_s)) I_{\{R_s\geq 0\}} e^{-Cs} ds + \\ + \int_0^t (2+\varepsilon)(X_s - Y_s)^{1+\varepsilon}(g(s, X_s) - g(s, Y_s)) I_{\{R_s\geq 0\}} e^{-Cs} dW_s \\ + \int_0^t \frac{(2+\varepsilon)(1+\varepsilon)}{2}(X_s - Y_s)^{\varepsilon}(g(s, X_s) - g(s, Y_s))^2 I_{\{R_s\geq 0\}} e^{-Cs} ds.\end{aligned}$$

Note that $\gamma_\varepsilon(0, R_0) = 0$ by assumption.

Taking expectation of the above expression leads to the following:

$$
\begin{aligned}
& \mathbf{E}\gamma_\varepsilon(t, R_t) = \\
= \; & \mathbf{E}[\int_0^t \{(2+\varepsilon)(X_s - Y_s)^{1+\varepsilon}(f(s, X_s) - \tilde{f}(s, Y_s)) + \\
+ \; & \frac{(2+\varepsilon)(1+\varepsilon)}{2}(X_s - Y_s)^\varepsilon (g(s, X_s) - g(s, Y_s))^2 - \\
- \; & C(X_s - Y_s)^{2+\varepsilon}\} I_{\{R_s \geq 0\}} e^{-Cs} ds].
\end{aligned}
$$

Taking limit as $\varepsilon \to 0$ and using Lebesgue dominated convergence theorem results in

$$
\begin{aligned}
& e^{-Ct}\mathbf{E}R_t^2 I_{\{R_t \geq 0\}} = \\
= \; & \mathbf{E}[\int_0^t \{2(X_s - Y_s)(f(s, X_s) - \tilde{f}(s, Y_s)) + \\
+ \; & (g(s, X_s) - g(s, Y_s))^2 - C(X_s - Y_s)^2\} I_{\{R_s \geq 0\}} e^{-Cs} ds].
\end{aligned}
$$

By assumption, the integrand is non-positive, therefore

$$
e^{-Ct}\mathbf{E}R_t^2 I_{\{R_t \geq 0\}} \leq 0.
$$

But random variable under the expectation is non-negative, thus

$$
R_t^2 I_{\{R_t \geq 0\}} = 0 \text{ (a.s.)}
$$

which is equivalent to

$$
R_t \leq 0 \text{ (a.s.)}.
$$

□

Remark 1. Condition (25) is very similar to the monotonity condition used in

[17] and will be studied more precisely below.

One does not't have to limit considerations to direct comparison of processes. In fact, the same method can be used to establish indirect comparison: for this consider $F^2(x)\mathbf{I}_{\{F(x)>0\}}$ for $x \in \mathbf{R}^d$. A special case $d = 2$ and $F(x_1, x_2) = (x_1 - x_2)$ corresponds to the standard one-dimensional comparison theorems.

As usual, consider a standard probability basis with a non-decreasing process a_t, a n-dimensional continuous local martingale M_t, a jump measure μ with a compensator ν. Furthermore, assume that predictable characteristics $< M >_t$ and ν_t are absolutely continuous with respect to a_t, namely that there exist a $n \times n$ predictable matrix m and a predictable process η such that

$$\begin{aligned} d < M^i, M^j > &= m_{ij} da \\ d\nu &= \eta da \end{aligned}$$

respectively.

Consider a d-dimensional process:

$$dX_t = f(X_{t-})da_t + g(X_{t-})dM_t + h(u, X_{t-})d(\mu - \nu) \tag{26}$$

where $f(x) = f(x, t, \omega)$ and $g(x) = g(x, t, \omega)$ are function from $\mathbf{R}^d \times \mathbf{R}_+ \times \Omega$ to $\mathbf{R}^d$ and real-valued $d \times n$ matrices respectively and $h(u, x) = h(u, x, t, \omega)$ is a function from $\mathbf{R} \backslash \{0\} \times \mathbf{R}^d \times \mathbf{R}_+ \times \Omega$ to $\mathbf{R}^d$. As before, unnecessary variables will be omitted whenever possible for short.

Theorem 2. Let $F(x) : \mathbf{R^d} \to \mathbf{R}$ be a C^2 function such that $F(X_0) \leq 0$ with probability 1. Furthermore, assume that for all x

$$I_{\{F(x+h(x))>0\}} = I_{\{F(x)>0\}}. \tag{27}$$

Assume that there exists a $\delta > 0$ and a non-negative predictable process C such that for any x with $0 < F(x) \leq \delta$ the following inequality holds

$$
\begin{aligned}
2F(x)(F'(x), f(x)) + (F'(x)g(x), m_s F'(x)g(x)) +& \\
+F(m_s \cdot F''(x)g(x), g(x)) +& \\
+\int_{|u|<1} (F(x+h(x,u))^2 - F^2(x) - (F'(x), h(x,u)))\eta_s \le C_s F^2(x),& \qquad (28)
\end{aligned}
$$

where $F'(x)$ and $F''(x)$ are vector of first derivatives and matrix of second derivatives respectively and (a,b) denotes a scalar product.

Then for any t

$$
F(X_t) \le 0 \qquad (a.s.)
$$

Proof. Define a stopping time

$$
\tau = \inf\{t : F(X_t) > \delta\}.
$$

Since $F(X_0) \le 0$ then $\tau > 0$ (a.s.)

Applying Ito's formula to $F^2(X_t)\exp[-\int_0^t C_s da_s]\mathbf{I}_{\{F(X_t)>0\}}$ results in:

$$
\begin{aligned}
F^2(X_t)e^{-\int_0^t C_s da_s}\mathbf{I}_{\{F(X_t)>0\}} = \text{local martingale} +& \\
+\int_0^t \mathbf{I}_{\{F>0\}} e^{-\int_0^s C_u da_u}(2F(F', f) - C_s F^2) da_s +& \\
+\int_0^t e^{-\int_0^s C_u da_u}\mathbf{I}_{\{F>0\}} \sum_{i,j} (F_i F_j + F F_{ij}) g_i g_j d < M^i, M^j >_s +& \\
+\int_0^t e^{-\int_0^s C_u da_u} \int_{|u|<1} F(X+h)^2 \mathbf{I}_{\{F(X+h)>0\}} - (F^2 + (F', h))\mathbf{I}_{\{F>0\}} d\nu_s.& \quad (29)
\end{aligned}
$$

Then, using uniform continuity of $< M >$ and ν with respect to a_s and taking expectation get:

$$
\begin{aligned}
\mathbf{E}e^{-\int_0^t C_s da_s} F^2(X_t)\mathbf{I}_{\{F(X_t)>0\}} = \mathbf{E}\int_0^t e^{-\int_0^s C_u da_u}\{2F(F', f) + \\
+(F'g, m_s F'g) + F(m_s \cdot F''g, g) + \\
+\int_{|u|<1} (F(X+h)^2 - F^2 - (F', h))\eta_s - C_s F^2\}\mathbf{I}_{\{F>0\}} da_s.
\end{aligned} \tag{30}
$$

The integrand in (30) is non-positive on $(0, \tau)$ by assumption, therefore

$$
\mathbf{E}e^{-\int_0^t C_s da_s} F^2(X_t)\mathbf{I}_{\{F(X_t)>0\}} \leq 0. \tag{31}
$$

But expression under the sign of expectation is non-negative, therefore (31) is possible only if $e^{-\int_0^t C_s da_s} F^2(X_t)\mathbf{I}_{\{F(X_t)>0\}} = 0$ (a.s.) what in turn means that $F(X_t) \leq 0$ (a.s.) for any $t < \tau$.

For any ω such that $\tau(\omega) < \infty$ above considerations yield $F(X_{\tau-}) \leq 0$ (a.s.). If there is a jump at time τ, then

$$
F(X_\tau) = F(X_{\tau-}) + h(X_{\tau-})
$$

which is again less than zero by (27). But by definition of τ and right-continuity $F(X_\tau) \geq \delta$. Contradiction implies that $\tau = \infty$ (a.s.). $\square$

Remark 2. Theorem 2 requires that $I_{\{F(x+h(x))>0\}} = I_{\{F(x)>0\}}$ for all x. When $d = 2$, $h(x_1, x_2) = (h(x_1), h(x_2))$ and $F(x_1, x_2) = (x_1 - x_2)$ it can be achieved by assuming that $x + h(x)$ is a non-decreasing function for all x and $h(0) = 0$ (this assumption is used in [14] as well).

Remark 3. Even if not evident from the proof, two processes can be compared only if they have identical diffusions. Indeed when there are no jumps, condition (28) can be expressed as:

$$
2F(F', f) + (F'g, mF'g) + F(m \cdot F''g, g) \leq 0
$$

whenever $F < 0$.

Letting F approach zero results in a necessary condition on the quadratic form:

$$(F'g, mF'g) \leq 0$$

whenever $F = 0$, which for $F(x, y) = (x - y)$ becomes

$$m(g(x) - \tilde{g}(x))^2 \leq 0.$$

Similar restrictions can be obtained for h when considering discontinuous process.

Theorem 2 provides a way to establish direct and indirect path-wise comparisons of stochastic processes. Replacing the absolute value function in standard proofs with its smooth alternatives results in two things. First, the proof is shortened and simplified. Second, the theorem requires a different type of condition, which in the case of path-wise comparisons of one-dimensional diffusion processes becomes very similar to the monotonity condition used in [17, 25] to establish existence and uniqueness of solution. This fact demonstrates another connection between two areas of my interest. Later in part 3.4 two types of conditions for path-wise comparisons of stochastic processes will be examined, as was the case with existence and uniqueness theorem.

It should be noted that path-wise comparison of d-dimensional stochastic processes using the above approach is not possible. One can, of course, consider a series of functions $F_k = x_k - y_k$ where index k runs from 1 to d, but this approach seems rather cumbersome. Instead, one can use a multidimensional comparison theorem proven in a standard fashion.

Extension of the comparison theorem to multidimensional case demands an additional condition on the drift coefficient. It was done by Geiß and Manthey [16]. Considering two systems of diffusion equations

$$dX_t^i = f_i(t, X_t)dt + \sum_{k=1}^{n} g_{ik}(t, X_t)dW_t^k$$

and

$$dY_t^i = \tilde{f}_i(t, Y_t)dt + \sum_{k=1}^{n} g_{ik}(t, Y_t)dW_t^k$$

they use the following condition on the drifts

$$f_i(t, x) < \tilde{f}_i(t, y)$$

whenever $x_i = y_i$ and $x_k \leq y_k$ for $k \neq i$, along with a standard in this case

$$|g_{ik}(x) - g_{ik}(y)| \leq \rho(|x - y|).$$

Later Ding and Wu [11] proved a multidimensional comparison theorem for continuous inequalities using local time approach.

None of the two above mentioned works represents the most general case. The next chapter will present a more general path-wise comparison theorem for multidimensional stochastic differential equations with respect to semimartingales. The extension is to add discontinuity represented by a jump measure.

3.3 Multidimensional comparison theorem[3]

This chapter presents the proof of multidimensional path-wise comparison theorem for solutions of stochastic differential equations with respect to semimartingales with a jump component.

The main semimartingale is multidimensional in this chapter as well. Its components are a d-dimensional continuous process $A = (A^1, A^2 ... A^d)$ with $A_i \in \mathcal{A}_{loc}$, a d-dimensional continuous local martingale $M = (M^1, M^2 ... M^d)$ and a n-dimensional jump measure $\mu = (\mu^1, \mu^2 ... \mu^n)$ with continuous compensators $(\nu^1, \nu^2 ... \nu^n)$.

[3] A version of this chapter has been published in [30]

The processes to be compared are given as solutions of stochastic differential equations:

$$\begin{aligned} dX_t^i &= \sum_{j=1}^{d} f_{ij}(X_{t-})dA_t^j + \sum_{j=1}^{d} g_{ij}(X_{t-}^i)dM_t^j + h_i(u, X_{t-}^i)d(\mu_t^i - \nu_t^i) \\ d\tilde{X}_t^i &= \sum_{j=1}^{d} \tilde{f}_{ij}(\tilde{X_{t-}})dA_t^j + \sum_{j=1}^{d} g_{ij}(\tilde{X_{t-}^i})dM_t^j + h_i(u, \tilde{X}_{t-}^i)d(\mu_t^i - \nu_t^i) \\ i &= 1...n \end{aligned} \tag{32}$$

where f_{ij}, g_{ij} and h_i depend on t and ω and are continuous in (t, x). Existence of solutions for both equations can be ensured in a standard way and is assumed here.

