Optional Processes and their Applications in Mathematical Finance, Risk Theory and Statistics

by

Andrey Pak

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematical Finance

Department of Mathematical and Statistical Sciences

University of Alberta

© Andrey Pak, 2021

Abstract

This thesis is dedicated to the study of the general class of random processes, called optional processes, and their various applications in Mathematical Finance, Risk Theory, and Statistics.

First, different versions of a comparison theorem and a uniqueness theorem for a general class of optional stochastic differential equations are stated and proved using a local time approach. Furthermore, these results are applied to the pricing of financial derivatives.

Second, the estimates of N. V. Krylov for distributions of stochastic integrals by means of Lebesgue norm of a measurable function are well-known and are widely used in the theory of stochastic differential equations and controlled diffusion processes. These estimates are generalized for optional semimartingales. After that, they are applied to extend the change of variables formula for a general class of functions from Sobolev space. It is also shown how to use the obtained estimates for the investigation of mean-square convergence of solutions of optional SDE's.

Furthermore, an optional semimartingale risk model for the capital process of a company is introduced and exhaustively investigated. A general approach to the calculation of ruin probabilities of such models is shown and supported by diverse examples.

The main object of the final part of this thesis is a general regression model in an optional setting – when an observed process is an optional semimartingale depending on an unknown parameter. The cases when the model consists of a one-dimensional and a multi-dimensional unknown parameter are studied separately. The main results include the proof of strong consistency of least squares estimates and the property of fixed accuracy of sequential least squares estimates. It is expected that the proposed general regression models will further be developed and applied in the context of modern mathematical finance.

Preface

The thesis is devoted to the development of one of the most general and promising approaches of modern stochastic analysis, namely, the theory of optional processes. By doing so, in addition to purely theoretical generalizations, this work has a primary goal to show how these methods and techniques are applied in the areas such as mathematical finance, statistics, and risk theory. The origins of the theory of optional processes belong to the papers dated 1970-80s. In the last decade, many fundamental works of that period were revisited from the purely theoretical point of view and, mainly, from the perspective of the applications of optional processes. As a result, this research was adequately exposed in the first monograph of its kind [4] by M. Abdelghani and A. Melnikov.

Some of the research conducted for this thesis forms part of a research collaboration, led by Professor A. Melnikov at the University of Alberta, with Dr. M. Abdelghani from Morgan Stanley, NY. Versions of Chapter 3 and Chapter 5 have been published in [6], [3] and [9], respectively. The author of the thesis was mainly responsible for giving a detailed analysis of the proposed concepts, transforming them into actual theorems, and providing the basic steps of their proofs, applications, and examples needed for illustration of the results. Dr. Abdelghani was involved in the early stages of the concept formation with further contribution to the manuscript editing. Dr. Melnikov was the actual supervisory author who was permanently involved in this research proposing project concepts and leading ideas as well as methods for their realization.

Acknowledgements

First, I would like to express my sincere gratitude to my supervisor Professor Alexander Melnikov for the constant support, excellent guidance, and motivation throughout my PhD program. I also would like to thank Dr. Mohamed Abdelghani for his support and motivation, and Prof. Christoph Frei for support and insightful comments during my program.

Second, I am grateful for the financial support that I have received from the Department of Mathematical and Statistical Sciences. This doctoral thesis was partially supported by the Natural Sciences and Engineering Research Council of Canada discovery grant RES0043487.

Finally, I would like to thank my parents Alexey and Yuliya, and my sister Alissa for their belief in me, support, and love. Also thanks to all my friends in Edmonton who made my journey through my PhD study much more enjoyable.

Contents

Chapter 1

Introduction

Let a triplet $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space, i.e., if $A \subseteq B \in \mathcal{F}$ and $P(B) = 0$ then $A \in \mathcal{F}$. This probability space equipped with a non-decreasing family of σ-algebras (filtration/information flow) $\mathbf{F} = (\mathcal{F}_t)_{t\in[0,\infty)}$, where $\mathcal{F}_t \subseteq \mathcal{F},\ \mathcal{F}_s \subseteq \mathcal{F}_t,\ s \leq t$, forms a stochastic basis - a fundamental notion in stochastic analysis. Usually, the following conditions are assumed on the filtration $\mathbf{F}$:

- $\mathbf{F}$ is right-continuous, i.e., $\mathcal{F}_t = \mathcal{F}_{t+}$ for all t, where $\mathcal{F}_{t+} = \cap_{t<s}\mathcal{F}_s$;
- $\mathbf{F}$ is complete, i.e., each $\mathcal{F}_t$ contains P-null sets of $\mathcal{F}$.

These assumptions are collectively known as "usual conditions" on the filtration $\mathbf{F}$.

A theory of stochastic processes is well-developed under "usual conditions". This theory is widely applied in different areas such as mathematical finance, mathematical statistics, risk theory, stochastic differential equations, and others. Although "usual conditions" is considered as a golden standard in stochastic analysis, one can immediately give an example illustrating the existence of a stochastic basis where such assumptions do not hold.

Example 1.0.1 *Suppose*

$$X_t = \mathbf{1}_{t>t_0} 1_A,$$

where A is $\mathcal{F}$-measurable with $0 < P(A) < 1$. Let $\mathcal{F}_t$ be a natural filtration of X, $\mathcal{F}_t = \sigma(X_s, s \leq t)$.

Then $\mathbf{F}$ is not right-continuous at t_0, i.e., $A \notin \mathcal{F}_{t_0}$, but $A \in \mathcal{F}_{t_0+}$, and it is not possible to make it right-continuous in a useful way.

A reasonable question arises: why are "usual conditions" so predominant in the studying of random processes? The majority of important results in the stochastic analysis are proved under "usual conditions", and for many of them, these conditions are vital. One of the most prominent illustrations is the following theorem.

Theorem 1.0.1 *If the filtration $\mathbf{F}$ is right-continuous, then every martingale admits a cadlag modification.*

If the assumption of right-continuity is omitted then the above theorem no longer works. However, in this case, Dellacherie and Meyer [26] proved the existence and uniqueness (up to indistinguishability) of optional modifications for martingales.

Theorem 1.0.2 *Let X be a bounded random variable. There exists a modification (X_t) of the martingale $(\mathbf{E}[X|\mathcal{F}_t])$ such that (X_t) is an optional process and, for any stopping time T,*

$$X_T \mathbf{1}_{(T<\infty)} = \mathbf{E}[X\mathbf{1}_{(T<\infty)}|\mathcal{F}_T] \quad a.s., \tag{1.1}$$

where $\mathbf{1}_{(T<\infty)}$ is the characteristic function of the set $(T < \infty)$.

If another optional modification $(\tilde{X}_t)$ exists satisfying (1.1), then (X_t) and $(\tilde{X}_t)$ are indistinguishable.

Further, Galchuk [30] extended the above result for any integrable random variable X. These two fundamental works served as the inception of the development of optional stochastic analysis. Dellacherie first started to use terminology - "unusual probability spaces" (or "unusual stochastic basis") - for the (filtered) probability spaces without "usual conditions", to which I stick for the rest of this work.

Furthermore, the theory of stochastic processes on the usual stochastic basis is mostly devoted to the studying of semimartingales, a large class of adapted random processes admitting modifications with right-continuous left limits (RCLL) paths. However, there are many stochastic processes that are neither right nor left-continuous. It was shown that optional processes on unusual probability spaces, in particular optional semimartingales, are not necessarily right or left-continuous processes but have right and left limits (RLL, sometimes called laglad processes).

The theory of optional processes was developed by many mathematicians such as Lepingle [63], Horowitz [46], Lenglart [62]. In these works, a theory of the stochastic analysis of optional processes on *un*usual probability spaces was constructed. Most of the foundations of stochastic calculus of optional processes were formulated by Gal'chuk in his series of works [30], [32], [33]. Further research in this direction was done by Gasparyan [35], [37]–[39], Kuhn and Stroh [58]. Recently, this direction received a new impulse mostly by the works of Abdelghani and Melnikov [4], [7], [8], [10]–[13], Abdelghani, Melnikov and Pak [3], [5], [6], [9], and Melnikov and Pak [74]. Apparently, this direction attracts substantial attention and many works have appeared during the last couple of years, to mention a few, [43], [44], [49]. From a theoretical point of view, the investigation of such processes is interesting because it allows for the generalization of different existing results for a richer class of processes, filling the gaps in theory and, consequently, unification of special cases under a general holistic approach. On the other hand, from a practical point of view, the optional processes have a promising potential in different applications.

The main goal of this dissertation is to develop new results in the theory of optional processes and apply the methods of optional processes in mathematical finance, statistics, and risk theory.

The rest of the thesis is organized as follows.

In Chapter 2, a brief introduction to the theory of optional processes is provided including the canonical decomposition of optional semimartingales and the change of variables formula.

In Chapter 3, different versions of *comparison theorem* and also a *uniqueness theorem* for a general class of optional stochastic differential equations are stated and proved. Furthermore, these results are applied to the pricing of financial derivatives.

In Chapter 4, the so-called *Krylov estimates* for distributions of stochastic integrals by means of L_d-norm of a measurable function are generalized for a class of optional processes called *optional semimartingales*. Corresponding applications of this result are illustrated.

In Chapter 5, a very general *optional semimartingale risk model* for the capital process of a company is introduced and exhaustively investigated. A general approach to the calculation of ruin probabilities of such models is shown and supported by diverse examples.

The main object of investigation in Chapter 6 and Chapter 7 is a very general *regression model in optional setting* – when an observed process is an optional semimartingale depending on an unknown parameter. Chapter 6 considers the one-dimensional optional regression model, while Chapter 7 studies the multi-dimensional one. The main results are devoted to the proof of the strong consistency of structural least squares estimates and the property of fixed accuracy of sequential least squares estimates.

Chapter 2

Preliminaries

Here, we provide a brief introduction to the theory of optional processes. All results in this chapter are required for the subsequent chapters and are presented without proofs for conciseness, however, a comprehensive exposition of optional processes and their different applications can be found in [4].

2.1 Optionality and predictability

Let us introduce $\mathcal{O}(\mathbf{F})$ and $\mathcal{P}(\mathbf{F})$ as the optional and predictable σ-algebras on $(\Omega, [0, \infty))$, respectively. $\mathcal{O}(\mathbf{F})$ is generated by all $\mathbf{F}$-adapted processes whose trajectories are right-continuous and have left limits (or cadlag, RCLL). $\mathcal{P}(\mathbf{F})$ is generated by all $\mathbf{F}$-adapted processes whose trajectories are left-continuous and have right limits.

A random process $X = (X_t)$, $t \in [0, \infty)$, is said to be *optional* if it is $\mathcal{O}(\mathbf{F})$-measurable. In general, an optional process have right and left limits but is not necessarily right- or left-continuous in $\mathbf{F}$.

A random process (X_t), $t \in [0, \infty)$, is predictable if X is $\mathcal{P}(\mathbf{F})$-measurable. As well, in general a predictable process has right and left limits but may not necessarily be right- or left- continuous in $\mathbf{F}$.

Denote $\mathcal{P}$ and $\mathcal{O}$ sets of predictable and optional processes, respectively. For either optional or predictable processes the following processes can be defined: $X_- = (X_{t-})_{t\geq 0}$ and $X_+ = (X_{t+})_{t\geq 0}$, $\Delta X = (\Delta X_t)_{t\geq 0}$ such that $\Delta X_t = X_t - X_{t-}$ and $\Delta^+ X = (\Delta^+ X_t)_{t\geq 0}$ such that $\Delta^+ X_t = X_{t+} - X_t$.

Since the random processes that we are going to work with possibly have left and right limits only, it is necessary to introduce the following definition.

Definition 2.1.1 *A random process* (X_t), $t \in [0, \infty)$, *is called* ***strongly predictable*** *if* X *is* $\mathcal{P}(\mathbf{F})$*-measurable and* X_+ *is* $\mathcal{O}(\mathbf{F})$*-measurable. Denote a space of strongly predictable processes* $\mathcal{P}_s$.

2.2 General stopping times

Since in general $\mathbf{F} \neq \mathbf{F}_+ = (\mathcal{F}_{t+})_{t\geq 0}$, where $\mathcal{F}_{t+} = \cap_{t<s}\mathcal{F}_s$, there are two distinct notions of stopping times (s.t.'s) with respect to $\mathbf{F}$ and $\mathbf{F}_+$ on unusual stochastic basis.

Definition 2.2.1 *The random variable* $T : \Omega \to [0, \infty]$ *is a stopping time (s.t.) if the set* $\{T \leq t\} \in \mathcal{F}_t$ *for all* $t \in [0, \infty)$.

Definition 2.2.2 *The random variable* $T : \Omega \to [0, \infty]$ *is a wide (broad) sense stopping time if the set* $\{T \leq t\} \in \mathcal{F}_{t+}$ *for all* $t \in [0, \infty)$.

As on a usual stochastic basis, there are notions of predictable and totally inaccessible stopping times on an unusual stochastic basis.

Definition 2.2.3 *A stopping time* T *is called predictable if there exists a sequence of wide sense stopping times* $(S_n), n \in \mathbb{N}$, *such that* $\lim S_n = T$ *a.s. and* $S_n < T$ *a.s. on the set* $\{T > 0\}$ *for all* $n \in \mathbb{N}$.

Definition 2.2.4 *A stopping time T is called totally inaccessible if $P(S = T < \infty) = 0$ for every predictable stopping time S.*

In addition to these, there exist totally inaccessible wide sense stopping times.

Definition 2.2.5 *A wide sense stopping time T is called totally inaccessible wide sense s.t. if $P(S = T < \infty) = 0$ for every stopping time S.*

Analogously, it is possible to define predictable wide sense stopping times, but it turns out that this is unnecessary because the set of predictable s.t.'s and the set of predictable wide sense s.t.'s are equal (see [33]). Denote by $\mathcal{T}$ and $\mathcal{T}_+$ a set of s.t.'s and a set of wide sense s.t.'s, respectively; $\mathcal{T}^p, \mathcal{T}^i, \mathcal{T}^i_+$ - set of predictable s.t.'s, set of totally inaccessible s.t.'s and totally inaccessible wide sense s.t.'s, respectively. It is immediately seen that $\mathcal{T}^p \subseteq \mathcal{T} \subseteq \mathcal{T}_+$.

On unusual stochastic basis, defined above three canonical types of stopping times have the following properties: Predictable stopping times, $S \in \mathcal{T}^p$, are such that $(S \leq t)$ is $\mathcal{F}_{t-}$ measurable for all t; Totally inaccessible stopping times, $T \in \mathcal{T}$, are such that $(T \leq t)$ is $\mathcal{F}_t$ measurable for all t, however, we note that $(T < t)$ is not necessarily $\mathcal{F}_t$ measurable since $\mathcal{F}_t$ is not right continuous; Finally, totally inaccessible stopping times in the broad sense, $U \in \mathcal{T}_+$, are such that $(U \leq t)$ is $\mathcal{F}_{t+}$ measurable for all t, but since $\mathcal{F}_{t+}$ is right continuous, $(U < t)$ is also $\mathcal{F}_{t+}$ measurable.

Definition 2.2.6 *Suppose S and T are maps $\Omega \to [0, \infty]$ and $S \leq T$ a.s. The stochastic interval denoted by $[S, T[$ is the set*

$$\{(t, \omega) \in [0, \infty[\times \Omega : S(\omega) \leq t < T(\omega)\}.$$

The stochastic intervals $[S, T],]S, T]$ and $]S, T[$ are defined similarly.

The stochastic interval

$$[T, T] = \{(t, \omega) \in [0, \infty[\times \Omega : T(\omega) = t\}$$

is denoted by $[T]$, *and is called the graph of* T. *Through out the thesis, we use the notation introduced above for wide sense stopping times. Whenever we use this notation for the deterministic times, we mean a usual interval of real numbers.*

Definition 2.2.7 *The sequence of* $(T_n)_{n\in\mathbb{N}} \subset \mathcal{T}_+$ *exhausts the jumps of the process* $(X_t)_{t\in[0,\infty)}$ *if for any* $T \in \mathcal{T}_+$ *for which the set* $[T] \cap (\cup_n [T_n]) = \emptyset$ *we have* $\Delta X_T = \Delta^+ X_T = 0$ *a.s. on* $\{t < \infty\}$.

Below, a fundamental result that gives foundations for many results in the optional stochastic analysis is presented. In particular, it is used in the canonical decomposition of optional semimartingales.

Theorem 2.2.1 *(see [33]) Suppose* $X = (X_t)_{t\in[0,\infty)}$ *is an optional process whose paths have left and right limits a.s.. Then there exist sequences* (S_n), (T_n) *and* (U_n), $n \in \mathbb{N}$, *of predictable s.t.'s, totally inaccessible s.t.'s and totally inaccessible wide sense s.t.'s respectively, exhausting all jumps of the process* X *and having the following properties: the graphs of these s.t.'s are mutually non-intersecting within each sequence.*

2.3 Processes of finite variation

Here we consider the processes which do not vary a lot, i.e., processes having paths of finite variation for almost all ω. Let us begin with the definitions of several spaces of such processes.

Definition 2.3.1 *A process $A = (A_t)$, $t \in \mathbb{R}_+$, is of finite variation if it has finite variation on every segment $[0,t]$, $t \in \mathbb{R}_+$ a.s., that is $\mathbf{Var}(A)_t < \infty$, for all $t \in \mathbb{R}_+$ a.s. where*

$$\mathbf{Var}(A)_t = \sum_{0 \le s < t} |\Delta^+ A_s| + \int_{]0,t]} |dA_s^r|.$$

Denote by $\mathcal{V}$ the set of all $\mathbf{F}$-adapted processes of finite variation.

Definition 2.3.2 *A process $A = (A_t)_{t \ge 0}$ is increasing if it is non-negative, $\mathbf{F}$-adapted, its trajectories do not decrease. Let $\mathcal{V}^+$ be the collection of increasing processes.*

We know that every increasing process is of finite variation, i.e., $\mathcal{V}^+ \subset \mathcal{V}$.

Definition 2.3.3 *An increasing process A is integrable if $\mathbf{E}A_\infty < \infty$. The collection of such processes is denoted by $\mathcal{A}^+$.*

Definition 2.3.4 *A process $A = (A_t)_{t \ge 0}$ of finite variation belongs to the space $\mathcal{A}$ of integrable finite variation processes if $\mathbf{E}\left[\mathbf{Var}(A)_\infty\right] < \infty$.*

A process $X = (X_t)_{t \ge 0}$ belongs to the space $\mathcal{J}_{loc}$ if there is a localizing sequence of wide sense s.t.'s, (R_n), $n \in \mathbb{N}$, $R_n \in \mathcal{T}_+$, $R_n \uparrow \infty$ a.s. such that $X\mathbf{1}_{[0,R_n]} \in \mathcal{J}$ for all n, where $\mathcal{J}$ is a space of processes and $\mathcal{J}_{loc}$ is an extension of $\mathcal{J}$ by localization.

In general, the spaces $\mathcal{V}$, $\mathcal{A}$, $\mathcal{A}^+$ can be extended to $\mathcal{V}_{loc}$, $\mathcal{A}_{loc}$, $\mathcal{A}^+_{loc}$ respectively by localization. It is well-known that $\mathcal{V} = \mathcal{V}_{loc}$, and the relationships $\mathcal{A} \subseteq \mathcal{A}_{loc} \subseteq \mathcal{V}$ and $\mathcal{A}^+ \subseteq \mathcal{A}^+_{loc} \subseteq \mathcal{V}^+$ hold.

A finite variation process A can be decomposed to $A = A^r + A^g = A^c + A^d + A^g$ where A^c is continuous, A^r is right-continuous, A^d is discrete right-continuous, A^g is discrete left-continuous such that

$$A_t^d = \sum_{0 < s \le t} \Delta A_s \quad \textit{and} \quad A_t^g = \sum_{0 \le s < t} \Delta^+ A_s,$$

where the series converge absolutely.

2.4 Optional Martingales

Definition 2.4.1 *A process $M = (M_t)_{t\geq 0}$ is an (square integrable) optional martingale (supermartingale, submartingale) if*

- *M is $\mathcal{O}(\mathbf{F})$-measurable,*
- *there exists an (square) integrable $\mathcal{F}_\infty$-measurable random variable $\hat{M}$ such that*

$$M_T = \mathbf{E}[\hat{M}|\mathcal{F}_T]$$

(respectively, $M_T \geq \mathbf{E}[\hat{M}|\mathcal{F}_T]$, $M_T \leq \mathbf{E}[\hat{M}|\mathcal{F}_T]$) a.s. on the set $(T < \infty)$ for any $T \in \mathcal{T}$.

Let $\mathcal{M}$ (resp. $\mathcal{M}^2$) denote the set of optional martingales (resp. square integrable optional martingales). The space $\mathcal{M}$ is extended to a space of local optional martingale, $\mathcal{M}_{loc}$, and the space $\mathcal{M}^2$ is extended to a space locally square integrable optional martingales, $\mathcal{M}^2_{loc}$, respectively.

Definition 2.4.2 *A process $M = (M_t)_{t\geq 0}$ is called an optional local (locally square integrable) martingale if there exists a sequence $(R_n, M^{(n)}), n \in \mathbb{N}$, where $R_n \in \mathcal{T}_+$, $R_n \uparrow \infty$ a.s. and $M^{(n)}$ is a (square integrable) optional martingale, such that $M = M^{(n)}$ on the stochastic interval $[0, R_n]$ and the random variable $M_{R_{n+}}$ is integrable for any $n \in \mathbb{N}$.*

If $M \in \mathcal{M}_{loc}$ then it can be decomposed to

$$M = M^r + M^g \quad where \quad M^r = M^c + M^d,$$

M^c is continuous, M^d is right-continuous and M^g is left-continuous optional local martingales. M^d and M^g are orthogonal to each other and to any continuous (local) martingale. Moreover, M^d and M^g can be written as

$$M_t^d = \sum_{0<s\le t} \Delta M_s \;\; and \;\; M_t^g = \sum_{0\le s<t} \Delta^+ M_s.$$

Now, we discuss the extension of the quadratic variation process for the square integrable optional martingales.

Lemma 2.4.1 *1) Suppose $X \in \mathcal{M}^2$. There exists a unique increasing strongly predictable process $\langle X\rangle \in \mathcal{A}$ such that $\mathbf{E}X_T^2 = \mathbf{E}\langle X\rangle_T$ for every s.t. T, or, equivalently, $X^2 - \langle X\rangle \in \mathcal{M}$, where $\langle X\rangle = \langle X^g\rangle + \langle X^c\rangle + \langle X^d\rangle$.*

2) If $X, Y \in \mathcal{M}^2$, then there exists a unique strongly predictable process $\langle X, Y\rangle \in \mathcal{A}$ such that $XY - \langle X, Y\rangle \in \mathcal{M}$, where

$$\langle X, Y\rangle = \frac{1}{2}[\langle X+Y\rangle - \langle X\rangle - \langle Y\rangle].$$

Definition 2.4.3 *Suppose $X \in \mathcal{M}^2$ and X^c is its continuous part. We define $[X, X]$ to be the process*

$$[X, X]_t = \langle X^c\rangle_t + \sum_{s\le t}(\Delta X_s)^2 + \sum_{s<t}(\Delta^+ X_s)^2, \quad t \in \mathbb{R}_+.$$

It is not difficult to show that $[X, X]$ is increasing, $\mathbf{F}$-adapted and integrable, and that $X^2 - [X, X] \in \mathcal{M}$.

Further, setting the polarization property of quadratic variation processes

$$[X, Y] = \frac{1}{2}([X+Y, X+Y] - [X, X] - [Y, Y])$$

for $X, Y \in \mathcal{M}^2$, we have $XY - [X, Y] \in \mathcal{M}$. Using Definition 2.4.3, we get that

$$[X, Y]_t = \langle X^c, Y^c\rangle_t + \sum_{s\le t} \Delta X_s \Delta Y_s + \sum_{s<t} \Delta^+ X_s \Delta^+ Y_s.$$

Lemma 2.4.2 *A non-negative optional local martingale X is a supermartingale.*

Proof. Let $X \in \mathcal{M}_{loc},\ X \geq 0$. Then by defintion of optional local martingale there exist $X^n \in \mathcal{M},\ X^n \geq 0$, and $R_n \in \mathcal{T}_+,\ R_n \uparrow \infty$ a.s. such that for any $n \geq 1:\ X = X^n \mathbf{1}_{[0,R_n]}$. Next, for any $t \geq s$ and $A \in \mathcal{F}_s$ we have

$$\begin{aligned}
\mathbf{E} X_t \mathbf{1}_A &= \lim_{n\to\infty} \mathbf{E} X_t \mathbf{1}_A \mathbf{1}_{(t \leq R_n)} \\
&= \lim_{n\to\infty} \mathbf{E} X_t^n \mathbf{1}_A \mathbf{1}_{(t \leq R_n)} \\
&\leq \lim_{n\to\infty} \mathbf{E} X_t^n \mathbf{1}_A \mathbf{1}_{(s \leq R_n)} \\
&= \lim_{n\to\infty} \mathbf{E} X_s^n \mathbf{1}_A \mathbf{1}_{(s \leq R_n)} \\
&= \mathbf{E} X_s \mathbf{1}_A.
\end{aligned}$$

Hence, the process X_t is a non-negative supermartingale (in usual sense). ■

2.5 Optional Semimartingales

On an unusual stochastic basis, the most general processes that have good properties to work with are optional semimartingales. Optional semimartingales are linear combinations of finite variation processes and local optional martingales. Since local optional martingales and processes of finite variation are in general neither right-continuous nor left-continuous, optional semimartingales possess the same continuity characteristics. Therefore, their structure is complicated.

Definition 2.5.1 *The stochastic process X is called an optional semimartingale if*

$$X = X_0 + M + A,$$

where $M \in \mathcal{M}_{loc},\ A \in \mathcal{V},\ A_0 = M_0 = 0$ and X_0 is an $\mathcal{F}_0$-measurable finite random variable.

Definition 2.5.2 *An optional semimartingale X is called special optional semimartingale if the above decomposition exists with a strongly predictable process $A \in \mathcal{A}_{loc}$.*

Let $\mathcal{S}$ denote the set of optional semimartingales and $\mathcal{S}_p$ the set of special optional semimartingales. If $X \in \mathcal{S}_p$ then the semimartingale decomposition is unique. By the optional martingale decomposition and the decomposition of predictable processes [32], [33] we can decompose a semimartingale further to

$$\begin{aligned} X =& X_0 + X^r + X^g, \text{ with} \\ X^r =& A^r + M^r, \\ M^r =& M^c + M^d, \\ A^r =& A^c + A^d, \\ X^g =& A^g + M^g, \end{aligned} \tag{2.1}$$

where A^c, A^d, and A^g are finite variation processes that are continuous, discrete right-continuous, and discrete left-continuous, respectively; $M^c \in \mathcal{M}^c_{loc},\ M^d \in \mathcal{M}^d_{loc},\ M^g \in \mathcal{M}^g_{loc}$ is a continuous, a discrete right-continuous and a left-continuous local martingale, respectively. This decomposition is useful for defining integration with respect to optional semimartingales.

2.6 Integration with respect to optional semimartingales

A stochastic integral with respect to optional semimartingale is defined as

$$\begin{aligned} \varphi \circ X_t &= \int_{[0,t]} \varphi_s dX_s = \int_{]0,t]} \varphi_{s-} dX^r_s + \int_{[0,t[} \varphi_s dX^g_{s+} \quad \text{, where} \\ \int_{]0,t]} \varphi_{s-} dX^r_s &= \int_{]0,t]} \varphi_{s-} dA^r_s + \int_{]0,t]} \varphi_{s-} dM^r_s \quad \text{and} \\ \int_{[0,t[} \varphi_s dX^g_{s+} &= \int_{[0,t[} \phi_s dA^g_{s+} + \int_{[0,t[} \phi_s dM^g_{s+}. \end{aligned}$$

The stochastic integral with respect to the finite variation processes or strongly predictable processes A^r over $]0,t]$ and A^g over $[0,t[$ are interpreted as usual, in the Lebesgue sense. The integral $\int_{]0,t]} \varphi_{s-} dM_s^r$ over $]0,t]$ is the usual stochastic integral with respect to cadlag local martingale whereas $\int_{[0,t[} \phi_s dM_{s+}^g$ over $[0,t[$ is the Gal'chuk stochastic integral (see [32]) with respect to left continuous local martingale. In general, the stochastic integral with respect to optional semimartingale X can be defined as a bilinear form $(\varphi^r, \varphi^g) \circ X_t$, $\varphi^r \in \mathcal{P}$ and $\varphi^g \in \mathcal{O}$ such that

$$Y_t = (\varphi^r, \varphi^g) \circ X_t = \varphi^r \cdot X_t^r + \varphi^r \odot X_t^g,$$
$$\varphi^r \cdot X^r = \int_{]0,t]} \varphi_s^r dX_s^r, \quad \varphi^g \odot X^g = \int_{[0,t[} \varphi_s^g dX_{s+}^g,$$

where Y is again an optional semimartingale. Note that the stochastic integral over optional semimartingales is defined on a much larger space of integrands, the product space of predictable and optional processes, $\mathcal{P} \times \mathcal{O}$. From now on, we are going to use the operator "$\circ$" to denote the stochastic optional integral, the operator "$\cdot$" to denote the regular stochastic integral with respect to RCLL semimartingales, and the operator "$\odot$" for the Galchuk stochastic integral $g \odot X^g$ with respect to left continuous semimartingales.

The properties of optional stochastic integral are: First, isometry is satisfied with

$$\left(f^2 \cdot [X^r, X^r]\right)^{1/2} \in \mathcal{A}_{loc} \quad and \quad \left(g^2 \odot [X^g, X^g]\right)^{1/2} \in \mathcal{A}_{loc}.$$

The quadratic variations are defined as

$$\begin{aligned}
[X, X] &= [X^r, X^r] + [X^g, X^g] \quad where \\
[X^r, X^r]_t &= \langle X^c, X^c \rangle_t + \sum_{0<s\leq t} (\Delta X_s)^2 \quad and \\
[X^g, X^g]_t &= \sum_{0\leq s<t} \left(\Delta^+ X_s\right)^2;
\end{aligned}$$

Linearity is also satisfied with $(f^1+f^2, g^1+g^2)\circ X_t = (f^1,g^1)\circ X_t + (f^2,g^2)\circ X_t$ for any (f^1,g^1) and (f^2,g^2) in the space $\mathcal{P}(\mathbf{F})\times\mathcal{O}(\mathbf{F})$; $\Delta^+ X^g$ is $\mathcal{O}(\mathbf{F}_+)$-measurable with its martingale part satisfying $\mathbf{E}\left[\Delta^+ M^g_T \mathbf{1}_{(T<\infty)}|\mathcal{F}_T\right]=0$ a.s. for any stopping time T in the broad sense and ΔX^r is $\mathcal{O}(\mathbf{F})$-measurable with its martingale part satisfying $\mathbf{E}\left[\Delta M^r_T \mathbf{1}_{(T<\infty)}|\mathcal{F}_T\right]=0$ a.s. for any stopping time T; Moreover, orthogonality is as such that $X^r\perp X^g$ are orthogonal. in the sense that their product is an optional local martingale; Also, differentials are independent: $\Delta Y = f\Delta X^r$ and $\Delta^+ Y = g\Delta^+ X^g$; Lastly, for any semimartingale Z the quadratic projection is $[Y,Z] = f\cdot[X^r,Z^r] + g\odot[X^g,Z^g]$.

2.7 Random Measures and their Compensators

Consider the Lusin space $(E,\mathcal{E})$, where $E=\mathbb{R}\setminus\{0\}$ and $\mathcal{E}=\mathcal{B}(E)$ is the Borel σ-algerba in E. Also, define the spaces

$$\widetilde{\Omega}=\Omega\times\mathbb{R}_+\times E,\quad \widetilde{E}=\mathbb{R}_+\times E,\quad \widetilde{\mathcal{E}}=\mathcal{B}(\mathbb{R}_+)\times\mathcal{E}$$
$$\tilde{\mathcal{O}}(\mathbf{F})=\mathcal{O}(\mathbf{F})\times\mathcal{E},\quad \widetilde{\mathcal{O}}(\mathbf{F}_+)=\mathcal{O}(\mathbf{F}_+)\times\mathcal{E},\quad \widetilde{\mathcal{P}}=\mathcal{P}\times\mathcal{E}.$$

Definition 2.7.1 *A non-negative random set function $\mu(\omega,\Gamma),\omega\in\Omega,\ \Gamma\in\widetilde{\mathcal{E}}$, is called a random measure on $\widetilde{\mathcal{E}}$ if $\mu(\cdot,\Gamma)\in\mathcal{F}$ for any $\Gamma\in\widetilde{E}$ and $\mu(\omega,\cdot)$ is a σ-finite measure on $(\widetilde{E},\widetilde{\mathcal{E}})$ for each $\omega\in\Omega$.*

A random measure is called integer-valued if $\mu(\omega,\Gamma)\in\{0,1,\ldots,+\infty\}$ and $0\le\mu(\omega,\{t\}\times E)\le 1$ for all (ω,Γ).

For a non-negative function $f\in\mathcal{F}\times\widetilde{\mathcal{E}}$ and measure μ let us form the process $f*\mu$, where

$$f*\mu_t=\int_{[0,t]\times E} f(\omega,s,x)\mu(\omega,ds,dx),\qquad t<\infty.$$

The random measure μ is $\mathcal{O}(\mathbf{F}_+)$-optional if the process $f * \mu$ is $\mathcal{O}(\mathbf{F}_+)$-measurable for any non-negative $\widetilde{O}(\mathbf{F})$-measurable f. Similarly, the random measure μ is $\mathcal{O}(\mathbf{F})$-optional (for brevity, optional) if $f * \mu \in \mathcal{O}$ for any non-negative $\widetilde{\mathcal{O}}(\mathbf{F})$-measurable f. The random measure μ is predictable if $f * \mu \in \mathcal{P}$ for any non-negative $\widetilde{\mathcal{P}}$-measurable f.

Let μ be an optional measure. On $(\widetilde{\Omega}, \widetilde{\mathcal{O}}(\mathbf{F}))$ consider

$$\mathbf{E} f * \mu_\infty, \quad f \in \widetilde{O}(\mathbf{F}), \quad f \geq 0.$$

Lemma 2.7.1 *Let the optional measure μ be such that the measure $\mathbf{E} f * \mu_\infty$ is $\widetilde{\mathcal{P}}$-σ-finite (i.e., the restriction of $\mathbf{E} f * \mu_\infty$ to $(\widetilde{\Omega}, \widetilde{\mathcal{P}})$ is σ-finite). Then there exists a unique (up to indistinguishability) predictable measure $\nu = \nu(\omega, dt, dx)$ such that for any function $f \in \widetilde{\mathcal{P}}$, $f \geq 0$, one has $\mathbf{E} f * \mu_\infty = \mathbf{E} f * \nu_\infty$.*

The measure ν can be written in the form $\nu(\omega, dt, dx) = dA_t(\omega) K(\omega, t, dx)$, where A is an increasing predictable right-continuous process, $K(\omega, t, dx)$ is the kernel of the space $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(E, \mathcal{E})$.

If the measure μ does not load any predictable s.t.'s whatever, then the same is true for ν, and the process $f * \nu$ is continuous for any $f \in \widetilde{\mathcal{P}}, f \geq 0$. Moreover, for any $S \in \mathcal{T}^p$ and any $f \in \widetilde{\mathcal{P}}$, $f \geq 0$, on $(S < \infty)$,

$$\mathbf{E}\left[\int_E f(S, x) \mu(\{S\}, dx) \Big| \mathcal{F}_{S-}\right] = \int_E f(S, x) \nu(\{S\}, dx).$$

The process $f * \nu$ is the dual projection for the process $f * \mu$, $f \in \widetilde{P}$, $f \geq 0$. If $f * \mu \in \mathcal{A}_{loc}$, then the process

$$f * \mu - f * \nu \in \mathcal{M}_{loc}.$$

In the case of integer-valued μ, outside a set of P-null measure, $0 \leq \nu(\omega, \{t\} \times E) \leq 1$ for all $t \in \mathbb{R}_+$.

The measure ν is called the compensator of the measure μ.

Let η be an $\mathcal{O}(\mathbf{F}_+)$-optional measure. On $\left(\widetilde{\Omega}, \widetilde{\mathcal{O}}(\mathbf{F}_+)\right)$ consider

$$\mathbf{E} f * \mu_\infty, \quad f \in \widetilde{O}(\mathbf{F}_+), \quad f \geq 0.$$

Lemma 2.7.2 *Let the $\mathcal{O}(\mathbf{F}_+)$-optional measure η be such that the measure $\mathbf{E} f * \eta_\infty$ is $\widetilde{\mathcal{O}}$-σ-finite. Then:*

a) there exists a unique (up to indistinguishability) optional measure $\lambda = \lambda(\omega, dt, dx)$ such that for any function $f \in \widetilde{\mathcal{O}}(\mathbf{F})$, $f \geq 0$, one has $\mathbf{E} f * \eta_\infty = \mathbf{E} f * \lambda_\infty$.

The measure λ can be written in the form $\lambda(\omega, dt, dx) = dA_t(\omega) K(\omega, t, dx)$, where A is an increasing optional right-continuous process, $K(\omega, t, dx)$ is the kernel of the space $(\Omega \times \mathbb{R}_+, \mathcal{O})$ into $(E, \mathcal{E})$.

b) If the measure η does not load any s.t.'s whatever, then the same is true for λ, and the process $f * \lambda$ is continuous for any $f \in \widetilde{\mathcal{O}}(\mathbf{F}), f \geq 0$.

c) The process

$$f * \lambda_+ = \left(\int_{[0,t] \times E} f \lambda(ds, dx) \right)$$

is the dual optional projection for the process

$$f * \eta_+ = \left(\int_{[0,t] \times E} f \eta(ds, dx) \right)$$

for $f \in \widetilde{\mathcal{O}}(\mathbf{F})$, $f \geq 0$. In particular, for any $T \in \mathcal{T}$ and any $f \in \widetilde{\mathcal{O}}(\mathbf{F})$, $f \geq 0$ on $(T < \infty)$

$$\mathbf{E} \left[\int_E f(T, x) \eta(\{T\}, dx) \middle| \mathcal{F}_T \right] = \mathbf{E} \left[\int_E f(T, x) \lambda(\{T\}, dx) \middle| \mathcal{F}_T \right] \ a.s..$$

If $f * \eta_+ \in \mathcal{A}_{loc}$, then the process

$$f * \eta_t - f * \lambda_t = \int_{[0,t[\times E} f \eta(ds, dx) - \int_{[0,t[\times E} f \lambda(ds, dx) \in \mathcal{M}_{loc}.$$

d) If the measure η is integer-valued, then there exists a modification of the measure λ such that outside a set of P-null measure, $0 \leq \lambda(\omega, \{t\} \times E) \leq 1$ for any $t \in \mathbb{R}_+$.

The measure λ is called the compensator of the $\mathcal{O}(\mathbf{F}_+)$-optional measure η.

2.8 Canonical Decomposition of Optional Semimartingales

Let Y be a one-dimensional optional semimartingale and $(S_n)_{n\geqslant 1}$, $(T_n)_{n\geqslant 1}$, $(U_n)_{n\geqslant 1}$ be sequences of predictable, totally inaccessible stopping times and totally inaccessible wide sense stopping times, respectively, exhausting all jumps of the process Y, i.e. the set $\{\Delta Y \neq 0\} \cup \{\Delta^+ Y \neq 0\}$, such that the graphs of these stopping times do not intersect within each sequence. Define integer random measures on $(\mathbb{R}_+ \times E, \tilde{\mathcal{E}})$

$$
\begin{aligned}
\mu^d(\Gamma) &= \sum_{n\geqslant 1} \mathbf{1}_\Gamma(T_n, \beta^d_{T_n}), & \mu^g(\Gamma) &= \sum_{n\geqslant 1} \mathbf{1}_\Gamma(U_n, \beta^g_{U_n}),\\
p^d(\Gamma) &= \sum_{n\geqslant 1} \mathbf{1}_\Gamma(S_n, \beta^d_{S_n}), & p^g(\Gamma) &= \sum_{n\geqslant 1} \mathbf{1}_\Gamma(S_n, \beta^g_{S_n}),\\
\eta(\Gamma) &= \sum_{n\geqslant 1} \mathbf{1}_\Gamma(T_n, \beta^g_{T_n}),
\end{aligned}
$$

where $\mathbf{1}_\Gamma(\cdot)$ is an indicator function of a set $\Gamma \in \tilde{\mathcal{E}}$, $\beta^d_t = \Delta Y_t$ if $\Delta Y_t \neq 0$ and $\beta^d_t = \delta$ if $\Delta Y_t = 0$, $\beta^g_t = \Delta^+ Y_t$ if $\Delta^+ Y_t \neq 0, \beta^g_t = \delta$ if $\Delta^+ Y_t = 0$, $t > 0$.

Under the *un*usual conditions on probability space Gasparyan [37, Theorem 1] showed that Y can be decomposed as follows

$$
\begin{aligned}
Y_t = Y_0 + a_t + m_t &+ \int_{]0,t]\times E} u\mathbf{1}_{(|u|\leq 1)} d(\mu^d - \nu^d) + \int_{[0,t[\times E} u\mathbf{1}_{(|u|\leq 1)} d(\mu^g - \nu^g)\\
&+ \int_{]0,t]\times E} u\mathbf{1}_{(|u|>1)} d\mu^d + \int_{[0,t[\times E} u\mathbf{1}_{(|u|>1)} d\mu^g\\
&+ \int_{]0,t]\times E} u dp^d + \int_{[0,t[\times E} u dp^g + \int_{[0,t[\times E} u d\eta
\end{aligned}
$$

or in short notation

$$
Y = Y_0 + a + m + \sum_{j=r,g} \left[u\mathbf{1}_{(|u|\leq 1)} * (\mu^j - \nu^j) + u\mathbf{1}_{(|u|>1)} * \mu^j + u * p^j \right] + u * \eta, \quad (2.2)
$$

where Y_0 is $\mathcal{F}_0$-measurable random variable, $a \in \mathcal{A}_{loc}, a_0 = 0$, and $m \in \mathcal{M}^2_{loc}, m_0 = 0$, ($a$ and m are both continuous); ν^j are the compensators of μ^j.

2.9 Change of variables formulas

The Ito's formula is a basic instrument in stochastic calculus. In optional settings we have a generalization of the Ito's formula - a change of variables formula, which will be applied extensively in the next chapters.

Theorem 2.9.1 *(Gal'chuk-Lenglart formula [33] [62])*

Suppose X is an n-dimensional optional semimartingale, i.e.,

$$X = (X^1, ..., X^n),$$

where X^i is an optional semimartingale, $i = 1, ..., n$, and $f(x) = f(x^1, ..., x^n)$ is a twice continuously differentiable function on $\mathbb{R}^n$. Then the process $f(x)$ is an optional semimartingale, and for all $t \in \mathbb{R}_+$

$$\begin{aligned} f(X_t) =& f(X_0) + \sum_{i=1}^{n} \int_{]0,t]} \frac{\partial f}{\partial x_i}(X_{s-}) d(A^{ir} + M^{ir})_s \\ &+ \frac{1}{2} \sum_{i,j=1}^{n} \int_{]0,t]} \frac{\partial^2 f}{\partial x_i \partial x_j} f(X_{s-}) d\langle M^{ic}, M^{jc} \rangle_s \\ &+ \sum_{0<s\leq t} \left[f(X_s) - f(X_{s-}) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s-}) \Delta X^i_s \right] \\ &+ \sum_{i=1}^{n} \int_{[0,t[} \frac{\partial f}{\partial x_i}(X_s) d(A^{ig} + M^{ig})_{s+} \\ &+ \sum_{0\leq s<t} [f(X_{s+}) - f(X_s) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_s) \Delta^+ X^i_s], \end{aligned} \tag{2.3}$$

where $A^{ir}, M^{ir}, M^{ic}, A^{ig}$, and M^{ig} are from the decomposition of X^i in (2.1).

After comparing Ito's formula with Gal'chuk-Lenglart formula, we can immediately notice that there are two additional last terms in (2.3) that appeared due to jumps from the right.

Let me now consider the following optional semimartingale

$$X = X_0 + f \cdot a + g \cdot m + \sum_{j=r,g} \left[U h_j * (\mu^j - \nu^j) + (k_j + l_j) * p^j \right] + (r + w) * \eta, \quad (2.4)$$

where X_0 is $\mathcal{F}_0$-measurable random variable; $U = U(u) = \mathbf{1}_{(|u| \leq 1)}(u)$; $a \in \mathcal{A}_{loc}$ is continuous increasing, $m, \mu^j, \nu^j, p^j, \eta$ as in (2.2). Further, for convenience we use the notation $f = f(\omega, t, x)$, $g = g(\omega, t, x)$, $h_d = h(\omega, t, u, x)$, $h_g = h(\omega, t, u, x)$ and similarly for k_j, l_j, w and r whenever this does not lead to confusion.

Assume the following

- For $j = r, g$,

$$|f(X)| \cdot a \in \mathcal{A}_{loc},$$

$$g(X) \cdot \langle m \rangle \in \mathcal{A}_{loc},$$

$$|h(X)|^2 * \nu^j \in \mathcal{A}_{loc},$$

$$|l_j(X)| * p^j \in \mathcal{A}_{loc},$$

$$\left[|k_j(X)|^2 * p^j \right]^{1/2} \in \mathcal{A}_{loc},$$

$$|r(X)| * \eta \in \mathcal{A}_{loc},$$

$$\left[|w(X)|^2 * \eta \right]^{1/2} \in \mathcal{A}_{loc},$$

  and $E[k_r(S, \beta^d_S, X_0) | \mathcal{F}_{S-}] = 0$ a.s. for any predictable stopping time S on $\{S < \infty\}$ and $E[k_g(T, \beta^g_T, X_0) | \mathcal{F}_T] = 0$, $E[w(T, \beta^g_T, X_0) | \mathcal{F}_T] = 0$ a.s., for any stopping time T on $\{T < \infty\}$.

- $f(\omega, s, x)$ and $g(\omega, s, x)$ are defined on $(\Omega \times \mathbb{R}_+ \times \mathbb{R})$ and $\mathcal{P} \times \mathcal{B}(\mathbb{R})$-measurable;

- $Uh_d(\omega, s, u, x)$ is defined on $(\Omega \times \mathbb{R}_+ \times E \cap (|u| \leq 1) \times \mathbb{R})$ and $\mathcal{P} \times \mathcal{B}(E \cap (|u| \leq 1)) \times \mathcal{B}(\mathbb{R})$-measurable;

- $Uh_g(\omega, s, u, x)$ is defined on $(\Omega \times \mathbb{R}_+ \times E \cap (|u| \leq 1) \times \mathbb{R})$ and $\mathcal{O} \times \mathcal{B}(E \cap (|u| \leq 1)) \times \mathcal{B}(\mathbb{R})$-measurable;

- $k_d(\omega, s, u, x)$, $l_d(\omega, s, u, x)$ are defined on $(\Omega \times \mathbb{R}_+ \times E \times \mathbb{R})$ and $\tilde{\mathcal{P}} \times \mathcal{B}(\mathbb{R})$-measurable;

- $k_g(\omega, s, u, x)$, $l_g(\omega, s, u, x)$, $r(\omega, s, u, x)$, $w(\omega, s, u, x)$ are defined on $(\Omega \times \mathbb{R}_+ \times E \times \mathbb{R})$ and $\tilde{\mathcal{O}} \times \mathcal{B}(\mathbb{R})$-measurable.

These integrability and measurability assumptions guarantee well-posedness of the integrals in (2.4).

Theorem 2.9.2 *Let X be an optional semimartingale given in* (2.4). *Let $f(x) = f(x^1, ..., x^n)$ be a twice continuously differentiable function on $\mathbb{R}^n$. Then the process $f(x)$ is an optional semimartingale, and for all $t \in \mathbb{R}_+$ it has the following represen-*

tation

$$
\begin{aligned}
f(X_t) =& f(X_0) + \sum_{i=1}^{n} \int_{]0,t]} \frac{\partial f}{\partial x_i}(X_{s-}) d(a^i + m^i)_s \\
&+ \frac{1}{2} \sum_{i,j=1}^{n} \int_{]0,t]} \frac{\partial^2 f}{\partial x_i \partial x_j} f(X_{s-}) d\langle m, m\rangle_s \\
&+ [f(X_- + h_d) - f(X_-)]\, U * (\mu^d - \nu^d)_t \\
&+ [f(X + h_g) - f(X)]\, U * (\mu^g - \nu^g)_t \\
&+ \left[f(X_- + h_d) - f(X_-) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s-}) h_d^i \right] U * \nu_t^d \\
&+ \left[f(X + h_g) - f(X) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_s) h_g^i \right] U * \nu_t^g \\
&+ [f(X_- + (k_d + l_d)) - f(X_-)] * p_t^d \\
&+ [f(X + (k_g + l_g)) - f(X)] * p_t^g \\
&+ [f(X + (r + w)) - f(X)] * \eta_t.
\end{aligned}
$$

Chapter 3

Optional SDE's

This chapter is devoted to the study of *comparison* of solutions of stochastic equations of optional semimartingales on *un*usual probability space, and the study of *pathwise uniqueness* of these solutions using local times.

A comparison theorem for stochastic equations with respect to continuous semimartingales was proved by Melnikov [68] who developed the Yamada method [89]. Later a similar result was given by Yan [90] using the local time technique. The case of SDEs with integer-valued random measures where the coefficients are not Lipschitz but satisfy weaker conditions similar to those of Yamada were considered by Gal'chuk [31]. Interesting applications of path-wise comparison theorem to mathematical finance were given in [53] and further developed in [52]. Recently, a comparison theorem for optional semimartingales on *un*usual probability space was given in [11] when coefficients satisfy the Yamada conditions. Therefore, our goal here is to study comparison of optional SDEs on *un*usual stochastic basis under a more general condition, *local time condition*, placed on the diffusion coefficient. Besides, we extend a version of the comparison theorem for solutions of SDEs with different diffusion coefficients (see [29], [78]) to the laglad jump-diffusion case.

Even though, the stochastic calculus of optional semimartingales is well developed, little is done in showing pathwise uniqueness of solutions of stochastic equations driven by optional semimartingales on *un*usual probability spaces, except, for the works of Gasparyan [38] and Abdelghani and Melnikov [13] on the existence and uniqueness of strong solutions under Lipschitz conditions and monotonicity conditions, respectively. On the other hand, Perkins [79] proved the pathwise uniqueness of solutions of stochastic equations of continuous semimartingales using the local time technique. As a result, we consider the questions of pathwise uniqueness under one-sided Lipschitz continuity on a drift function and local time condition on the diffusion coefficient for *laglad* optional semimartingale using the method of local time, which was not done before.

Besides a purely theoretical interest, the topic is motivated by the needs of the energy market. In many electricity markets, retailers buy electricity at an unregulated price and sell it to consumers at a regulated price. Therefore, the occurrence of price spikes due to sudden changes in electricity demand or supply in these markets represents a major source of risk to retailers. Hence, accurate modeling of price spikes is important. As a result, we have modeled spikes in spot price in a way so that each upward jump is accompanied by an immediate downward jump. The flexibility, modeling capacity, and accuracy of laglad processes can not be achieved by using cadlag processes, because they are right-continuous and, consequently, can not have immediate downward jumps. Moreover, even if we use a sequence of right jumps, it is hard to control times at which downward jumps happen after an upward jump, and, thus, even if we tried to model "almost" immediate downward jumps after upward jumps for cadlag processes, we would not succeed.

3.1 Existence and Uniqueness of solutions of optional SDE's

Before starting the investigation of comparison of optional SDE's, we provide a short exposition of results on their existence and uniqueness.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a filtration $\mathbf{F} = (\mathcal{F}_t)_{t\geq 0}$ which is not complete, right- or left-continuous.

Let Y be a one-dimensional optional semimartingale which has the representation (2.2). Consider the following SDE

$$\begin{aligned} X = X_0 + f(X)\cdot a + g(X)\cdot m + (r+w)(X) * \eta \\ + \sum_{j=r,g} \left[U h_j(X) * (\mu^j - \nu^j) + (k_j + l_j)(X) * p^j \right], \quad (3.1) \end{aligned}$$

where X_0 is $\mathcal{F}_0$-measurable random variable; $U = U(u) = \mathbf{1}_{(|u|\leq 1)}(u)$; $a \in \mathcal{A}_{loc}$ is continuous increasing, $m, \mu^j, \nu^j, p^j, \eta$ as in (2.2). Further, for convenience we use the notation $f(X) = f(\omega, t, X_{t-})$, $g(X) = g(\omega, t, X_{t-})$, $h_d(X) = h(\omega, t, u, X_{t-})$, $h_g(X) = h(\omega, t, u, X_t)$ and similarly for k_j, l_j, w and r whenever this does not lead to confusion.

To guarantee well-posedness of the integrals in (3.1) we make the following assumptions.

Assumptions 1.

- For $j = r, g$,

$$|f(X)| \cdot a \in \mathcal{A}_{loc},$$

$$(g(X))^2 \cdot \langle m \rangle \in \mathcal{A}_{loc},$$

$$|h(X)|^2 * \nu^j \in \mathcal{A}_{loc},$$

$$|l_j(X)| * p^j \in \mathcal{A}_{loc},$$

$$\left[|k_j(X)|^2 * p^j\right]^{1/2} \in \mathcal{A}_{loc},$$

$$|r(X)| * \eta \in \mathcal{A}_{loc},$$

$$\left[|w(X)|^2 * \eta\right]^{1/2} \in \mathcal{A}_{loc},$$

  and $E[k_r(S, \beta^d_S, X_0)|\mathcal{F}_{S-}] = 0$ a.s. for any predictable stopping time S on $\{S < \infty\}$ and $E[k_g(T, \beta^g_T, X_0)|\mathcal{F}_T] = 0$, $E[w(T, \beta^g_T, X_0)|\mathcal{F}_T] = 0$ a.s., for any stopping time T on $\{T < \infty\}$.

- $f(\omega, s, x)$ and $g(\omega, s, x)$ are defined on $(\Omega \times \mathbb{R}_+ \times \mathbb{R})$ and are $\mathcal{P} \times \mathcal{B}(\mathbb{R})$-measurable;

- $h_d(\omega, s, u, x), k_d(\omega, s, u, x),\ l_d(\omega, s, u, x)$ are defined on $(\Omega \times \mathbb{R}_+ \times E \times \mathbb{R})$ and $\tilde{\mathcal{P}} \times \mathcal{B}(\mathbb{R})$-measurable;

- $h_g(\omega, s, u, x), k_g(\omega, s, u, x),\ l_g(\omega, s, u, x),\ r(\omega, s, u, x),\ w(\omega, s, u, x)$ are defined on $(\Omega \times \mathbb{R}_+ \times E \times \mathbb{R})$ and $\tilde{\mathcal{O}} \times \mathcal{B}(\mathbb{R})$-measurable.

For convenience we state here sufficient conditions for the existence and the uniqueness of the strong solution of (3.1).

Definition 3.1.1 *Let Assumptions 1 hold. We say that the functions $f, g, h_j, k_j, l_j, r, w$ in* (3.1) *satisfy the* $\mathbf{L(Y, X_0)}$ ***conditions*** *if:*

(L1) there exists non-negative functions $F, G, H^j, L^j, K^j, R, W,\ j = r, g,$ such that

(a) $F(\omega, s)$, $G(\omega, s)$ *are* $\mathcal{P}$*-measurable;* $H^d(\omega, s, u)$ *is* $\mathcal{P} \times \mathcal{B}(E \cap (|u| \leq 1))$*-measurable;*

$H^g(\omega, s, u)$ *is* $\mathcal{O} \times \mathcal{B}(E \cap (|u| \leq 1))$*-measurable;*

$L^d(\omega, s, u), K^d(\omega, s, u)$ *are* $\tilde{\mathcal{P}}$*-measurable;*

$L^g(\omega, s, u), K^g(\omega, s, u), R(\omega, s, u), W(\omega, s, u)$ *are* $\tilde{\mathcal{O}}$*-measurable.*

(b) $F \cdot a_t + G \cdot \langle m \rangle_t + \sum_{j=r,g} [H^j U * \nu^j + (K^j + L^j) * \lambda^j] + (R+W) * \zeta < \infty$ *a.s. for any* $t > 0$, *where* λ^j *and* ζ *are compensators of* p^j *and* η, *respectively.*

(c) *for any* $x, y \in \mathbb{R}$, $s \in \mathbb{R}_+$ *and* $j = r, g$,

$$|f(x) - f(y)| \cdot a_t \leq (F|x-y|) \cdot a_t$$
$$(g(x) - g(y))^2 \cdot \langle m \rangle_t \leq \left(G|x-y|^2\right) \cdot \langle m \rangle_t$$
$$|h_j(x) - h_j(y)| * \nu_t^j \leq \left(H^j|x-y|\right) * \nu_t^j,$$
$$|l_j(x) - l_j(y)| * \lambda_t^j \leq \left(L^j|x-y|\right) * \lambda_t^j,$$
$$|k_j(x) - k_j(y)|^2 * \lambda_t^j \leq \left(K^j|x-y|^2\right) * \lambda_t^j,$$
$$|r(x) - r(y)| * \zeta_t \leq (R|x-y|) * \zeta_t,$$
$$|w(x) - w(y)|^2 * \zeta_t \leq \left(W|x-y|^2\right) * \zeta_t$$

(L2)

$$(g(X_0))^2 \cdot \langle m \rangle_t + f(X_0) \cdot a_t + \left[r(X_0) + (w(X_0))^2\right] * \zeta$$
$$+ \sum_{j=r,g} \left[(h_j(X_0))^2 U * \nu^j + \left[(k_j(X_0))^2 + l_j(X_0)\right] * \lambda^j\right] < \infty$$

a.s. for any $t > 0$.

Theorem 3.1.1 *(see [38, Theorem 1], [39, Theorem 3.3.1]) Let* Y *be an optional semimartingale and suppose that* $f, g, h_j, k_j, l_j, r, w$ *satisfy the* $\mathbf{L}(\mathbf{Y}, \mathbf{X_0})$ *conditions. Then the strong solution of* (3.1) *exists and is unique.*

Remark 3.1.1 *Results in this chapter can be easily generalized to the eq.* (3.1) *with an additional term* $\mathbf{1}_{(|u|>1)} h'_j * \mu^j$ *due to its simple structure (see [11, Lemma 3.2]).*

3.2 Local Times for optional processes

Next, we discuss the notion of a local time for an optional semimartingale which was first introduced in ([62], VI.3.4). This concept is crucial for our proof of comparison of solutions and pathwise uniqueness. A local time at a of an optional semimartingale X is denoted by $L_t^a(X)$ and given by

$$\begin{aligned} L_t^a(X) &= |X_t - a| - |X_0 - a| - \int_0^t \operatorname{sign}(X_{s-} - a) dX_s \\ &- \sum_{0<s\le t} [|X_s - a| - |X_{s-} - a| - \operatorname{sign}(X_{s-} - a)\Delta X_s] \\ &- \sum_{0\le s<t} \left[|X_{s+} - a| - |X_s - a| - \operatorname{sign}(X_s - a)\Delta^+ X_s\right] \end{aligned}$$

where $sign(x) = 1$ if $x > 0$ and $sign(x) = -1$ if $x \le 0$.

Definition 3.2.1 *(cf. [17, Definition 1], [16, Section 2]).*

We say that a coefficient function g of equation (3.1) *satisfies the **LT** condition, if for any two solutions X^1 and X^2 of equation (3.1), the local time at level* 0 *satisfies,*

$$\forall t \ge 0 \quad L_t^0(X^1 - X^2) = 0. \tag{3.2}$$

A generalization of the change of variables formula (Theorem 2.9.2) can be obtained using the local time.

Theorem 3.2.1 *If F is the difference of two convex functions and F' is its left derivative and let ρ be the signed measure which is the second derivative of F. Then* *we have*

$$
\begin{aligned}
F(X_t) =& F(X_0) + \int_0^t F'(X)dX_s + \frac{1}{2}\int_{-\infty}^{\infty} L_t^a \rho(da) \\
&+ \sum_{0<s\leq t} [F(X_s) - F(X_{s-} - F'(X_{s-})\Delta X_s] \\
&+ \sum_{0\leq s<t} \left[F(X_{s+}) - F(X_s - F'(X_s)\Delta^+ X_s\right]
\end{aligned} \tag{3.3}
$$

Proof. First note that, if F is linear function, then by the change of variables formula (see (2.3)), the result holds for $\rho = 0$. For a general function F, let $g(x) = \frac{1}{2}\int |x - y|\rho(dy)$. Then g is a convex function, $g'(x) = \int_{\mathbb{R}} \operatorname{sign}(x-y)\rho(dy)$ and $g''(x) = \rho(dy)$. Therefore $F - g$ has second derivative 0, hence is linear, and so the result holds for the function $F - g$. By linearity, it remains to show that the result holds for the function g. By integrating $\frac{1}{2}|X_t - a|$ with respect to $\rho(da)$, we have that

$$
\begin{aligned}
g(X_t) &= \frac{1}{2}\int_{\mathbb{R}} |X_t - a|\rho(da) \\
&= \frac{1}{2}\int_{\mathbb{R}} |X_0 - a|\rho(da) + \frac{1}{2}\int_{\mathbb{R}} \left(\int_0^t \operatorname{sign}(X_{s-} - a)dX_s + L_t^a\right)\rho(da) \\
&+ \int_{\mathbb{R}} \sum_{0<s\leq t} \left[|X_s - a| - |X_{s-} - a| - \operatorname{sign}(X_{s-} - a)\Delta X_s\right]\rho(da) \\
&+ \int_{\mathbb{R}} \sum_{0\leq s<t} \left[|X_{s+} - a| - |X_s - a| - \operatorname{sign}(X_s - a)\Delta^+ X_s\right]\rho(da) \\
&= g(X_0) + \frac{1}{2}\int_{\mathbb{R}} \left(\int_0^t \operatorname{sign}(X_{s-} - a)dX_s + L_t^a\right)\rho(da) \\
&+ \sum_{0<s\leq t} [g(X_s) - g(X_{s-}) - g'(X_{s-})\Delta X_s] \\
&+ \sum_{0\leq s<t} \left[g(X_{s+}) - g(X_s) - g'(X_s)\Delta^+ X_s\right]
\end{aligned}
$$

By Fubini's theorem,

$$
\frac{1}{2}\int_{\mathbb{R}} \left(\int_0^t \operatorname{sign}(X_{s-} - a)dX_s + L_t^a\right)\rho(da) = \int_0^t g'(X_{s-})dX_s + \frac{1}{2}\int_{\mathbb{R}} L_t^a \rho(da).
$$

Combining these yields the result. ■

Corollary 3.2.1 *(Occupation-Density (cf. [61]) Let X be an optional semimartingale with local time $(L^a)_{a\in\mathbb{R}}$. Let g be a bounded Borel measurable function. Then a.s.*

$$\int_{-\infty}^{\infty} L_t^a g(a)da = \int_0^t g(X_{s-})d\langle X\rangle_s^c. \tag{3.4}$$

Proof. By first assuming that $g = F''$, where F is a convex and twice continuously differentiable function, and simply comparing the change of variables formula (2.3) with the formula in (3.3), we obtain the equality (3.4). Since this formula holds for positive continuous function g, by monotone class argument it must hold, up to a $\mathbf{P}$-null set, for any bounded, Borel measurable function g. ■

Now we present the formula (3.3) using jump measures.