Theorem 3. Let functions f_{ij}, g_{ij}, and h_i be such that:

$$\tilde{f}_{ij}(\tilde{X}_0) > f_{ij}(X_0) \text{ (a.s)} \tag{33}$$

$$\text{for all } i \text{ and } j$$

$$\tilde{f}_{ij}(\tilde{x_1}, ..., \tilde{x}_{i-1}, x_i, \tilde{x}_{i+1}, ..) > f_{ij}(x_1, ...x_i, ..) \tag{34}$$

$$\text{whenever} \quad \tilde{x}_k \geq x_k,$$

$$h_i(y) \geq h_i(x) \tag{35}$$

$$\text{for all} \quad y \geq x$$

and

$$|g_{ij}(s, y) - g_{ij}(s, x)| \leq G_s \rho(|y - x|) \tag{36}$$

$$|h_i(s, x) - h_i(s, y)| \leq G_s \rho(|x - y|) \tag{37}$$

where the function $\rho(x)$ satisfies (5) and (6) and G_s is a non-negative predictable process with

$$\mathbf{E}\int_0^\infty G_s^2 d < M_s^j, M_s^k > < \infty$$

and

$$\mathbf{E}\int_0^\infty G_s^2 d\nu_s^i < \infty.$$

Also assume that the processes

$$\int |g_{kl}(X)| d < M^i, M^j >$$

$$\int |f_{kl}(X)| dA^i$$

$$\int |\tilde{f}_{kl}(\tilde{X})| dA^i$$

belong to $\mathcal{A}_{loc}$.

Then for $i = 1...n$ and any t

$$\tilde{X}_t^i \geq X_t^i \text{ (a.s.)}.$$

Proof. Define stopping times

$$T_i = \inf(t > 0 : f_{ij}(X_t) > \tilde{f}_{ij}(\tilde{X}_t) \text{ for at least one } j).$$

Denote $T = \min(T_i)$ and $\tau = T \wedge t$.

Recall a sequence of twice continuously differentiable functions $\varphi_m(x)$ that approximate absolute value of x. It was used in part 2.2 to prove existence and uniqueness of solution. For now m will be fixed.

Since f_{ij} are continuous, X_t and $\tilde{X}_t$ are right-continuous and

$$f_{ij}(X_0) < \tilde{f}_{ij}(\tilde{X}_0) \text{ (a.s.)}$$

by assumption, then $T_i > 0$ (a.s.) for all i and, therefore $\tau > 0$ (a.s.).

Applying Ito's formula to $\varphi_n(\tilde{X}^i_\tau - X^i_\tau)$ gets

$$\varphi_m(\tilde{X}^i_\tau - X^i_\tau) = \text{local martingale} + I_1 + 1/2 I_2 + I_3. \tag{38}$$

Where

$$I_1 = \int_0^\tau \varphi'_m(\tilde{X}^i_{s-} - X^i_{s-}) \sum_{j=1}^d (\tilde{f}_{ij}(\tilde{X}_{s-}) - f_{ij}(X_{s-}))dA^j_s$$

$$\begin{aligned} I_2 &= \int_0^\tau \varphi''_m(\tilde{X}^i_{s-} - X^i_{s-}) \times \\ &\times \sum_{j,k=1}^d (g_{ij}(\tilde{X}^i_{s-}) - g_{ij}(X^i_{s-}))(g_{ik}(\tilde{X}^i_{s-}) - g_{ik}(X^i_{s-}))d < M^j_s, M^k_s > \end{aligned}$$

$$\begin{gathered} I_3 = \int_0^\tau [\varphi_m(\tilde{X}^i_{s-} - X^i_{s-} + h_i(u, \tilde{X}^i_{s-}) - h_i(u, X^i_{s-})) - \\ -\varphi_m(\tilde{X}^i_{s-} - X^i_{s-}) - \varphi'_m(\tilde{X}^i_{s-} - X^i_{s-})(h_i(u, \tilde{X}^i_{s-}) - h_i(u, X^i_{s-}))]d\nu^i_s. \end{gathered}$$

The exact structure of the local martingale in (38) is irrelevant to the proof so it is not specified. As before, it is possible to assume that all terms in (38) admit expectations without loss of generality.

Taking expectations of both sides of (38) yields

$$\mathbf{E}\varphi_m(\tilde{X}^i_\tau - X^i_\tau) = \mathbf{E}I_1 + 1/2\mathbf{E}I_2 + \mathbf{E}I_3. \tag{39}$$

Studying every term one by one results in the following conclusions:

since

$$\tilde{f}_{ij}(\tilde{X}_{s-}) - f_{ij}(X_{s-}) \geq 0$$

on $(0, \tau)$ by construction and

$$\varphi(\cdot) \leq 1$$

the first integral can be estimated as

$$\mathbf{E}I_1 \leq \mathbf{E}\int_0^\tau \sum_{j=1}^d (\tilde{f}_{ij}(\tilde{X}_{s-}) - f_{ij}(X_{s-}))dA_s^j = \mathbf{E}(\tilde{X}_\tau^i - X_\tau^i).$$

For the second term recall that the function $\psi_m = \varphi_m''$ is equal to zero outside of the interval (b_m, b_{m-1}). Condition (36) can be used to arrive at

$$\mathbf{E}I_2 \leq 1/2 \sum_{j,k=1}^d \max_{b_m \leq x \leq b_{m-1}} [\psi_m(|x|)\rho^2(|x|)]\mathbf{E}\int_0^\tau |G_s^2| d < M_s^j, M_s^k > \leq$$

$$\leq 1/m \sum_{j,k=1}^d \mathbf{E}\int_0^\infty |G_s^2| d < M_s^j, M_s^k > .$$

The expectation in the above term is less than ∞ by assumption, therefore

$$\mathbf{E}I_2 \to 0 \text{ as } m \to \infty.$$

When dealing with the third term, denote $h(u, \tilde{X}_{s-}^i) - h(u, X_{s-}^i)$ by Δh_i for short. This expression will be estimated using condition (37). Taylor's approximation formula implies that there exists $0 \leq \alpha \leq 1$ such that

$$\mathbf{E}I_3 = 1/2\mathbf{E}\int_0^\tau \varphi_m''(\tilde{X}^i_{s-} - X^i_{s-} + \alpha\Delta h_i)(\Delta h_i)^2 d\nu^i_s \le$$
$$\le 1/2\mathbf{E}\int_0^\tau |H_s(u)|^2\rho^2(|\tilde{X}^i_{s-} - X^i_{s-}|)\varphi_m''(\tilde{X}^i_{s-} - X^i_{s-} + \alpha\Delta h_i)d\nu^i_s.$$

Using lemma 1 results in

$$\rho^2(|\tilde{X}^i_{s-} - X^i_{s-}|)\varphi_m''(\tilde{X}^i_{s-} - X^i_{s-} + \alpha\Delta h_i) \le$$
$$\le m^{-1}\rho^2(b_{m-1})\rho^{-2}(b_{m-1} - \varepsilon_m)G^2(s),$$

and condition (6) implies that $\mathbf{E}I_3$ converges to 0 as $m \to \infty$ as well.

Using the above considerations and taking limits as $m \to \infty$ in (39) results in the following conclusions

$$\mathbf{E}|\tilde{X}^i_t - X^i_t| \le \mathbf{E}(\tilde{X}^i_t - X^i_t),$$

therefore

$$\tilde{X}^i_t \ge X^i_t \text{ (a.s.)} \tag{40}$$

on $(0, \tau)$ which completes the first part of the proof.

Now define stopping times $\theta_i = \inf(r > \tau : \tilde{X}^i_r < X^i_r\ )$ and denote $\theta = \min(\theta_i)$.

For those ω when $\theta(\omega) = \infty$ then the proof is complete. Otherwise

$$\tilde{X}^i_{\theta-} \ge X^i_{\theta-}$$

for all i by the definition of θ and right-continuity of X_t and $\tilde{X}_t$. If there is no jump at time θ, then

$$\tilde{X}^i_{\theta-} = \tilde{X}^i_\theta \quad \ge \quad X^i_\theta = X^i_{\theta-}.$$

Otherwise,

$$\tilde{X}^i_\theta = \tilde{X}^i_{\theta-} + h_i(u, \tilde{X})$$
$$X^i_\theta = X^i_{\theta-} + h_i(u, X),$$

and it follows from (35) that

$$\tilde{X}^i_\theta \geq X^i_\theta \text{ (a.s.)} \tag{41}$$

for all $i = 1, .., n$.

Now fix i and limit considerations below to a set

$$B_i = \{\omega | \theta(\omega) = \theta_i(\omega) < \infty\}.$$

It follows from the definition of θ_i and right-continuity of $\tilde{X}_t$ and X_t that

$$\tilde{X}^i_{\theta_i} \leq X^i_{\theta_i}$$

for all ω. Together with (41) it implies that

$$\tilde{X}^i_{\theta_i} = X^i_{\theta_i} \text{ (a.s.)}.$$

.

For $j \neq i$ it is true that

$$\theta_j \geq \theta,$$

therefore, $\tilde{X}^j_{\theta_i} \geq X^j_{\theta_i}$. Condition (34) implies that

$$\tilde{f}_{ij}(\tilde{X}_{\theta_i}) > f_{ij}(X_{\theta_i}) \text{ (a.s.)}$$

for all j. Defining a stopping time

$$\eta_i = \inf(t > \theta_i : \tilde{f}_{ij}(\tilde{X}_t) < f_{ij}(X_t) \text{ for at least one } j)$$

it follows from above that

$$\eta_i > \theta_i \text{ (a.s.)}.$$

Reproducing the argument used in the first part of this proof for $\varphi_m(\tilde{X}^i - X^i)$ on the time interval $[\theta_i, \eta_i)$ results in the following conclusion:

there exist a stopping time $\tau_i > \theta_i$, such that

$$\tilde{X}_t^i \geq X_t^i$$

on $\theta_i < t \leq \tau_i$ which contradicts the definition of θ_i. Therefore, $B_i = \emptyset$ (a.s.) which in turn implies that $\theta = \infty$ (a.s.). $\square$

The multidimensional path-wise comparison theorem goes hand in hand with other similar results in this area. It provides a way of handling processes with complex inter-dependence structures that can potentially arise in various applications.

As demonstrated by theorems 2 and 3, distinct techniques can be used to establish path-wise inequality-like relationships between stochastic processes. These two methods rely on different conditions as well and the question is which would be more general in what scenario. This question will be answered in the next part.

3.4 Relationships between results

Before getting to the topic of this part, it is useful to note that technically the method proposed in theorem 2 does not require continuity of coefficients f, g and h. This approach can be used to compare processes with discontinuous coefficients, but it must be said that existence and uniqueness of solutions becomes a problem in this case.

The continuous coefficient case still holds some differences between the monotonity condition and the classical results of [14, 50] and others. For the purpose of demonstration considerations will be limited to a simple diffusion case, although similar results can be obtained in a more general setting. To be more precise, assume that the main semimartingale has no jumps, $a_t = t$ and $M_t = W_t$ is a Wiener process. Monotonity condition (25) is then reduced to:

$$2(x-y)(f(x)-\tilde{f}(y)) + (g(x)-g(y))^2 \leq C(x-y)^2$$

for $x > y$ and some non-negative constant C.

Connection between two types of conditions will be demonstrated in the form of two lemmas.

Lemma 5. Consider process

$$\begin{aligned} dX_t = &\ f(X_t)dt + g(X_t)dW_t \\ d\tilde{X}_t = &\ \tilde{f}(\tilde{X}_t)dt + g(\tilde{X}_t)dW_t \end{aligned} \tag{42}$$

where continuous functions f, $\tilde{f}$ and g are such that:

$$f(x) < \tilde{f}(x) \tag{43}$$

$$|g(x)-g(y)| \leq \rho(|x-y|) \tag{44}$$

with $\rho(x)$ satisfying (5).

Furthermore, define $r(x) = \frac{\rho^2(x)}{x}$ and assume that

$$\lim_{x \to 0} r(x) = 0. \tag{45}$$

Then the monotonity condition is satisfied.

Proof. The method used in theorem 1 allows to assume, without loss of generality, that all coefficients are uniformly continuous.

Fix y and define

$$k_y(x) = \frac{\rho^2(x-y)}{2(x-y)}.$$

From conditions (43), (45) and continuity of f and $k_y(x)$ it follows that $\tilde{f}(y) \geq f(x) + k_y(x)$ for some δ and $0 < x - y < \delta$.

Using definition of $k_x(y)$ gets

$$\tilde{f}(y) \geq f(x) + \frac{\rho^2(x-y)}{2(x-y)}$$

which is equivalent to

$$(f(x) - \tilde{f}(y)) + \frac{\rho^2(x-y)}{2(x-y)} \leq 0.$$

Multiplying the above inequality by $2(x-y)$ its sign will be preserved, since it is sufficient to consider the case $x > y$. Therefore

$$2(x-y)(f(x) - \tilde{f}(y)) + \rho^2(x-y) \leq 0$$

and according to (5)

$$2(x-y)(f(x) - \tilde{f}(y)) + (g(x) - g(y))^2 \leq$$
$$\leq 2(x-y)(f(x) - \tilde{f}(y)) + \rho^2(x-y) \leq 0.$$

$\square$

Lemma 6. Let functions f, $\tilde{f}$ and g satisfy (25). Then for all x

$$f(x) \leq \tilde{f}(x).$$

Furthermore, there exists a non-negative increasing function ρ, such that

$$\begin{aligned} |g(x) - g(y)| &\leq \rho(|x-y|) \\ \int_{0+} \rho^{-2}(x)dx &= \infty \end{aligned} \tag{46}$$

Proof. Dividing both sides of the monotonity condition (with $C = 0$) by $(x-y)$ results in:

$$f(x) + \frac{(g(x) - g(y))^2}{2(x-y)} \leq \tilde{f}(y).$$

Taking limit as $y \to x$ from the left yields:

$$f(x) + \lim_{y \to x-} \frac{(g(x) - g(y))^2}{2(x-y)} \leq \tilde{f}(x).$$

Two conclusions can be made from the above inequality. First, the second term on the left is non-negative, therefore

$$f(x) \leq \tilde{f}(x).$$

.

Second,

$$0 \leq \lim_{y \to x-} \frac{(g(x) - g(y))^2}{2(x-y)} \leq \tilde{f}(x) - f(x),$$

therefore the limit is finite (which is consistent with (46)). $\square$

As can be seen from lemma 5 the monotonity condition is more general whenever

$$\lim_{y \to x} \frac{(g(x) - g(y))^2}{x - y} = 0$$

for all x. The fact can be further demonstrated by a simple example.

Example 1. Let $g(x) = c$, $f(x) = 0$ and $\tilde{f}(x) = x\mathbf{I}_{\{x>0\}}$. Here the standard $\tilde{f}(x) > f(x)$ fails, but the monotonity condition

$$2(x - y)x\mathbf{I}_{\{x>0\}} \leq 0$$

is clearly satisfied for $x < y$.

In conclusion of this chapter, if

$$\lim_{y \to x+} \frac{\rho(x - y)^2}{2(x - y)} = 0$$

then the monotonity condition in more general than classical. If the limit is not equal to zero then the monotonity condition might not be satisfied.

4 Financial applications of comparison theorems

4.1 Introduction and motivations

This final section of my work is devoted to practical usages of various theoretical results studied before. It is focused on mathematical finance as the main field, although one can imagine similar techniques being put to use in other areas of applied mathematics.

The well-known Black-Scholes model assumes, among other things, that the stock price volatility is constant. Had this been true, the Black-Scholes implied volatility for options on one stock would have been the same across different strikes and maturities. This, however, is not the case. Most derivative markets exhibit persistent patterns of volatilities varying by strike. In some markets, those patterns form a smile. In others, such as equity index options markets, it is more of a skewed curve. This has motivated the name *volatility skew*. Another dimension to this problem is that of volatilities varying by expiration. A three-dimensional graph indicating implied volatilities by both strike and expiration is called *volatility surface*. Some of these effects were studied by Rubinshtein in [43, 44].

The above mentioned works can be used as evidence of the fact that models, more complicated than that of Black and Scholes need to be studied and used in practice. Multiple steps can be undertaken in this direction, but the first and, perhaps, the most intuitive one would be to allow volatility to be a deterministic function of the stock price level and time

$$dS_t = S_t(\mu dt + \sigma(t, S_t)dW_t),$$

where μ is a constant and W_t is a Brownian motion.

Before proceeding with studies of various properties of such models one has to ensure existence of (preferably unique) solution of the above stochastic differential equation. This can be achieved by the means of the existence and uniqueness theorem, studied in chapter 2 of my work. It places one condition on the diffusion function (see theorem 1) limiting growth of $\sigma(t, x)$. Apart from technical conditions there are virtually no restrictions on the form of dependence of σ on the stock price, so one can specify the model to capture many desired features.

Note also that one can easily replace the constant drift coefficient with another deterministic function of time and stock price. This, however, will not add anything to many financial applications as those are usually studied under the martingale measure and are, therefore, independent of the particular form of the real measure drift.

The feature that distinguishes the above mentioned approach from many others is completeness. The unique martingale measure $\mathbf{Q}$ is determined with the help of Gyrsanov's theorem: its density is defined as

$$\frac{d\mathbf{Q}_t}{d\mathbf{P}_t} = \exp[-\int_0^t \frac{\mu - r}{\sigma(u, S_u)} dW_u - \int_0^t \frac{(\mu - r)^2}{2\sigma^2(u, S_u)} du]$$

while the shifted process

$$d\bar{W}_t = \frac{\mu - r}{\sigma(t, S_t)} dt + dW_t$$

is a $\mathbf{Q}$-Brownian motion. Of course, one has to assume here that $\sigma(t, S_t) > 0$ (a.s.) for all t.

Completeness is an important property of financial models as far as applications are concerned. It implies that all contingent claims admit a unique fair price and can be perfectly hedged. In fact, this property will be preserved in a more complex model with several sources of randomness as demonstrated by the following proposition.

Proposition 2. Consider a stock price process

$$dS_t = S_t(\mu dt + \sum_{k=1}^{d} \sigma_k(t, S_t) dW_t^k),$$

where W_t^k are independent Brownian motions. Assume interest rate r is constant and denote a set of martingale measures by $\mathbf{M}$.

Then for a measurable function $f(x)$ the following quantity

$$\mathbf{E}_{\mathbf{Q}} f(S_t)$$

is identical for every $\mathbf{Q} \in \mathbf{M}$.

Proof. Let $\mathbf{Q}$ be in $\mathbf{M}$. By the definition of a martingale measure and Gyrsanov's theorem the stock price dynamics can be expressed as

$$dS_t = S_t(rdt + \sum_{k=1}^{d} \sigma_k(t, S_t) d\bar{W}_t^k),$$

where $\bar{W}_t^k$ are independent $\mathbf{Q}$-Brownian motions. Thus, the $\mathbf{Q}$ measure distribution of S_t is completely defined by r and functions $\sigma_k(t, x)$.

But $\mathbf{E}_{\mathbf{Q}} f(S_t)$ is completely defined by the $\mathbf{Q}$-distribution of S_t and is therefore dependent only on $\sigma_k(t, x)$ and not the particular choice of martingale measure. $\square$

In fact, the same proposition can be used to prove that not only will every stock option have a unique price, but admit a replicating strategy. Here one only needs to cite the work [24]. It proves that a contingent claim is attainable if it has a unique price under any martingale measure.

In general, modelling stock prices via stochastic differential equations with non-constant diffusion functions allows to extend the scope of available processes beyond normally (or log-normally) distributed ones. Various features of observable quantities can be captured in the model preserving uniqueness of contingent claim prices and availability of perfect hedges.

Complications of financial market models, however, come at a price of significant increases in computational difficulties. Vast majority of models will not admit explicit solutions. This leads to necessity of development of complex and high precision numerical methods, which of course is aided by rapid progress in the modern computer and software industries. Some of those are studied in [41], Monte Carlo methods and its variations are also reflected in [8]. Finite difference methods can be applied to solve Black-Scholes type partial differential equations

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 C}{\partial x^2} \quad + \quad rx\frac{\partial C}{\partial x} = rC$$

describing the option price function $C(t,x)$ with appropriate initial and boundary condition. An example of such method is presented in [1].

In some instances, however, numerical methods may still not be fast enough. Another issue is convergence and stability, not obvious in some cases. Comparison of stochastic processes can be viewed as an alternative to numerical methods for complicated models. The leading idea is to estimate the main processes with another, for which computations can be performed explicitly (or at least simplified). In this way, the approach represent a certain trade-off between time and precision. Unlike numerical methods, the sign of the error, no matter how large, is known: if the stock price process is estimated from above then that estimate will produce higher option prices (provided the options pay-off function is non-decreasing).

Comparison of stochastic processes has been studied in financial theory framework in [3, 29] among others. Both papers focus mostly on theory, deriving properties of option prices and not providing means of explicit calculations. A method of option price estimation, which relies only on some properties of the stock price is presented in [45]. Another interesting work on related matter is [34], where stochastic dominance is used to find option price bounds in discrete markets. My aim in this part of the thesis is to develop practically useful methods and demonstrates their power on specific numerical examples. Comparison theorems will play a role of theoretical tool in those methods.

4.2 Overview of examples

This subsection provides an overview of all the examples and can serve as an easy reference. The example section is divided into three main parts.

First part includes examples 1 and 2 and is devoted to direct applications of comparison theorems to option pricing and hedging. It demonstrates how path-wise and mean comparison theorems can be used to build explicit estimates for European option prices in models where exact calculations are either tedious or not possible at all. Moreover, the same methodology is applied to replication of the underlying contingent claim.

The next part presents two methodologies for estimating the stock's cumulative distribution function. It starts with description of two general approaches, which are based on similar principals as the comparison theorem. The methodologies are demonstrated by two specific examples. Relative performance and effectiveness of both approaches is studied.

The last part consists of more theoretical examples. It demonstrates other, not so obvious usages of comparisons in financial applications.

4.3 Direct applications of comparison theorems to option pricing

One of the more important questions in the theory of mathematical finance is pricing and hedging of options. But even a small step away from geometric Brownian motion can result in significant complications in the call-price formula or, more likely, no explicit formulas at all. This statement can be demonstrated by the constant elasticity of variance model, derived by Cox and Ross in [9]. The proposed stock price dynamics is described by the following stochastic differential equation

$$\begin{aligned} dS_t &= S_t(\mu dt + \sigma S_t^{\alpha-1} dW_t) \\ S_0 &= s, \end{aligned}$$

where μ and $\sigma > 0$ are constants and α is between 0 and 1.

The conditional density of S_T can be (without going into too many details) expressed as

$$f(S_T, T|S_t, t) = h(T)\exp[q(T)S_T]I_u(p(T)\sqrt{S_T}),$$

where $I_u(x)$ is a modified Bessel function of the first kind of order u, resulting in the call option pricing formula

$$\begin{aligned} P(S,t) &= \\ &= S\sum_{n=0}^{\infty} a_n(t)S^{n+v}G(n+2,c(t)) - K\sum_{n=0}^{\infty} b_n(t)S^{n+1+v}G(n+1,c(t)), \end{aligned}$$

where $G(m, x)$ is a complementary Gamma(m,1) distribution function.

The first example in this section deals with the above mentioned CEV model.

Example 1. *Constant elasticity of variance model estimation*[4].

Let the main process be given by

$$\begin{aligned} dS_t &= S_t(rdt + \sigma S_t^{\alpha-1} dW_t) \\ S_0 &= s \end{aligned}$$

with $\frac{1}{2} \leq \alpha < 1$.

Remark. The original model considers the range of α to be from 0 to 1. However, to avoid various technical difficulties this example is build for the case $\alpha \geq 0.5$ only. This condition can also be used to ensure existence and uniqueness of solution.

To construct an upper estimate consider a process

$$\begin{aligned} dY_t &= r[(1-\alpha)Y_t + \frac{s^{1-\alpha}}{\sigma}]dt + dW_t \\ Y_0 &= 0 \end{aligned}$$

This stochastic differential equation describes a normally distributed Ornstein-Uhlenbeck process with mean

$$\frac{s^{1-\alpha}}{\sigma(1-\alpha)}(e^{r(1-\alpha)t} - 1)$$

and variance

$$\frac{(\exp[2r(1-\alpha)t] - 1)}{2r(1-\alpha)}$$

and can be solved explicitly. The process of this type was considered in [37] and is also used in [48] as an interest rate model. Its exact distribution can be found in [27].

[4] A version of this example has been submitted for publication in [33]

Now let $G(y) = (\sigma(1-\alpha)y + s^{1-\alpha})^{\frac{1}{1-\alpha}}$. Taking derivatives yields:

$$\begin{aligned} G'(y) &= \sigma G^{\alpha}(y) \\ G''(y) &= \alpha\sigma^2 G^{2\alpha-1}(y). \end{aligned}$$

Note, that $\alpha \geq 0.5$ implies $\frac{1}{1-\alpha} \geq 2$, therefore G is a C^2 function. Hence, one can use it in conjunction with Ito's formula.