Theorem 3.2.2 *Let X be an optional semimartingale given in* (3.1)*. Let function F be convex on $\mathbb{R}$. Then $F(X) = (F(X)_t)$ is an optional semimartingale and has the following representation*

$$F(X_t) = F(X_0) + \int_{0+}^t F'(X_{s-})dX_s^c + \frac{1}{2}\int_{-\infty}^{\infty} L_t^a \rho(da) \tag{3.5}$$

$$\begin{aligned}
&+ \left[F(X_- + h_d(X_-)) - F(X_-)\right] U * (\mu^d - \nu^d)_t \\
&+ \left[F(X + h_g(X)) - F(X)\right] U * (\mu^g - \nu^g)_t \\
&+ \left[F(X_- + h_d(X_-)) - F(X_-) - F'(X_-)h_d(X_-)\right] U * \nu_t^d \\
&+ \left[F(X + h_g(X)) - F(X) - F'(X)h_g(X)\right] U * \nu_t^g \\
&+ \left[F(X_- + (k_d + l_d)(X_-)) - F(X_-)\right] * p_t^d \\
&+ \left[F(X + (k_g + l_g)(X)) - F(X)\right] * p_t^g \qquad (3.6) \\
&+ \left[F(X + (r + w)(X)) - F(X)\right] * \eta_t.
\end{aligned}$$

Proof. Apply the change of variables formula (3.3) to the optional semimartingale

(2.2)

$$
\begin{aligned}
F(X_t) =& F(X_0) + \int_{0+}^t F'(X_{s-})d(a+m)_s + \frac{1}{2}\int_{-\infty}^{\infty} L_t^a \rho(da) \\
&+ F'(X_-)\left[h_d U * (\mu^d - \nu^d) + [k_d + l_d] * p^d + [r+w] * \eta\right] \\
&+ \sum_{0<s\le t} [F(X_s) - F(X_{s-} - F'(X_{s-})\Delta X_s] \\
&+ F'(X)\left[h_g U * (\mu^g - \nu^g) + [k_g + l_g] * p^g\right] \\
&+ \sum_{0\le s<t} \left[F(X_{s+}) - F(X_s - F'(X_s)\Delta^+ X_s\right]
\end{aligned}
$$

Let us transform sums $\sum_{0<s\le t}[\cdot]$ and $\sum_{0\le s<t}[\cdot]$. Define $C_t = \sum_{0<s\le t}[\cdot]$, $B_t = \sum_{0\le s<t}$ and represent them in the form $C_t = \sum_{i=1}^2 C_t^i,\ B_t = \sum_{j=1}^3 B_t^j$, where

$$
C_t^1 = \sum_{T_n\le t}[\cdot]_{T_n}\mathbf{1}_{|\Delta X_{T_n}|\le 1},\quad C_t^2 = \sum_{S_n\le t}[\cdot]_{S_n},
$$
$$
B_t^1 = \sum_{U_n<t}[\cdot]_{U_n}\mathbf{1}_{|\Delta^+ X_{U_n}|\le 1},\quad B_t^2 = \sum_{S_n<t}[\cdot]_{S_n},\quad B_t^3 = \sum_{T_n<t}[\cdot]_{T_n}.
$$

In ([62], Sec. 3, Theorem 5) it is shown that $C^1 - C^2,\ B^1 - B^3$ belong to $\mathcal{V}$ in case $F(x) = x^+$. The same result follows for the general function F from the facts that $x^+ - x = x^-$ and $|x| = x^+ + x^-$ using the same approach as in Theorem A.1. We are going to rewrite this sums using stochastic integrals.

1. Define $\sigma_N^d = \inf(t > 0 : \int_0^t |dC_s^1| \ge N \textit{ or } |X_{t-}| > N)$, $\sigma_N^g = \inf(t > 0 : \int_0^{t-} |dB_{s+}^1| > N \textit{ or } |X_{t+}| > N)$. For any N, σ_N^d is $\mathbf{F}$-s.t., σ_N^g is $\mathbf{F}_+$-s.t. and $\sigma_N^d \uparrow \infty, \sigma_N^g \uparrow \infty$ a.s. as $n \to \infty$, $\int_{[0,\sigma_N^d]} |dC_t^1| = \int_{[0,\sigma_N^d[} |dC_t^1| + \Delta C^1_{\sigma_N^d} \le N + K$, where $K = \max_{|x|\le N+1}(2F(x) + |F'(x)|)$, such that $|X_{\sigma_N^d}| \le |X_{\sigma_N^d -}| + |\Delta X_{\sigma_N^d}| \le N+1$, thus, $E\int_{[0,\sigma_N^d]} dC_t^1| < \infty$. Moreover, $\int_{[0,\sigma_N^g]} |dB_{t+}^1| \le N$, e.g., $B^1, C^1 \in \mathcal{A}_{loc}$. Functions $F(X_- + h_d) - F(X_-) - h_d F'(X_-)$ and $F(X + h_g) - F(X) - h_g F'(X)$ are $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{O}}$ measurable respectively. From [27] it is easy to see that $\Delta X_{T_n} = h_d(T_n)$, $\Delta^+ X_{T_n} = h_g(T_n)$, e.g. $F(X_{T_n}) = F(X_{T_n -} + h_d(T_n))$, $F(X_{T_{n+}}) = F(X_{T_n} + h_g(T_n))$. Next, from

$B^1, C^1 \in \mathcal{A}_{loc}$ and properties of stochastic integrals with respect to μ^j it follows that

$$C^1 = \left[F^{(d)}_{h_d} - h_d F'(X_-)\right] U * (\mu^d - \nu^d) + \left[F^{(d)}_{h_d} - h_d F'(X_-)\right] U * \nu^d,$$
$$B^1 = \left[F^{(g)}_{h_g} - h_g F'(X)\right] U * (\mu^g - \nu^g) + \left[F^{(g)}_{h_g} - h_g F'(X)\right] U * \nu^g,$$

where $F^{(d)}_{h_j} = F(X_- + h_d) - F(X_-)$, $F^{(g)}_{h_g} = F(X + h_g) - F(X)$. First terms in the above sums belong to $\mathcal{M}^d_{loc}$ and $\mathcal{M}^g_{loc}$,respectively, while second terms belong to $\mathcal{A}_{loc}$. Now, using generalized mean value theorem for convex functions [82] we have for any n on $\{\sigma^d_N < \infty)\}$ and $\{\sigma^g_N < \infty)\}$ respectively

$$\begin{aligned} &|F^{(d)}_{h_d}(T_n, \Delta X_{T_n})|^2 \mathbf{1}_{(|h_d| \le 1, T_n \le \sigma^d_N)} \le \\ &\qquad (\max_{|x| \le N+1} |F'(x)|) |h_d(T_n, \Delta X_{T_n})|^2 \mathbf{1}_{(|h_d| \le 1, T_n \le \sigma^d_N)}, \\ &|F^{(g)}_{h_g}(U_n, \Delta^+ X_{U_n})|^2 \mathbf{1}_{(|h_g| \le 1, U_n \le \sigma^g_N)} \le \\ &\qquad (\max_{|x| \le N+1} |F'(x)|) |h_g(U_n, \Delta^+ X_{U_n})|^2 \mathbf{1}_{(|h_d| \le 1, U_n \le \sigma^g_N)}, \end{aligned}$$

taking into account that $|X_{U_n}| \le |X_{U_{n+}}| + |\Delta^+ X_{U_n}| \le N+1$ in the second inequality. By summing up along n each of these inequalities we obtain

$$|F^{(j)}_{h_j}|^2 U * \mu^j_{\sigma^j_N} \le \left(\max_{|x| \le N+1} |F'(x)|\right) |h_j|^2 U * \mu^j_{\sigma^j_N}.$$

These inequalities and assumptions about h_i imply $|F^{(j)}_{h_j}|^2 U * \mu^j \in \mathcal{A}_{loc}$ and, thus, by properties of stochastic integrals with respect to jump measures $\mu^j - \nu^j$ we get $|F^{(j)}_{h_j}|^2 U * (\mu^j - \nu^j) \in \mathcal{M}^{2,j}_{loc}$. By Cauchy-Schwartz inequality we deduce

$$|F'(X_-)U|^2 * \mu^d \in \mathcal{A}_{loc} \text{ and } |F'(X)U|^2 * \mu^g \in \mathcal{A}_{loc},$$

thus

$$F'(X_-)U * (\mu^d - \nu^d) \in \mathcal{M}^{2,d}_{loc} \text{ and } F'(X)U * (\mu^g - \nu^g) \in \mathcal{M}^{2,g}_{loc}.$$

Finally, we decompose C^1 and B^1 in the following form

$$\begin{aligned}
C^1 =& F_{h_d}^{(d)} U * (\mu^d - \nu^d) - [h_d F'(X_-)]\, U * (\mu^d - \nu^d) \\
&+ \left[F_{h_d}^{(d)} - h_d F'(X_-) \right] U * \nu^d, \\
B^1 =& F_{h_g}^{(g)} U * (\mu^g - \nu^g) - [h_g F'(X)]\, U * (\mu^g - \nu^g) \\
&+ \left[F_{h_g}^{(g)} - h_g F'(X) \right] U * \nu^g.
\end{aligned}$$

where the first two terms in each formula are in $\mathcal{M}_{loc}^{2,d}$ and $\mathcal{M}_{loc}^{2,g}$ respectively, and last terms are in $\mathcal{A}_{loc}$.

2. Since the processes

$$\sum_{S_n \le t} F'(X_{S_{n-}}) \Delta X_{S_n}, \quad \sum_{S_n < t} F'(X_{S_n}) \Delta^+ X_{S_n}, \quad \sum_{T_n \le t} F'(X_{T_n}) \Delta^+ X_{T_n},$$

are semimartingales, then we represent the processes C^2, B^2 and B^3 as

$$\begin{aligned}
C^2 &= F_{k_d + l_d}^{(d)} * p^d - [k_d + l_d] F'(X_-) * p^d, \\
B^2 &= F_{k_g + l_g}^{(g)} * p^g - [k_g + l_g] F'(X) * p^g, \\
B^3 &= F_{r+w}^{(g)} * \eta - [r + w] F'(X) * \eta,
\end{aligned}$$

where all terms on the right side are semimartingales. By plugging $\sum_{i=1}^2 C_t^i$, $\sum_{j=1}^3 B_t^j$ instead of $\sum_{0<s\le t}[\cdot]$ and $\sum_{0\le s<t}[\cdot]$, respectively, we get the required change of variables formula. ■

Theorem 3.2.3 *Let X be an optional semimartingale satisfying*

$$\sum_{0<s\le t} |\Delta X_s| + \sum_{0\le s<t} |\Delta^+ X_s| < \infty.$$

Then there exists a version of $\mathcal{B}(\mathbb{R}) \times \mathcal{P}$ version of $(a, t, \omega) \to L_t^a(\omega)$ which is everywhere jointly right continuous in a and continuous in t.

Proof. The process $J_t = \sum_{0<s\le t} \Delta X_s + \sum_{0\le s<t} \Delta^+ X_s$ is a finite variation optional semimartingale, and $Y = X - J$ is a continuous optional semimartingale. We let $Y = M + A$ be the (unique) decomposition of Y, with $M_0 = A_0 = 0$. Then $X = M + A + J$. Further define

$$\begin{aligned} S_t^a = &\sum_{0<s\le t} \mathbf{1}_{(X_{s-}>a)}(X_s - a)^- + \mathbf{1}_{(X_{s-}\le a)}(X_s - a)^+ \\ &+ \sum_{0\le s<t} \mathbf{1}_{(X_s>a)}(X_{s+} - a)^- + \mathbf{1}_{(X_s\le a)}(X_{s+} - a)^+ \end{aligned}$$

Observe that $|S_t^a| \le \sum_{0<s\le t} |\Delta X_s| + \sum_{0\le s<t} |\Delta^+ X_s| < \infty$. By the change of variables formula

$$\begin{aligned} (X_t - a)^+ - (X_0 - a)^+ = &\int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dA_s + \int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dM_s \\ &+ \int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dJ_s + S_t^a + \frac{1}{2} L_t^a. \end{aligned}$$

We have

$$\begin{aligned} \lim_{a\to b, a<b} \int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dA_s &= \int_{0+}^t \mathbf{1}_{(X_{s-}>b)} dA_s \\ \lim_{a\to b, a>b} \int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dA_s &= \int_{0+}^t \mathbf{1}_{(X_{s-}\ge b)} dA_s \end{aligned}$$

where the convergence is uniform in t. We have similar results for

$$\int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dJ_s$$

and also for S_t^a, because it is dominated by $\sum_{0<s\le t} |\Delta X_s| + \sum_{0\le s<t} |\Delta^+ X_s| < \infty$. Since we already know that $\int_{0+}^t \mathbf{1}_{(X_{s-}>a)} dM_s$ is continuous, the proof is complete. ■

3.3 Comparison Theorem via local times

Let us investigate comparison of solutions of stochastic differential equations driven by optional semimartingales. In this section we consider two processes given by SDE's of the same type as equation (3.1):

$$X^i = X_0^i + f^i(X^i)\cdot a + g(X^i)\cdot m \\ + \sum_{j=r,g}\left[Uh_j(X^i) * (\mu^j - \nu^j) + (k_j^i + l_j^i)(X^i) * p^j\right] + (r^i + w^i)(X^i) * \eta, \quad i = 1,2. \tag{3.7}$$

We are going to present a general version of a Comparison Theorem with **LT** condition on g and the following conditions on functions $f^i, h_j, k_j^i,\ l_j^i, r^i, w^i,\ i = 1,2:$

D Conditions. Suppose that

(D1) $X_0^2 \geq X_0^1$;

(D2) $f^2(s,x) \geq f^1(s,x)$ for any $s \in \mathbb{R}_+,\ x \in \mathbb{R}$;

(D3) For any $s \in \mathbb{R}_+,\ u \in E,\ x, y \in \mathbb{R}, y \geq x$

$$y + h_j(s,u,y) \geq x + h_j(s,u,x),$$
$$y + k_j^2(s,u,y) + l_j^2(s,u,y) \geq x + k_j^2(s,u,x) + l_j^2(s,u,x),$$
$$y + r^2(s,u,y) + w^2(s,u,y) \geq x + r^2(s,u,x) + w^2(s,u,x);$$

(D4) For any $s \in \mathbb{R}_+,\ u \in E,\ x \in \mathbb{R}$

$$k_j^2(s,u,x) + l_j^2(s,u,x) \geq k_j^1(s,u,x) + l_j^1(s,u,x),$$
$$r^2(s,u,x) + w^2(s,u,x) \geq r^1(s,u,x) + w^1(s,u,x);$$

Theorem 3.3.1 *Suppose that $f^i, g, h_j, k_j^i, l_j^i, r^i, w^i$ in (3.7) satisfy $\mathbf{L}(\mathbf{Y}, \mathbf{X_0^i})$, $i = 1,2$, and conditions* ***D*** *and* ***LT*** *hold. Then there exist unique strong solutions X^1 and X^2, and $X_t^1 \leq X_t^2$ for all $t \in \mathbb{R}_+$ a.s. ($X^1 \leq X^2$).*

Proof. Let $Y := X^1 - X^2$ and

$$\begin{aligned}
I_1 :=&\mathbf{1}_{(Y_->0)}(f^1(X^1) - f^2(X^2)) \cdot a_t,\\
I_2 :=&\mathbf{1}_{(Y_->0)}(g(X^1) - g(X^2)) \cdot m_t\\
&+ \left[(Y_- + h_d(X^1) - h_d(X^2))^+ - Y_-^+\right] U * (\mu^d - \nu^d)_t\\
&+ \left[(Y + h_g(X^1) - h_g(X^2))^+ - Y^+\right] U * (\mu^g - \nu^g)_t,\\
I_3 :=& \left[(Y_- + h_d(X^1) - h_d(X^2))^+ - Y_-^+ - (h_d(X^1) - h_d(X^2))\mathbf{1}_{(Y_->0)}\right] U * \nu_t^d\\
&+ \left[(Y + h_g(X^1) - h_g(X^2))^+ - Y^+ - (h_g(X^1) - h_g(X^2))\mathbf{1}_{(Y>0)}\right] U * \nu_t^g,\\
I_4 :=& \left[(Y_- + k_d^1(X^1) + l_d^1(X^1) - k_d^2(X^2) - l_d^2(X^2))^+ - Y_-^+\right] * p_t^d\\
&+ \left[(Y + k_g^1(X^1) + l_g^1(X^1) - k_g^2(X^2) - l_g^2(X^2))^+ - Y^+\right] * p_t^g,\\
I_5 :=& \left[(Y + r^1(X^1) + w^1(X^1) - r^2(X^2) - w^2(X^2))^+ - Y^+\right] * \eta_t.
\end{aligned}$$

By Theorem 3.2.2, Y^+ is expressed in the following form

$$Y_t^+ = Y_0^+ + \frac{1}{2}L_t^0(Y) + I_1 + I_2 + I_3 + I_4 + I_5. \tag{3.8}$$

After using (D1) and **LT** conditions, equation (3.8) becomes

$$Y_t^+ = I_1 + I_2 + I_3 + I_4 + I_5.$$

Next, we examine each of $I_1 - I_5$ separately. By (D2) and (L1) conditions,

$$\begin{aligned}
I_1 =&\mathbf{1}_{(Y_->0)}\left[(f^1(X^1) - f^2(X^1)) + (f^2(X^1) - f^2(X^2))\right] \cdot a_t\\
\leq& F Y_-^+ \cdot a_t.
\end{aligned}$$

Now, consider I_3

$$\begin{aligned}
I_3 =& \left[(Y_- + h_d(X^1) - h_d(X^2))^+ - (Y_- + h_d(X^1) - h_d(X^2))\mathbf{1}_{(Y_->0)}\right] U * \nu_t^d\\
&+ \left[(Y + h_g(X^1) - h_g(X^2))^+ - (Y + h_g(X^1) - h_g(X^2))\mathbf{1}_{(Y>0)}\right] U * \nu_t^g
\end{aligned}$$

Further, applying the identity $I = I^+ - I^-$ to the terms $Y_- + h_d(X^1) - h_d(X^2)$ and $Y + h_g(X^1) - h_g(X^2)$, we get

$$\begin{aligned} I_3 =& \Big[(Y_- + h_d(X^1) - h_d(X^2))^- \mathbf{1}_{(X^1_- > X^2_-)} \\ &+ (Y_- + h_d(X^1) - h_d(X^2))^+ \mathbf{1}_{(X^1_- \leq X^2_-)} \Big] U * \nu^d_t \\ &+ \Big[(Y + h_g(X^1) - h_g(X^2))^- \mathbf{1}_{(X^1 > X^2)} \\ &+ (Y + h_g(X^1) - h_g(X^2))^+ \mathbf{1}_{(X^1 \leq X^2)} \Big] U * \nu^g_t. \end{aligned}$$

It follows from (D3), that $I_3 = 0$.

Using (D4), we get

$$\begin{aligned} I_5 =& \mathbf{1}_{(Y>0)} \left[r^2(X^1) + w^2(X^1) - r^2(X^2) - w^2(X^2)) \right] * \eta_t \\ &+ \big[(Y + r^1(X^1) + w^1(X^1) - r^2(X^1) - w^2(X^1) \\ &+ r^2(X^1) + w^2(X^1) - r^2(X^2) - w^2(X^2))^+ \\ &- Y^+ - \mathbf{1}_{(Y>0)} \left(r^2(X^1) + w^2(X^1) - r^2(X^2) - w^2(X^2)) \right) \big] * \eta_t \\ \leq& \mathbf{1}_{(Y>0)} \left[r^2(X^1) + w^2(X^1) - r^2(X^2) - w^2(X^2) \right] * \eta_t \\ &+ \big[(Y + r^2(X^1) + w^2(X^1) - r^2(X^2) - w^2(X^2))^+ \\ &- \mathbf{1}_{(Y>0)} \left(Y + r^2(X^1) + w^2(X^1) - r^2(X^2) - w^2(X^2)) \right) \big] * \eta_t. \end{aligned}$$

Due to (D3) and (L1) conditions

$$I_5 \leq RY^+ * \eta_t + \mathbf{1}_{(Y>0)} \left[w^2(X^1) - w^2(X^2) \right] * \eta_t.$$

Repeating the same calculations for I_4, we obtain that

$$\begin{aligned} I_4 \leq& L^d Y^+_- * p^d + L^g Y^+ * p^g \\ &+ \mathbf{1}_{(Y_->0)} \left[k^2_d(X^1) - k^2_d(X^2) \right] * p^d_t \\ &+ \mathbf{1}_{(Y>0)} \left[k^2_g(X^1) - k^2_g(X^2) \right] * p^g_t \end{aligned}$$

By combining all the estimates for $I_1,\ I_3 - I_5$ with (3.3), we have

$$Y_t^+ \le M_t + Y^+ \circ C_t,$$

where $\circ$ means an optional stochastic integral (see [4, Section 7.1]), $C := F \cdot a_t + W * \eta_t + L^d * p^d + L^g * p^g$ is a non-negative increasing process, and

$$\begin{aligned} M_t :=& I_2 + \mathbf{1}_{(Y>0)} \left[w^2(X^1) - w^2(X^2) \right] * \eta_t \\ &+ \mathbf{1}_{(Y_->0)} \left[k_d^2(X^1) - k_d^2(X^2) \right] * p_t^d + \mathbf{1}_{(Y>0)} \left[k_g^2(X^1) - k_g^2(X^2) \right] * p_t^g, \\ M_0 =& 0. \end{aligned}$$

Using Assumptions 1, we have

$$\begin{gathered} \mathbf{1}_{(Y_->0)}(g(X^1) - g(X^2))^2 \cdot \langle m \rangle_t \le \\ 2(g(X^1))^2 \cdot \langle m \rangle_t + 2(g(X^2))^2 \cdot \langle m \rangle_t \in \mathcal{A}_{loc}, \\ \left[(Y_- + h_d(X^1) - h_d(X^2))^+ - Y_-^+ \right]^2 U * \nu_t^d \le \\ 2 \left[h_d(X^1) \right]^2 * \nu_t^d + 2 \left[h_d(X^2) \right]^2 * \nu_t^d \in \mathcal{A}_{loc}, \\ \left[(Y + h_g(X^1) - h_g(X^2))^+ - Y^+ \right]^2 U * \nu_t^g \le \\ 2 \left[h_g(X^1) \right]^2 * \nu_t^g + 2 \left[h_g(X^2) \right]^2 * \nu_t^g \in \mathcal{A}_{loc}, \\ \left[\mathbf{1}_{(Y>0)} \left[w^2(X^1) - w^2(X^2) \right]^2 * \eta_t \right]^{1/2} \le \\ \left[2 \left[w^2(X^1) \right]^2 * \eta_t \right]^{1/2} + \left[2 \left[w^2(X^2) \right]^2 * \eta_t \right]^{1/2} \in \mathcal{A}_{loc}, \\ \left[\mathbf{1}_{(Y_->0)} \left[k_d^2(X^1) - k_d^2(X^2) \right]^2 * p_t^d \right]^{1/2} \le \\ \left[2 \left[k_d^2(X^1) \right]^2 * p_t^d \right]^{1/2} + \left[2 \left[k_d^2(X^2) \right]^2 * p_t^d \right]^{1/2} \in \mathcal{A}_{loc}, \\ \left[\mathbf{1}_{(Y>0)} \left[k_g^2(X^1) - k_g^2(X^2) \right]^2 * p_t^g \right]^{1/2} \le \\ \left[2 \left[k_g^2(X^1) \right]^2 * p_t^g \right]^{1/2} + \left[2 \left[k_g^2(X^2) \right]^2 * p_t^g \right]^{1/2} \in \mathcal{A}_{loc}. \end{gathered}$$

Thus, M is an optional local martingale (see [4, Section 7.4.2, p. 234]).

Now, by Grownwall lemma (see [10, Lemma 3.2])

$$Y_t^+ \leq \mathcal{E}_t(C)M_t.$$

Since C_t is an increasing process, $\mathcal{E}_t(C) \geq 0$. Thus, $M_t \geq 0$ since $Y_t^+ \geq 0$. Therefore, M_t is a non-negative optional local martingale and, consequently, by Lemma 2.4.2 it is a non-negative supermartingale starting from 0. It follows that $M_t = 0$ for all $t \in \mathbb{R}_+$ a.s. Hence, $Y^+ \leq 0$, and $X^1 \leq X^2$. ■

Next, we show generality of **LT** condition imposed on function g.

Definition 3.3.1 *(Yamada condition (see [89, Theorem 1.1], [11, Theorem 3.2]). We say that a coefficient function g of equation* (3.1) *satisfies the Yamada condition, if there exists a non-negative non-decreasing function $\rho(u)$ on $\mathbb{R}_+$ and a $\mathcal{P}$-measurable non-negative function G such that*

$$|g(x) - g(y)| \leq \rho(|x-y|)G(s),$$
$$G^2 \cdot \langle m\rangle_s < \infty \ a.s., \int_0^\epsilon \rho^{-2}(u)du = \infty \ \textit{for any } s \in \mathbb{R}_+, \epsilon > 0, x, y \in \mathbb{R}.$$

Lemma 3.3.1 *If g satisfies Yamada condition then **LT** condition holds.*

Proof. Using the formula of occupation density we have

$$\begin{aligned}\int_0^\infty L_t^a(X^1 - X^2)\rho^{-2}(a)da &= \int_0^t \mathbf{1}_{(X^1 - X^2 > 0)}\rho^{-2}(X_s^1 - X_s^2)d\langle X^c\rangle_s \\ &= \int_0^t \mathbf{1}_{(X^1 - X^2 > 0)}\rho^{-2}(X_s^1 - X_s^2) \\ &\quad \times \left[g(X^1) - g(X^2)\right]^2 d\langle m\rangle_s \\ &< \infty\end{aligned}$$

Thus, since $a \mapsto L_t^a(X^1 - X^2)$ is right-continuous and $\int_0^\epsilon \rho^{-2}(u)du = \infty, \forall\epsilon > 0$, it follows that g satisfies **LT** condition. ■

Example 3.3.1 *We provide an example of the function $g(x)$ which satisfies **LT** condition but does not satisfy Yamada condition. Let $g(x) = 1+[\log(|x|^{-1} \vee 2)]^{-p}$ ($p > 0$), it can be shown (see [79, Example 3]) that Yamada condition does not hold. On the other hand, by the mean value theorem for some $c > 0$, and for all $0 < y < \frac{1}{4}$,*

$$(g(x+y) - g(x))^2 \leq cy|\log y|(1 + [\log(|x|^{-1} \vee 2)]^{-p})^2|x|^{-1} \\ \times \left(\log\left(|x|^{-1}\right)\right)^{-(2p+1)} \mathbf{1}_{[-3/4,1/2]}(x)$$

Let $dX_t^1 = g(X_t^1)dW_t$ and $dX_t^2 = g(X^2)dW_t$ be two strong solutions (W is a Wiener process), then by using the formula of occupation density and the above inequality we prove that

$$\begin{aligned}\int_0^{1/4} \frac{1}{a \log a} L_t^a(X^1 - X^2)da &= \int_0^t \mathbf{1}_{(\frac{1}{4} > X^1 - X^2 > 0)} \frac{[g(X^1) - g(X^2)]^2}{(X_s^1 - X_s^2)\log(X_s^1 - X_s^2)} ds \\ &\leq c \int_0^t \frac{(1 + [\log(|X^2|^{-1} \vee 2)]^{-p})^2}{|X^2| \left(\log\left(|X^2|^{-1}\right)\right)^{(2p+1)}} \mathbf{1}_{[-3/4,1/2]}(X^2)ds \\ &< \infty\end{aligned}$$

*because the expression under the integral sign is Lebesgue integrable over compacts. Thus, since $a \mapsto L_t^a(X^1 - X^2)$ is right-continuous and $\int_0^{1/4} \frac{1}{a \log a} da = \infty$ it follows that g satisfies **LT** condition.*

Remark 3.3.1 *The local time technique allows us to prove the comparison theorem in a short and concise way. As shown in Example 3.3.1, **LT** condition in Theorem 3.3.1 is generally weaker than Yamada condition in [11]. Notice further, that conditions on functions f^i, l_j^i and r^i are weakened in the sense that inequalities in (D2) and (D4) are not strict as the ones given in [11, Theorem 3.1]. In addition, we have not used conditions (A4) and (A8) from [11, Theorem 3.1] in our proof.*

Remark 3.3.2 *Note that the condition on the function g in (L1)-c) guarantees fulfillment of the (**LT**) condition by Lemma 3.3.1, but not other way around.*

3.4 Comparison of solutions of SDEs with different jump-diffusions

In this section we expand the comparison theorem proved for SDEs with different diffusions in [29] to the optional jump-diffusion case. For the sake of brevity, here we want to compare two processes following a simplified version of SDE (3.7) in the form of

$$X_t^i = X_0^i + f^i(X^i) \cdot a_t + g_i(X^i) \cdot m_t + \sum_{j=r,g} \left[U h_j^i(X^i) * (\mu^j - \nu^j)_t \right], \quad i = 1, 2, \quad (3.9)$$

with initial condition $X_0^i = x_0^i$. Since now g_i and h_j^i can be all different, we should put stronger conditions on f^i, g_i, h_j^i and $x_0^i,\ i = 1, 2$.

Denote

$$\begin{aligned} F_i(z) &:= \int_{x_0^i}^{z} \frac{dx}{g_i(x)}, \\ \tilde{f}^i(z) &:= \frac{f^i(z)}{g_i(z)} - \frac{1}{2} g_i'(z)\alpha - \sum_{j=r,g} \int_E \left[\frac{h_j^i(z,u)}{g_i(z)} \right] U \tilde{\nu}^j(du), \\ \tilde{h}_j^i(z) &:= \int_z^{z+h_j^i(z)} \frac{dx}{g_i(x)}. \end{aligned}$$

where α and $\tilde{\nu}^j$ are given below (see B1).

Let us introduce the following

B Conditions.

(B1) (structural conditions): There exist densities

$$\begin{aligned} \alpha &= \frac{d\langle m \rangle_t}{da_t}, \\ \nu^d(\omega, (0,t], \Gamma) &= \int_{0+}^{t} \tilde{\nu}^d(\omega, s, \Gamma) da_s, \\ \nu^g(\omega, [0,t), \Gamma) &= \int_{0}^{t-} \tilde{\nu}^g(\omega, s, \Gamma) da_{s+}. \end{aligned}$$

(B2) g_i $i = 1, 2$, is positive and continuously differentiable in z such that for any $s \in \mathbb{R}_+$, $z \in \mathbb{R}$

$$F_1(s, z) \geq F_2(s, z) \tag{3.10}$$

(B3) For any $s \in \mathbb{R}_+$, $u \in E$, $z, y \in \mathbb{R}, y \geq z$

$$\begin{aligned} \tilde{f}^1(s, z) &\leq \tilde{f}^2(s, y), \\ \tilde{h}_j^1(s, u, z) &\leq \tilde{h}_j^2(s, u, y), \\ z + \tilde{h}_j^2(s, u, z) &\leq y + \tilde{h}_j^2(s, u, y). \end{aligned}$$

(B4) $\tilde{f}^2(z)$ and $\tilde{h}_j^2(z)$ are Lipschitz continuous.