Define $Z_t = G(Y_t)$. With that in mind the SDE for Y_t can be expressed as

$$dY_t = (rZ_t^{1-\alpha}/\sigma)dt + dW_t.$$

By construction $Z_0 = s$ and by Ito's formula

$$dZ_t = (rZ_t + \frac{\alpha\sigma^2}{2}Z_t^{2\alpha-1})dt + \sigma Z_t^{\alpha}dW_t$$

and (provided $Z_t^{2\alpha-1}$ is non-negative) the one-dimensional path-wise comparison theorem yields:

$$S_t \leq Z_t \qquad (a.s.), \tag{47}$$

therefore

$$e^{-rT}\mathbf{E}f(S_T) \leq e^{-rT}\mathbf{E}f(Z_T)$$

for non-decreasing pay-off function $f(x)$.

The process Z_t represents a function of normally distributed Y_t and allows for explicit calculations of expectations. Similar estimate can be derived for another well-known stochastic process, the Cox-Ingersoll-Ross interest rate model [10].

For the purpose of demonstration, consider a European call option on stock in the square-root model ($\alpha = 0.5$). In this case $Z_t^{2\alpha-1} = 1$ so comparison (47) holds. The upper bound for the call price can then be calculated as expectation

$$e^{-rT}\mathbf{E}((\sigma(a\xi+b)/2+\sqrt{s})^2-K)^+,$$

where $a=\sqrt{(e^{rT}-1)/r}$, $b=2\sqrt{s}(e^{rT/2}-1)/\sigma$ and ξ is a standard normal variable. The expression can be simplified to

$$C(T,s)=e^{-rT}\mathbf{E}((\sigma a\xi+e^{rT/2}\sqrt{s})^2-K)^+. \tag{48}$$

For numerical demonstration fix parameters $r=0.05$ and $s=20$. The results of using explicit pricing formula as in [9] and the proposed estimate (48) for two different values of σ are presented in tables 1 and 2. The spread between estimated and exact price, measured as estimate/(true value)-1 is also presented. As can be seen from the two tables, derived method of estimation is very accurate.

At the same time precision decreases with increasing σ. This result can be expected from the form of $\alpha\sigma^2/2$, which is the difference between equations describing dynamics of dZ_t and dS_t. Higher values of σ should result in larger difference between the stock price and its estimate, which can be observed from the two tables.

Similar estimation procedure can be used to build a hedging strategy. Fix time $u\le T$. Of course, (47) means that

$$e^{-rT}\mathbf{E}(f(S_T)|\mathcal{F}_u)\le e^{-rT}\mathbf{E}(f(Z_T)|\mathcal{F}_u)$$

where the right-hand side can be computed explicitly and used as an estimate for the discounted option price at time u. This, however, is not the best approach.

First, computed in this way estimate will depend on Y_t (or, equivalently, W_t). Since neither is observable, it can not be used for hedging purposes. Second reason is that precision can actually be improved.

Table 1: Call option price estimates and spreads for $\sigma = 0.35$ for Example 1

Maturity $T = 1$

strike K	22	20	15
estimated price	0.275	1.239	5.7614
exact price	0.2667	1.217	5.7316
percentage spread	3.14	1.82	0.52

Maturity $T = 1.5$

strike K	22	20	15
estimated price	0.5853	1.7088	6.128
exact price	0.5668	1.674	6.0838
percentage spread	3.27	2.08	0.73

Maturity $T = 2$

strike K	22	20	15
estimated price	1.8867	2.958	6.733
exact price	1.7656	2.8	6.511
percentage spread	6.86	5.7	3.4

Table 2: Call option price estimates and spreads for $\sigma = 0.7$ for Example 1

Maturity $T = 1$

Strike K	22	20	15
estimated price	0.884	1.84	5.87
exact price	0.8376	1.766	5.755
percentage spread	5.57	4.28	2.03

Maturity $T = 1.5$

Strike K	22	20	15
estimated price	1.3935	2.42	6.308
exact price	1.311	2.306	6.138
percentage spread	6.29	4.9	2.78

Maturity $T = 2$

Strike K	22	20	15
estimated price	1.8867	2.958	6.733
exact price	1.7656	2.8	6.511
percentage spread	6.86	5.7	3.4

On the time interval $[u, T]$ consider a process

$$\begin{aligned} dZ_t^{(u)} &= (rZ_t^{(u)} + \frac{\alpha\sigma^2}{2}(Z_t^{(u)})^{2\alpha-1})dt + \sigma(Z_t^{(u)})^{\alpha}dW_t \\ Z_u^{(u)} &= S_u. \end{aligned}$$

By the comparison theorem for any $u < t$

$$S_t \leq Z_t^{(u)} \qquad (a.s.),$$

therefore

$$e^{-r(T-u)}\mathbf{E}(f(S_T)|\mathcal{F}_u) \leq e^{-r(T-u)}\mathbf{E}(f(Z_T^{(u)})|\mathcal{F}_u)$$

But, as before $Z_t^{(u)}$ is equal to $(\sigma(1-\alpha)Y_t^{(u)} + S_u^{1-\alpha})^{\frac{1}{1-\alpha}}$, where $Y_t^{(u)}$ is a process satisfying

$$\begin{aligned} dY_t^{(u)} &= r[(1-\alpha)Y_t^{(u)} + \frac{S_u^{1-\alpha}}{\sigma}]dt + dW_t \\ Y_u^{(u)} &= 0 \end{aligned}$$

and its distribution (conditional on S_u) is normal. Therefore, $e^{-r(T-u)}\mathbf{E}(f(Z_T^{(u)})|\mathcal{F}_u)$ can be computed explicitly as $C(T-u, S_u)$, where $C(t, x)$ is defined in (48). This function can be used to estimate the option's price at any moment in time and to build a hedging strategy.

A few things should be noted at this point. First fact to consider is that constructed processes $Z_t^{(u)}$ are decreasing in u. To prove that consider $v > u$ and time interval $[v, T]$ where both $Z_t^{(u)}$ and $Z_t^{(v)}$ are defined. As proved above, $Z_v^{(u)} \geq S_v = Z_v^{(v)}$ so at time v the two processes satisfy to the necessary inequality. The corresponding SDEs for the two processes on $t \geq v$ are identical, therefore it remains to prove that the process Z_t is increasing in initial value. That can be achieved by looking at the definition of Z_t as an increasing function of Y_t and Z_0, while by comparison theorem Y_t itself is increasing in s. Thus, for any $t \geq v$

$$Z_t^{(u)} \geq Z_t^{(v)} \text{ (a.s.).} \tag{49}$$

This inequality immediately points to the fact that using processes $Z_T^{(u)}$ at time u instead of original Z_T to estimate the option's price gives improvement in precision. This fact is consistent with the numerical example presented in Tables 1 and 2, which show improved precision for lower maturities.

Second, by construction $C(0, x) = (x - K)^+$, therefore the strategy represents a perfect hedge. However, since different processes $Z_T^{(u)}$ are used as upper boundaries for S_T at different times the discounted option price estimate $e^{-ru}C(T-u, S_u)$ is not a martingale. This means that the strategy is not self-financing. In fact, inequality (49) can be used to prove that for non-decreasing pay-offs the proposed hedging strategy is a strategy with consumption.

To do so, denote the discounted strategy value $e^{-rt}C(T-t, S_t)$ by L_t and consider two times $v > u$. By construction

$$L_u = e^{-rT}\mathbf{E}(f(Z_T^{(u)})|\mathcal{F}_u)$$

and

$$L_v = e^{-rT}\mathbf{E}(f(Z_T^{(v)})|\mathcal{F}_v).$$

Therefore, using inequality (49) and properties of conditional expectations yields

$$\begin{aligned} \mathbf{E}(L_v|\mathcal{F}_u) &= e^{-rT}\mathbf{E}(f(Z_T^{(v)})|\mathcal{F}_u) \leq \\ &\leq e^{-rT}\mathbf{E}(f(Z_T^{(u)})|\mathcal{F}_u) = L_u. \end{aligned}$$

Thus, the process L_t is a supermartingale. It can be represented as a sum of a martingale (which in turn is a discounted value of some self-financing strategy) and a decreasing process. This second component can be understood as a consumption process, which continuously removes excess hedging capital

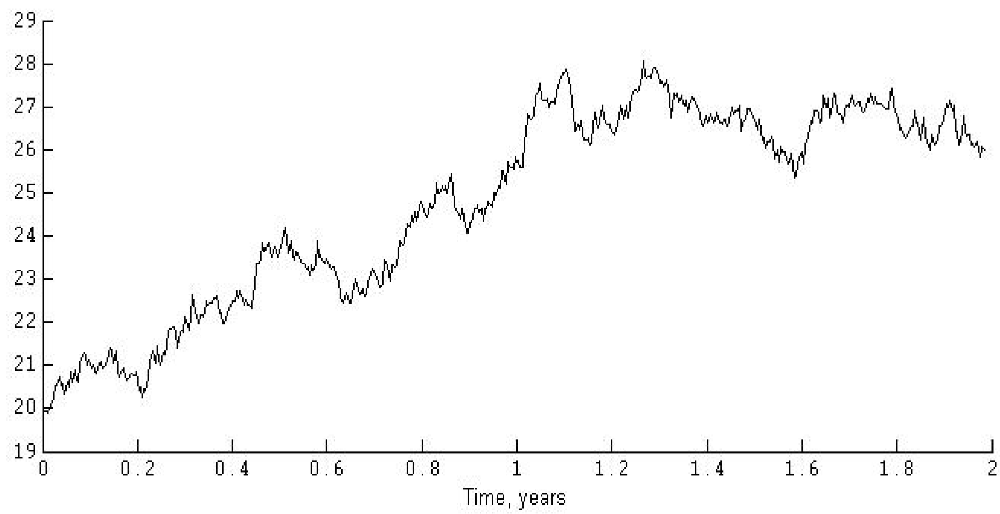

Figure 1: *Simulated sample CEV process with* $S_0 = 50$*,* $\alpha = 0.5$*,* $\mu = 0.14$ *and* $\sigma = 0.7$ *for Example 1*

from the strategy until it exactly matches the option's value at maturity.

Another topic to discuss in this set-up is how do estimation errors compare to those arising from model misidentification.

Consider the following scenario: an investor observes a CEV-distributed stock price process and wants to have explicit pricing and hedging formulas. One way is to use the method presented in this example. An alternative would be to assume that the observed process is actually log-normally distributed, which will result in pricing errors but provide means of explicit calculations. The question is: which method produces better results?

To follow the second proposed path investor will need to estimate the volatility of the observed stock price process as if it was log-normally distributed. Since the actual distribution is not log-normal, this estimated volatility will depend on the sample path. For the purpose of demonstration I used 50 simulations of the stock price path over a period of 2 years, one of which is presented in Figure 1. A scatter plot of annualized volatilities for all 50 simulations is presented in Figure 2.

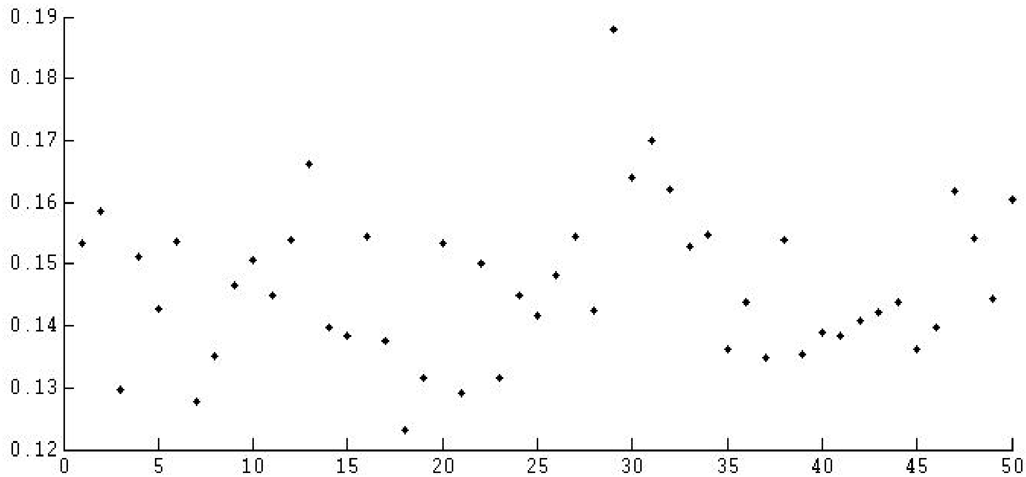

Figure 2: *Annualized volatilities of simulated CEV processes for Example 1*

As expected, simulation results produce a range of volatility parameters. Most of them lie in an interval $[13\%, 17\%]$. With this range of volatilities call option prices can be calculated using the Black-Scholes formula.

Looking at a cross section of results presented in Table 3 it can be said that usage of comparison theorems has a clear advantage over model misidentification for pricing in- and at-the money options. Not only in terms of precision, but also because the sing of error is known in advance. At the same time deep out-of-the money calls are better priced using model misidentification, especially since there is very little variability in calculated prices for different volatility values.