Theorem 3.4.1 *Suppose that f^i, g_i, h_j^i in (3.9) satisfy $\mathbf{L}(\mathbf{Y}, \mathbf{X_0^i}), i = 1, 2,$ and conditions **B** hold. Then there exist unique strong solutions X^1 and X^2, and $X^1 \leq X^2$.*

Proof. We transform the processes X^i with the help of the change of variables formula and structural conditions, $i = 1, 2$,

$$\begin{aligned} \tilde{X}_t^i :=& \int_{x_0^i}^{X_t^i} \frac{dx}{g_i(x)} \\ =& \frac{f^i(X_{t-}^i)}{g_i(X_{t-}^i)} \cdot a_t + m_t - \frac{1}{2} g_i'(X_{t-}^i) \cdot \langle m \rangle_t \\ &+ \int_{X_{t-}^i}^{X_{t-}^i + h_d^i(X_{t-}^i)} \frac{dx}{g_i(x)} U * \mu_t^d + \int_{X_t^i}^{X_t^i + h_g^i(X_t^i)} \frac{dx}{g_i(x)} U * \mu_t^g \\ &- \frac{h_d^i(X_{t-}^i)}{g_i(X_{t-}^i)} U * \nu_t^d - \frac{h_g^i(X_t^i)}{g_i(X_t^i)} U * \nu_t^g \\ =& \tilde{f}^i(F_i^{-1}(\tilde{X}_{t-}^i)) \cdot a_t + m_t + \tilde{h}_d^i(F_i^{-1}(\tilde{X}_{t-}^i)) U * \mu_t^d + \tilde{h}_g^i(F_i^{-1}(\tilde{X}_t^i)) U * \mu_t^g \end{aligned}$$

From the condition (B2), it follows that $F_1^{-1}(z) \leq F_2^{-1}(z)$ and, consequently, for any

z

$$\tilde{f}^1(F_1^{-1}(z)) \leq \tilde{f}^2(F_2^{-1}(z)), \tag{3.11}$$

$$\tilde{h}_j^1(F_1^{-1}(z)) \leq \tilde{h}_j^2(F_2^{-1}(z)) \tag{3.12}$$

$$z + \tilde{h}_j^2(s,u,z) \leq y + \tilde{h}_j^2(s,u,y). \tag{3.13}$$

by applying (B3).

Functions $\tilde{f}^2(F_2^{-1}(x))$ and $\tilde{h}_j^2(F_2^{-1}(x))$ are obviously Lipschitz continuous since F_2^{-1} is continuously differentiable transformation and (B4).

Now, we cannot directly use Theorem 3.3.1 because

$$\tilde{h}_j^1(F_1^{-1}(z)) \neq \tilde{h}_j^2(F_2^{-1}(z))$$

in general. Instead, notice that

$$\begin{aligned}
Y^+ :=& (\tilde{X}_t^1 - \tilde{X}_t^2)^+ \\
=& \mathbf{1}_{Y_- > 0}\Big[\tilde{f}^1(F_1^{-1}(\tilde{X}_{t-}^1)) - \tilde{f}^2(F_2^{-1}(\tilde{X}_{t-}^1)) \\
& + \tilde{f}^2(F_2^{-1}(\tilde{X}_{t-}^1)) - \tilde{f}^2(F_2^{-1}(\tilde{X}_{t-}^2))\Big] \cdot a_t \\
& + \left[\Big(Y + \tilde{h}_d^1(F_1^{-1}(\tilde{X}_{t-}^1)) - \tilde{h}_d^2(F_2^{-1}(\tilde{X}_{t-}^2))\Big)^+ - Y^+\right] U * \mu_t^d \\
& + \left[\Big(Y + \tilde{h}_g^1(F_1^{-1}(\tilde{X}_t^1)) - \tilde{h}_g^2(F_2^{-1}(\tilde{X}_t^2))\Big)^+ - Y^+\right] U * \mu_t^g \\
=& I_1 + I_2 + I_3.
\end{aligned}$$

By (3.11) and Lipschitz continuity, $I_1 \leq const.Y^+ \cdot a_t$. Next, applying the same approach as in finding inequality for I_5 in the proof of Theorem 3.3.1 and using (3.12), (3.13) and Lipschitz continuity, we get $I_2 \leq const.Y^+ U * \mu_t^d$ and $I_3 \leq const.Y^+ U * \mu_t^g$. Consequently, by Gronwall Lemma we prove that $\tilde{X}^1 \leq \tilde{X}^2$.

This, together with (B2), implies that

$$\int_{x_0^2}^{X_t^1} \frac{dx}{g_2(x)} \leq \int_{x_0^1}^{X_t^1} \frac{dx}{g_1(x)} \leq \int_{x_0^2}^{X_t^2} \frac{dx}{g_2(x)}$$

Since $g_2(x) > 0$ we conclude that $X^1 \leq X^2$. ■

Remark 3.4.1 *Proceeding with the same technique for jump measures as in the above proof, Theorem 3.4.1 can be directly extended to solutions of* (3.1).

Let us give specific examples.

Example 3.4.1 *Let X^1 and X^2 follows the equation* (3.9) *with $m_t = W_t, a_t = t$. For $A \in \mathcal{B}(\mathbb{R}_+)$ and $\Gamma \in \mathcal{E}$ the Poisson random measures μ^d and μ^g are defined as*

$$\begin{aligned} \mu^d (A \times \Gamma) &: = \# \left\{ \left(t, \Delta L_t^d\right) \in A \times \Gamma | t > 0 \text{ such that } \Delta L_t^d \neq 0 \right\} \\ \mu^g (A \times \Gamma) &: = \# \left\{ \left(t, \Delta^+ L_t^g\right) \in A \times \Gamma | t > 0 \text{ such that } \Delta^+ L_t^g \neq 0 \right\} \end{aligned}$$

where L_t^1 and L_t^2 are a Poisson process and a left-continuous modification of a Poisson process with constant intensities $\gamma^d = 1$ and $\gamma^g = 2$, respectively, and compensators $\nu^d = \gamma^d t$ and $\nu^g = \gamma^g t$. Furthermore, we assume that L^1 and L^2 are independent. We have

$$\begin{aligned} &f^1(z) = 0, && g_1(z) = 1, && x_0^1 = 0; \\ &f^2(z) = \frac{0.15cos(z)}{(1 - 0.3sin(z))^3}, && g_2(z) = (1 - 0.3sin(z))^{-1}, && x_0^2 = 0; \end{aligned}$$

Firstly, we consider the case with no jumps, i.e. $h_j^i = 0, \ i = 1, 2$. It is easy to check that all assumptions of Theorem 3.4.1 hold, and, thus, $X^1 \leq X^2$.

Next, let us assume that $h_j^i = 1, \ i = 1, 2$. Intuitively, by adding the compensated jumps with the same magnitude we anticipate the same result as in the continuous case. However, condition (B3) of Theorem 3.4.1 is clearly not satisfied. Therefore, $X^1 \leq X^2$ is not necessarily true.

Finally, assume that $h_j^1 = 0.7, h_j^2 = 1$ *and* $f^1(z) = -1.8$*, other functions stay the same. Since*

$$\tilde{f}^1(z) = -3.9,$$

$$\tilde{f}^2(y) = -3(1 - 0.3sin(y))$$

$$\tilde{h}_j^1(z) = \int_z^{z+0.7} \frac{dx}{1} = 0.7,$$

$$\tilde{h}_j^2(y) = \int_y^{y+1} \frac{dx}{(1 - 0.3sin(x))^{-1}} = 1 + 0.3cos(y+1) - 0.3cos(y)$$

then (B3) holds. In addition, $\tilde{f}^2$ *and* $\tilde{h}_j^2$ *are obviously Lipschitz continuous. Thus, by Theorem 3.4.1,* $X^1 \leq X^2$.

Example 3.4.2 *Let* X^1 *and* X^2 *follows the equation* (3.9) *with the same* $a_t, m_t, \mu_t^j, \lambda_t^j$, *as in Example 3.2. and*

$$\begin{array}{llll} f^1(z) = -0.5e^{-2z}, & g_1(z) = e^{-z}, & h_j^1 = 0, & x_0^1 = 0; \\ f^2(z) = 0.3 + z + z^3, & g_2(z) = 1 + z^2, & h_j^2 = 1, & x_0^2 = 0; \end{array}$$

It is not hard to show that all assumptions of Theorem 3.4.1 hold, and, thus, $X^1 \leq X^2$.

3.5 Approximation of option price bounds using comparison property

The comparison theorems considered in the previous sections can be applied to find boundaries of option prices in case of so-called Constant Elasticity of Variance (CEV) model. This idea was introduced in [53] and developed further in [52]. Here we extend it to jump-diffusion CEV model and solve the problem formulated in [52]. Option pricing for jump-diffusion financial market models were also considered in the context of imperfect (quantile and efficient) hedging (see, for example, [50], [54]).

CEV model was proposed by Cox and Ross [23]. It is often used in mathematical finance to capture leverage effects and stochasticity of volatility. It is also widely used by practitioners in the financial industry for modeling equities and commodities. Consider a more general version of the jump-diffusion CEV model [91] where the stock price is said to satisfy the following integral equation,

$$\begin{aligned} S_t =& \rho \int_0^t S_{s-} ds + \sigma S^\alpha \cdot W_t + S_{t-} U * \left(\mu^1 - \nu^1\right)_t + S_t U * \left(\mu^2 - \nu^2\right)_t, \\ S_0 =& s, \end{aligned} \tag{3.14}$$

where ρ and σ are constants. W_t is a Wiener process, $\mu^1 - \nu^1$ is a compensated measure of left jumps and $\mu^2 - \nu^2$ is a compensated measure of right jumps. For $B \in \mathcal{B}(\mathbb{R}_+)$ and $\Gamma \in \mathcal{E}$ the jump measures are defined as follows

$$\begin{aligned} \mu^1 (B \times \Gamma) \;\; &: \;\; = \# \left\{ \left(t, \Delta L_t^1\right) \in B \times \Gamma | t > 0 \textit{ such that } \Delta L_t^1 \neq 0 \right\} \\ \mu^2 (B \times \Gamma) \;\; &: \;\; = \# \left\{ \left(t, \Delta^+ L_t^2\right) \in B \times \Gamma | t > 0 \textit{ such that } \Delta^+ L_t^2 \neq 0 \right\} \end{aligned}$$

where L_t^1 and L_t^2 are a Poisson process and a left-continuous modification of a Poisson process with constant intensities γ^1 and γ^2, respectively, and L^1 and L^2 are independent. Hence, $\nu^1 = \gamma^1 t$ and $\nu^2 = \gamma^2 t$.

Consider the function

$$\begin{aligned} F(x) &= \frac{1}{\sigma} \int_s^x u^{-\alpha} du = \frac{x^{1-\alpha} - s^{1-\alpha}}{\sigma(1-\alpha)}, \text{ and find} \\ F'(x) &= \frac{x^{-\alpha}}{\sigma}, \quad F''(x) = \frac{-\alpha x^{-\alpha-1}}{\sigma}. \end{aligned}$$

where $0 < \alpha < 1$.

Denote $X_t = F(S_t)$ and apply the change of variables formula (see Theorem 2.9.2).

We have

$$
\begin{aligned}
X_t =& F'(S)\left[\rho\int_0^t S_{s-}ds+\sigma S^{\alpha}\cdot W_t\right]+\int_0^t \frac{\sigma^2}{2}F''(S_{s-})S_{s-}^{2\alpha}ds\\
&+\left[F(2S_{t-})-F(S_{t-})\right]U*(\mu^1-\nu^1)_t\\
&+\left[F(2S_t)-F(S_t)\right]U*(\mu^2-\nu^2)_t\\
&+\left[F(2S_{t-})-F(S_{t-})-S_{t-}F'(S_{t-})\right]U*\nu^1_t\\
&+\left[F(2S_t)-F(S_t)-S_tF'(S_t)\right]U*\nu^2_t\\
=&\int_0^t\left[kS_{s-}^{1-\alpha}-\frac{\sigma\alpha}{2}S_{s-}^{\alpha-1}\right]ds+W_t\\
&+cS_{t-}^{1-\alpha}U*(\mu^1-\nu^1)_t+cS_t^{1-\alpha}U*(\mu^2-\nu^2)_t\\
=&\int_0^t\left[k\left(X_{s-}\sigma(1-\alpha)+s^{1-\alpha}\right)-\frac{\sigma\alpha}{2}\left(X_{s-}\sigma(1-\alpha)+s^{1-\alpha}\right)^{-1}\right]ds+W_t\\
&+c\left(X_{t-}\sigma(1-\alpha)+s^{1-\alpha}\right)U*(\mu^1-\nu^1)_t\\
&+c\left(X_t\sigma(1-\alpha)+s^{1-\alpha}\right)U*(\mu^2-\nu^2)_t
\end{aligned}
$$

where $k:=\frac{\rho}{\sigma}+(2^{1-\alpha}-1-\frac{1}{\sigma})\lambda_1U+(2^{1-\alpha}-1-\frac{1}{\sigma})\lambda_2U$ and $c:=\frac{2^{1-\alpha}-1}{\sigma(1-\alpha)}$.

With the Comparison Theorem 3.3.1, we can give an estimate of the process X_t from above by a new process Y_t, satisfying the equation,

$$
\begin{aligned}
Y_t &= \int_0^t\left[k\left(Y_{s-}\sigma(1-\alpha)+s^{1-\alpha}\right)\right]ds+W_t\\
&\quad +c\left(Y_{t-}\sigma(1-\alpha)+s^{1-\alpha}\right)U*(\mu^1-\nu^1)_t\\
&\quad +c\left(Y_t\sigma(1-\alpha)+s^{1-\alpha}\right)U*(\mu^2-\nu^2)_t,\\
Y_0 &= 0.
\end{aligned}
$$

The process Y is an Ornstein-Uhlenbeck process with left and right jumps. The explicit solution for the above non-homogeneous linear stochastic integral equation is given by the following formula (see [10, Theorem 3.1])

$$
Y_t=\mathcal{E}_t(H)\left[\mathcal{E}(H)^{-1}\tilde{G}_t\right]
$$

where $\mathcal{E}$ is an optional stochastic exponent and

$$H_t = k\sigma(1-\alpha)t + (2^{1-\alpha}-1)U*(\mu^1-\nu^1)_t + (2^{1-\alpha}-1)U*(\mu^2-\nu^2)_t,$$

$$\tilde{G}_t = s^{1-\alpha}\left(k - c(1-2^{\alpha-1})\gamma^1 U - c(1-2^{\alpha-1})\gamma^2 U\right)t$$
$$+c(s2)^{1-\alpha}U*\left[(\mu^1-\nu^1)_t + (\mu^2-\nu^2)_t\right] + W_t.$$

Applying the comparison theorem to X_t and Y_t yields that, $Y_t \geq X_t = F(S_t)$ a.s. Since $F(x)$ is monotonically increasing function we have

$$S_t \leq F^{-1}(Y_t) \qquad a.s. \tag{3.15}$$

Now let us consider a function f with an option payoff $f(S_T)$, where f is increasing. Assuming zero interest rates, the price of such option is given by $\tilde{\mathbf{E}}f(S_T)$ for an appropriate martingale measure $\tilde{\mathbf{P}}$ (see [10, Section 4], where the existence of $\tilde{P}$ is discussed). Using inequality (3.15) we have that $\tilde{\mathbf{E}}f(S_T) \leq \tilde{\mathbf{E}}f(F^{-1}(Y_T))$ and thus we obtain an estimate for the option price for which $\tilde{\mathbf{E}}f(F^{-1}(Y_T))$ is easier to compute.

3.6 Pathwise Uniqueness via local times

In the last section of this chapter, we demonstrate how the local time technique used in the proof of the Theorem 3.3.1 can similarly be utilized to prove the pathwise uniqueness of (3.1). We say that solution of (3.1) is pathwise unique if whenever X and Z are any two solutions of (3.1) defined on the same stochastic basis $(\Omega, \mathcal{F}, \mathbf{F}, P)$ with the same $m \in \mathcal{M}_{loc}$ and the same measures μ^j, p^j, η such that $X_0 = Z_0$ a.s., then $X_t = Z_t$ for all t a.s.

To prove the pathwise uniqueness of solutions of (3.1) we require several assumptions on its coefficient functions. To begin with, we present a slightly more general condition than one-sided Lipschitz condition (cf. [59, Chapter 1.1]).

Definition 3.6.1 (One Sided Lipschitz Condition) *We say that a coefficient function f of equation* (3.1) *satisfies one-sided Lipschitz condition with respect to x if there exists predictable function $G(\omega, s)$ such that for any $x, y \in \mathbb{R}, s \in \mathbb{R}_+, \omega \in \Omega$*

$$(x-y)(f(\omega,s,x)-f(\omega,s,y)) \le G(\omega,s)(x-y)^2.$$

Let us introduce the following conditions:

C Conditions.

We say that the functions $f, g, h_j, k_j, l_j, r, w$ satisfy the **C** conditions if:

(C1) f is one-sided Lipschitz continuous,

(C2) g satisfies **LT** condition,

(C3) there exists non-negative functions

$$H^d(\omega,s,u) \in \mathcal{P} \times \mathcal{B}(E \cap (|u| \le 1)),$$

$$H^g(\omega,s,u) \in \mathcal{O} \times \mathcal{B}(E \cap (|u| \le 1)),$$

$$L^d(\omega,s,u), K^d(\omega,s,u) \in \tilde{\mathcal{P}},$$

$$L^g(\omega,s,u), K^g(\omega,s,u), R(\omega,s,u), W(\omega,s,u) \in \tilde{\mathcal{O}}$$

such that for any $x, y \in \mathbb{R},\ s \in \mathbb{R}_+, u \in E \cap (|u| \le 1), \omega \in \Omega$:

$$|h_j(\omega,s,u,x) - h_j(\omega,s,u,y)| \le H^j(\omega,s,u)|x-y|, \text{ and any } u \in E$$

$$|l_j(\omega,s,u,x) - l_j(\omega,s,u,y)| \le L^j(\omega,s,u)|x-y|,$$

$$|k_j(\omega,s,u,x) - k_j(\omega,s,u,y)| \le K^j(\omega,s,u)|x-y|,$$

$$|r(\omega,s,u,x) - r(\omega,s,u,y)| \le R(\omega,s,u)|x-y|,$$

$$|w(\omega,s,u,x) - w(\omega,s,u,y)| \le W(\omega,s,u)|x-y|,$$

$$G \cdot a_t + [L^d + K^d] * p^d_t + [L^g + K^g] * p^g_t$$

$$+2H^dU * \nu^d_t + 2H^gU * \nu^g_t + [R+W] * \zeta_t < \infty,$$

and the process

$$\begin{aligned}&G\cdot a_t+[L^d+K^d]*p^d_t+[L^g+K^g]*p^g_t\\&+2H^dU*\nu^d_t+2H^gU*\nu^g_t+[R+W]*\zeta_t\end{aligned}$$

is increasing.

Theorem 3.6.1 *Suppose that functions f, g, h_j, l_j, k_j, w and r satisfy $\boldsymbol{C}$ conditions, then if the solution of equation* (3.1) *exists then it is pathwise unique.*

Proof. Assume that there are two solutions X and Z of equation (3.1), and let $Y := X - Z$. Applying the formula (3.5) to Y and using identity $|Y| = 2Y^+ - Y$, we get

$$\begin{aligned}|Y_t| =&\mathrm{sign}(Y_-)\,[(f(X)-f(Z))\cdot a_t+(g(X)-g(Z))\cdot m_t]+L^0_t(Y)\\&+[|Y_-+h_d(X)-h_d(Z)|-|Y_-|]\,U*(\mu^d-\nu^d)_t\\&+[|Y+h_g(X)-h_g(Z)|-|Y|]\,U*(\mu^g-\nu^g)_t\\&+[|Y_-+h_d(X)-h_d(Z)|-|Y_-|-(h_d(X)-h_d(Z))\mathrm{sign}(Y_-)]\,U*\nu^d_t\\&+[|Y+h_g(X)-h_g(Z)|-|Y|-(h_g(X)-h_g(Z))\mathrm{sign}(Y)]\,U*\nu^g_t\\&+[|Y_-+l_d(X)-l_d(Z)+k_d(X)-k_d(Z)|-|Y_-|]*p^d_t\\&+[|Y+l_g(X)-l_g(Z)+k_g(X)-k_g(Z)|-|Y|]*p^g_t\\&+[|Y+r(X)-r(Z)+w(X)-w(Z)|-|Y|]*\eta_t\end{aligned}$$

Let

$$\begin{aligned}M_t :=&\mathrm{sign}(Y_-)(g(X)-g(Z))\cdot m_t\\&+[|Y_-+h_d(X)-h_d(Z)|-|Y_-|]\,U*(\mu^d-\nu^d)_t\\&+[|Y+h_g(X)-h_g(Z)|-|Y|]\,U*(\mu^g-\nu^g)_t\in\mathcal{M}_{loc}.\end{aligned}$$

By using one-sided Lipschitz condition on the drift coefficient function f, **LT** condition and simple algebraic inequalities, we have

$$\begin{aligned}|Y| \leq & |Y_-|G \cdot a_t + M_t \\ & + \sum_{j=r,g} [|l_j(X) - l_j(Z)| + |k_j(X) - k_j(Z)|] * p_t^j \\ & + 2|h_j(X) - h_j(Z)|U * \nu_t^j \\ & + [|r(X) - r(Z)| + |w(X) - w(Z)|] * \eta_t\end{aligned} \tag{3.16}$$

Further, we apply (C3) to (3.16) and get

$$\begin{aligned}|Y_t| \leq & M_t + |Y_-|G \cdot a_t + |Y_-|[L^d + K^d] * p_t^d + |Y|[L^g + K^g] * p_t^g \\ & + 2|Y_-|H^d U * \nu_t^d + 2|Y|H^g U * \nu_t^g + |Y|[R + W] * \eta_t\end{aligned} \tag{3.17}$$

If we now define a process $C_t := G \cdot a_t + [L^d + K^d] * p_t^d + [L^g + K^g] * p_t^g + 2H^d U * \nu_t^d + 2H^g U * \nu_t^g + [R + W] * \zeta_t$, then equation (3.17) can be rewritten as

$$|Y_t| \leq M_t + |Y| \circ C_t.$$

Finally, by Gronwall lemma (see [10, Lemma 3.2]) we get that $|Y_t| \leq \mathcal{E}_t(C)M_t$. Since C_t is increasing process, $\mathcal{E}_t(C) \geq 0$. Thus, $M_t \geq 0$ because $|Y| \geq 0$. Therefore, M_t is a non-negative optional local martingale and, consequently, by Lemma 2.4.2 it is a non-negative supermartingale starting from 0. It follows that $M = 0$. Hence, $Y_t = 0$ for all t a.s., and the pathwise uniqueness follows. ■

Remark 3.6.1 *Existence and uniqueness theorems for differential equations under one-sided Lipschitz condition on the drift coefficient function were explored by several authors (see, for example, [45], [13]). One-sided Lipschitz continuity is weaker than Lipschitz continuity, and an example illustrating this relation is a function $f(x) = e^{-x}$.*

Chapter 4

Krylov's Etimates for optional semimartingales

The estimates of N. V. Krylov have a great importance in the theory of controlled diffusion processes and stochastic differential equations (see [55],[57]). Anulova and Pragarauskas (see [14]) generalized this result to the Ito processes with Poisson random measures. Melnikov in [69] proved Krylov type estimates for continuous semimartingales on probability spaces under usual conditions. In this paper, we do not assume our probability space satisfy such technical conditions; and, our goal is to generalize Krylov's estimates to the class of optional semimartingales - *laglad* processes defined on a complete probability space such that the underlying filtration is not necessarily left nor right continuous nor complete.

Using these estimates Krylov obtained a generalization of Ito's formula for functions which have generalized derivatives up to and including the second order (see [55], [57]). In the same way, by using our obtained estimates we extend the change of variables formula for optional semimartingales (see [33], [37]) for class C^2 to the Sobolev class of functions W_d^2. Furthermore, we show how Krylov's estimates can be applied to the mean square convergence of optional solutions of SDE's under quite general assumptions on their coefficients.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a given complete space and $\mathbf{F} = (\mathcal{F}_t)_{t\in[0,\infty)}$ be a corresponding filtration on it. The family $\mathbf{F}$ is not assumed right- or left-continuous, and it is not assumed to be complete.

$\mathcal{A}_{loc}$ ($\mathcal{A}^c_{loc}$) is the set of all (continuous) processes $A = (A_t, \mathcal{F}_t)_{t\geq 0}$ having locally integrable variation, with $A_0 = 0$.

$\mathcal{M}^c_{loc}$ ($\mathcal{M}^{2,c}_{loc}$) is the set of all continuous optional local (square integrable) martingales $M = (M_t, \mathcal{F}_t)_{t\geq 0}$, $M_0 = 0$.

For functions $f : \mathbb{R}^d \to \mathbb{R}$ we set

$$f_{x_i} = \frac{\partial f}{\partial x_i}, \quad f_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad f_x = (f_{x_1}, ..., f_{x_d}).$$

For vectors $x = (x_i, ..., x_d), y = (y_1, ..., y_d) \in \mathbb{R}^d$: $|x| = \sum_{i=1}^d |x_i|$, $(x, y) = x_1 y_1 + ... + x_d y_d$. For a square matrix A: trA is the trace of A and $\det A$ is the determinant of A.

$L_d(U)$ is the space of measurable functions f, defined in the region $U \subset \mathbb{R}^d$, $d \geq 1$ such that

$$\|f\|_{d,U} = \left(\int_U |f(x)|^d \, dx \right)^{1/d} < \infty.$$

$B(\Gamma)$ denotes the set of bounded Borel functions on Γ with the norm

$$\|f\|_{B(\Gamma)} = \sup_{x\in\Gamma} |f(x)|$$

Let D be a bounded region in $\mathbb{R}^d$, and let $u(x)$ be a function in $\bar{D}$. We write $u \in W^2(D)$ ($u \in \bar{W}^2(D)$) if there exists a sequence of functions $u^n \in C^2(\bar{D})$ such that

$$\|u - u^n\|_{B(\bar{D})} \to 0, \quad \|u^n - u^m\|_{W^2(D)} \to 0 \quad \left(\|u^n - u^m\|_{\bar{W}^2(D)} \to 0 \right)$$

as $n, m \to \infty$, where

$$\|f\|_{W^2(D)} = \sum_{i,j=1}^{d} \|f_{x_i x_j}\|_{d,D} + \sum_{i=1}^{d} \|f_{x_i}\|_{d,D} + \|f\|_{B(\bar{D})}$$

$$\left(\|f\|_{\bar{W}^2(D)} = \|f\|_{W^2(D)} + \sum_{i=1}^{d} \|f_{x_i}\|_{2d,D} \right).$$

Definition 4.0.1 *Let $D \subset \mathbb{R}^d$, let v and h be Borel functions locally summable in D. The function h is said to be a generalized derivative (in the region D) of the function v of order n in the direction of coordinate vectors $r_1, ..., r_n$ and this function h is denoted by $v_{x_{r_1} \dots x_{r_n}}$ if for each $\phi \in C_0^\infty(D)$*

$$\int_D \phi(x) h(x) dx = (-1)^n \int_D v(x) \phi_{x_{r_1} \dots x_{r_n}} dx$$

The properties of generalized derivatives are well known (see [76]). We will apply some of them without proofs. Note first that a generalized derivative can be defined uniquely almost everywhere. The function $u \in W^2(D)$ has generalized derivatives up to and including the second order, and these derivatives belong to $L_d(D)$.

4.1 Krylov's estimates

We will consider the following form of a d-dimensional optional semimartingale X, for $i = 1, ..., d$,

$$\begin{aligned} X_t^i = X_0^i + a_t^i + m_t^i + \int_{0+}^{t} \int_{0<z\leq 1} z(\mu^r - \nu^r)(ds, dz) + \int_{0+}^{t} \int_{z>1} z\mu^r(ds, dz) \\ + \int_{0}^{t-} \int_{0<z\leq 1} z(\mu^g - \nu^g)(ds, dz) + \int_{0}^{t-} \int_{z>1} z\mu^g(ds, dz), \quad (4.1) \end{aligned}$$

where X_0^i is $\mathcal{F}_0$-measurable random variable, $a^i \in \mathcal{A}_{loc}^c$ and $m^i \in \mathcal{M}_{loc}^{2,c}$. The jump measures μ^r and μ^g are defined on $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \times \mathcal{E})$ as follows

$$\mu^r(\Gamma) = \sum_{n \geqslant 1} \mathbf{1}_\Gamma(T_n, \Delta X_{T_n}), \qquad \mu^g(\Gamma) = \sum_{n \geqslant 1} \mathbf{1}_\Gamma(U_n, \Delta^+ X_{U_n}),$$

where $(T_n)_{n\geqslant 1}$, $(U_n)_{n\geqslant 1}$ are sequences of totally inaccessible stopping times and totally inaccessible wide sense stopping times, respectively; $\mathbf{1}_\Gamma(\cdot)$ is an indicator function of a set $\Gamma \in \mathcal{B}(\mathbb{R}_+) \times \mathcal{E}$. The processes ν^j are respective compensators of $\mu^j,\ j = r, g$.

The following facts from the theory of parabolic partial differential equations are necessary for our proof. We take an auxiliary non-negative smooth function $\varphi(x),\ \varphi(x) = 0$ for $x \geq 1$, and $\int_{\mathbb{R}^d} \varphi(x)dx = 1$. For $\epsilon > 0$ we set $\varphi^\epsilon(x) = \epsilon^{-1}\varphi(x\epsilon^{-1})$.

Lemma 4.1.1 *(see [56]) For each $\lambda > 0, \epsilon > 0$ and for every continuous function $f : \mathbb{R}^d \to [0, \infty)$ with compact support, there exists a smooth function $u^\epsilon : \mathbb{R}^d \to [0, \infty)$ ($u^\epsilon(x) = \int_{\mathbb{R}^d} u(y)\varphi^\epsilon(x - y)dy$, see [56] for detailed explanation of $u(x)$) with the properties:*

(a) for each $l \in \mathbb{R}^d$

$$\sum_{i,j=1}^{d} u^\epsilon_{x_i x_j} l_i l_j \leq \lambda u^\epsilon |l|^2;$$

(b) $|u^\epsilon_x| \leq \sqrt{\lambda} u^\epsilon$

(c) for all symmetric nonnegative definite $d \times d$ matrices A

$$\sum_{i,j=1}^{d} A_{ij} u^\epsilon_{x_i x_j} - \lambda(trA + 1)u^\epsilon \leq -(\det A)^{1/d} f^\epsilon,$$

$\left(f^\epsilon(x) = \int_{\mathbb{R}^d} f(y)\varphi^\epsilon(x - y)dy\right)$;

(d) for all $p \geq d, x \in \mathbb{R}^d$

$$|u^\epsilon(x)| \leq N(p, d, \lambda)\|f\|_{p,\mathbb{R}^d}.$$

We now present the main result of this chapter.