To summarize this example, application of comparison theorems allows to find an explicit price estimate for any European option in the constant elasticity of variance model, as well as build a perfect hedging (but not self-financing) strategy. The method presents a viable alternative to usage of numerical methods and simplified models for the purpose of pricing and hedging options.

End of example.

Table 3: Estimation errors versus model misidentification for Example 1

Maturity $T = 1$

Strike K	22	20	15
comparison theorem estimate	0.884	1.84	5.87
range of Black Scholes prices	0.66..0.97	1.57..1.866	5.73..5.756
exact price	0.8376	1.766	5.755

Maturity $T = 2$

Strike K	22	20	15
comparison theorem estimate	1.8867	2.958	6.733
range of Black Scholes prices	1.5..1.956	2.547..2.929	6.446..6.516
exact price	1.7656	2.8	6.511

Example 1 studies a well known CEV model and develops an explicit estimate for the European call option's price. Not to focus on one process, further examples will look at stock price models, not used in literature. Nevertheless, the models appear to be reasonable and allow to demonstrate the main ideas, which is the sole purpose of this part of my work.

One of the first comparison theorems is that of Hajek [18]. The result looks very natural when applied to mathematical finance: option prices for convex pay-offs (such as call option) are increasing functions of volatility. In fact, the work [19] provides extension of price monotonicity in volatility to path-dependent options.

The next example constructs a double sided estimate, derived with the help of two types of comparison theorems.

Example 2. *Two-sided estimate*[5]

Let the main process be described by

$$dS_t \quad = \quad rS_t dt + (aS_t + b\sqrt{S_t})dW_t$$

where a, b, and r are non-negative constants.

[5] A version of this example has been published in [31]

An upper estimate for S_t is given by a stochastic differential equation

$$
\begin{aligned}
dZ_t &= (rZ_t + (r + a^2/4)\frac{b}{a}\sqrt{Z_t} + b^2/4)dt + (aZ_t + b\sqrt{Z_t})dW_t \\
Z_0 &= S_0.
\end{aligned}
$$

By the one-dimensional comparison theorem $Z_t \geq S_t$ (a.s.)

The advantages of this particular estimate are derived from the fact that Z_t can be expressed as $Z_t = (\frac{b}{a} - X_t)^2$, where

$$
dX_t = (r/2 - a^2/8)X_t dt + a/2 X_t dW_t,
$$

which is log-normally distributed.

For the lower estimate use the process

$$
\begin{aligned}
dY_t &= rY_t dt + aY_t dW_t \\
Y_0 &= S_0.
\end{aligned}
$$

Since $ax + b\sqrt{x} \geq ax$ for any x, the mean comparison theorem implies that any convex pay-off option on S_t is priced above the same option on Y_t.

Altogether, if a pay-off function $h(x)$ is non-decreasing and convex then

$$
\mathbf{E}h(Y_t) \leq \mathbf{E}h(S_t) \leq \mathbf{E}h(Z_t).
$$

For call option price the lower bound is computed using the Black-Scholes formula, while the upper estimate is

$$
C(T, S_0) = e^{-rT}\mathbf{E}((\frac{b}{a} - (\sqrt{S_0} + \frac{b}{a})\exp[(\frac{r}{2} - \frac{a^2}{4})T + \frac{a}{2}\sqrt{T}\xi])^2 - K)^+ \quad (50)
$$

with ξ - a standard normal random variable.

As in the previous example one can look at the structure of both estimates to determine how precision is related to the parameters of the initial model. Clearly the quality of both estimates depends on the size of the parameter b. Smaller values of b result in less significant differences between stochas-

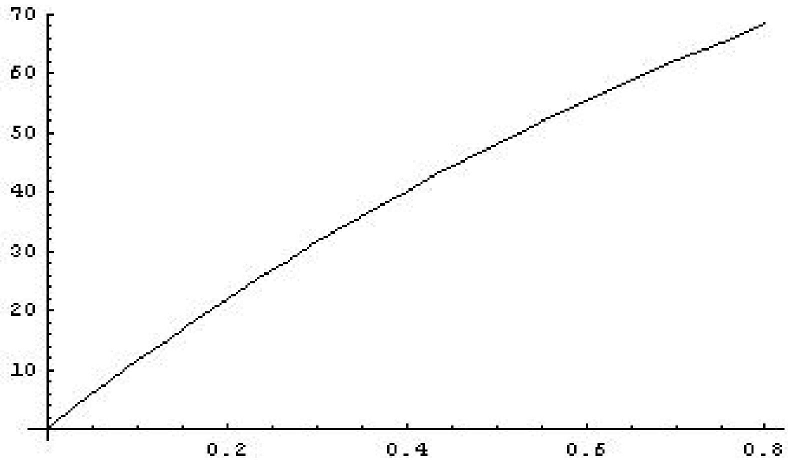

Figure 3: *Percentage spread for call option price estimates as a function of the parameter b with $S_0 = 50$, $K = 55$, $r = 0.04$, $T = 1$ and $a = 0.2$ for Example 2*

tic differential equations for the two estimates when compared with that of S_t. Therefore, one can conclude that relative percentage spread, measured as 2(upper-lower)/(upper+lower), is an increasing function of b and eventually decreases to 0 when $b = 0$ (since all three processes become log-normal in that case). This theoretical observation is supported by figure 3.

At the same time it is not entirely obvious how precision relates to the size of the parameter a. For the upper estimate Z_t the extra term $(r + a^2/4)\frac{b}{a}\sqrt{Z_t}$ is not monotone in a. The structure of the lower estimate Y_t suggests that higher values of a will result in better precision, since the foregone term $b\sqrt{S_t}$ will have relatively smaller impact.

As demonstrated by figure 4 relative percentage spread is a decreasing function of a. This fact can be attributed to improved precision effect of the lower estimate being overwhelming.

Finally, from figure 5 precision of estimation increases with K, suggesting that the method is best used for in the money call options.

As in example 1, the same methodology can be used to build two replicating strategies. For call option the upper estimate will be $C(T - t, S_t)$ defined in (50), while lower estimate is provided by the Black Scholes formula. Both will

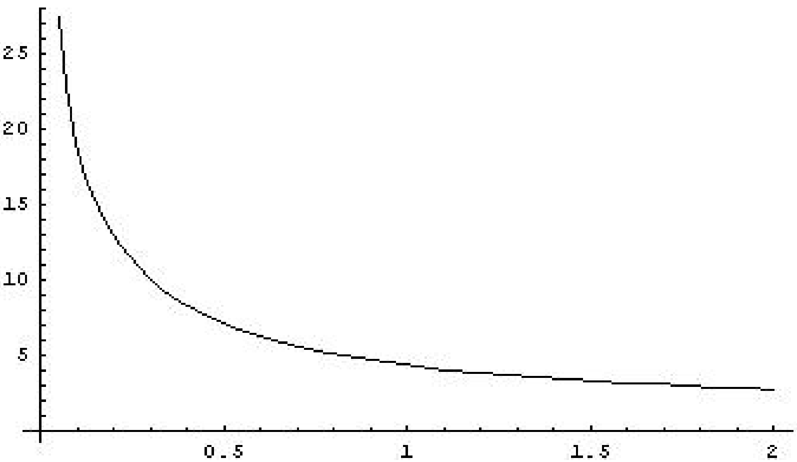

Figure 4: *Percentage spread for call option price estimates as a function of the parameter a with* $S_0 = 50$*,* $K = 47$*,* $r = 0.04$*,* $T = 1$ *and* $b = 0.25$ *for Example 2*

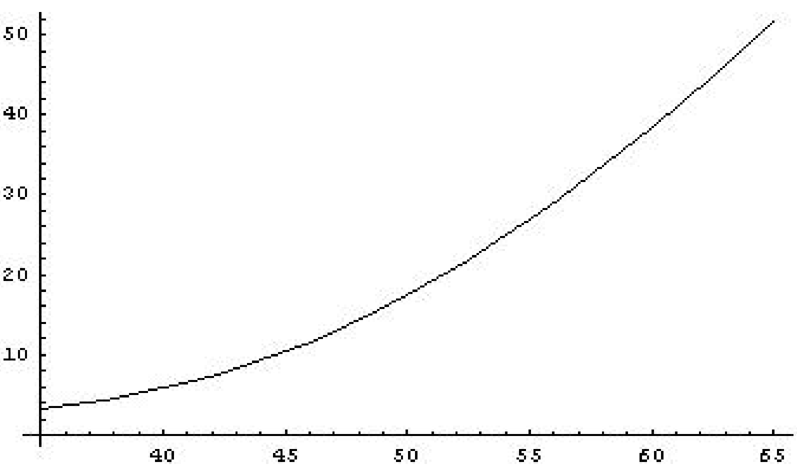

Figure 5: *Percentage spread for call option price estimates as a function of the strike price K with* $S_0 = 50$*,* $r = 0.04$*,* $T = 1$*,* $a = 0.2$ *and* $b = 0.25$ *for Example 2*

represent perfect hedges, but will not be self-financing.

End of example.

Two previous examples demonstrate usefulness of comparison theorems in option price estimations. The full range of possibilities was examined, showing advantages of estimates on specific examples. The next chapter will focus on how to build estimates in more general cases.

4.4 Estimating cumulative distributions

The previous part, and example 1 in particular, demonstrate that comparison theorems can be very powerful in practice. One criticism of such approach, however, is that it is far from uniform. Direct application of path-wise comparison theorem requires individualistic analysis of every model: both example 1 and 2 were constructed by finding a suitable process to dominate the stock price. In some cases this might be difficult.

In most practical applications estimating distribution of the process in question is sufficient. Mathematically speaking, the quantities of the type

$$\mathbf{E}I_{\{X_t \geq z\}}$$

can be of great interest. Risk management would be the first and most obvious application. Since risk measures, such as value at risk and conditional value at risk rely on knowledge of probability distribution, their estimates can be derived with the help of comparison methods. Option prices can also be estimated with the help of approximate distribution. Consider a put option on stock. Its fair price is given by

$$P = e^{-rT}\mathbf{E}(K - S_T)^+ = e^{-rT}\int_0^K (K - x)d\mathbf{P}(S_T \leq x).$$

Integration by parts formula yields

$$\begin{aligned} P &= -e^{-rT}K\mathbf{P}(S_1 \leq 0) + e^{-rT}\int_0^K \mathbf{P}(S_1 \leq x)dx = \\ &= e^{-rT}\int_0^K \mathbf{P}(S_1 \leq x)dx. \end{aligned} \tag{51}$$

This part of my thesis present two alternative methods of estimating distribution, which can be applied to a variety of stochastic processes.

4.4.1 Method of transformations[6]

The first approach is based on path-wise comparison theorem. Indeed, since $I_{\{x \geq z\}}$ is a monotone function, inequality of the type

$$X_t \leq Y_t \text{ (a.s.)}$$

can be used as a sufficient condition to ensure that

$$\mathbf{E} I_{\{X_t \leq z\}} \geq \mathbf{E} I_{\{Y_t \leq z\}}.$$

Unlike example 1 and 2, the method is more general. It will be presented for a specific process, but identical technique can be used in different settings.

Consider a stock price model

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t \frac{a + S_t}{b + S_t} dW_t \\ S_0 &> 0. \end{aligned}$$

An estimate for S_t will be constructed as a function $F(X_t)$ of a normally distributed process

$$\begin{aligned} dX_t &= \alpha dt + \sigma dW_t \\ X_0 &= 0 \end{aligned}$$

By Ito's formula

$$dF(X_t) = (\alpha F'(X_t) + \frac{1}{2}\sigma^2 F''(X_t))dt + \sigma F'(X_t) dW_t.$$

[6] A version of this chapter has been submitted for publication in [33]

The path-wise comparison theorem requires that both processes have the same diffusion coefficient. Therefore, $F(x)$ has to solve the following differential equation

$$y' = y\frac{a+y}{b+y} \tag{52}$$

In an implicit form the solution can be found to be

$$\frac{b}{a}\ln(y) + \frac{a-b}{a}\ln(y+a) - c = x,$$

which effectively represents $F^{-1}(y) = x$. Note that the inverse function is monotone. The constant c is calculated from initial condition $F^{-1}(S_0) = 0$.

The value of the parameter α is determined from inequality on drifts. For a lower estimate it is

$$\alpha y' + \frac{1}{2}\sigma^2 y'' \leq \mu y.$$

Second derivative can be calculated from the explicit expression for y' and is equal to

$$y'' = y\frac{(y^2+2by+ab)(a+y)}{(b+y)^3},$$

meaning that α has to satisfy

$$y[\alpha\frac{a+y}{b+y} + \frac{\sigma^2}{2}\frac{(y^2+2by+ab)(a+y)}{(b+y)^3}] \leq \mu y. \tag{53}$$

An upper estimate can be built in the same way with the inequality sign reversed.

The estimated distribution function can now be constructed. Since $F(X_t) \leq S_t$, then

$$\mathbf{P}(S_t \leq z) \leq \mathbf{P}(F(X_t) \leq z) = \mathbf{P}(X_t \leq F^{-1}(z)) = \Phi(\frac{F^{-1}(z) - \alpha t}{\sigma\sqrt{t}}). \quad (54)$$

Here $\Phi(x)$ denotes the standard normal cumulative distribution function.

Notice that without knowing the actual expression for $F(x)$ it is difficult to find the option price estimate as

$$\mathbf{E}f(Y_t) = \mathbf{E}f(F(X_t)).$$

However, estimating the stock's distribution is possible. The put option's price can then be approximated from (51) by numerically integrating (54). Quantities such as VaR can also be estimated using simple numerical procedures.

The method presented above can be applied to many different models, as long as (52) can be solved, either explicitly or implicitly.

4.4.2 Method of substitution

The second approach requires less restrictive condition and, therefore, is more general. It follows similar strategy as theorem 2. While not based on any comparison theorem per se, it is inspired by the work [4], which was briefly described in part 3.1.

To demonstrate the idea, consider a log-normally distributed process

$$X_t = X_0 \exp\left((a - \frac{b^2}{2})t + bW_t\right)$$

and define

$$v(x,t) = \mathbf{E}(I_{\{X_T > z\}} | X_t = x).$$

The function v can be computed explicitly to be

$$v(x,t) = \Phi(\frac{\ln(x/z) + (a - b^2/2)\tau}{b\sqrt{\tau}}),$$

where $\tau = T - t$. At the same time, by Ito's formula $v(x,t)$ satisfies

$$v_t(x,t) + axv_x(x,t) + \frac{b^2}{2}x^2 v_{xx}(x,t) = 0. \tag{55}$$

The leading idea in this methodology is to substitute $v(S_T, T)$ instead of $I_{\{S_T > z\}}$ and estimate

$$\mathbf{E}v(T, S_T).$$

Since by construction

$$v(x,T) = I_{\{x \geq z\}}$$

the two quantities have identical expectations. At the same time $v(x,t)$ is twice continuous differentiable with respect to x and differentiable with respect to t, which means that Ito's formula can be applied to $v(S_t, t)$.

Let S_t be given by

$$\begin{aligned} dS_t &= f(S_t)dt + g(S_t)dW_t \\ S_0 &= X_0. \end{aligned}$$

Using the definition of v and Ito's formula produces

$$\mathbf{E}I\{S_T > z\} = v(X_0, 0) +$$
$$+\mathbf{E}\int_0^T (v_t(S_t,t) + f(S_t)v_x(S_t,t) + \frac{1}{2}g(S_t)^2 v_{xx}(S_t,t))dt$$

and using (55) allows to get rid of the v_t term:

$$\mathbf{E}I\{S_t > z\} = \mathbf{E}I\{S_T > z\} +$$
$$+\mathbf{E}\int_0^T ((f(S_t) - aS_t)v_x(S_t,t) + \frac{1}{2}(g(S_t)^2 - b^2S_t^2)v_{xx}(S_t,t))dt. \tag{56}$$

Calculating exact expressions for v_x and v_{xx} yields:

$$\begin{aligned} v_x &= \frac{1}{bx\sqrt{\tau}}\varphi(\frac{\ln(x/z) + (a - b^2/2)\tau}{b\sqrt{\tau}}) \\ v_{xx} &= [-\frac{1}{bx^2\sqrt{\tau}} - \frac{1}{b^3x^2\tau\sqrt{\tau}}(\ln(x/z) + (a - b^2/2)\tau)]\varphi(), \end{aligned}$$

where $\varphi(x)$ is a standard normal density.

Altogether, the integrand in (56) can be expressed as

$$\frac{2b^2S_t\tau(f(S_t) - aS_t) - (\ln(\frac{S_t}{z}) + (a + b^2/2)\tau)(g^2(S_t) - b^2S_t^2)}{2b^3S_t^2\tau\sqrt{\tau}}\varphi().$$

To establish comparison between $\mathbf{P}(S_T > z)$ and explicitly calculated $v(X_0, 0)$ it is sufficient to ensure that the integral in (56) is non-negative (or for a different sided comparison non-positive) regardless of the value of Y_t. It will be so if

$$2b^2x\tau(f(x) - ax) - (\ln(x/z) + (a + b^2/2)\tau)(g^2(x) - b^2x^2)$$

is constant in sign for all $0 \le \tau \le T$ and x.

Further manipulations reduce the above expression to

$$\tau[2b^2x(f(x)-ax)-(a+b^2/2)(g^2(x)-b^2x^2)]-\ln(\frac{x}{z})(g^2(x)-b^2x^2).$$

When functions f and g do not depend on t it can be useful to separate the term with τ. It means that two terms

$$\begin{aligned}2b^2xf(x)-(a+b^2/2)g^2(x)+x^2(b^4/2-ab^2)\\ -\ln(\frac{x}{z})(g^2(x)-b^2x^2)\end{aligned}$$

have to be of the same sign for all x.

First conclusion to be made is that the expression $(g^2(x)-b^2x^2)$ has to change sign as x goes from $< z$ to $> z$. Continuity implies that $g(z)=bz$ which can be used to determine the only applicable value of $b=g(z)/z$. With that in mind

$$\ln(\frac{x}{z})x^2(g^2(x)/x^2-g^2(z)z^2)$$

will be constant in sign if $g(x)/x$ is monotone.

The first inequality can then be solved to find the value of the parameter a. The result can be summarized in the form of the following proposition.

Proposition 3. Let S_t be a non-negative process such that

$$dS_t \;=\; f(S_t)dt+g(S_t)dW_t.$$

If $g(x)/x$ is a non-increasing function and $f(x)$ is such that

$$a=\inf_{x\geq 0}\frac{2f(x)xg^2(z)-g^2(z)(g^2(x)z^2-g^2(z)x^2)/(2z^2)}{g^2(x)z^2+g^2(z)x^2}>-\infty$$

then

$$\mathbf{P}(S_T > z) \geq \Phi(\frac{\ln(S_0/z) + (a - 0.5g^2(z)/z^2)T}{g(z)\sqrt{T}} z). \tag{57}$$

Proof. With a as defined and $g(x)/x$ non-increasing inequalities

$$\begin{aligned} 2b^2xf(x) - (a + b^2/2)g^2(x) + x^2(b^4/2 - ab^2) &\geq 0 \\ -\ln(\frac{x}{z})(g^2(x) - b^2x^2) &\geq 0 \end{aligned} \tag{58}$$

will hold for $b = g(z)/z$, which will produce the desired estimate. □

If $g(x)/x$ is increasing, the other sided estimate can be constructed in the same fashion.

4.4.3 Numerical demonstrations

Now both methods will be demonstrated, and compared, on numerical examples.

Example 3. Consider a stock price model with risk-neutral dynamics as

$$\begin{aligned} dS_t &= 0.06S_tdt + 0.35S_t\frac{30 + S_t}{20 + S_t}dW_t \\ S_0 &= 37. \end{aligned}$$

The first approach allows to build two estimates. Inequality (53) is satisfied with

$$\alpha_1 = \mu\frac{b}{a} - \frac{\sigma^2 a}{2b} = -0.051875$$

which gives an upper estimate

$$\mathbf{P}(S_t \le z) \le \Phi(\frac{2\ln[z\sqrt{z+30}/(37\sqrt{57})]/3 + 0.051875t}{0.35\sqrt{t}}) = F_U(z)$$

Similarly, other sided estimate can be built. For this inequality (53) has to be reversed, and will be satisfied with

$$\alpha_2 = \mu - \frac{\sigma^2}{2} = -0.00125.$$

A lower estimate then is

$$\mathbf{P}(S_t \le z) \ge \Phi(\frac{2\ln[z\sqrt{z+30}/(37\sqrt{57})]/3 + 0.00125t}{0.35\sqrt{t}}) = F_D(z).$$

The second method can be used to obtain another estimate from above. Here $g(x)/x = \sigma\frac{30+x}{20+x}$ which is a decreasing function so the first condition of proposition 3 holds. Denote the ratio $\frac{30+x}{20+x}$ by $h(x)$ for short. The second condition can, after some simplification, be expressed as

$$a = -\sigma^2 h^2(z)/2 + \inf_{x\ge 0} \frac{2\mu + \sigma^2 h^2(z)}{1 + h^2(x)/h^2(z)} > -\infty$$

which is clearly satisfied. Since $h(x)$ is decreasing, minimum will be attained at $x = 0$, thus

$$a = -0.06125h^2(z) + \frac{0.12 + 0.1225h^2(z)}{1 + 2.25/h^2(z)}. \tag{59}$$

The distribution of S_t can then be estimated from above by

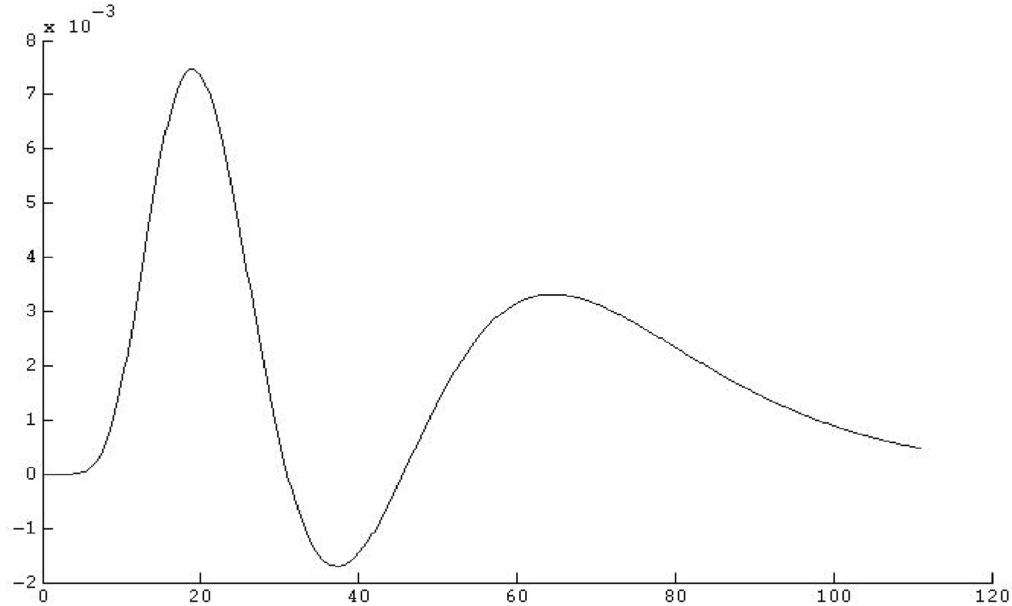

Figure 6: *Difference between two upper estimates for Example 3*

$$\mathbf{P}(S_T \le z) \le 1 - \Phi(\frac{20+z}{30+z}\frac{\ln(S_0/z) + (a - 0.06125(\frac{30+z}{20+z})^2)T}{0.35\sqrt{T}}) = F_A(z). \quad (60)$$

Figure 6 demonstrates the difference between the two upper estimate, calculated as $F_A(z) - F_U(z)$. It shows that for most values of z the first method produces better estimate, although using $\min\{F_A(z), F_U(z)\}$ is advantageous. Another graph, figure 7, measures the spread between the upper and the lower estimates, calculated as $2\frac{F_U(z)-F_D(z)}{F_U(z)+F_D(z)}$.