Theorem 4.1.1 *Let $V \in \mathcal{A}^c_{loc}$ be an increasing process, and suppose the characteristics $a^i, \langle m^i \rangle, \nu^j, j = r, g,$ of X satisfy the structural conditions:*

There exist densities ($dV \times dP$-a.s.)

$$\alpha^i = \frac{da^i}{dV}, \beta^{ik} = \frac{\langle m^i, m^k \rangle}{dV}, \beta = [\beta^{ik}], \; i, k = 1, 2, ..., d, \tag{4.2}$$

and measures $\bar{\nu}^j(\omega, t, \Gamma), \Gamma \in \mathcal{E}, j = r, g,$ such that for all $t > 0$ ($dV \times dP$-a.s.)

$$\nu^r(\omega, (0, t], \Gamma) = \int_0^t \bar{\nu}^r(\omega, s, \Gamma) dV_s, \quad \nu^g(\omega, [0, t), \Gamma) = \int_0^{t-} \bar{\nu}^g(\omega, s, \Gamma) dV_s, \tag{4.3}$$

$$|\alpha_t| + \sum_{j=r,g} \int_{0<|z|\le 1} z^2 \bar{\nu}^j(\omega, t, dz) + \sum_{j=r,g} \int_{|z|>1} \bar{\nu}^j(\omega, t, dz) \le C(\omega, t), \tag{4.4}$$

where $C(\omega, t)$ is a predictable function such that $k_\infty = \mathbf{E} \int_0^\infty e^{-\phi_t} C(\omega, t) dV_t < \infty$.

Then for any measurable function $f \ge 0, \; \lambda > 0, \; p \ge d$

$$\mathbf{E} \int_0^\infty e^{-\lambda \int_0^t [\frac{1}{2} tr \beta_s + 1] dV_s} (\det \beta_t)^{1/d} f(X_{t-}) dV_t \le N(k_\infty, \lambda, d, p) \|f\|_{p, \mathbb{R}^d}. \tag{4.5}$$

Proof. We follow the same approach as in [69]. First, consider continuous non-negative function $f = f(x)$ with compact support. Denote $\phi_t = \lambda \int_0^t [\frac{1}{2} tr \beta_s + 1] dV_s$. Applying the integration by parts formula (see Lemma 5.2.3) to $u^\epsilon(X_t) e^{-\phi_t}$, we get

$$u^\epsilon(X_t) e^{-\phi_t} - u^\epsilon(X_0) = \int_{0+}^t e^{-\phi_s} du^{\epsilon, r}(X) + \int_0^{t-} e^{-\phi_s} du^{\epsilon, g}(X) + \int_{0+}^t u^\epsilon(X_{s-}) de^{-\phi_s},$$

where $u^{\epsilon,r}(X)$ and $u^{\epsilon,g}(X)$ are right- and left-continuous part of $u(X)$, respectively.

Next, using the change of variables formula to find $u^\epsilon(X_t)$ and $e^{-\phi_t}$, we find that

$$\begin{aligned}
&u^\epsilon(X_t)e^{-\phi_t} - u^\epsilon(X_0) \\
&\quad = \int_{0+}^{t} e^{-\phi_s}\Bigg\{\frac{1}{2}\sum_{i,j=1}^{d} u^\epsilon_{x_i,x_j}(X_{s-})d\langle m^i, m^j\rangle_s + \sum_{i=1}^{d} u^\epsilon_{x_i}(X_{s-})d(a^i_s + m^i_s) \\
&\qquad + \int_{0<z\le 1} \left[u^\epsilon(X_{s-}+z) - u^\epsilon(X_{s-})\right](\mu^r - \nu^r)(ds, dz) \\
&\qquad + \int_{z>1} \left[u^\epsilon(X_{s-}+z) - u^\epsilon(X_{s-})\right]\mu^r(ds, dz) \\
&\qquad + \int_{0<z\le 1} \left[u^\epsilon(X_{s-}+z) - u^\epsilon(X_{s-}) - \sum_{i=1}^{d} u^\epsilon_{x_i}(X_{s-})z\right]\nu^r(ds, dz)\Bigg\} \\
&\qquad + \int_{0}^{t-} e^{-\phi_s}\Bigg\{\int_{0<z\le 1} \left[u^\epsilon(X_s+z) - u^\epsilon(X_s)\right](\mu^g - \nu^g)(ds, dz) \\
&\qquad + \int_{z>1} \left[u^\epsilon(X_s+z) - u^\epsilon(X_s)\right]\mu^g(ds, dz) \\
&\qquad + \int_{0<z\le 1} \left[u^\epsilon(X_s+z) - u^\epsilon(X_s) - \sum_{i=1}^{d} u^\epsilon_{x_i}(X_s)z\right]\nu^g(ds, dz)\Bigg\} \\
&\qquad - \int_{0+}^{t} e^{-\phi_s}\lambda\left(\frac{1}{2}tr\beta_s + 1\right)u^\epsilon(X_{s-})dV_s
\end{aligned} \tag{4.6}$$

Let $\{\sigma_n\}_{n\ge1}, \{\tau_n\}_{n\ge1}$ and $\{\xi_n\}_{n\ge1}$ be localizing sequences for

$$\int_{0+}^{t} e^{-\phi_s}\sum_{i=1}^{d} u^\epsilon_{x_i}(X_{s-})dm_i,$$

$$\int_{0+}^{t} e^{-\phi_s}\int_{0<z\le 1}\left[u^\epsilon(X_{s-}+z) - u^\epsilon(X_{s-})\right](\mu^r - \nu^r)(dz, ds)$$

and

$$\int_{0}^{t-} e^{-\phi_s}\int_{0<z\le 1}\left[u^\epsilon(X_s+z) - u^\epsilon(X_s)\right](\mu^g - \nu^g)(dz, ds),$$

respectively.

Define $\forall n \ge 1, R_n := t\wedge\sigma_n\wedge\tau_n\wedge\xi_n, R_n \in \mathcal{T}_+,\ R_n \uparrow \infty$ a.s. as $n \uparrow \infty$ and $t \uparrow \infty$. Taking expectation and applying structural conditions (4.2)-(4.3), we obtain from (4.6):

$$
\begin{aligned}
&\mathbf{E}u^\epsilon(X_{R_n})e^{-\phi_{R_n}} - \mathbf{E}u^\epsilon(X_0) \\
&\quad =\mathbf{E}\int_{0+}^{R_n} e^{-\phi_s}\Bigg\{\frac{1}{2}\sum_{i,j=1}^d u^\epsilon_{x_i,x_j}(X_{s-})(\beta_s)_{ij} - \lambda\left(\frac{1}{2}tr\beta_s + 1\right)u^\epsilon(X_{s-}) \\
&\qquad + \sum_{i=1}^d \alpha_s^i u^\epsilon_{x_i}(X_{s-}) \\
&\qquad + \int_{0<z\le 1}\left[u^\epsilon(X_{s-}+z) - u^\epsilon(X_{s-}) - \sum_{i=1}^d u^\epsilon_{x_i}(X_{s-})z\right]\bar{\nu}^r(dz) \\
&\qquad + \int_{0<z\le 1}\left[u^\epsilon(X_{s}+z) - u^\epsilon(X_{s}) - \sum_{i=1}^d u^\epsilon_{x_i}(X_{s})z\right]\bar{\nu}^g(dz) \\
&\qquad + \int_{z>1}\left[u^\epsilon(X_{s-}+z) - u^\epsilon(X_{s-})\right]\bar{\nu}^r(dz) \qquad (4.7)\\
&\qquad + \int_{z>1}\left[u^\epsilon(X_{s}+z) - u^\epsilon(X_{s})\right]\bar{\nu}^g(dz)\Bigg\}dV_s \\
&\quad :=\mathbf{E}\int_{0+}^{R_n} e^{-\phi_s}\left\{I_1+I_2+I_3+I_4+I_5+I_6+I_7\right\}dV_s \qquad (4.8)
\end{aligned}
$$

Using the properties (a)-(d) of the function u^ϵ in Lemma 4.1.1, we have

$$
\begin{aligned}
u^\epsilon(X_0) \le& N_0\|f\|_{p,\mathbb{R}^d};\\
I_1 + I_2 =& \frac{1}{2}\sum_{i,j=1}^d u^\epsilon_{x_i,x_j}(X_{s-})(\beta_s)_{ij} - \lambda\left(\frac{1}{2}tr\beta_s + 1\right)u^\epsilon(X_{s-})\\
\le& -\left(det\frac{1}{2}\beta_s\right)^{1/d} f^\epsilon(X_{s-});\\
I_3 =& \sum_{i=1}^d \alpha_s^i u^\epsilon_{x_i}(X_{s-})\\
\le& |\alpha_s|\|f\|_{p,\mathbb{R}^d}.
\end{aligned}
$$

Next, with the help of Taylor's decomposition for multivariate functions and

Lemma 4.1.1-(d), we obtain

$$
\begin{aligned}
I_5 =& \int_{0<z\leq 1} \left[u^\epsilon(X_s+z) - u^\epsilon(X_s) - \sum_{i=1}^{d} u^\epsilon_{x_i}(X_s) z \right] \bar{\nu}^g(dz) \\
=& \int_{0<z\leq 1} \int_0^1 (1-\theta) \sum_{i,j=1}^{d} u^\epsilon_{x_i,x_j}(X_s+\theta z) z^2 d\theta \bar{\nu}^g(dz) \\
\leq& \frac{\lambda}{2} \int_{0<z\leq 1} z^2 \bar{\nu}^g(dz) \sup_{x\in\mathbb{R}^d} u^\epsilon(x) \\
\leq& \frac{\lambda N_2}{2} \|f\|_{p,\mathbb{R}^d} \int_{0<z\leq 1} z^2 \bar{\nu}^g(dz),
\end{aligned}
$$

where $\theta \in [0,1]$ is an auxiliary parameter.

We can also find similar inequality for the integral I_4:

$$
I_4 \leq \frac{\lambda N_3}{2} \|f\|_{p,\mathbb{R}^d} \int_{0<z\leq 1} z^2 \bar{\nu}^r(dz).
$$

Using property (d) of Lemma 4.1.1, we get

$$
\begin{aligned}
I_7 =& \int_{z>1} \left[u^\epsilon(X_s+z) - u^\epsilon(X_s) \right] \bar{\nu}^g(dz) \\
\leq& \int_{z>1} \bar{\nu}^g(dz) \sup_{x\in\mathbb{R}^d} u^\epsilon(x) \\
\leq& N_4 \|f\|_{p,\mathbb{R}^d} \int_{z>1} \bar{\nu}^g(dz).
\end{aligned}
$$

Similarly,

$$
I_6 \leq N_5 \|f\|_{p,\mathbb{R}^d} \int_{z>1} \bar{\nu}^r(dz).
$$

It follows from obtained inequalities and the relation (4.8) that

$$
\begin{aligned}
\mathbf{E}\int_0^{R_n} e^{-\phi_s}(\det\beta_s)^{1/d} f^\epsilon(X_{s-})dV_s \leq & N(\lambda,d,p)\|f\|_{p,\mathbb{R}^d} \\
& \times \mathbf{E}\int_0^{R_n} e^{-\phi_s}\Bigg[|\alpha_s| \\
& + \sum_{j=r,g}\int_{0<|z|\leq 1} z^2\bar{\nu}^j(\omega,s,dz) \\
& + \sum_{j=r,g}\int_{|z|>1}\bar{\nu}^j(\omega,s,dz)\Bigg]dV_s.
\end{aligned} \tag{4.9}
$$

After applying condition (4.4) to (4.9), it becomes

$$
\begin{aligned}
\mathbf{E}\int_0^{R_n} e^{-\phi_s}(\det\beta_s)^{1/d} f^\epsilon(X_{s-})dV_s &\leq N(\lambda,d,p)\|f\|_{p,\mathbb{R}^d}\mathbf{E}\int_0^{R_n} e^{-\phi_s}C(\omega,s)dV_s \\
&\leq N(k_\infty,\lambda,d,p)\|f\|_{p,\mathbb{R}^d}.
\end{aligned}
$$

Finally, we let $n\uparrow\infty$ then $t\uparrow\infty$ and $\epsilon\downarrow 0$ and reach estimate (4.5). Extension of the estimate to the Borel measurable function f is standard (see, for example, [14]).

■

Corollary 4.1.1 *If, in addition to the structural conditions* (4.2)-(4.3) *of Theorem 4.1.1, there exist constants* $0 < c_1 \leq c_2 < \infty$ *such that for all* $x \in \mathbb{R}^d$, $c_1|x|^2 \leq (\beta_t x, x) \leq c_2|x|^2$ *(*$dV \times dP$*-a.s.) and*

$$
\begin{aligned}
|\alpha_t| &\leq \frac{K}{2} tr\beta_t, \\
\int_{0<|z|\leq 1} z^2\bar{\nu}^j(\omega,t,dz) &\leq \frac{K}{2} tr\beta_t, \\
\int_{|z|>1}\bar{\nu}^j(\omega,t,dz) &\leq \frac{K}{2} tr\beta_t.
\end{aligned} \tag{4.10}
$$

Then for any measurable function $f\geq 0$, $\lambda>0$, $p\geq d$

$$
\mathbf{E}\int_0^\infty e^{-\lambda\int_0^t[\frac{1}{2}tr\beta_s+1]dV_s} f(X_{t-})dV_t \leq N(K,\lambda,d,p,c_1,c_2)\|f\|_{p,\mathbb{R}^d}.
$$

4.2 Application: Change of Variables formula with Generalized Derivatives

Change of variables formula is an essential tool of Stochastic Calculus. In this section, we prove that in some cases the change of variables formula remains valid for functions whose generalized derivatives are ordinary functions.

Theorem 4.2.1 *Let X_0 be fixed, $X_0 \in \mathbb{R}^d$. Let τ_D be the first exit time of the process X_t in* (4.1) *from a bounded region $D \subset \mathbb{R}^d$, and let $\tau \in \mathcal{T}, \tau < \tau_D$. Suppose that X satisfies assumptions of Corollary 4.1.1.*

Then for any $v \in \bar{W}^2(D)$

$$
\begin{aligned}
&v(X_\tau)e^{-\phi_\tau} - v(X_0) \\
&\quad = \int_{0+}^{\tau} e^{-\phi_s} \Bigg\{ \frac{1}{2} \sum_{i,j=1}^{d} v_{x_i,x_j}(X_{s-}) d\langle m^i, m^j \rangle_s - \lambda \left(\frac{1}{2} tr\beta_s + 1 \right) v(X_{s-}) dV_s \\
&\qquad + \sum_{i=1}^{d} v_{x_i}(X_{s-}) da_s^i + \sum_{i=1}^{d} v_{x_i} dm_s^i \\
&\qquad + \int_{0<z\le 1} [v(X_{s-}+z) - v(X_{s-})](\mu^r - \nu^r)(dz, ds) \\
&\qquad + \int_{0<z\le 1} \left[v(X_{s-}+z) - v(X_{s-}) - \sum_{i=1}^{d} v_{x_i}(X_{s-}) z \right] \nu^r(dz, ds) \\
&\qquad + \int_{z>1} [v(X_{s-}+z) - v(X_{s-})]\, \mu^r(dz, ds) \Bigg\} \\
&\qquad + \int_{0}^{\tau-} e^{-\phi_s} \Bigg\{ \int_{0<z\le 1} [v(X_s+z) - v(X_s)](\mu^g - \nu^g)(dz, ds) \\
&\qquad + \int_{0<z\le 1} \left[v(X_s+z) - v(X_s) - \sum_{i=1}^{d} v_{x_i}(X_s) z \right] \nu^g(dz, ds) \\
&\qquad + \int_{z>1} [v(X_s+z) - v(X_s)]\, \mu^g(dz, ds) \Bigg\} \quad (a.s. \ on \ \{0 \le \tau\}).
\end{aligned}
\tag{4.11}
$$

Proof. Let a sequence $v^n \in C^2(\bar{D})$ be such that

$$\|v - v^n\|_{B(D)} \to 0, \ \|v - v^n\|_{W^2(D)} \to 0,$$
$$\||(v_x - v_x^n)|^2\|_{d,D} \to 0.$$

For convenience rewrite (4.11) as following

$$\begin{aligned}
&v(X_\tau)e^{-\phi_\tau} - v(X_0) \\
&\quad = \int_{0+}^{\tau} e^{-\phi_s} \left\{ I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 \right\} \\
&\quad\quad + \int_0^{\tau-} e^{-\phi_s} \left\{ I_8 + I_9 + I_{10} \right\}.
\end{aligned}$$

We prove that the right side of (4.11) makes sense. For $s < \tau$

$$\begin{aligned}
I_1 + I_2 + I_3 = \Big[&\frac{1}{2} \sum_{i,j=1}^{d} v_{x_i,x_j}(X_{s-})\beta^{ij} - \lambda \left(\frac{1}{2} tr\beta_s + 1 \right) v(X_{s-}) \\
&+ \sum_{i=1}^{d} v_{x_i}(X_{s-})\alpha^i \Big] dV_s
\end{aligned}$$

From this, using Theorem 4.1.1 we obtain

$$\begin{aligned}
\mathbf{E} \int_{0+}^{\tau} e^{-\phi_s} \Bigg| &\frac{1}{2} \sum_{i,j=1}^{d} v_{x_i,x_j}(X_{s-})\beta^{ij} - \lambda \left(\frac{1}{2} tr\beta_s + 1 \right) v(X_{s-}) \\
&+ \sum_{i=1}^{d} v_{x_i}(X_{s-})\alpha^i \Bigg| dV_s \\
\leq & N\mathbf{E} \int_{0+}^{\tau} e^{-\phi_s} \left[\sum_{i,j=1}^{d} |v_{x_i,x_j}(X_{s-})| + |v(X_{s-})| + \sum_{i=1}^{d} |v_{x_i}(X_{s-})| \right] dV_s \\
\leq & N\|v\|_{W^2(D)},
\end{aligned}$$

where N depends on $\lambda, p, d, K, c_1, c_2$.

Similarly,

$$\begin{aligned}
\mathbf{E} \left| \int_{0+}^{\tau} e^{-\phi_s} \sum_{i}^{d} v_{x_i}(X_{s-}) dm_s^i \right|^2 \leq & N\mathbf{E} \int_{0+}^{\tau} e^{-2\phi_s} \left| v_x(X_{s-}) \right|^2 dV_s \\
\leq & N\||v_x|^2\|_{d,D}.
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{E}\left|\int_{0+}^{\tau} e^{-\phi_s} \int_{0<z\leq 1} [v(X_{s-}+z) - v(X_{s-})]\mu^r(dz,ds)\right| \\
&\leq \mathbf{E}\int_{0+}^{\tau} e^{-\phi_s} \int_{0<z\leq 1} |v(X_{s-}+z) - v(X_{s-})|\bar{\nu}^r(dz)dV_s \\
&\leq \mathbf{E}\int_{0+}^{\tau} e^{-\phi_s} \int_{0<z\leq 1} \left[|v(X_{s-}+z)| + |v(X_{s-})|\right]\bar{\nu}^r(dz)dV_s \\
&\leq N\|v\|_{B(D)}.
\end{aligned}$$

Using the same technique, integrals I_7, I_8 and I_{10} are well-defined.

Since

$$\begin{aligned}
&\mathbf{E}\int_{0+}^{\tau} e^{-\phi_s} \int_{0<z\leq 1} \left|\sum_{i=1}^{d} v_{x_i}(X_{s-})z\right| \nu^r(dz,ds) \\
&\leq N\mathbf{E}\int_{0+}^{\tau} e^{-\phi_s} \int_{0<z\leq 1} |v_x(X_{s-})|z\bar{\nu}^r(dz)dV_s \\
&\leq N\|v_x\|_{d,D},
\end{aligned}$$

I_6 is well-defined. The same holds for I_9.

Further, we apply the change of variables formula for optional semimartingales (see Theorem 2.9.2) to the expression $v^n(x_t)e^{-\phi_t}$. Then, we have almost surely

$$
\begin{aligned}
& v^n(X_\tau)e^{-\phi_\tau} - v^n(X_0) \\
&\quad = \int_{0+}^{\tau} e^{-\phi_s}\Bigg\{ \frac{1}{2}\sum_{i,j=1}^{d} v^n_{x_i,x_j}(X_{s-})d\langle m^i, m^j\rangle_s - \lambda\left(\frac{1}{2}tr\beta_s + 1\right) v^n(X_{s-})dV_s \\
&\qquad + \sum_{i=1}^{d} v^n_{x_i}(X_{s-})da^i_s + \sum_{i=1}^{d} v^n_{x_i} dm^i_s \\
&\qquad + \int_{0<z\leq 1} [v^n(X_{s-}+z) - v^n(X_{s-})](\mu^r - \nu^r)(dz, ds) \\
&\qquad + \int_{0<z\leq 1} \left[v^n(X_{s-}+z) - v^n(X_{s-}) - \sum_{i=1}^{d} v^n_{x_i}(X_{s-})z \right] \nu^r(dz, ds) \\
&\qquad + \int_{z>1} [v^n(X_{s-}+z) - v^n(X_{s-})]\, \mu^r(dz, ds) \Bigg\} \\
&\qquad + \int_{0}^{\tau-} e^{-\phi_s}\Bigg\{ \int_{0<z\leq 1} [v^n(X_s+z) - v^n(X_s)](\mu^g - \nu^g)(dz, ds) \\
&\qquad + \int_{0<z\leq 1} \left[v^n(X_s+z) - v^n(X_s) - \sum_{i=1}^{d} v^n_{x_i}(X_s)z \right] \nu^g(dz, ds) \\
&\qquad + \int_{z>1} [v^n(X_s+z) - v^n(X_s)]\, \mu^g(dz, ds) \Bigg\}
\end{aligned} \tag{4.12}
$$

We pass to the limit in equality (4.12) as $n \to \infty$. By the Sobolev Theorem (see [83]) $v^n \to v$ uniformly in each finite region. From estimates similar to the estimates we found earlier it easily follows that the right side of (4.12) tends to the right side of (4.11). ■

4.3 Application: Convergence of optional solutions of SDE

In this section we consider a sequence of solutions $(X_t^n)_{t\in[0,T]},\ n = 0, 1, 2, ...$ satisfying the following d-dimensional SDE's, respectively,

$$\begin{aligned}X_t^n =& X_0 + \int_0^t b^n(X_s^n)ds + \int_0^t \sigma^n(X_s^n)dW_s \\ &+ \int_{0+}^t \int_E c^n(X_{s-}^n, z)(\mu^r - \nu^r)(ds, dz) + \int_0^{t-} \int_E h^n(X_s^n, z)(\mu^g - \nu^g)(ds, dz),\end{aligned}$$

$n = 0, 1, 2, \ldots,$ where $E = \mathbb{R} \setminus \{0\}$, W_t is a d-dimensional Wiener process, μ^r and μ^g are, respectively, right-continuous and left-continuous modifications of 1-dimensional Poisson measures with corresponding compensators ν^r and ν^g, and $b^n,\ c^n, h^n \in \mathbb{R}^d$, and $\sigma^n \in \mathbb{R}^{d\times d}$. Hereafter, we write $X_t = X_t^0, b = b^0$ and so on.

Theorem 4.3.1 *Assume that*

(a) $|b^n(x)|^2+|\sigma^n(x)|^2+\int_E |c^n(x, z)|^2\nu^r(dz)+\int_E |h^n(x, z)|^2\nu^g(dz) \leq k$ *and* $\mathbf{E}|X_0|^2 < k_0$*, where* k_0 *and* $k \geq 0$ *are constants;*

$$(x - y) \cdot (b(x) - b(y)) \leq F(s)\rho(|x - y|^2),$$

$$\begin{aligned}\sum_{i,j=1}^d |\sigma_{ij}(x) - \sigma_{ij}(y)|^2 + \int_E |c(x, z) - c(y, z)|^2\nu^r(dz)& \\ + \int_E |h(x, z) - h(y, z)|^2\nu^g(dz) &\leq F(s)\rho(|x - y|^2),\end{aligned}$$

where $0 \leq F(s)$ *satisfies that* $\forall t \geq 0\ \int_0^t F(s)ds < \infty$*, and* $\rho(u)$ *is strictly increasing, continuous, and concave such that* $\rho(0) = 0,\ \rho(u) > 0,$ *as* $u > 0$*; and* $\int_{0+} du/\rho(u) = \infty$*;*

(b)

$$\left\||b^n(x)-b(x)|^2\right\|_{p,\mathbb{R}^d}+\left\||\sigma^n(x)-\sigma(x)|^2\right\|_{p,\mathbb{R}^d}$$
$$+\left\|\int_E |c^n(x,z)-c(x,z)|^2\nu^r(dz)\right\|_{p,\mathbb{R}^d}$$
$$+\left\|\int_E |h^n(x,z)-h(x,z)|^2\nu^g(dz)\right\|_{p,\mathbb{R}^d}\to 0,$$

as $n\to\infty$, where $p\geq d+1$;

(c)there exists $k_1>0$ and $k_2>0$ such that for all $x\subset\mathbb{R}^d$,

$$k_1|x|^2\leq(\beta x,x)\leq k_2|x|^2,$$

where $\beta=\sigma^n\sigma^{n*},\ n=1,2,\ldots;$

(d) $lim_{n\to\infty}\mathbf{E}|X_0^n-X_0|^2=0$.

Then we have $\forall t\geq 0$

$$\lim_{n\to\infty}\mathbf{E}|X_t^n-X_t|^2=0.$$

We will need the following two lemmas to prove the main theorem.

Lemma 4.3.1 *(see [81], Lemma 116) If for all $t\geq 0$ a real non-random funcion y_t satisfies*

$$0\leq y_t\leq\int_0^t\rho(y_s)ds<\infty,$$

where $\rho(u)$ defined on $u\geq 0$, is non-negative, increasing such that $\rho(0)=0,\ \rho(u)>0$, as $u>0$; and $\int_{0+}du/\rho(u)=\infty$, then

$$y_t=0,\forall t\geq 0.$$

Lemma 4.3.2 *Suppose $\mathbf{E}|X_0|^2<k_0$ and*

$$|b(x)|^2+|\sigma(x)|^2+\int_E|c(x,z)|^2\nu^r(dz)+\int_E|h(x,z)|^2\nu^g(dz)\leq k,$$

where $k_0,k\geq 0$ are constants. Then $\mathbf{E}\sup_{t\in[0,T]}|X_t|\leq k_T$ for some constant k_T.

Proof. First, note that

$$
\begin{aligned}
|X_t|^2 \leq & 4\Bigg[|X_0|^2 + \left|\int_0^t b(X_s)ds\right|^2 + \left|\int_0^t \sigma(X_s)dW_s\right|^2 \\
& + \left|\int_{0+}^t c(X_{s-},z)(\mu^r - \nu^r)(ds,dz)\right|^2 + \left|\int_0^{t-} c(X_s,z)(\mu^g - \nu^g)(ds,dz)\right|^2\Bigg].
\end{aligned}
$$

Next, using Doob's inequality and optional stochastic integral properties we obtain

$$
\begin{aligned}
\mathbf{E}\sup_{t\in[0,T]}\left|\int_0^t b(X_s)ds\right|^2 \leq & \mathbf{E}\sup_{t\in[0,T]}\int_0^t |b(X_s)|^2\, ds \leq kT, \\
\mathbf{E}\sup_{t\in[0,T]}\left|\int_0^t \sigma(X_s)dW_s\right|^2 \leq & 4\mathbf{E}\int_0^t |\sigma(X_s)|^2\, ds \leq 4kT, \\
\mathbf{E}\sup_{t\in[0,T]}\left|\int_{0+}^t \int_E c(X_{s-},z)(\mu^r - \nu^r)(ds,dz)\right|^2 \leq & 8\mathbf{E}\int_{0+}^t \int_E |c(X_{s-},z)|^2\, \nu^r(ds,dz) \\
\leq & 8kT, \\
\mathbf{E}\sup_{t\in[0,T]}\left|\int_0^{t-} \int_E h(X_s,z)(\mu^g - \nu^g)(ds,dz)\right|^2 \leq & 8\mathbf{E}\int_0^{t-} \int_E |h(X_s,z)|^2\, \nu^g(ds,dz) \\
\leq & 8kT.
\end{aligned}
$$

Thus, we conclude that

$$
\mathbf{E}\sup_{t\in[0,T]} |X_t|^2 \leq k_T,
$$

where $k_T = 4(21kT + k_0)$. ■

Proof of Theorem. By the change of variables formula

$$
\begin{aligned}
\mathbf{E}|X_t^n - X_t|^2 - \mathbf{E}|X_0^n - X_0|^2 =& 2\mathbf{E}\int_0^t (X_t^n - X_t)(b^n(X_s^n) - b(X_s))ds \\
&+ \mathbf{E}\int_0^t \sum_{i,j=1}^d |\sigma_{ij}(X_s^n) - \sigma_{ij}(X_s)|^2 ds \\
&+ \mathbf{E}\int_{0+}^t \int_E |c^n(X_{s-}^n, z) - c(X_{s-}, z)|^2 \nu^r(ds, dz) \\
&+ \mathbf{E}\int_0^{t-} \int_E |h^n(X_s^n, z) - h(X_s, z)|^2 \nu^g(ds, dz) \\
=& \sum_{i=1}^4 I^n(i).
\end{aligned}
$$

For the process X_t^n, we have

$$
\begin{aligned}
V_t =& t, \\
\alpha_t =& b^n(X_t^n), \\
\beta_t =& \sigma^n \sigma^{n*}(X_t^n), \\
\bar{\nu}^r(dz) =& |c(X_t^n, z)|^2 \nu^r(dz), \\
\bar{\nu}^g(dz) =& |h(X_t^n, z)|^2 \nu^g(dz).
\end{aligned}
$$

Furthermore,

$$
k_\infty = \int_0^\infty e^{-\nu \int_0^t [1/2 tr \sigma^n \sigma^{n*}(X_s^n) + 1] ds} k dt \leq k \int_0^\infty e^{-\nu t} dt = \frac{k}{\nu} < \infty.
$$

Therefore, condition (4.4) is satisfied. Thus, by the Krylov's estimate (Corollary 4.1.1) and the assumption (b) we get

$$
\begin{aligned}
I^n(1) \leq& 2\mathbf{E}\int_0^t (X_t^n - X_t)(b^n(X_s^n) - b(X_s^n))ds \\
&+ 2\mathbf{E}\int_0^t (X_t^n - X_t)(b(X_s^n) - b(X_s))ds \\
\leq& \mathbf{E}\int_0^t |X_t^n - X_t|^2 ds + N \left\| |b^n(x) - b(x)|^2 \right\|_{p,\mathbb{R}^d} \\
&+ 2\int_0^t F(s)\rho\left(\mathbf{E}|X_s^n - X_s|^2\right) ds,
\end{aligned}
$$

where N depends on $\lambda, p, d, K, T, k_1, k_2$.