Using the upper and lower estimates for the stock price distribution, put option prices can be approximated using (51) and numerical integration. Results for several options are presented in table 4. The exact prices, found as numerical solutions of Black-Scholes type partial differential equation

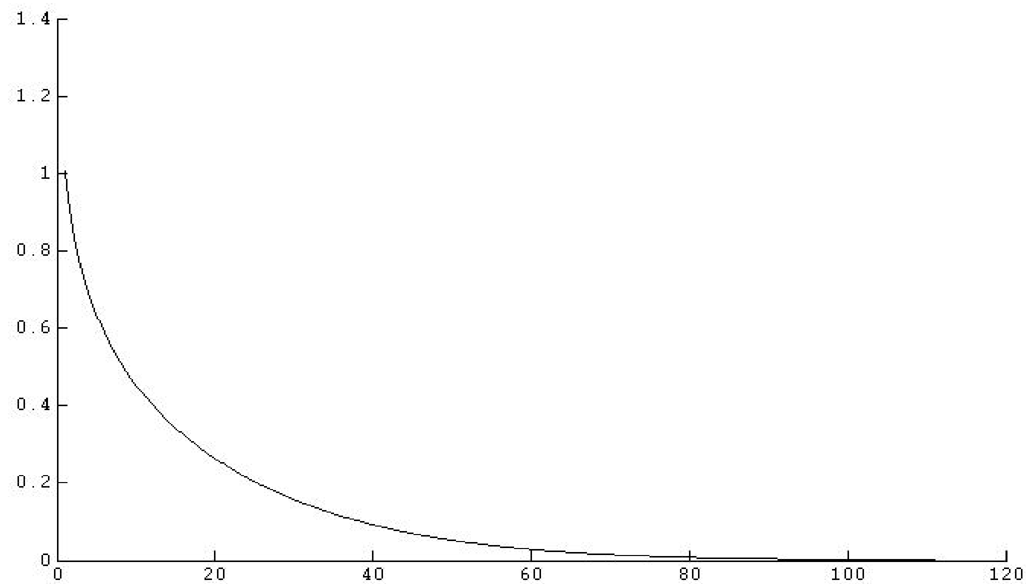

Figure 7: *Percentage spread between upper and lower estimates for Example 3*

$$
\begin{aligned}
\frac{\partial P}{\partial t} + 0.06125(x\frac{30+x}{20+x})^2\frac{\partial^2 P}{\partial x^2} &+ 0.06x\frac{\partial P}{\partial x} = rP \\
P(T,x) &= (K-x)^+
\end{aligned}
$$

are also presented. As can be seen from those numbers, the methods produce reasonably accurate estimates.

At the same time, put option pay-off is concave, meaning that the mean comparison theorem can be used in this example as well. Since

$$0.35x \le 0.35x\frac{30+x}{20+x} \le 0.525x$$

the theorem yields

$$P_{BS}(37,1,0.35,K) \le P \le P_{BS}(37,1,0.525,K)$$

where $P_{BS}(S_0,T,\sigma,K)$ is a Black Scholes put option price. Results of such calculations are also presented in table 4. One can see that in this case the mean comparison theorem is not accurate enough. Option price estimates derived using the distribution estimation methodology are more accurate.

In practice one can improve precision by making certain assumptions about the lowest attainable stock price. For example, assuming that $S_t \ge 10$ for $t \le 1$, parameters α and a of the upper estimates need to be chosen to satisfy (53) and (58) on $[10,\infty)$. With this range of stock prices the put option price estimate for maturity 1 and strike 32 will be 2.8775 and overpricing error is reduced to just 5.4%. While not 100% correct mathematically, this price estimate (or the way it was obtained to be precise) can make sense in practice.

End of example.

Two methods of estimating cumulative distributions were presented above.

Table 4: Put option price estimates and spreads for Example 3

Maturity $T = 0.5$

Strike K	32	37	43
exact price	1.6641	3.6791	7.2503
estimated prices	1.599	3.5708	7.0919
	1.9143	4.0642	7.4027
percentage spreads	3.91	2.94	2.18
	15.04	10.47	2.10
mean comparison estimates	0.92	2.54	5.79
	2.22	4.26	7.6

Maturity $T = 1$

Strike K	32	37	43
exact price	2.7302	4.8269	8.1472
estimated prices	2.5724	4.6146	7.8671
	3.1273	5.4324	8.99
percentage spreads	5.78	4.40	3.44
	14.54	12.54	10.34
mean comparison estimates	2.00	3.99	7.31
	4.03	6.41	9.89

Both approximate the stock price process with the help of normally distributed process. However, the two methods do so in different ways. As a result, they rely on different conditions, provide different results and, more importantly, work best in different scenarios.

The next example is designed to demonstrate some of those differences. It is build for a stock price model that works quite well with one method, but not very well with the other.

Example 4. Let the stock price process satisfy

$$\begin{aligned} dS_t &= 0.15 S_t dt + 0.25 S_t (2 - 1.05^{-S_t}) dW_t \\ S_0 &= 30. \end{aligned}$$

Estimating this process with the method of transformations, derived in 4.4.1 would be very difficult. It would require solving a differential equation

$$y' = 0.25 y (2 - 1.05^{-y}),$$

which can only be done numerically. Using this estimate in applications would in turn mean a sequence of numerical procedures.

At the same time the method of substitution as in 4.4.2 can be used to build a simple estimate. Using proposition 3, $g(x)/x = 0.25(2 - 1.05^{-x})$ is an increasing function. Therefore a lower estimate for the distribution of S_t can be constructed if the parameter a is chosen as

$$\begin{aligned} a &= \sup_{x \geq 0} \frac{2f(x)xg^2(z) - g^2(z)(g^2(x)z^2 - g^2(z)x^2)/(2z^2)}{g^2(x)z^2 + g^2(z)x^2} = \\ &= -0.03125(2 - 1.05^{-z})^2 + \sup_{x \geq 0} \frac{0.5 + 0.0625(2 - 1.05^{-z})^2}{1 + (2 - 1.05^{-x})^2 (2 - 1.05^{-z})^{-2}} = \\ &= -0.03125(2 - 1.05^{-z})^2 + \frac{0.5 + 0.0625(2 - 1.05^{-z})^2}{1 + (2 - 1.05^{-z})^{-2}}. \end{aligned}$$

The distribution of S_t can be estimated by

$$\mathbf{P}(S_T \leq z) \geq 1 - \Phi(\frac{\ln(S_0/z) + (a - 0.03125(2 - 1.05^{-z})^2)T}{0.25(2 - 1.05^{-z})\sqrt{T}}) = F(z, T). \quad (61)$$

As in example 3, this estimate can be used for approximate pricing of European options. Another application of distribution estimation is risk measures.

One can estimate Value at Risk for S_t. Indeed, a 5% VaR for S_1 is defined as

$$\{z : \mathbf{P}(S_1 \leq z) = 0.05\}$$

hence

$$\{z : \mathbf{P}(S_1 \leq z) = 0.05\} \leq \{z : F(z, 1) = 0.05\}.$$

The estimate is obtained by solving $F(z, 1) = 0.05$ numerically to yield

$$VaR_{5\%}(S_1) \leq 18.47.$$

Another numerical procedure suggests that the exact value of $VaR_{5\%}(S_1)$ is 14.8 which demonstrates a decent precision of the obtained estimate.

End of example.

4.5 Other applications

4.5.1 Multidimensional processes comparison

The purpose of this example is to demonstrate on a particular process the mechanics of the multidimensional comparison theorem. The multidimensional case was studied separately and it is therefore important to devote at least one specific example of its applications.

Consider a 2-dimensional process (X_t, Y_t) which satisfies a system of SDEs:

$$\begin{aligned} dX_t &= \frac{X_t(1-X_t)}{Y_t(X_t+Y_t)}dt + 2X_t dW_t \\ dY_t &= X_t dt + 2Y_t dW_t \\ (X_0, Y_0) &= (2, 2). \end{aligned}$$

Assuming both X_t and Y_t are non-negative, an upper estimate for this 2-dimensional process will be provided by a pair (A_t, B_t) defined by

$$\begin{aligned} dA_t &= -\frac{A_t^3}{A_t B_t + B_t^2}dt + 2A_t dW_t \\ dB_t &= \frac{A_t B_t}{B_t + A_t}dt + 2B_t dW_t \\ (A_0, B_0) &= (2, 2). \end{aligned}$$

The first claim of this example is that for any t the following holds

$$\begin{aligned} X_t &\geq A_t \\ Y_t &\geq B_t \end{aligned}$$

with probability 1.

According to the multidimensional comparison theorem it is necessary to show that

$$\frac{X_t(1-X_t)}{Y_t(X_t+Y_t)} \geq -\frac{A_t^3}{A_tB_t+B_t^2}$$

whenever $A_t = X_t$ and $Y_t \geq B_t$, and

$$X_t \geq \frac{A_tB_t}{B_t+A_t}$$

whenever $A_t \leq X_t$ and $Y_t = B_t$.

First notice that $-\frac{a^3}{b(a+b)}$ is an increasing function with respect to b and

$$-\frac{a^3}{b(a+b)} \leq -\frac{a^3}{y(a+y)} \text{ when } y \geq b.$$

The knowledge of relationship between B_t and Y_t is important at this stage, demonstrating that 2-dimensional processes can be compared using multidimensional theorem only.

The first claim then follows from a simple inequality

$$-x^3 \leq x(1-x)$$

for any non-negative x.

Similarly for the second claim start with the fact that $\frac{ab}{b+a}$ is increasing in a and

$$\frac{ab}{b+a} \leq \frac{xb}{b+x} \text{ when } x \geq a.$$

The second inequality in this chain is

$$\frac{xy}{y+x} \leq x$$

for non-negative x.

The system

$$
\begin{aligned}
dA_t &= -\frac{A_t^3}{A_tB_t + B_t^2}dt + 2A_t dW_t \\
dB_t &= \frac{A_tB_t}{B_t + A_t}dt + 2B_t dW_t \\
(A_0, B_0) &= (2, 2)
\end{aligned}
$$

has an advantage over the original one in that it can be solved explicitly to

$$
\begin{aligned}
A_t &= \frac{t+2}{t+1}\exp[2W_t - t] \\
B_t &= (t+2)\exp[2W_t - t].
\end{aligned}
$$

4.5.2 Application to theoretical properties studies

Another area where comparison of stochastic processes can be used is modelling. With its help one can design a several asset model with specific relationships between prices. The simplest example is the required non-negativity of stock prices. While this particular feature may be easily imposed, it is not clear how to ensure more complicated conditions are satisfied. The next two examples are derived to demonstrate the advantages provided by comparison theorems in this respect.

Example 5. Suppose that a model with two processes X_t and Y_t needs to be derived such that $X_t^2 + Y_t^2 \leq 1$ (a.s.) for any time t. Let X and Y be solutions of

$$\begin{aligned} dX_t &= f_1(X_t, Y_t)dt + g_1(X_t)dW_t \\ dY_t &= f_2(X_t, Y_t)dt + g_2(Y_t)dW_t. \end{aligned}$$

Applying the comparison theorem with $F(x, y) = x^2 + y^2 - 1$ results in the following condition

$$\begin{aligned} &2(x^2 + y^2 - 1)(2xf_1(x, y) + 2yf_2(x, y)) + (2xg_1(x) + 2yg_2(y))^2 + \\ &\qquad +2(x^2 + y^2 - 1)(g_1^2(x) + g_2^2(y)) \leq C(x^2 + y^2 - 1)^2, \end{aligned}$$

for constant C whenever $1 \leq x^2 + y^2 \leq 1 + \delta$ for some positive δ.

The first step in finding appropriate f_1, f_2, g_1 and g_2 is to look at the necessary condition for the quadratic form: if $x^2 + y^2 = 1$ then the above inequality becomes

$$4(xg_1(x) + yg_2(y))^2 \leq 0,$$

meaning that $xg_1(x) + yg_2(y)$ should be equal to 0 whenever $x^2 + y^2 = 1$. One way to achieve that is to make $g(x) = x(1 - x^2) = -g_2(x)$.

With this particular choice of diffusion coefficients the term $xg_1(x) + yg_2(y)$ can be simplified to $(x^2 + y^2 - 1)(y^2 - x^2)$

The main inequality now becomes

$$2(x^2+y^2-1)(2xf_1(x,y)+2yf_2(x,y)+x^2(1-x^2)^2+y^2(1-y^2)^2)+\\ +4(x^2+y^2-1)^2(y^2-x^2)^2 \le C(x^2+y^2-1)^2.$$

Dividing by (x^2+y^2-1) and letting x^2+y^2 approach 1 yields:

$$2xf_1(x,y)+2yf_2(x,y)+x^2y^4+x^4y^2 \le 0$$

whenever $x^2+y^2=1$. This can be simplified further to

$$2xf_1(x,y)+2yf_2(x,y)+x^2y^2 \le 0.$$

The above inequality holds when $f_1(x,y)=-xy^2/4$ and $f_2(x,y)=-x^2y/4$. With this choice of drift coefficients the main inequality is

$$2(x^2+y^2-1)(-x^2y^2+x^2(1-x^2)^2+y^2(1-y^2)^2)+\\ +4(x^2+y^2-1)^2(y^2-x^2)^2 \le C(x^2+y^2-1)^2,$$

which can be simplified to

$$2(x^2+y^2-1)^2[3x^4-5x^2y^2-x^2-y^2+3y^4+4(y^2-x^2)^2] \le C(x^2+y^2-1)^2.$$

The expression in the square brackets is clearly bounded if $x^2+y^2 \le 1+\delta$, therefore the comparison theorem holds and $X_t^2+Y_t^2 \le 1$ (a.s.) when the initial conditions satisfy the same inequality.