Similarly,

$$
\begin{aligned}
I^n(2) \leq & 2N \left\| |\sigma^n(x) - \sigma(x)|^2 \right\|_{p,\mathbb{R}^d} + 2 \int_0^t F(s)\rho\left(\mathbf{E}|X_s^n - X_s|^2\right) ds, \\
I^n(3) \leq & 2N \left\| \int_E |c^n(x,z) - c(x,z)|^2 \nu^r(dz) \right\|_{p,\mathbb{R}^d} \\
& + 2 \int_{0+}^t F(s)\rho\left(\mathbf{E}|X_{s-}^n - X_{s-}|^2\right) ds, \\
I^n(4) \leq & 2N \left\| \int_E |h^n(x,z) - h(x,z)|^2 \nu^g(dz) \right\|_{p,\mathbb{R}^d} \\
& + 2 \int_0^{t-} F(s)\rho\left(\mathbf{E}|X_s^n - X_s|^2\right) ds.
\end{aligned}
$$

Consequently, applying the assumptions (c) and (e) we have

$$
\mathbf{E}|X_t^n - X_t|^2 \leq \mathbf{E} \int_0^t |X_t^n - X_t|^2 ds + 8 \int_0^{t-} F(s)\rho\left(\mathbf{E}|X_s^n - X_s|^2\right) ds.
$$

Notice that by Lemma 4.3.2 for every $n = 0, 1, 2, ...$

$$
\mathbf{E} \sup_{t \in [0,T]} |X_t^n| \leq k_T < \infty
$$

Therefore, using Fatou's lemma, it follows that

$$
\begin{aligned}
\limsup_{n \to \infty} \mathbf{E}|X_t^n - X_t|^2 \leq & \int_0^t \limsup_{n \to \infty} \mathbf{E}|X_t^n - X_t|^2 ds \\
& + 8 \int_0^t F(s)\rho\left(\limsup_{n \to \infty} \mathbf{E}|X_s^n - X_s|^2\right) ds.
\end{aligned}
$$

Thus, by Lemma 4.3.1

$$
\lim_{n \to \infty} \mathbf{E}|X_t^n - X_t|^2 = 0.
$$

■

References

[5] M. Abdelghani, A. V. Melnikov, and A. A. Pak, “An optional semimartingale approach to risk theory,” Submitted, 2020.

[14] S. Anulova and H. Pragarauskas, “Weak markov solutions of stochastic equations,”Lithuanian Mathematical Journal, vol. 17, no. 2, pp. 141–155, 1977.

[26] C. Dellacherie and P.-A. Meyer, “Un nouveau théoreme de projection et de section,” in Séminaire de Probabilités IX Université de Strasbourg, Springer, 1975, pp. 239–245.

[33] L. I. Galchuk, “Optional martingales,” Matematicheskii Sbornik, vol. 154, no. 4,pp. 483–521, 1980.

[32] L. I. Gal’chuk, “Stochastic Integrals with Respect to Optional Semimartingales and Random Measures,” Theory of Probability & Its Applications, vol. 29, no. 1,pp. 93–108, Mar. 1985. doi:10.1137/1129008.

[37] K. Gasparyan, “On change of variables formula for optional semimartingales,” Uch. Zap. Erevan. Gos. Univ., vol. 1, pp. 14–25, 1984.

[55] N. V. Krylov, “On Ito stochastic integral equations,” Teor. Veroyatnost. i Primen., vol. 14, no. 2, pp. 340–348, 1969.

[56] N. V. Krylov, "Some estimates for the density of the distribution of a stochastic integral," Izv. Akad. Nauk SSSR Ser. Mat, vol. 38, pp. 228–248, 1974.

[57] N. V. Krylov, Controlled diffusion processes. Springer Science & Business Media, 2008, vol. 14.

[69] A. V. Mel'nikov, "Stochastic equations and Krylov's estimates for semimartingales," Stochastics: An International Journal of Probability and Stochastic Processes, vol. 10, no. 2, pp. 81–102, 1983.

[76] S. M. Nikol'skii, Approximation of functions of several variables and embedding theorems. Springer Science & Business Media, 2012, vol. 205.

[81] S. Rong, Theory of stochastic differential equations with jumps and applications: mathematical and analytical techniques with applications to engineering. Springer Science & Business Media, 2006.

[83] S. L. Sobolev, Some applications of functional analysis in mathematical physics. American Mathematical Soc., 2008, vol. 90.

Chapter 5

An Optional Semimartingales Approach to Risk Theory

Mathematical risk theory is concerned with the study of stochastic models of risk in finance and insurance. In a basic risk model the value of a risk-portfolio is the sum of opposing cash-flows: premium payments that increase the value of the portfolio and claim payouts that decrease the value of the portfolio. Premium payments are received to cover liabilities – expected losses from claim payouts and other costs. Claims is a result of risk events that occur at random times. Usually, claims' cash-flow is modelled by point processes. An important problem in risk theory is the calculation of the probability of ruin and time to ruin. Ruin probability is the probability that the value of risk-portfolio will ever become negative. Time to ruin is the time it takes for passage below 0.

In [40], Gerber first applied martingale methods in risk theory. Since then these methods have become a standard technique, and a vast amount of papers have appeared, where martingale methods have been used to analyze increasingly complicated risk models. As noted in [41], risk models can be generalized in the following ways:

1. The model includes inflation and interest;

2. The occurrence of claims may be described by a more general point process than the Poisson process.

While papers, including [25], [24] and [28], consider the general risk models of the type 1, the works, for example, of [42] and [47] are mainly focused on the generalisations of type 2. A recent comprehensive review of the literature can be found in [15].

In this chapter we mainly concentrate on the paper of [84], who first used a classical semimartingale theory to find bounds for ruin probability, and we present a new formulation of risk theory based on the general theory of optional semimartingales on *un*usual probability spaces. We derive an optional local martingales representation and use it to compute the probability of ruin for a very general risk model which, in fact, encompasses two types of the generalizations mentioned before. Similar to other works on probability of ruin, our illustrative examples belong to the field of insurance, however, we must note that the proposed model can be applied in other fields as well (see [4]).

5.1 Optional Risk Model

Let $\left(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P}\right)$ be an *un*usual probability space on which risk processes lie. Elements of a risk portfolio include but are not limited to the following components: premium payments, returns on investments, payments of liabilities, costs and claims. These elements are inherently random. For example, returns may jump up, some premium payments may not be paid, costs may increase and claims and liabilities exacerbate. So, let us consider a risk process whose flow can be summarized by the following equation

$$R_t = u + B_t + N_t + D_t + L_t, \tag{5.1}$$

where $u > 0$ is the initial capital and $B_0 = W_0 = D_0 = L_0 = 0$.

The process B is a continuous predictable process of finite variation characterizing a stable flow of income payments including premiums and other sources, N is a continuous local martingale representing a random perturbation, D and L are right-continuous and left continuous jump processes, respectively. The process L may model some substantial gains or losses in returns on investment. The process D includes a sum of negative jumps representing accumulated claims. In addition, D may also consist of jumps formed by non-anticipated sharp falls or rises in returns on investment. All these processes are optional and adapted to the filtration $\mathcal{F}_t$.

Let us consider an example. Assume that A_t is a capital process of some company at time t, and

$$A_t = u + ct + \sigma W_t + \sum_{k=1}^{N_t^r} Y_k - \sum_{i=1}^{N_t^g} Z_i$$

where c and σ are some constant parameters, W is a Wiener process, N^r and N^g are a Poisson process and left-continuous modification of a Poisson process with intensities λ^r and λ^g respectively, Z_i and Y_k denote the left and right jump sizes respectively with some specified distribution. In this case, the process ct is included in the income process B, $N_t = \sigma W_t$, while $L_t = \sum_{i=1}^{N_t^g} Z_i$ and $\sum_{k=1}^{N_t^r} Y_k$ is a part of the process D.

Let $\mu^r(\omega, dt, dx)$ and $\mu^g(\omega, dt, dx)$ be random measures that describe jumps of the process D_t and L_t, respectively, i.e., on $(\mathcal{B}(\mathbb{R}_+) \times \mathcal{B}(\mathbb{R}_0))$, $(\mathbb{R}_0 = \mathbb{R} \setminus \{0\})$ define

$$\mu^r(dt, dx) = \sum_{0<s} \mathbf{1}_{\{\Delta D_s \neq 0\}} \delta_{(s, \Delta D_s)}(dt, dx),$$
$$\mu^g(dt, dx) = \sum_{0\leq s} \mathbf{1}_{\{\Delta^+ L_s \neq 0\}} \delta_{(s, \Delta^+ L_s)}(dt, dx),$$

where $\mathbf{1}_\omega$ is indicator function of a set ω and $\delta_{(s,y)}(dt, dx)$ is the Dirac measure. We assume that

$$\int_{]0,t]}\int_{\mathbb{R}_0} x\ d\mu^r \in \mathcal{A}_{loc} \text{ and } \int_{[0,t[}\int_{\mathbb{R}_0} x\ d\mu^g \in \mathcal{A}_{loc}. \tag{5.2}$$

This assumption implies that the processes

$$\int_{]0,t]}\int_{\mathbb{R}_0} xd(\mu^r - \nu^r) \quad \text{and} \quad \int_{[0,t[}\int_{\mathbb{R}_0} xd(\mu^g - \nu^g)$$

are optional local martingales (see [32], Lemma 3.1, 3.3), where ν^r, ν^g are compensators of μ^r, μ^g, respectively. By Doob-Meyer decomposition of optional semimartingales, R is a special optional semimartingale adapted to $\mathcal{F}_t$.

5.2 The Laplace optional cumulant function and its properties

To find an upper bound of probability of ruin, we first obtain the martingale characterization of optional semimartingales by means of stochastic exponentials. Consider an optional semimartingale X (e.g. risk process) with the local characteristics $(a, \langle X^c\rangle, \nu^r, \nu^g)$ and the following representation (see [32])

$$\begin{aligned} X_t = u + a_t + X_t^c + \int_{]0,t]}\int_{|x|\le 1} xd(\mu^r - \nu^r) + \int_{[0,t[}\int_{|x|\le 1} xd(\mu^g - \nu^g) \\ + \int_{]0,t]}\int_{|x|>1} xd\mu^r + \int_{[0,t[}\int_{|x|>1} xd\mu^g, \end{aligned} \tag{5.3}$$

where $a_t \in \mathcal{P}_s$, $X_t^c \in \mathcal{M}_{loc}^c$, $\int_{]0,t]}\int_{|x|\le 1} xd(\mu^r - \nu^r) \in \mathcal{M}_{loc}^r$ and $\int_{[0,t[}\int_{|x|\le 1} xd(\mu^g - \nu^g) \in \mathcal{M}_{loc}^g$.

In particular, for the risk process R:

$$a_t = B_t + \int_{]0,t]}\int_{|x|\le 1} xd\nu^r + \int_{[0,t[}\int_{|x|\le 1} xd\nu^g, \quad \langle X^c\rangle_t = \langle N\rangle_t .$$

We introduce a (Laplace) optional cumulant function for X:

$$G_t(z) = -za_t + \frac{z^2}{2}\langle X^c\rangle_t + \int_{]0,t]}\int_{\mathbb{R}_0}(e^{-zx} - 1 + zx\mathbf{1}_{(|x|\le 1)})d\nu^r_s + \int_{[0,t[}\int_{\mathbb{R}_0}(e^{-zx} - 1 + zx\mathbf{1}_{(|x|\le 1)})d\nu^g_{s+} \quad (5.4)$$

with the corresponding optional stochastic exponential $(\mathcal{E}(G(z)) = \mathcal{E}_t(G(z)))_{t\ge 0}$ (see [10]).

Let us discuss when the optional stochastic cumulant function $G_t(z)$ in (5.4) is well-defined. Let

$$I_+ = \int_{]0,t]}\int_0^{\infty}(e^{-zx} - 1 + zx\mathbf{1}_{(|x|\le 1)})d\nu^r_s,$$
$$I_- = \int_{]0,t]}\int_{-\infty}^{0}(e^{-zx} - 1 + zx\mathbf{1}_{(|x|\le 1)})d\nu^r_s.$$

Using Taylor's formula we have

$$\begin{aligned} I_- &\le \int_{]0,t]}\int_{-1}^{0}|e^{-zx} - 1 + zx|d\nu^r_s + \int_{]0,t]}\int_{-\infty}^{-1}|e^{-zx} - 1|d\nu^r_s \\ &\le \frac{z^2}{2}\int_{]0,t]}\int_{-1}^{0}x^2 d\nu^r_s + \int_{]0,t]}\int_{-\infty}^{-1}e^{-zx}d\nu^r_s; \\ I_+ &\le \int_{]0,t]}\int_{0}^{1}|e^{-zx} - 1 + zx|d\nu^r_s + \int_{]0,t]}\int_{1}^{\infty}|e^{-zx} - 1|d\nu^r_s \\ &\le \frac{z^2}{2}\int_{]0,t]}\int_{0}^{1}x^2 d\nu^r_s + \nu^r_t([1,+\infty)). \end{aligned}$$

The same inequalities hold for the integral $\int_{[0,t[}\int_{\mathbb{R}_0}(e^{-zx} - 1 + zx\mathbf{1}_{(|x|\le 1)})d\nu^g_{s+}$ in (5.4). We can see that $G_t(z)$ is well-defined, if, in addition to the assumption $\nu^j_t([1,+\infty)) < \infty, j = r, g$, there exists $z_0 > 0$ such that

$$\int_{]0,t]}\int_{-\infty}^{-1}e^{-zx}d\nu^r < \infty, \quad \int_{[0,t[}\int_{-\infty}^{-1}e^{-zx}d\nu^g < \infty \quad (5.5)$$

almost surely for all $t > 0$ and $0 < z \le z_0$.

Denote

$$T(z) = \inf(t : |\mathcal{E}_t(G(z))| = 0)$$

and

$$Z_t = e^{-z(X_t - X_0)}\mathcal{E}_t^{-1}(G(z))\mathbf{1}_{(|\mathcal{E}_t(G(z))|>0)}.$$

Now, we formulate the crucial result of this section.

Theorem 5.2.1 *For every $z \in [0, z_0]$ the process $Z(z) = (Z_{t\wedge T}(z))_{t\geq 0}$ is an optional local martingale.*

To prove Theorem 5.2.1 we need the following lemmas.

Lemma 5.2.1 *(see [32]) Let X be an $\mathcal{O}(\mathbf{F}_+)$-measurable process. There exists a unique (up to indistinguishability) process $U \in \mathcal{M}^g_{loc}$ with the property $\Delta^+U = X$ if and only if the following conditions are satisfied*

(a) the $\mathcal{O}(\mathbf{F})$-optional projection of X is zero,

(b) the process $(\sum_{s\leq t} X_s^2)^{1/2} \in \mathcal{A}_{loc}$.

Lemma 5.2.2 *Let $M \in \mathcal{M}^g_{loc}$ and $Y \in \mathcal{V} \cap \mathcal{O}$. Then*

$$\sum_{s<t} \Delta^+Y_s\Delta^+M_s = \int_{[0,t[} \Delta^+Y_s dM_{s+}.$$

Proof. sWe use a simple fact:

$$\Delta^+U := \Delta^+\left(\sum_{s<t} \Delta^+Y_s\Delta^+M_s - \int_{[0,t[} \Delta^+Y_s dM_s\right) = 0.$$

Notice that integral in the above expression is well-defined. Let $X := \Delta^+U$ as in the Lemma 5.2.1. It is easily seen that all sufficient conditions of Lemma 5.2.1 are satisfied and, thus, $U \in \mathcal{M}^g_{loc}$. On the other hand, since $\Delta^+U = 0$, it follows that U is continuous. Therefore, $U = 0$ or

$$\sum_{s<t} \Delta^+Y_s\Delta^+M_s = \int_{[0,t[} \Delta^+Y_s dM_{s+}.$$

■

Lemma 5.2.3 *If X is a semimartingale and $Y^r \in \mathcal{V} \cap \mathcal{P}$ and $Y^g \in \mathcal{V} \cap \mathcal{O}$, then* $[X,Y] = \int_{]0,t]} \Delta Y_s dX^r_s + \int_{[0,t[} \Delta^+ Y_s dX^g_{s+}$ *and*

$$X_tY_t - X_0Y_0 = \int_{]0,t]} Y_s dX^r_s + \int_{]0,t]} X_{s-} dY^r_s + \int_{[0,t[} Y_{s+} dX^g_{s+} + \int_{[0,t[} X_s dY^g_{s+}$$

Proof.

$$\begin{aligned}
[X,Y]_t &= \sum_{s \le t} \Delta Y_s \Delta X_s + \sum_{s<t} \Delta^+ Y_s \Delta^+ X_s \\
&= \sum_{s \le t} \Delta Y_s \Delta (M^r + A^r)_s + \sum_{s<t} \Delta^+ Y_s \Delta^+ (M^g + A^g)_s \\
&= \int_{]0,t]} \Delta Y_s dM^r_s + \int_{]0,t]} \Delta Y dA^r_s \\
&+ \int_{[0,t[} \Delta^+ Y_s dM^g_{s+} + \int_{[0,t[} \Delta^+ Y_s dA^g_{s+} \\
&= \int_{]0,t]} \Delta Y_s dX^r_s + \int_{[0,t[} \Delta^+ Y_s dX^g_{s+}
\end{aligned}$$

holds because of Lemma 5.2.2 and Proposition 4.49 in [48]. Next, using integration by parts

$$\begin{aligned}
X_tY_t - X_0Y_0 &= \int_{]0,t]} Y_{s-} dX^r_s + \int_{]0,t]} X_{s-} dY^r_s \\
&\quad + \int_{[0,t[} Y_s dX^g_{s+} + \int_{[0,t[} X_s dY^g_{s+} + [X,Y]_t \\
&= \int_{]0,t]} Y_s dX^r_s + \int_{]0,t]} X_{s-} dY^r_s \\
&\quad + \int_{[0,t[} Y_{s+} dX^g_{s+} + \int_{[0,t[} X_s dY^g_{s+}
\end{aligned}$$

■

Proof of Theorem 5.2.1. In all considerations below we will fix the parameter z and write T, G,... instead of $T(z), G(z)$,... In addition, we define $t := t \wedge T$ for convenience.

Next,

$$Z_t = e^{-z(X_t - X_0)} \mathcal{E}^{-1}_t(G),$$

where

$$\mathcal{E}_t^{-1}(G) = \exp\left(-\int_{]0,t]} \frac{dG_s^r}{1+\Delta G_s} - \int_{[0,t[} \frac{dG_{s+}^g}{1+\Delta^+ G_s}\right)$$
$$\times \prod_{0<s\le t}\left(1-\frac{\Delta G_s}{1+\Delta G_s}\right) e^{\frac{\Delta G_s}{1+\Delta G_s}} \prod_{0\le s<t}\left(1-\frac{\Delta^+ G_s}{1+\Delta^+ G_s}\right) e^{\frac{\Delta^+ G_s}{1+\Delta^+ G_s}}.$$

By this representation it follows that $\mathcal{E}_t^{-1}(G)$ is the solution of Doleans equation (see [32])

$$\mathcal{E}_t^{-1}(G) = 1 - \int_{]0,t]} \mathcal{E}_{s-}^{-1}(G)\frac{dG_s^r}{1+\Delta G_s} - \int_{[0,t[} \mathcal{E}_s^{-1}(G)\frac{dG_{s+}^g}{1+\Delta^+ G_s} \tag{5.6}$$

Using (5.3), (5.4) and change of variables formula for optional semimartingales (see Theorem 2.9.2) we get

$$\begin{aligned} L_t =& e^{-rz(X_t-X_0)} \\ =& 1 + \int_{]0,t]} L_{s-} dG_s^r + \int_{[0,t[} L_s dG_{s+}^g - z\int_{]0,t]} L_{s-} dX_s^c \\ &+ \int_{]0,t]}\int_{\mathbb{R}_0} L_{s-}(e^{-zx}-1)d(\mu^r-\nu^r)_s \\ &+ \int_{[0,t[}\int_{\mathbb{R}_0} L_s(e^{-zx}-1)d(\mu^g-\nu^g)_{s+}. \end{aligned} \tag{5.7}$$

By (5.6) and (5.7) and Lemma 5.2.3, we find

$$
\begin{aligned}
Z_t =& L_t \mathcal{E}_t^{-1}(G) \\
=& 1 + \int_{]0,t]} \mathcal{E}_s^{-1}(G) dL_s^r + \int_{]0,t]} L_{s-} d(\mathcal{E}_s^{-1}(G))^r \\
&+ \int_{[0,t[} \mathcal{E}_{s+}^{-1}(G) dL_{s+}^g + \int_{[0,t[} L_s d(\mathcal{E}_{s+}^{-1}(G))^g \\
=& 1 + \int_{]0,t]} \frac{\mathcal{E}_{s-}^{-1}(G)}{1+\Delta G_s} \left[L_{s-} dG_s^r - z L_{s-} dX_s^c + \int_{\mathbb{R}_0} L_{s-}(e^{-zx}-1) d(\mu^r - \nu^r)_s \right] \\
&+ \int_{[0,t[} \frac{\mathcal{E}_s^{-1}(G)}{1+\Delta^+ G_s} \left[L_s dG_{s+}^g + \int_{\mathbb{R}_0} L_s (e^{-zx}-1) d(\mu^g - \nu^g)_{s+} \right] \\
&- \int_{]0,t]} \mathcal{E}_{s-}^{-1}(G) L_{s-} \frac{dG_s^r}{1+\Delta G_s} - \int_{[0,t[} \mathcal{E}_s^{-1}(G) L_s \frac{dG_{s+}^g}{1+\Delta^+ G_s} \\
=& 1 - z \int_{]0,t]} \frac{Z_{s-}}{1+\Delta G_s} dX_s^c + \int_{]0,t]} \int_{\mathbb{R}_0} \frac{Z_{s-}(e^{-zx}-1)}{1+\Delta G_s} d(\mu^r - \nu^r)_s \\
&+ \int_{[0,t[} \int_{\mathbb{R}_0} \frac{Z_s(e^{-zx}-1)}{1+\Delta^+ G_s} d(\mu^g - \nu^g)_{s+}
\end{aligned}
$$

This implies $(Z_{t\wedge T})_{t\geq 0}$ is an optional local martingale (for each $z \in [0, z_0]$). ■

Along with the stochastic exponential of the cumulant process $\mathcal{E}(G)$ there is a usual exponent of the cumulant process e^G. We know that if $\Delta G > -1$ and $\Delta^+ G > -1$ then $\mathcal{E}(G)$ can be represented as

$$
\begin{aligned}
\mathcal{E}_t(G) = \exp \Bigg\{ G_t &+ \sum_{0<s\leq t} (\log(1+\Delta G_s) - \Delta G_s) \\
&+ \sum_{0\leq s<t} (\log(1+\Delta^+ G_s) - \Delta^+ G_s) \Bigg\}
\end{aligned}
$$

From above we observe that if $\Delta G = 0$ and $\Delta^+ G = 0$, or equivalently, G is a continuous process, then

$$\mathcal{E}(G) = e^G.$$

Let us present sufficient conditions for the cumulant process G to be continuous.

Lemma 5.2.4

- *If $\Delta X_T = 0$ a.s. on the set $\{T < \infty\}$ for every predictable time T, i.e. $T \in \mathcal{T}^p$, then $\Delta G = 0$.*
- *If $\Delta^+ X_T = 0$ a.s. on the set $\{T < \infty\}$ for every totally inaccessible time T, i.e. $T \in \mathcal{T}$, then $\Delta^+ G = 0$.*

Proof. From the first condition it follows that $\Delta G_t = \int_{\mathbb{R}_0} (e^{-rx} - 1)\nu^r(\{t\}, dx) = 0$ (proof without usual hypothesis on the filtration is the same as in [48], II.1.19). By Lemma 3.3 in [32], it follows from the second condition that $\Delta^+ G = \int_{\mathbb{R}_0} (e^{-rx} - 1)\nu^g(\{t\}, dx) = 0$. ■

5.3 Probability of Ruin

Here we apply the theory developed in the previous section to study probability of ruin of the optional semimartingale R introduced by formula (5.1). Given R, our main goal is to evaluate the ruin probability $\mathbf{P}(\tau < \infty)$, where $\tau = \inf\{t > 0 : R_t < 0\}$.

Let us assume that there exists $z_0 > 0$ such that

$$\int_{]0,t]} \int_{-\infty}^{-1} e^{-zx} d\nu^r < \infty, \quad \int_{[0,t[} \int_{-\infty}^{-1} e^{-zx} d\nu^g < \infty \tag{5.8}$$

almost surely for all $t > 0$ and $0 < z \leq z_0$. Note that, in literature (see [15], p.338) such risk process R is called as a process with light-tailed negative jumps.

Let us define the optional cumulant process for the risk process R: for all $t > 0$ and $z \in (0, z_0]$

$$G_t(z) = -zB_t + \frac{z^2}{2} \langle N \rangle_t + \int_{]0,t]} \int_{\mathbb{R}_0} (e^{-zx} - 1) d\nu^r + \int_{[0,t[} \int_{\mathbb{R}_0} (e^{-zx} - 1) d\nu^g.$$

Define a process

$$M_t(z) = \exp[-z(R_t - u)]\mathcal{E}_t^{-1}(G(z)). \tag{5.9}$$

It follows from Theorem 5.2.1 that the process $M_t(z)$ is an optional local martingale for every z in $[0, z_0]$, if $\Delta G > -1$ and $\Delta^+ G > -1$.

We use the optional local martingale M in a similar way as in [84], which is standard in risk theory. We know that a non-negative optional local martingale is a supermartingale (see Lemma 2.4.2). Thus, since $M_0(z) = 1$, it follows that

$$\begin{aligned}
1 \geq & \mathbf{E}(M_{\tau \wedge t}(z)) \\
= & \mathbf{E}(M_\tau(z)\mathbf{1}_{\tau \leq t} + M_t(z)\mathbf{1}_{\tau > t}) \\
\geq & \mathbf{E}(M_\tau(z)\mathbf{1}_{\tau \leq t}) \\
= & \mathbf{E}(M_\tau(z) | \tau \leq t)\mathbf{P}(\tau \leq t)
\end{aligned}$$

for every z in $[0, z_0]$ and $t > 0$.

Hence, for all $z \in [0, z_0]$

$$\begin{aligned}
\mathbf{P}(\tau \leq t) \leq & \frac{1}{\mathbf{E}(M_\tau | \tau \leq t)} \\
= & \frac{1}{\mathbf{E}(e^{-z(R_\tau - u)}\mathcal{E}_\tau^{-1}(G(z)) | \tau \leq t)}.
\end{aligned}$$

Since $R_\tau \leq 0$ on $\{\tau \leq t\}$, we get

$$\mathbf{P}(\tau \leq t) \leq \frac{e^{-zu}}{\mathbf{E}(\mathcal{E}_\tau^{-1}[G(z)] | \tau \leq t)}, \tag{5.10}$$

for all $z \in [0, z_0]$.

At this point, if we want to find a better estimate of the upper bound of the ruin probability in (5.10), we are required to impose additional assumptions.

Firstly, suppose that

$$-G'_\tau(0) = B_\tau + \int_{]0,\tau]} \int_{\mathbb{R}_0} x d\nu_s^r + \int_{[0,\tau[} \int_{\mathbb{R}_0} x d\nu_{s+}^g > 0. \tag{5.11}$$

Note that differentiation with respect to z under the integral sign with respect to ν^j, $j = r, g$, in (5.11) is possible because zero is an interior point in the range of z-values for which the integral exists.

If we assume that $\Delta G = 0$ and $\Delta^+ G = 0$ (see Remark 5.3.3 below), then

$$\mathcal{E}_t(G(z)) = \exp(G_t(z)). \tag{5.12}$$

Now, using (5.12) and Jensen's inequality, we get from (5.10) that

$$\begin{aligned} \mathbf{P}(\tau \leq t) &\leq e^{-zu} \left[\mathbf{E}\left(\mathcal{E}_\tau^{-1}[G(z)]|\tau \leq t\right)\right]^{-1} \\ &\leq e^{-zu}\mathbf{E}(\mathcal{E}_\tau[G(z)]|\tau \leq t) \\ &= \mathbf{E}(\exp[-zu + G_\tau(z)]|\tau \leq t), \end{aligned} \tag{5.13}$$

for all $z \in [0, z_0]$ and $t > 0$.

The function $G_\tau(z)$ is a strictly convex function of z with $G_\tau(0) = 0$. Further, due to assumption (5.11), the function $\mathbf{E}(\exp[-zu + G_\tau(z)]|\tau \leq t)$ is strictly convex function of z decreasing from 1 at $z = 0$. Furthermore, under assumption (5.8) with $z_0 = \infty$, the function $\mathbf{E}(\exp[-zu + G_\tau(z)]|\tau \leq t)$ increases to $+\infty$ as $z \to \infty$. Hence, there exists a unique $z^* \in [0, z_0]$ for which this function attains its minimum, and by (5.13) we have

$$\mathbf{P}(\tau \leq t) \leq \frac{e^{-z^* u}}{\mathbf{E}(\exp[-G_\tau(z^*)]|\tau \leq t)}. \tag{5.14}$$

It seems the estimate (5.14) is the best upper bound for ruin probability obtained from (5.10).

We can find a more explicit estimate of the ruin probability if we can choose z-value $z_t > 0$, for which the denominator of (5.10) equals one. Let us note that this z-value is not necessarily unique, and it often depends on t. In this case we get

$$\mathbf{P}(\tau \leq t) \leq e^{-uz_t}.$$

Further, if z_t exists for all $t > 0$ and if $\hat{z} = \lim_{t\to\infty} z_t$ exists, then

$$\mathbf{P}(\tau < \infty) \leq e^{-u\hat{z}}.$$

Finally, we summarize our results in the following theorem.

Theorem 5.3.1 *Given the optional risk model* (5.1), *suppose assumptions* (5.2) *and*

$$\nu_t^j([1,+\infty)) < \infty, j = r, g,$$

hold and there exists an $z_0 > 0$ *such that*

$$\int_{]0,t]}\int_{|x|>1} e^{-zx} d\nu^r < \infty, \quad \int_{[0,t[}\int_{|x|>1} e^{-zx} d\nu^g < \infty$$

almost surely for all $t > 0$ *and* $0 < z \leq z_0$.

1. *If* $\Delta G > -1$ *and* $\Delta^+ G > -1$ *then*

$$\mathbf{P}(\tau \leq t) \leq \frac{e^{-zu}}{\mathbf{E}(\mathcal{E}_\tau^{-1}[G(z)]|\tau \leq t)}$$

   for all $z \in [0, z_0]$ *and* $t > 0$.

2. *If* $G'_\tau(0) < 0$, $\Delta G = 0$ *and* $\Delta^+ G = 0$ *then there exists a unique* $z^* \in [0, z_0]$ *for which the right-hand side of*

$$\mathbf{P}(\tau \leq t) \leq \frac{e^{-z^* u}}{\mathbf{E}(\exp[-G_\tau(z^*)]|\tau \leq t)}$$

   attains its minimum.

3. *If* $z_t > 0$ *exists for which* $\mathbf{E}(\mathcal{E}_\tau^{-1}[G(z_t)]|\tau \leq t) = 1$ *then*

$$\mathbf{P}(\tau \leq t) \leq e^{-uz_t}.$$

   Moreover, if z_t *exists for all* $t > 0$ *and if* $\hat{z} = \lim_{t\to\infty} z_t$ *exists, then*

$$\mathbf{P}(\tau < \infty) \leq e^{-u\hat{z}}. \tag{5.15}$$

Remark 5.3.1 *Notice that the inequality* (5.15) *is an analogue of the classical Cramer-Lundberg bound (see, for example, [41]) obtained in a very general optional setting.*

Remark 5.3.2 *If the process D represents the accumulated claims (all jumps are downwards) while the process $L = 0$, then the condition* (5.11) *becomes*

$$-G'_t(0) = B_t + \int_{]0,t]} \int_{-\infty}^{0} x d\nu^r > 0 \ for \ all \ t > 0$$

and is known as net-profit condition in risk theory. This means that the insurance company adopts the wise premium policy such that premiums follow the claims' intensity.

Remark 5.3.3 *By Lemma 5.2.4, condition $\Delta G = \int (e^{-zx} - 1)\nu^r(\{t\}, dx) = 0$ is satisfied if $\Delta R_T = 0$ a.s. on the set $\{T < \infty\}$ for every predictable stopping time T (i.e. quasi-left-continuity of R (see [65])). From the point of view of risk theory it means that claims cannot be predicted beforehand. On the other hand, by Lemma 5.2.4, condition $\Delta^+ G = \int (e^{-zx} - 1)\nu^g(\{t\}, dx) = 0$ is satisfied if $\Delta^+ R_T = 0$ a.s. on the set $\{T < \infty\}$ for every stopping time T measurable with respect to the filtration* $\mathbf{F}$.

5.4 Particular Models and Examples

Let us apply our results to different risk models.

Example 5.4.1 *We consider a particular type of the general risk model studied in the previous section. Specifically, we assume that*

$$R_t = u + B_t + \int_0^t \sigma_s dW_s + \sum_{i=1}^{N_t^g} Z_i - \sum_{k=1}^{N_t^r} Y_k$$

where B is a continuous process of finite variation, W is a standard Wiener process, and σ is a predictable process. The process N_s^g and N_s^r are left- and right-continuous *counting processes with intensities λ_s^g and λ_s^r, respectively, such that $N_t^j - \int_0^t \lambda_s^j ds,\ j = r, g$, are optional local martingales. The positive random variables Y_k and Z_i assumed to be mutually independent. The distribution of the claim Y_k depend on the time at which k'th jump occurs, but is otherwise non-random and independent of the N^r process. Thus, the Y_k-s can depend on the N^r-process only through the time-dependence of the distributions of the Y_k-s. The same holds for random variables Z_i-s and the process N^g. An example of time-dependence is when the claims are subject to inflation or interest force (see, for example, [21]).*

Under these assumptions $\nu^r(dt, dx) = \lambda_t^r(1-F_t^r(-dx))dt$ and $\nu^g(dt, dx) = \lambda_t^g F_t^g(dx))dt$, where F_t^r and F_t^g are the respective distributions of Y_k and Z_i at time t. Hence

$$\begin{aligned} G_t(z) = &- zB_t + \frac{z^2}{2}\int_{[0,t]} \sigma_s^2 ds \\ &+ \int_{]0,t]} [\varphi_s^r(-z) - 1]\lambda_s^r ds + \int_{[0,t[} [\varphi_s^g(z) - 1]\lambda_s^g ds \end{aligned}$$

where $\varphi_s^j(z) = \int e^{-zx} dF_s^j(x)$ is the Laplace transform of $F_s^j,\ j = r, g$.

For this model, from condition (5.11) we obtain

$$B_t + \int_0^t \mu_s^g \lambda_s^g ds > \int_0^t \mu_s^r \lambda_s^r ds,$$

for all $t > 0$ where μ_s^g and μ_s^r denotes the mean of Z_i and Y_k at time s, respectively.

We will next discuss simple situations where the ruin probability can easily be evaluated. We suppose that for each $t > 0$ there exists a distribution function $\tilde{F}^j$ such that $F_s^r(x) \geq \tilde{F}_t^r(x)$ and $F_s^g(x) \leq \tilde{F}_t^g(x)$ for all $x > 0$ and all $s \leq t,\ j = r, g$. Under these conditions, $\mu_s^j \leq \tilde{\mu}_t^j,\ j = r, g$, for $s \leq t$ and

$$\begin{aligned} \int_0^s [\varphi_u^r(-z) - 1]\lambda_u^r du \leq [\tilde{\varphi}_t^r(-z) - 1]\Lambda_t^r, \\ \int_0^s [\varphi_u^g(z) - 1]\lambda_u^g du \leq [\tilde{\varphi}_t^g(z) - 1]\Lambda_t^g. \end{aligned} \tag{5.16}$$

where $\tilde{\mu}_t^j$ *denotes the mean value of* $\tilde{F}_t^j$, $\tilde{\varphi}_t^j(z) = \int e^{-zx} d\tilde{F}_t^j(x)$, *and* $\Lambda_s^j = \int_0^t \lambda_s^j ds$ *is integrated intensity of* N^j, $j = r, g$.

We assume that a company adopts a policy such that for some constant $c > 1$

$$B_t \geq c\left(\tilde{\mu}_t^r \int_0^t \lambda_s^r ds - \tilde{\mu}_t^g \int_0^t \lambda_s^g ds\right). \tag{5.17}$$

for $s \leq t$. *If, moreover,* σ_s^2 *is bounded by a constant* ζ_t^2 *for* $s \leq t$, (5.16) *and* (5.17) *implies that*

$$\begin{aligned} G_s(z) &\leq [-zc\tilde{\mu}_t^r + \tilde{\varphi}_t^r(-z) - 1]\Lambda_s^r + [-zc\tilde{\mu}_t^g + \tilde{\varphi}_s^g(z) - 1]\Lambda_{s+}^g + \frac{1}{2}z^2\zeta_t^2 s \\ &= (g_t^r(z), g_t^g(z)) \circ \Lambda_s + \frac{1}{2}z^2\zeta_t^2 s \end{aligned}$$

for all $s \leq t$ *where* $g_t^r(z) = -zc\tilde{\mu}_t^r + \tilde{\varphi}_t^r(-z) - 1$, $g_t^g(z) = -zc\tilde{\mu}_t^g + \tilde{\varphi}_s^g(z) - 1$ *and* $\Lambda_s = \Lambda_s^r + \Lambda_s^g$. *Under the conditions imposed* $g_t^j(z)$ *is convex,* $g_t^j(0) = 0$ *and* $(g_t^j)'(0) < 0$, *so there is a range* $[0, z_t]$ *of* z*-values for which* $g_t^j(z) \leq 0$, $j = r, g$. *For* $z \in [0, z_0]$ *it follows from* (5.10) *that*

$$\mathbf{P}(\tau \leq t) \leq \frac{e^{-zu + \frac{1}{2}z^2\zeta_t^2 t}}{\mathbf{E}(\exp[-(g_t^r(z), g_t^g(z)) \circ \Lambda_\tau] | \tau \leq t)} \tag{5.18}$$

The Laplace transform of Λ *is rarely known, but when the Laplace transform of* Λ *is known, it is sometimes possible to proceed in a way analogous to the derivation of the upper bound* (5.21) *demonstrated in Example 5.4.2. Quite generally we can use that* $-zu + \frac{1}{2}z^2\zeta_t^2 t$ *has a minimum at* $z = u/(t\zeta_t^2)$, *which implies*

$$\mathbf{P}(\tau \leq t) \leq e^{-\frac{1}{2}u^2/(\zeta_t^2 t)}$$

provided $u/(t\zeta_t^2) \leq z_t$.

In general, we have the result

$$\mathbf{P}(\tau \leq t) \leq \exp\left(-z_t u + \frac{1}{2}z_t^2\zeta_t^2 t\right).$$

Example 5.4.2 *Consider the special case, namely, the classical compound Poisson risk model with additional random positive left-continuous jumps of size Z_i and perturbed by a Wiener process W*

$$R_t = u + ct + \sigma W_t + \sum_{i=1}^{N_t^g} Z_i - \sum_{k=1}^{N_t^r} Y_k \tag{5.19}$$

where c is the premium rate, N^r and N^g are a Poisson process and left-continuous modification of a Poisson process with intensities λ^r and λ^g, respectively. The Z_i's and Y_k's are positive, independent identically distributed random variables with distribution functions F^g and F^r, respectively. We assume that $W, N^r, N^g, \{Z_i\}$ and $\{Y_k\}$ are all mutually independent.

In this particular case, $B_t = ct$, $\langle N \rangle_t = \sigma^2 t$, $\nu^r(\omega; dt, dx) = \lambda^r(1 - F^r(-dx))dt$ and $\nu^g(\omega; dt, dx) = \lambda^g F^g(dx)dt$, so

$$\begin{aligned} G_t(z) =& g(z)t \\ =& \left(-zc + \frac{1}{2}\sigma^2 z^2 + \lambda^r \left[\varphi_{F^r}(-z) - 1\right] + \lambda^g \left[\varphi_{F^g}(z) - 1\right]\right) t, \end{aligned}$$

where $\varphi_{F^j}(z) = \int e^{-zx} dF^j(x)$ is the Laplace transform of F^j, $j = r, g$. Since the process R_t in this case is a process with independent increments , $M_t(z)$ in (5.9) is a martingale for every z in the domain of φ_{F^r} and φ_{F^g}. We see that $z_t = \hat{z}$ is the positive solution of $g(z) = 0$. Thus, for all $u \geq 0$

$$\mathbf{P}(\tau < \infty) \leq e^{-u\hat{z}}.$$

This bound was already obtained in the case with no positive jumps (see, for instance, [80], 13.2.1). We can sometimes obtain a more accurate upper bound of finite time probability of ruin. For $z \in [\hat{z}, z_0]$, $g(z) \geq 0$, thus by (5.10)

$$\mathbf{P}(\tau \leq t) \leq \exp[-zu + g(z)t] \quad \text{for all } z \in [\hat{z}, z_0]. \tag{5.20}$$

The right hand side of (5.20) *attains its minimum at* z^*, *which is given as the solution of* $g'(z^*) = u/t$, *provided there is a solution in* $[0, z_0]$. *Otherwise the minimum is attained at* $z^* = z_0$, *in which case* $g'(z^*) < u/t$. *If* $t \leq u/g'(\hat{z})$, *the convexity of* g *implies that* $z^* \geq \hat{z}$, *so*

$$\mathbf{P}(\tau \leq t) \leq \exp[-z^* u + g(z^*)t] \text{ for } t \leq u/g'(z). \tag{5.21}$$

Since $g(z^*) < (z^* - \hat{z})u/t$ *for* $z^* > \hat{z}$ *(using again the convexity of* g *and the fact that* $g(\hat{z}) = 0$*), the right hand side of* (5.21) *is strictly smaller than* $\exp(-u\hat{z})$ *when* $t < u/g'(\hat{z})$.

Example 5.4.3 *Consider a specific case of the model* (5.19) *with the following cumulative distribution functions of* Z_i *and* Y_i

$$F^g(x) = 1 - e^{-bx}, \quad b > 0, \quad F^r(x) = 1 - e^{-ax}, \quad a > 0.$$

Additionally, we assume that net profit condition is satisfied, $c > \frac{\lambda^r}{a} - \frac{\lambda^g}{b}$. *For* $z \in (0, a]$, *we obtain*

$$\begin{aligned} G_t(z) =& g(z)t \\ =& \left(-cz + \frac{1}{2}\sigma^2 z^2 + \lambda^r \left[\frac{a}{a-z} - 1\right] + \lambda^g \left[\frac{b}{b+z} - 1\right]\right) t, \end{aligned}$$

We rewrite $g(z) = \frac{zh(z)}{2(a-z)(b+z)}$, *where*

$$\begin{aligned} h(z) = -\sigma^2 z^3 + (\sigma^2(a-b) + 2c)z^2 \\ + (\sigma^2 ab - 2c(a-b) + 2(\lambda^r + \lambda^g))z + 2(\lambda^r b - \lambda^g a - cba). \end{aligned} \tag{5.22}$$

Thus, if the equation $h(z) = 0$ *has solution* $\hat{z} \in (0, a]$ *then, by Theorem 5.3.1, it follows that*

$$P(\tau < \infty) \leq e^{-\hat{z}u}.$$

This result generalizes the following special cases:

1. *(cadlag case, cf. [85]) If $c > 0, \sigma > 0$ and there is no positive jumps, i.e., $Z_i = 0$. Then net profit condition is $c > \frac{\lambda^r}{a}$, and (5.22) becomes*

$$h(z) = -\sigma^2 z^2 + (\sigma^2 a + 2c)z - 2(ca - \lambda^r).$$

Quadratic equation $h(z) = 0$ has exactly two real roots for $z \in (0, \infty)$, given by

$$\hat{z}_{\pm} = \frac{\sigma^2 a + 2c \pm \sqrt{\Delta}}{2\sigma^2},$$

where $\Delta := (\sigma^2 a - 2c)^2 + 8\sigma^2 \lambda^r$.

Since $(\sigma^2 a - 2c)^2 \leq \Delta \leq (\sigma^2 a + 2c)^2$, we have $\hat{z}_- \geq 0$ and $\hat{z}_+ \geq a$. If $\hat{z}_- \leq a$ then, by Theorem 5.3.1, we get for all $u \geq 0$

$$\mathbf{P}(\tau < \infty) \leq e^{-u\hat{z}_-}.$$

2. *(pure jumps case, cf. [75], 8.3) If $c = 0, \sigma = 0$, then net profit condition is $\frac{\lambda^g}{b} > \frac{\lambda^r}{a}$, and (5.22) becomes*

$$h(z) = 2(\lambda^r + \lambda^g))z + 2(\lambda^r b - \lambda^g a).$$

Then the equation $h(z) = 0$ has a unique real root in the interval $(0, a)$, given by

$$\hat{z} = \frac{\lambda^g a - \lambda^r b}{\lambda^g + \lambda^r}.$$

Therefore, by Theorem 5.3.1, for all $u \geq 0$

$$\mathbf{P}(\tau < \infty) \leq e^{-u\hat{z}}.$$

Chapter 6

One-dimensional optional regression models

Regression Analysis is an integral part of Mathematical Statistics. Developments in this area are important from both theoretical and applied points of view. In statistics of random processes a regression model is considered as a semimartingale where the drift depends on an unknown parameter and the martingale part presents the errors in observations. Such a viewpoint is very productive because it creates a possibility to study a variety of regression models (with discrete and continuous time) in an unified way, using martingale methods (see, for example, [65], [71]).

The standard martingale theory is well-developed under so-called "usual conditions", when filtration (information flow) is complete and right continuous. However, statistical data is usually delivered by a stochastic process, whose history (natural filtration) may not be right-continuous, and therefore such technical conditions may not be fulfilled (see [4]). This is the main reason why we need to consider regression models in more general setting which we call here the optional regression model. Optional semimartingales, on which our optional regression model is based, admit trajectories which are not right-continuous and arise when "usual conditions" are not assumed on filtered probability space. Up to our knowledge, currently there are no works devoted to the relaxing of these "usual conditions" and investigation of such general optional regression model.

Firstly, we focus on strong consistency of the proposed LS-estimate. In case of the observed process being cadlag (right-continuous with left limits) semimartingale this problem was extensively studied in [71] (also see [70]).

Secondly, we concentrate our attention on the sequential estimates with guaranteed accuracy. In comparison to the structural LS-estimates the sequential LS-estimates posses an advantage of having bounded variance. This type of estimates in cadlag case is well-established (see [72], [71], [34], [19]).

The chapter is organized in the following way: in section 5.1 we introduce the general regression model along with structural LS-estimates and auxiliary results. In section 5.2 we prove strong consistency of the proposed LS-estimates. In section 5.3 we consider sequential estimates, show that these estimates are unbiased and have a property of guaranteed accuracy under suitable conditions on regressor and error term. In addition, we investigate a problem related to hypothesis testing. Finally, in section 5.4 we present an extension of sequential LS-estimators for non-linear regression models and several illustrative examples.

6.1 Optional stochastic regression model

Suppose that on the fixed stochastic basis $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ without "usual conditions", we observe a one-dimensional process X.

Suppose the process X has the following form

$$X_t = f \circ a_t\theta + M_t, \tag{6.1}$$

where $f \circ a_t$ is an optional stochastic integral such that

$$f \circ a_t = \int_{]0,t]} f_s^r da_s^r + \int_{[0,t[} f_s^g da_{s+}^g,$$

$a = a^r + a^g \in \mathcal{A}_{loc}^+ \cap \mathcal{P}_s$, $M \in \mathcal{M}_{loc}$, f_t is a bilinear pair $f_t = (f_t^r, f_t^g)$, $f_t^r \in \mathcal{P}$, $f_t^g \in \mathcal{O}$, and $\theta \in \mathbb{R}$ is the unknown parameter which we need to estimate.

As the estimator of θ we consider the statistic

$$\theta_t = F_t^{-1} (f \circ X_t), \tag{6.2}$$

where $F_t := f^2 \circ a_t \in \mathcal{A}_{loc}^+ \cap \mathcal{P}_s$ is assumed to be non-zero (a.s.). This assumption is not restrictive because further we suppose that $F_t \to \infty$ (a.s.) to provide strong consistency of θ_t.

The structure of the estimator (6.2) is similar to estimator obtained by the method of Least Squares (LS) in classical regression analysis. Therefore, θ_t will be called the structural LS-estimator of θ. It is well known how to study its asymptotic behaviour with the help of the Strong Law of Large Numbers (SLLN). Liptser (1980) [66] proposed a very general form of SLLN for local martingales using a stochastic Kronecker's Lemma. For reader's convenience, let us reproduce this scheme in optional setting (see [35], [71]).

To prove Kronecker's Lemma in optional setting, we need the following result on sets of convergence of optional martingales. In what follows, we denote $\widetilde{D}$ the compensator of some increasing process D.

Lemma 6.1.1 *(see [36]) If $Y \in \mathcal{M}_{loc}$ then*

$$(\widetilde{D}_\infty < \infty) \subseteq (Y \to) \ a.s.,$$

where

$$D_t = \langle Y^c \rangle_t + \sum_{0<s\le t} \frac{(\Delta Y_s)^2}{1 + |\Delta Y_s|} + \sum_{0\le s<t} \frac{(\Delta^+ Y_s)^2}{1 + |\Delta^+ Y_s|}, \tag{6.3}$$

and $(Y \rightarrow)$ *is the set, on which there exists a finite random variable* $Y_\infty(\omega) = \lim_{t\to\infty} Y_t(\omega) < \infty$.

Now we present the following generalization of Kronecker's Lemma.

Lemma 6.1.2 *For processes* $N \in \mathcal{M}_{loc}$ *and* $A \in \mathcal{V}^+ \cap \mathcal{P}_s$ *the following relation holds*

$$(A_\infty = \infty) \cap (Y_t \rightarrow) \subseteq (A_t^{-1} N_t \rightarrow 0) \quad (a.s.) \quad (t \rightarrow \infty),$$

where

$$Y_t = \int_{]0,t]} (1 + A_s)^{-1} dN_s^r + \int_{[0,t[} (1 + A_{s+})^{-1} dN_{s+}^g. \tag{6.4}$$

Proof. From (6.4) it is easy to see that

$$\int_{]0,t]} (1 + A_s) dY_s^r + \int_{[0,t[} (1 + A_{s+}) dY_{s+}^g = N_t - N_0. \tag{6.5}$$

Using integration by parts formula (see Lemma 3.4, [5]) we obtain

$$(1 + A_t) Y_t = \int_{]0,t]} (1 + A_s) dY_s^r + \int_{[0,t[} (1 + A_{s+}) dY_{s+}^g + \int_{]0,t]} Y_{s-} dA_s^r + \int_{[0,t[} Y_s dA_{s+}^g.$$

Then from (6.5) we conclude that

$$\frac{N_t}{1 + A_t} = \frac{N_0 + Y_t}{1 + A_t} + \frac{1}{1 + A_t} (A_t Y_t - \int_{]0,t]} Y_{s-} dA_s^r - \int_{[0,t[} Y_s dA_{s+}^g).$$

Since $\sup_{t \geq 0} |Y_t| < \infty$ on the set $(A_\infty = \infty) \cap (Y \rightarrow)$, we have that $(1 + A_t)^{-1}(N_0 +$

$Y_t) \to 0$ a.s. as $t \to \infty$. On the other hand, we have for $u < t, v < t$

$$
\begin{aligned}
&\frac{1}{1+A_t}\left|A_tY_t - \int_{]0,t]} Y_{s-}dA^r_s - \int_{[0,t[} Y_s dA^g_{s+}\right| \\
&\quad = \frac{1}{1+A_t}\left|\int_{]0,t]} (Y_t - Y_{s-})dA^r_s + \int_{[0,t[} (Y_t - Y_s)dA^g_{s+}\right| \\
&\quad \leq \frac{1}{1+A_t}\left(\int_{]0,u]} |Y_t - Y_{s-}|dA^r_s + \int_{]u,t]} (|Y_\infty - Y_t| + |Y_\infty - Y_{s-}|)dA^r_s \right. \\
&\qquad \left. + \int_{[0,v[} |Y_t - Y_s|dA^g_{s+} + \int_{[v,t[} (|Y_\infty - Y_t| + |Y_\infty - Y_s|)dA^g_{s+}\right) \\
&\quad \leq 2\sup_{s\geq 0}|Y_s|(1+A_t)^{-1}(A^r_u + A^g_v) + |Y_\infty - Y_t| \\
&\qquad + (1+A_t)^{-1}\int_{]u,t]} |Y_\infty - Y_{s-}|dA^r_s + (1+A_t)^{-1}\int_{[v,t[} |Y_\infty - Y_s|dA^g_{s+}. \qquad (6.6)
\end{aligned}
$$

Using the fact that $\sup_{s\geq 0}|Y_s| < \infty$ on the set $(Y \to)$, we can choose for sufficiently large t appropriate values u and v on the set $(A_\infty = \infty) \cap (Y \to)$ to make the right side of (6.6) tend to zero. Consequently, the statement of the lemma follows. ■

Remark 6.1.1 *Although we proved Lemma 6.1.2 for the process $N \in \mathcal{M}_{loc}$, this proof also works for any optional semimartingale N.*

6.2 Strong Consistency

In this section we will show that the estimator θ_t in (6.2) is strongly consistent. The proof of the strong consistency is based on the SLLN in optional case.

Let $N \in \mathcal{M}_{loc}$ and

$$
N_t = N^c_t + \int_{]0,\infty]}\int_{\mathbb{R}_0} xd(\mu^r - \nu^r)_s + \int_{[0,\infty[}\int_{\mathbb{R}_0} xd(\mu^g - \nu^g)_{s+} \qquad (6.7)
$$

be the canonical decomposition of N, where $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$, N^c be a continuous part of N, μ^r and μ^g be random measures of right and left jumps of N, and ν^r and ν^g be their respective compensators.

Theorem 6.2.1 *Let $A \in \mathcal{V}^+ \cap \mathcal{P}_s$ and $A_\infty = \infty$ a.s. If $N \in \mathcal{M}_{loc}$ and for some $q \in [1,2]$*

$$\int_{]0,\infty]} \frac{d\langle N^c\rangle_s}{(1+A_s)^2} + \int_{]0,\infty]}\int_{\mathbb{R}_0} |1+A_s|^{-q}|x|^q d\nu_s^r + \int_{[0,\infty[}\int_{\mathbb{R}_0} |1+A_{s+}|^{-q}|x|^q d\nu_{s+}^g < \infty, \quad (6.8)$$

then

$$A_t^{-1}N_t \to 0 \text{ a.s. as } t \to \infty.$$

Proof. Using the fact that

$$\frac{x^2}{1+|x|} \leq |x|^q,\ q \in [1,2],$$

we get for $q \in [1,2]$

$$\begin{aligned}\tilde{D}_\infty =&\langle Y^c\rangle_\infty + \sum_{s\leq\infty} \frac{\widetilde{(\Delta Y_s)^2}}{1+|\Delta Y_s|} + \sum_{s<\infty}\frac{\widetilde{(\Delta^+ Y_s)^2}}{1+|\Delta^+ Y_s|}\\ \leq&\langle Y^c\rangle_\infty + \sum_{s\leq\infty}\widetilde{(\Delta Y_s)^q} + \sum_{s<\infty}\widetilde{(\Delta^+ Y_s)^q},\end{aligned}$$

where Y as defined in (6.4). Thus, from (6.8) it follows that

$$\tilde{D}_\infty \leq \int_{]0,\infty]} \frac{d\langle N^c\rangle_s}{(1+A_s)^2} + \int_{]0,\infty]}\int_{\mathbb{R}_0} |1+A_s|^{-q}|x|^q d\nu_s^r + \int_{[0,\infty[}\int_{\mathbb{R}_0} |1+A_{s+}|^{-q}|x|^q d\nu_{s+}^g < \infty.$$

By Lemma 6.1.1 and Lemma 6.1.2, $A_t^{-1}N_t \to 0$ a.s. as $t \to \infty$. ■

Theorem 6.2.2 *Suppose for the model (6.1) that $F_\infty = \infty$ and for some $q \in [1,2]$*

$$\int_{]0,\infty]} \frac{(f_s^r)^2 d\langle M^c\rangle_s}{(1+F_s)^2} + \int_{]0,\infty]}\int_{\mathbb{R}_0} |1+F_s^r|^{-q}|f_s^r|^q|x|^q d\nu_s^r + \int_{[0,\infty[}\int_{\mathbb{R}_0} |1+F_s^g|^{-q}|f_s^g|^q|x|^q d\nu_{s+}^g < \infty. \quad (6.9)$$

Then $\theta_t \to \theta$ (a.s.) as $t \to \infty$.

Proof. It is sufficient to note that

$$\theta_t - \theta = A_t^{-1} N_t,$$

where $A_t := F_t$ and $N_t := f \circ M_t$. By Theorem 6.2.1 we get immediately the statement of the theorem. ■

6.3 Sequential LS-estimators

Let us consider the model (6.1) with $M \in \mathcal{M}^2_{loc}(\mathbb{R})$. We assume that there exists a non-negative random variable ξ such that

$$\frac{d\langle M\rangle_t}{da_t} \leq \xi, \qquad f^2 \circ a_t \in \mathcal{A}^+_{loc} \cap \mathcal{P}_s. \tag{6.10}$$

Next, for fixed H we define

$$\tau_H = \inf\{t : f^2 \circ a_t \geq H\}, \tag{6.11}$$

with $\tau_H = \infty$ if the corresponding set is empty. We assumed that processes $f^r \in \mathcal{P}$, $f^g \in \mathcal{O}$ and $a \in \mathcal{P}_s$, consequently, by Theorem 2.4.14 in [4] τ_H is a wide sense stopping time.

On the set $\{\tau_H < \infty\}$ we define a random variable β_H by the relation

$$f^2 \circ a_{\tau_H -} + \beta_H \left((f^r_{\tau_H})^2 \Delta a_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ a_{\tau_H} \right) = H, \tag{6.12}$$

and on $\tau_H = \infty$ we put $\beta_H = 0$. Then $\beta_H \in [0,1]$ and is a $\mathcal{F}_{\tau_H}$-measurable random variable.

We consider the following statistic as an estimator of θ

$$\hat{\theta}_H = H^{-1} \left[f^2 \circ X_{\tau_H -} + \beta_H \left((f^r_{\tau_H})^2 \Delta X_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ X_{\tau_H} \right) \right]. \tag{6.13}$$

The next theorem shows that the statistic defined by means of (6.10)-(6.13) is an unbiased estimator of θ and has the property of guaranteed accuracy, i.e., bounded variance.

Theorem 6.3.1 *Suppose that assumptions* (6.10) *hold,* $\mathbf{E}\xi < \infty$, *and*

$$\mathbf{P}\{f^2 \circ a_\infty = \infty\} = 1. \tag{6.14}$$

Then for all $H > 0$

$$\mathbf{P}\{\tau_H < \infty\} = 1, \quad \mathbf{E}\hat{\theta}_H = \theta, \qquad \mathbf{Var}\hat{\theta}_H \le H^{-1}\mathbf{E}\xi.$$

Proof. First,

$$\begin{aligned}
\mathbf{P}\{\tau_H = \infty\} &= \mathbf{P}\{f^2 \circ a_\infty < H\} \\
&= 1 - \mathbf{P}\{f^2 \circ a_\infty \ge H\} \\
&\le 1 - \mathbf{P}\{f^2 \circ a_\infty = \infty\} \\
&= 1 - 1 = 0.
\end{aligned}$$

Thus,

$$\mathbf{P}\{\tau_H < \infty\} = 1 - \mathbf{P}\{\tau_H = \infty\} = 1.$$

Next, using (6.1) and (6.13), we obtain

$$\begin{aligned}
\hat{\theta}_H =& H^{-1}\left[f^2 \circ X_{\tau_H -} + \beta_H\left((f^r_{\tau_H})^2 \Delta X_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ X_{\tau_H}\right)\right] \\
=& H^{-1}\big[(f^2 \circ a_{\tau_H -})\theta + \beta_H\left((f^r_{\tau_H})^2 \Delta a_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ a_{\tau_H}\right)\theta \\
& + f^2 \circ M_{\tau_H -} + \beta_H\left((f^r_{\tau_H})^2 \Delta M_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ M_{\tau_H}\right)\big] \\
=& \theta + H^{-1} N_{\tau_H},
\end{aligned}$$

where

$$N_t = \mathbf{1}_{\{t<\tau_H\}} f^2 \circ M_t + \mathbf{1}_{\{t=\tau_H\}}\beta_H\left((f^r_{\tau_H})^2 \Delta M_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ M_{\tau_H}\right)$$

Since the process N_t is a stochastic integral with respect to the optional square integrable local martingale M, by the properties of optional stochastic integrals we have

$$\langle N\rangle_t = \mathbf{1}_{\{t<\tau_H\}} f^2 \circ \langle M\rangle_t + \mathbf{1}_{\{t=\tau_H\}} \beta_H^2 \left((f^r_{\tau_H})^2 \Delta \langle M\rangle_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ \langle M\rangle_{\tau_H} \right).$$

Hence by (6.10) and (6.12) we get

$$\langle N\rangle_{\tau_H} \leq \xi \left[f^2 \circ a_{\tau_H -} + \beta_H \left((f^r_{\tau_H})^2 \Delta a_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ a_{\tau_H} \right) \right] = \xi H.$$

Consequently, $N_{t\wedge\tau_h}$ is an optional square integrable martingale, and therefore

$$\mathbf{E} N_{\tau_H} = 0, \quad \mathbf{E} N^2_{\tau_H} \leq H \mathbf{E} \xi,$$

which proves the theorem. ■

Now, it is reasonable to discuss the following problem of distinguishing two hypotheses with simultaneous estimation of the parameter $\theta \in \mathbb{R}$:

$$H_0 : \; X_t = f \circ a_t \theta + M_t, \tag{6.15}$$

$$H_1 : \; X_t = M_t, \tag{6.16}$$

where $M \in \mathcal{M}^2_{loc}(\mathbb{R})$. Assuming that (6.10) is fulfilled, we define $\tau_H, \; \beta_H$ as in (6.11), (6.12) and

$$\phi_H(X) = H^{-1} \left[f^2 \circ X_{\tau_H -} + \beta_H \left((f^r_{\tau_H})^2 \Delta X_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ X_{\tau_H} \right) \right].$$

Theorem 6.3.2 *Suppose that in the problem* (6.15) *the parameter satisfies*

$$\theta \in \; R_\delta = \{\theta \in \mathbb{R} : |\theta| \geq \delta > 0\}$$

and condition (6.14) *is fulfilled for both hypotheses* H_0 *and* H_1, $\mathbf{E}\xi < \infty$. *Then for a given* $\epsilon > 0$ *the criterion*

$$\Delta(\tau_H) = \begin{cases} 0 \text{ if } |\phi_H(X)| \geq \delta/2, \\ 1 \text{ if } |\phi_H(X)| < \delta/2, \end{cases}$$

for $H \geq 4\delta^{-2}\epsilon^{-1}\mathbf{E}\xi$ *ensures distinguishability of the hypotheses* H_0 *and* H_1 *with probabilities of errors not exceeding* $\epsilon > 0$.