End of example.

The next example deals with a more common problem in finance which is bid-ask spread modelling. A typical approach in this setting would be to consider either constant or proportional spread. The proposed model is rather simple and, at the same time, allows for a stochastic spread.

Example 6. *Bid-Ask spread*

Consider a market with transaction costs where bid and ask prices are described by:

$$dS_t^a = (aS_t^a + bS_t^b + \alpha)dt + \sigma S_t^a dW_t$$
$$dS_t^b = (cS_t^a + dS_t^b + \alpha)dt + \sigma S_t^b dW_t.$$

It is quite easy to show that the necessary condition $S_t^a \geq S_t^b$ will hold if $a + b \geq c + d$. The above inequality will hold if for some constant K and all $x \geq y$

$$2(x-y)(cx + dy - ax - by) + \sigma^2(x-y)^2 \leq K(x-y)^2.$$

Let $K = c - a$. Then for $d - b \leq -K_1$ and, since x and y are non-negative,

$$(c-a)x + (d-b)y \leq -K_1 y + K_1 x = K_1(x-y).$$

Multiplying by $2(x-y)$ and adding $\sigma^2(x-y)^2$ to both sides one gets:

$$2(x-y)((c-a)x + (d-b)y) + \sigma^2(x-y)^2 \leq (2K_1 + \sigma^2)(x-y)^2$$

and making $K = 2K_1 + \sigma^2$ the desired inequality is obtained. The same technique can be used to show that if $S_0^a = S_0^b$ and $a + b = c + d$ then for any t $S_t^a = S_t^b$ (a.s.).

End of example.

When dealing with incomplete markets there exists more than one martingale measure and (generally) more than one arbitrage-free option price. Dependence of price on a particular choice of martingale measure is the question that can be studied. As a demonstration one can cite the work [19], where jump-diffusion model is studied along with different martingale measures, such as minimal entropy measure and others, showing what measures produce the smallest and the highest option prices. This theoretical result is derived solely from models properties and does not rely on any computations. The idea of using theoretical properties of the model to establish comparisons between option prices under different measures can also be found in [20] for stochastic volatility models.

Example 7. *Option on a non-traded asset*

Consider a market with two assets:

$$\begin{aligned} dX_t &= \alpha_t dt + \beta(t, X_t) dW_t^1 \\ dS_t &= S_t(\mu dt + \sigma_1 dW_t^1 + \sigma_2 dW_t^2) \end{aligned}$$

of which only S_t is traded, and an option $f(X_T)$. Such a market is incomplete, because there is only one equation identifying two parameters of a martingale measure (X_t is not traded and so does not have to be a martingale).

One way to look at pricing such an option is to add a second traded asset to make the market complete. Different assets, however, will yield different martingale measures and different option prices.

More precisely, consider a market completion:

$$d\bar{S}_t = \bar{S}_t(\bar{\mu} dt + \bar{\sigma}_1 dW_t^1 + \bar{\sigma}_2 dW_t^2)$$

The unique martingale measure $\bar{P}$ is defined by:

$$\begin{aligned} d\bar{W}^1 = dW^1 &+ \gamma_1 dt \\ d\bar{W}^2 = dW^2 &+ \gamma_2 dt \\ \sigma_1\gamma_1 + \sigma_2\gamma_2 + \mu &= r \\ \bar{\sigma}_1\gamma_1 + \bar{\sigma}_2\gamma_2 + \bar{\mu} &= r, \end{aligned}$$

from where: $\gamma_1 = \frac{(r-\mu)\bar{\sigma}_2 - (r-\bar{\mu})\sigma_2}{\sigma_1\bar{\sigma}_2 - \sigma_2\bar{\sigma}_1}$

Under this measure the dynamics of X_t can be expressed as:

$$dX_t = (\alpha_t - \gamma_1\beta(t, X_t))dt + \beta(t, X_t) d\bar{W}_t^1.$$

One can see that changing martingale measures (or additional assets) is equivalent to changing the drift of the main process X_t. Provided $\beta(t, x) > 0$

one can apply the comparison theorem to processes

$$dX_t^1 = (\alpha_t - \gamma\beta(t, X_t^1))dt + \beta(t, X_t^1)d\bar{W}_t^1$$

and

$$dX_t^2 = (\alpha_t - \gamma'\beta(t, X_t^2))dt + \beta(t, X_t^2)d\bar{W}_t^1$$

and see, that $\gamma < \gamma'$ implies $X_t^1 \geq X_t^2$.

But the distribution of X_t coincides with that of X_t^1 (or X_t^2) if $\gamma_1 = \gamma$ (or $\gamma_1 = \gamma'$ respectively). Therefore decreasing γ_1 is equivalent to increasing the option's price (for an increasing pay-off function).

End of example.

References

[1] Andersen, L., Andreasen, J.: Jump-diffusion processes: volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, **4(3)**, 231-262 (2000)

[2] Bassan B., Çinlar E., Scarsini M.: Stochastic Comparison of Itô Processes., *Stochastic Processes and their Applications,* **45**, 1-11 (1993)

[3] Bergenthum J., Ruschendorf L.: Comparison of Option Prices in Semimartingale Models. *Finance and Stochastic* **10**, 222-249 (2006)

[4] Bergenthum J., Ruschendorf L.: Comparison of Semimartingales and Lévy Processes. *Annals of Probability,* **35(1)**, 228-254 (2007)

[5] Black F.: The holes of Black-Scholes. *RISK,* **1(4)**, 30-33 (1988)

[6] Black F., Scholes M.: Pricing of Options on Corporate Liabilities. *Journal of Political Economy,* **81**, 637-659 (1973)

[7] Boyle P.: Options: A Monte Carlo Approach. *Journal of Financial Economics,* **4(3)**, 323-338 (1997)

[8] Boyle P., Broadie M., Glasserman P.: Monte Carlo Methods for Security Pricing. *Journal of Economic Dynamics and Control,* **21**, 1267-1321 (1997)

[9] Cox, J.C., Ross, S.A.: The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics*, **3**, 145-166 (1976)

[10] Cox, J.C., Ingersoll J.E., Ross, S.A.: A Theory of Term Structure of Interest Rates. *Econometrica*, **53(2)**, 385-407 (1985)

[11] Ding X., Wu R.: A New Proof for Comparison Theorems for Stochastic Differential Inequalities with Respect to Semimartingales. *Stochastic Processes and their Applications,* **78**, 155-171 (1998)

[12] Doléans-Dade C.: On the Existence and Unicity of Solutions of Stochastic Integral Equations. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* **36**, 93-101 (1976)

[13] Gal'chuk L.: Existence and uniqueness of a solution for stochastic equations with respect to semimartingales. *Theory Probability and Applications,* **23(4)**, 751-763 (1979)

[14] Gal'chuk L.I.: A Comparison Theorem for Stochastic Equations with Integral with Respect to Martingales and Random Measures. *Theory of Probability and Applications,* **27(3)**, 450-460 (1982)

[15] Gal'chuk, L.I., Davis M.H.A.: A Note on a Comparison Theorem for Equations with Different Diffusions. *Stochastics*, **6(2)**, 147-149, (1982)

[16] Geiss C., Manthey R.: Comparison Theorems for Stochastic Differential Equations in Finite and Infinite Dimensions. *Stochastic Processes and their Applications,* **53**, 23-35 (1994)

[17] Gyöngy I., Krylov N.: On stochastic equations with respect to semimartingales I. *Stochastics,* **4**, 1-21 (1980)

[18] Hajek B.: Mean Stochastic Comparison of Diffusions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* **68**, 315-329 (1985)

[19] Henderson V.: Price Comparison Results and Superreplication: an Application to Passport Options. *Applied Stochastic Models in Business and Industry,* **16**, 297-310 (2000)

[20] Henderson V., Hobson D., Howison S., Kluge, T.: A Comparison of Option Prices under Different Pricing Measures in a Stochastic Volatility Model with Correlation. *Review of Derivatives Research,* **8(1)**, 5-25 (2005)

[21] Heston, S.L.: A Closed-form Solution for Options with Stochastic Volatility and Applications to Bond and Currency Options. *The Review of Financial Studies*, **2(2)**, 327-343 (1993)

[22] Hobson, D.G., Rogers, L.C.G.: Complete Models with Stochastic Volatility. *Mathematical Finance*, **8(1)**, 27-48 (1998)

[23] Hull, J., White, A.: The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, **42(2)**, 281-300 (1987)

[24] Jacka S. D.: A Martingale Representation Result and an Application to Incomplete Financial Markets. *Mathematical Finance,* **2(4)**, 239-250 (1992)

[25] Jacod J.: Une condition d'existence et d'unicite pour les solutions fortes d 'equations. differentielles stochastiques. *Stochastics,* **4**, 23-38 (1980)

[26] Jacod J., Shiryaev A. Limit Theorems for Stochastic Processes. Berlin ; New York : Springer-Verlag (1987)

[27] Jamshidian, F: An Exact Bond Option Formula. *The Journal of Finance*, **44(1)**, 205-209 (1989)

[28] Kazamaki N.: On the existence of solutions of martingale integral equations. *Tohoku Mathematical Journal* **2**, 463-468 (1972)

[29] Karoui N., Jeanblanc-Picque M., Shreve S. E.: Robustness of the Black and Scholes Formula. *Mathematical Finance* **8(2)**, 93-126 (1998)

[30] Krasin V., Melnikov A.: On Comparison Theorem and its Applications to Finance, in *Optimality and Risk - Modern Trends in Mathematical Finance*, Springer-Verlag, 171-181 (2009)

[31] Krasin V.,Melnikov A.: Techniques for Pricing of Derivatives in the Theory of Mathematical Finance, *Survey in Applied and Industrial Mathematics*, **16(6)**, 994-1008 (2009)

[32] Krasin V., Melnikov A.: On Existence, Uniqueness and Comparison of Solutions of Stochastic Differential Equations, (2010), submitted

[33] Krasin V., Melnikov A.: On Option Price Estimation via Comparison of Stochastic Processes, (2010), submitted

[34] Levy H.: Upper and Lower Bounds of Put and Call Option Value: Stochastic Dominance Approach, *Journal of Finance* **40** (1985) 1197-1217.

[35] Melnikov A.V.:, On the Theory of Stochastic Equations in Components of Semimartingales. *Mat.Sb.,* **110(3)**, 414-427 (1979)

[36] Melnikov A.V.: On Solutions of Stochastic Equations with Driving Semimartingales. *Proceedings of the Third European Young Statisticians Meeting, Catholic University, Leuven,* 120-124 (1983)

[37] Merton, R.C.: Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory*, **3**, 373-413 (1971)

[38] Merton, R.C.: Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics* **3**, 125-144 (1976)

[39] O'Brien Ci. L.: A new Comparison Theorem for Solution of Stochastic Differential Equations, *Stochastics,* **3**, 245-249 (1980)

[40] Peng S., Zhu X.: Necessary and Sufficient Condition for Comparison Theorem of 1-dimensional Stochastic Differential Equations. *Stochastic Processes and their Applications,* **116**, 370-380 (2006)

[41] Platen R., Heath D.: A Benchmark Approach to Quantitative Finance. Springer-Verlag, Berlon-Heidelberg (2006)

[42] Protter P.E.: On the Existence, Uniqueness, Convergence, and Explosions of solutions of Systems of Stochastic Integral Equations. *Annals of Probability,* **5(2)**, 243-261 (1977)

[43] Rubinstein, M.: Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 Through August 31, 1978 *The Journal of Finance*, **40(2)**, 455-480 (1985)

[44] Rubinstein, M.: Implied Binomial Trees. *The Journal of Finance*, **49(3)**, 771-818 (1994)

[45] Schepper A.D., Heijnenm B.: Distribution-free option pricing, *Insurance Mathematics and Economics* **40**, 179-199 (2007)

[46] Situ R.: Reflecting Stochastic Differential Equations with Jumps and Applications. *Boca Raton : Chapman Hall/CRC* (2000)

[47] Skorokhod A.V.: Studies in the Theory of Random Process (1961) (in English: *Addison-Wesley, Reading, MA*, 1965)

[48] Vasicek O. A.: An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics,* **5(2)**, 177-188 (1977)

[49] Watanabe S., Yamada T.: On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics Kyoto University,* **11**, 155-167 (1971)

[50] Yamada T.: On Comparison Theorem for Solutions of Stochastic Differential Equations and its Applications *J. Math. Kyoto Univ.,* **13**, 497-512 (1973)

[51] Yamada T.: Sur une construction des solutions d'equations differentielles stochastiques dans le cas non-lipschitzien. *Séminaire de Probabilités,* **13**, 114-131 (1978)