Proof. From the definition of $\phi_H(X)$, if either H_0 or H_1 is true, then it follows, respectively, that

$$\phi_H(X) = \theta + \phi_H(M) \quad \text{or} \quad \phi_H(X) = \phi_H(M).$$

Note that by Theorem 6.3.1 under either of the hypotheses H_0, H_1

$$\mathbf{E}\phi_H(M) = 0, \quad \mathbf{E}\phi_H^2(M) \leq H^{-1}\mathbf{E}\xi. \tag{6.17}$$

Then in case when H_0 is true, applying Chebyshev's inequality we obtain that for $4\delta^{-2}\epsilon^{-1}\mathbf{E}\xi \leq H$

$$\begin{aligned} \mathbf{P}\{\omega : \Delta(\tau_H) \neq 1\} =& \mathbf{P}\{\omega : |\phi_H(X)| \geq \delta/2\} \\ =& \mathbf{P}\{\omega : |\phi_H(M)| \geq \delta/2\| \leq 4\delta^{-2}\mathbf{E}|\phi_H(M)|^2 \\ \leq& 4\delta^{-2}H^{-1}\mathbf{E}\xi \leq \epsilon. \end{aligned}$$

In case when H_1 is true, using the simple fact that for $\theta \in R_\delta$

$$\delta - |\phi_H(M)| \leq |\delta + \phi_H(M)| \leq |\theta + \phi_H(M)| = |\phi_H(X)|$$

implying that

$$\{\omega : |\phi_H(X)| < \delta/2\} \subseteq \{\omega : |\phi_H(M)| \geq \delta/2\},$$

we arrive to the following estimate of the probability of error:

$$\begin{aligned}\sup_{\theta\in R_\delta}\mathbf{P}\{\omega:\Delta(\tau_H)\neq 0\} =&\sup_{\theta\in R_\delta}\mathbf{P}\{\omega:|\phi_H(X)|<\delta/2\}\\ \le&\sup_{\theta\in R_\delta}\mathbf{P}\{\omega:|\phi_H(M)|<\delta/2\}\\ \le& 4\delta^{-2}\sup_{\theta}\mathbf{E}\phi_H^2(M)\le 4\delta^{-2}H^{-1}\mathbf{E}\xi\le\epsilon.\end{aligned}$$

■

6.4 Further extensions and examples

Let us show how the linear optional regression model (6.1) can be extended in a non-linear case (see, for example, [18]). The non-linear optional regression model has the following form

$$X_t = f\circ a_t g(\theta) + M_t, \tag{6.18}$$

where the processes f, a, M satisfy the same assumptions as in (6.10)-(6.12) and $g:\mathbb{R}\to\mathbb{R}$ is a continuous, bijective function with a continuous inverse g^{-1}. Let τ_H and β_H be as in (6.11) and (6.12), respectively. Then the sequential LS-estimator can be obtained by defining $\zeta := g(\theta)$ and realizing from (6.13) that

$$\begin{aligned}&\hat\zeta_H = g(\hat\theta_H) = H^{-1}\left[f^2\circ X_{\tau_H-}+\beta_H\left((f^r_{\tau_H})^2\Delta X_{\tau_H}+(f^g_{\tau_H})^2\Delta^+X_{\tau_H}\right)\right]\ \text{ or}\\ &\hat\theta_H = g^{-1}\left(H^{-1}\left[f^2\circ X_{\tau_H-}+\beta_H\left((f^r_{\tau_H})^2\Delta X_{\tau_H}+(f^g_{\tau_H})^2\Delta^+X_{\tau_H}\right)\right]\right).\end{aligned}\tag{6.19}$$

Now, suppose (6.14) holds, $g(\theta)$ is differentiable and

$$\int_{-\infty}^{\infty}(g^{-1}(x))^2\exp(-x^2/2)dx<\infty. \tag{6.20}$$

Using the same argument as in the proof of Theorem 6.3.1, we show that $\mathbf{P}\{\tau_H = \infty\} = 1$.

Next, note that

$$\mathbf{E}(\theta_H - \theta) = \mathbf{E}\left[g^{-1}(g(\theta) + H^{-1}N_{\tau_H}) - g^{-1}(g(\theta))\right],$$
$$\mathbf{E}(\theta_H - \theta)^2 = \mathbf{E}\left[g^{-1}(g(\theta) + H^{-1}N_{\tau_H}) - g^{-1}(g(\theta))\right]^2,$$

where

$$N_t = \mathbf{1}_{\{t<\tau_H\}} f^2 \circ M_t + \mathbf{1}_{\{t=\tau_H\}} \beta_H \left((f^r_{\tau_H})^2 \Delta M_{\tau_H} + (f^r_{\tau_H})^2 \Delta^+ M_{\tau_H}\right).$$

From the proof of Theorem 6.3.1 we already know that

$$N_{t\wedge\tau} \in \mathcal{M} \cap \mathcal{M}^2, \ \mathbf{E}N_{\tau_H} = 0 \text{ and } \mathbf{E}N^2_{\tau_H} \leq H\mathbf{E}\xi.$$

From assumption (6.14) it follows that $\langle N\rangle_\infty = \infty$. Thus, by Theorem 6.2.1 we have

$$\lim_{t\to\infty} \frac{N_t}{\langle N\rangle_t} = 0 \ (a.s.) \text{ and } \lim_{H\to\infty} \frac{N_{\tau_H}}{\langle N\rangle_{\tau_H}} = 0 \ (a.s.).$$

Since $\langle N\rangle_{\tau_H} \leq H\xi$, we get

$$\lim_{H\to\infty} \frac{N_{\tau_H}}{H} = 0 \ (a.s.).$$

Using the Skorokhod embedding theorem we obtain a

$$\begin{aligned}\mathbf{E}(\theta_H - \theta) &= \int_{-\infty}^{\infty} \left[g^{-1}(g(\theta) + H^{-1}x) - g^{-1}(g(\theta))\right] e^{-x^2/(2H)} dx \\ &= \int_{y\leq A} \left[g^{-1}(g(\theta) + H^{-1/2}y) - g^{-1}(g(\theta))\right] e^{-y^2/2} dy \\ &\quad + \int_{y>A} \left[g^{-1}(g(\theta) + H^{-1/2}y) - g^{-1}(g(\theta))\right] e^{-y^2/2} dy\end{aligned}$$

Applying condition (6.14) and (6.20) one can always choose numbers $A_0(\epsilon)$ and $H_0(\epsilon)$ such that for $H \geq H_0(\epsilon)$

$$\int_{y\leq A_0(\epsilon)} \left[g^{-1}(g(\theta) + H^{-1/2}y) - g^{-1}(g(\theta))\right] e^{-y^2/2} dy < \epsilon$$

and

$$\int_{y>A_0(\epsilon)} \left[g^{-1}(g(\theta)+H^{-1/2}y)-g^{-1}(g(\theta))\right] e^{-y^2/2}dy<\epsilon.$$

Hence, $\mathbf{E}(\theta_H-\theta)<2\epsilon$ for $H>H_0(\epsilon)$. This proves θ_H is asymptotically unbiased as $H\to\infty$.

It is not difficult to show similar calculations for $\mathbf{E}(\theta_{\tau_H}-\theta)^2$. From Cramer-Rao-Wolfovitz inequality it follows that θ_{τ_H} is asymptotically efficient.

On the other hand, we can show that θ_H is unbiased and efficient estimator under assumption that $g^{-1}(\theta)$ is differentiable and has a bounded first derivative. Using Mean Value theorem, we have

$$\begin{aligned}\mathbf{E}|\theta_H-\theta| =&\mathbf{E}\left[g^{-1}(g(\theta)+H^{-1}N_{\tau_H})-g^{-1}(g(\theta))\right]\\ \le&\mathbf{E}\sup_{\zeta\in[g(\theta),g(\theta)+H^{-1}N_{\tau_H}]}|(g^{-1})'(\zeta)|\frac{N_{\tau_H}}{H}\\ \le&0\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbf{E}|\theta_H-\theta|^2 \le&\mathbf{E}\left[\sup_{\zeta\in[g(\theta),g(\theta)+H^{-1}N_{\tau_H}]}|(g^{-1})'(\zeta)|\right]^2\frac{N_{\tau_H}^2}{H^2}\\ \le&K^2H^{-1}\mathbf{E}\xi,\end{aligned}$$

where K is the constant bound on $(g^{-1})'(\xi)$.

Let us now illustrate several examples.

Example 6.4.1 ***Non-linear regression model.*** *Consider the following non-linear model*

$$X_t=f\circ a_t\sqrt{\theta}+M_t,$$

where f,a,M satisfy assumption (6.10)-(6.12),(6.14).

The function $g(\theta) = \sqrt{\theta}, \theta \geq 0$, *is differentiable and its inverse* $g^{-1}(\theta) = \theta^2$ *clearly satisfies* (6.20)*:*

$$\int_{-\infty}^{\infty} (g^{-1}(x))^2 \exp(-x^2/2)dx = \int_{-\infty}^{\infty} x^4 \exp(-x^2/2)dx = 3\sqrt{2\pi} < \infty.$$

Thus, by the above discussion the sequential LS-estimator

$$\hat{\theta}_H = \left(H^{-1}\left[f^2 \circ X_{\tau_H -} + \beta_H \left((f^r_{\tau_H})^2 \Delta X_{\tau_H} + (f^g_{\tau_H})^2 \Delta^+ X_{\tau_H}\right)\right]\right)^2$$

is both asymptotically unbiased and asymptotically efficient as $H \to \infty$.

Example 6.4.2 ***Risk process.*** *Consider the following risk process*

$$X_t = ct + \sigma W_t - aN^r_t + bN^g_t, \tag{6.21}$$

where c, σ, a, b *are positive constants,* W *is a Wiener process,* N^r *and* N^g *are a Poisson process and left-continuous modification of a Poisson process, respectively. The constant* c *usually represents premium payments in risk theory, whereas* a *and* b *describe average value of claims and positive gains, respectively. The process* σW_t *is a random perturbation.*

We can rewrite the process X_t *as follows*

$$X_t = \theta t + M_t,$$

where $\theta := c - a\lambda^r + b\lambda^g$, $M_t := \sigma W_t - a(N^r_t - \lambda^r t) + b(N^g_t - \lambda^g t)$, λ^r *and* λ^g *are jump intensities of* N^r_t *and* N^g_t*, respectively.*

The structural LS estimator of θ *is*

$$\theta_t = \frac{X_t}{t}. \tag{6.22}$$

The condition (6.9) *of Theorem 3.2, i.e.,*

$$\sigma^2 \int_{]0,\infty]} \frac{ds}{(1+s)^2} + a^q \lambda^r \int_{]0,\infty]} (1+s)^{-q} ds < \infty,$$

holds for any $q \in (1,2]$. Thus, the estimator (6.22) is strongly consistent.

The sequential LS-estimators have the following form

$$\hat{\theta}_H = \frac{X_{\tau_H}}{H}.$$

All assumptions of Theorem 6.3.1 are obviously satisfied, and $\xi = \sigma^2 + a^2\lambda^r + b^2\lambda^g$. Consequently,

$$\mathbf{P}\{\tau_H < \infty\} = 1, \quad \mathbf{E}\hat{\theta}_H = \theta, \quad \mathbf{Var}\hat{\theta}_H \leq \frac{\sigma^2 + a^2\lambda^r + b^2\lambda^g}{H}.$$

We usually want the process X_t to be positive, so the estimator $\theta_t = c - a\lambda^r + b\lambda^g > 0$. This assertion is called as a net profit condition in risk theory.

Example 6.4.3 ***Ornstein–Uhlenbeck process.*** *In mathematical finance we often deal with Ornstein–Uhlenbeck type processes that possess mean reversion property, i.e.*

$$X_t = \int_0^t (\mu - X_{s-}) ds\theta + M_t,$$

where μ is a positive constant, and $M \in \mathcal{M}^2_{loc}$.

We assume that

$$F_t := \int_0^t (\mu - X_{s-})^2 ds \in \mathcal{A}^+_{loc} \cap \mathcal{P}_s \quad \text{and} \quad \frac{d\langle M \rangle_t}{dt} \leq \xi. \tag{6.23}$$

Then, the structural LS estimator of θ is

$$\theta_t = \frac{(\mu - X_-) \circ X_t}{\int_0^t (\mu - X_{s-})^2 ds}, \tag{6.24}$$

and sequential LS-estimators have the following form

$$\hat{\theta}_H = H^{-1}\left[(\mu - X_{s-})^2 \circ X_{\tau_H -}\right].$$

If the following condition

$$\int_{]0,\infty]} \frac{(\mu - X_{s-})^2 d\langle M^c \rangle_s}{(1+F_s)^2} + \int_{]0,\infty]} \int_{\mathbb{R}_0} |1 + F^r_s|^{-q} |\mu - X_{s-}|^q |x|^q d\nu^r_s < \infty$$

holds for some $q \in [1,2]$ and $F_\infty = \infty$, then by Theorem 3.2 the estimator (6.24) *is strongly consistent.*

Furthermore, by Theorem 6.3.1, we have

$$\mathbf{P}\{\tau_H < \infty\} = 1, \quad \mathbf{E}\hat{\theta}_H = \theta, \quad \mathbf{Var}\hat{\theta}_H \le H^{-1}\mathbf{E}\xi.$$

Example 6.4.4 *Finally, consider a regression model with well-known centered Gaussian martingale $M \in \mathcal{M}^2_{loc}$ and a deterministic function f_t,*

$$X_t = \int_0^t f_s ds\theta + M_t. \tag{6.25}$$

Then, the LS estimator of θ is

$$\theta_t = \frac{f \circ X_t}{\int_0^t f_s^2 ds}, \tag{6.26}$$

It can be shown in the same way as in [71] that strong consistency of (6.26) *follows only from the assumption of $\int_0^\infty f_s^2 ds = \infty$.*

We assume that

$$\frac{d\langle M\rangle_t}{dt} \le \xi, \tag{6.27}$$

where ξ is constant. Note that in case of centered Gaussian martingales $\langle M\rangle_t = \mathbf{E}M_t^2 < \infty$ is a deterministic function. Then sequential LS-estimators have the following form

$$\hat{\theta}_H = H^{-1}\left[f^2 \circ X_{\tau_H -}\right].$$

If, in addition,

$$\mathbf{P}\left\{\int_0^\infty f_s^2 ds = \infty\right\} = 1,$$

then, by Theorem 6.3.1, we have

$$\mathbf{P}\{\tau_H < \infty\} = 1, \quad \mathbf{E}\hat{\theta}_H = \theta, \quad \mathbf{Var}\hat{\theta}_H \le H^{-1}\xi.$$

6.5 Asymptotic behaviour of the trajectories of weighted LS-estimates

In this section we will need the following theorem of law of iterated logarithms in case of optional martingales.

Theorem 6.5.1 *(see [36]) Let* $N \in \mathcal{M}^2_{loc}$ *and* $\mathbf{E}\sup_t |\Delta X_t| < \infty$, $\mathbf{E}\sup_t |\Delta^+ X_t| < \infty$. *Then on the set* $(\langle N\rangle_\infty = \infty)$

$$\limsup_t \frac{N_t}{(2\langle N\rangle_t loglog\langle N\rangle_t)^{1/2}} \leq 1 \quad a.s.$$

Let us now consider a version of the model (6.1) where $a_t = \langle M\rangle_t$, i.e.

$$X_t = f \circ \langle M\rangle_t + M_t,$$

where $M \in \mathcal{M}^2_{loc}$.

Next, we define

$$Y_t := \theta_t - \theta = \frac{N_t}{\langle N\rangle_t},$$

where $N_t = f \circ M_t$.

Theorem 6.5.2 *Suppose that the following conditions are fulfilled:*

$$\mathbf{E}\sup_t |\Delta X_t| < \infty,\ \mathbf{E}\sup_t |\Delta^+ X_t| < \infty,$$

$$\langle N\rangle_\infty = \infty.$$

Then

$$\limsup_t \frac{|Y_t|\langle N\rangle_t^{1/2}}{(2\log\log\langle N\rangle_t)^{1/2}} \leq 1 \quad a.s.$$

Proof. Multiplying by $\langle N\rangle_t^{1/2}$ and dividing by $(2\log\log\langle N\rangle_t)^{1/2^{1/2}}$, we get

$$\frac{|Y_t|\langle N\rangle_t^{1/2}}{(2\log\log\langle N\rangle_t)^{1/2}} = \frac{N_t}{(2\langle N\rangle_t loglog\langle N\rangle_t)^{1/2}}.$$

Applying directly Theorem 6.5.1, the statement of the theorem follows. ■

References

[65] R. S. Liptser and A. N. Shiryaev, Theory of Martingales. Springer Netherlands, 1989.

[71] A. V. Mel'nikov, "Stochastic differential equations: Singularity of coefficients,regression models, and stochastic approximation,"Russian Mathematical Surveys, vol. 51, no. 5, p. 819, 1996.

[4] M. Abdelghani and A. V. Melnikov, Optional Processes. Chapman & Hall/CRC,2020.

[70] A. V. Mel'nikov, "On regression models with non-square integrable martingale-like errors," Publications mathématiques et informatique de Rennes, no. 1, pp. 97–107, 1987.

[72] A. V. Mel'nikov and A. A. Novikov, "Sequential inferences with guaranteed accuracy for semimartingales,"Prob. Theory Appl., vol. 33, pp. 480–494, 1988.

[34] L. Galtchouk and V. Konev, "On sequential estimation of parameters in semimartingale regression models with continuous time parameter," Annals of statistics, vol. 29, pp. 1508–1536, 2001.

[19] J. Bishwal, Parameter Estimation in Stochastic Differential Equations. Springer,Jan. 2008, vol. 1923, isbn: 978-3-540-74447-4. doi:10.1007/978-3-540-74448-1.

[66] R. S. Liptser, “A strong law of large numbers for local martingales,” Stochastics, vol. 3, no. 1-4, pp. 217–228, 1980.

[35] K. Gasparyan, “On asymptotic behavior of optional martingales,”Dokl. Nats. Akad. Nauk Armen., vol. 93, no. 5, pp. 212–219, 1992.

[5] M. Abdelghani, A. V. Melnikov, and A. A. Pak, “An optional semimartingale approach to risk theory,” Submitted, 2020.

[18] J. Bishwal, “Sequential maximum likelihood estimation in semimartingales,” Journal of Statistics and Applications, vol. 1, no. 2-4, pp. 15–25, 2006.

Chapter 7

Multi-dimensional optional regression models

Estimation of regression parameters is a classical statistical problem which has various applications in natural and social sciences. In classical statistics this problem was studied for regression models with discrete time. With the development of theory of stochastic processes it obtained a new formulation as a problem of parameter estimation in random processes and became one of the cornerstones in the statistics of random processes. The book of [67] brilliantly demonstrates this for the models with diffusion type processes.

The development of general theory of random processes and theory of martingales in particular opened new ways for construction of more general models embedding discrete time and continuous time models at the same time. Apparently, works of [1], [2] were the first of this kind. Further, a series of other works gave a new impulse to this direction by consistently examining regression models in the form of cadlag semimartingales, where the drift depends on an unknown parameter and the martingale part represents the errors in observations. The parameter estimation was usually implemented by utilizing least squares (LS) method (see, for example, [77], [22], [70], [60]). Strong consistency and asymptotic normality of the LS-estimates under very general conditions were established. Moreover, a mean-square guaranteed accuracy was proved for sequential LS-estimates (see [72]). The detailed review of the results of this theory for that period can be found in [71], [73]. Furthermore, in the work of [34] sequential estimation with prescribed accuracy was extended to the multivariate semimartingale regression models with out prior existed restrictions on the number of unknown parameters and the dimension of the observation process.

We also would like to point out that this area attracts an ongoing research interest, especially, from the point of view of its applications in finance, econometrics etc (see, for example, [88], [64]). Modern works on parameter estimation seem to focus on non-semimartingale models driven by fractional Brownian Motion (see, for example, [51]) and on special cases of semimartingale models (see, for example, [19], [20], [92]). The recent review of the historical development in sequential prescribed precision estimation can be found in [86, Section 4].

Martingale theory is well-developed under so-called "usual conditions", when filtration (information flow) is complete and right-continuous. However, statistical data and information are usually delivered by a stochastic process, whose history (natural filtration) may not be right-continuous, and therefore such technical conditions may not be fulfilled (see [4]). This is the main reason why we need to consider regression models in more general setting which we call here the optional semimartingale regression model. Optional semimartingales, on which our optional regression model is based, admit trajectories which are not right-continuous and arise when "usual conditions" are not assumed on filtered probability space. Up to our knowledge, currently there are no works devoted to the relaxing of these "usual conditions" and investigation of such general optional regression model.

The current chapter is a natural extension of [9] in which the authors studied LS estimates and their sequential versions for a 1-dimensional optional regression model. The first goal of this chapter is to derive the weighted LSE for a multivariate optional regression model and prove strong consistency of this estimator under conditions on regressors similar to ones given in [71]. The second goal is to investigate sequential LS-estimates in the multivariate optional semimartingales regression model. In the case, when the dimension of unknown parameter does not exceed the dimension of observable process, we adopt the approach proposed in [72]. In general case, when such restriction is not assumed, we solve the problem with the help of two-step procedure proposed by [34].

This chapter is organized in the following way. In Section 6.1, we introduce the optional semimartingale model. In Section 6.2, we derive the weighted LSE for this model. In Section 6.3, we prove strong consistency of the weighted LSE by applying SLLN. Section 6.4 presents the construction of the unbiased fixed accuracy estimators for multivariate optional regression models with the number of parameters less than or equal to the dimension of the observation process. The two-step sequential estimation procedure for the general case with an arbitrary number of parameters is studied in Section 6.5.

7.1 Model

Let $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ be a filtered probability space not necessarily satisfying the usual conditions: the filtration $\mathbf{F} = (\mathcal{F}_t)_{t\geq 0}$ is right-continuous, that is $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$, and the σ-algebra $\mathcal{F}_0$ contains all $\mathbf{P}$-null sets. Consider the observation process $X = (X_t)_{t\geq 0}$ specified by the stochastic regression model

$$X_t = X_0 + \Phi' \circ a_t\theta + m_t, \quad t \geq 0, \tag{7.1}$$

where the optional stochastic integral

$$\Phi' \circ a_t = \int_{]0,t]} (\Phi^r_s)' da^r_s + \int_{[0,t[} (\Phi^g_s)' da^g_{s+},$$

the prime denotes transposition; $X,\ m \in \mathbb{R}^n$; Φ^r and Φ^g are, respectively, predictable and optional $p \times n$ matrix (the matrix of stochastic regressors); $\theta \in \mathbb{R}^p$ is the vector of unknown parameters; $m_t = (m^1_t, \dots, m^n_t)', m_0 = 0$, is a noise which is a locally square integrable optional martingale with the trajectories having right and left limits but not necessarily righ- or left-continuous:

$$a_t = \langle m \rangle_t = tr(\langle m^i, m^j \rangle_t)_{1 \le i,j \le n} = \sum_{i=1}^n \langle m^i \rangle_t.$$

Let

$$B_t = \left(\frac{d\langle m^i, m^j \rangle_t}{da_t} \right)_{1 \le i,j \le n}, \qquad B^c_t = \left(\frac{d\langle m^{ic}, m^{jc} \rangle_t}{d\langle m^c \rangle_t} \right)_{1 \le i,j \le n},$$
$$B^d_t = \left(\frac{d\langle m^{id}, m^{jd} \rangle_t}{d\langle m^d \rangle_t} \right)_{1 \le i,j \le n}, \qquad B^g_t = \left(\frac{d\langle m^{ig}, m^{jg} \rangle_t}{d\langle m^g \rangle_t} \right)_{1 \le i,j \le n},$$

$$\langle m^c \rangle_t = \sum_{i=1}^n \langle m^{ic} \rangle_t, \qquad \langle m^d \rangle_t = \sum_{i=1}^n \langle m^{id} \rangle_t, \qquad \langle m^g \rangle_t = \sum_{i=1}^n \langle m^{ig} \rangle_t;$$

here $m = m^c + m^d + m^g$ is the orthogonal decomposition of the vector-valued martingale m into a continuous optional martingale, a right-continuous and left-continuous martingale parts. Note that $\langle m^{ic}, m^{jd} \rangle = \langle m^{ig}, m^{jd} \rangle = \langle m^{ic}, m^{jg} \rangle = 0$, for all $1 \le i, j \le n$,

$$\begin{aligned} a_t =& \langle m^c \rangle_t + \langle m^d \rangle_t + \langle m^g \rangle_t, \\ B_t d\langle m \rangle_t =& B^c_t d\langle m^c \rangle_t + B^d_t d\langle m^d \rangle_t + B^g_t d\langle m^g \rangle_t. \end{aligned}$$

Example 7.1.1 *Consider a special case of model* (7.1)*:*

$$X^i_t = \theta k t + m^i_t, \tag{7.2}$$

$$m^i_t = \sigma^i W^i_t + a^i(N^{ir}_t - \lambda^{ir} t) - b^i(N^{ig}_t - \lambda^{ig} t), \quad i = 1, \dots, n. \tag{7.3}$$

where $\sigma^i, a^i, b^i, \lambda^{ir}, \lambda^{ig}$ are constants, $k = \sum_{i=1}^n k^i,\ k^i := (\sigma^i)^2 + (a^i)^2\lambda^{ir} + (b^i)^2\lambda^{ig},\ W_t^i$ is a 1-dimensional Wiener process, N^{ir} is a Poisson process and N^{ig} is a left-continuous modification of a Poisson process, λ^{ir} and λ^{ig} are corresponding intensities of N^{ir} and N^{ig}. The processes W^i, N^{id} and N^{ig} are all mutually independent.

Here, m_t^i is an optional martingale with an orthogonal decomposition, $m_t^i = m^{ic} + m^{id} + m^{ig},\ m^{ic} = \sigma^i W_t^i,\ m^{id} = a^i(N_t^{ir} - \lambda^{ir}t),\ m^{ig} = b^i(N_t^{ig} - \lambda^{ig}t)$.

In this example, the process $\Phi = (1, \dots, 1)$ in (7.1) is an n-dimensional row vector of ones, the process

$$a_t \;=\; \sum_{i=1}^n \langle m^{ic}\rangle_t \,+\, \langle m^{id}\rangle_t \,+\, \langle m^{ig}\rangle_t \;=\; \sum_{i=1}^n \left[(\sigma^i)^2 + (a^i)^2\lambda^{ir} + (b^i)^2\lambda^{ig}\right] t \;=\; kt,$$

and the process B_t is a matrix with fixed values $(\frac{k^i}{k})_{i=1,\dots,n}$ on its main diagonal. In addition, we assume that B_t is invertible, and $\beta := \sum_{i,j=1}^n \beta^{ij} > 0$, where β^{ij} is a value from the inverse matrix $(B_t)^{-1}$.

The problem of estimation of θ in the model (7.2) can be important when one wants to estimate a cumulative drift of a portfolio of financial assets (index) modelled by the process (7.2), e.g., what is the main trend of a portfolio of some stocks. Another application can be the estimation of the drift parameter of model (7.2) when X_t describes capital processes of some firms (see Chapter 4). For instance, X_t^i can be a risk process of an insurance company where θkt describes a stable flow of income payments while N_t^r describes an accumulation of claims. For simple exposition of further results, we assumed constant coefficients in (7.2)-(7.3), though it is, of course, possible to consider deterministic or stochastic coefficients.

Example 7.1.2 *Consider the model (7.1) where the process $\Phi_t = (\Phi_t^r, \Phi_t^g)$ is deterministic and m_t is an optional Gaussian martingale. We know that the distribution of m_t is fully characterized by its covariance function which is a deterministic func-*

tion. We would like to emphasize optionality of m_t by pointing out that its covariance function has right and left jumps. As a result, the process a_t is a deterministic process which has jumps Δa_t and $\Delta^+ a_t$.

7.2 Weighted LSE for optional semimartingale models in continuous time

Let $(X_t)_{t\geq 0}$ be an observation process starting from X_0 and specified by the stochastic differential equation (7.1). Let $\widetilde{W} = (\widetilde{W}_t^r, \widetilde{W}_t^g),\ t \geq 0$, $\widetilde{W}_t^r \in \mathcal{P}$ and $\widetilde{W}_t^g \in \mathcal{O}$ be some symmetric positive definite weight matrices of size $n \times n$. Let $a = a^c + a^d + a^g$ be the decomposition of the increasing process $a \in \mathcal{P}_s$ into continuous and purely discontinuous parts:

$$a_t^d = \sum_{0<s\leq t} \Delta a_s, \quad a_t^g = \sum_{0\leq s<t} \Delta^+ a_s, \quad a_t^c = a_t - a_t^d - a_t^g,$$
$$\Delta a_s = a_s - a_{s-}, \quad \Delta^+ a_s = a_{s+} - a_s.$$

Introduce a loss function

$$\begin{aligned} L_\delta(\theta) = &\sum_{t_k<T} \left(\Delta \widetilde{X}_{t_k}^d - (\Phi_{t_k}^r)'\theta \Delta a_{t_k}^d\right)' \widetilde{W}_{t_k}^r \left(\Delta \widetilde{X}_{t_k}^d - (\Phi_{t_k}^r)'\Delta a_{t_k}^d\right) \\ &+ \sum_{t_k\leq T} \left(\Delta^+ \widetilde{X}_{t_k}^g - (\Phi_{t_k}^g)'\theta \Delta a_{t_k}^g\right)' \widetilde{W}_{t_k}^g \left(\Delta \widetilde{X}_{t_k}^g - (\Phi_{t_k}^g)'\Delta a_{t_k}^g\right) \\ &+ \sum_{t_k<T} (\Delta a_{t_k}^c)^{\oplus} \left(\Delta \widetilde{X}_{t_k}^c - (\Phi_{t_k}^r)'\theta \Delta a_{t_k}^c\right)' \widetilde{W}_{t_k}^r \left(\Delta \widetilde{X}_{t_k}^c - (\Phi_{t_k}^r)'\Delta a_{t_k}^c\right), \end{aligned}$$

where $\delta = \{t_0, t_1, \ldots\},\ 0 = t_0 < t_1 < \ldots < t_n = T$, is some partition of the interval $[0, T]$;

$$x^{\oplus} = \begin{cases} x^{-1}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0; \end{cases}$$

$$
\begin{aligned}
\Delta\widetilde{X}^d_{t_k} &= \widetilde{X}^d_{t_{k+1}} - \widetilde{X}^d_{t_k}, & \Delta\widetilde{X}^g_{t_k} &= \widetilde{X}^g_{t_{k+1}} - \widetilde{X}^g_{t_k}, & \Delta\widetilde{X}^c_{t_k} &= \widetilde{X}^c_{t_{k+1}} - \widetilde{X}^c_{t_k}, \\
\widetilde{X}^d_t &= \int_{]0,t]} \mathbf{1}_{\{\Delta a_s \neq 0\}} dX^r_s, & \widetilde{X}^g_t &= \int_{]0,t]} \mathbf{1}_{\{\Delta^+ a_s \neq 0\}} dX^g_{s+}, & \widetilde{X}^c_t &= X^r_t - \widetilde{X}^d_t, \\
\Delta a^c_{t_k} &= a^c_{t_{k+1}} - a^c_{t_k}, & \Delta a^d_{t_k} &= a^d_{t_{k+1}} - a^d_{t_k}, & \Delta a^g_{t_k} &= a^g_{t_{k+1}} - a^g_{t_k};
\end{aligned}
$$

$\mathbf{1}_A$ is the indicator of a set A.

For the fixed partition δ and weight matrix $\widetilde{W}$ one can find an estimator $\hat{\theta}_\delta$ which minimizes the loss function $L_\delta(\theta)$. We have

$$
\begin{aligned}
\nabla_\theta L_\delta(\theta) = &- 2 \sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} \Delta\widetilde{X}^d_{t_k} \Delta a^d_{t_k} + 2 \sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} (\Phi^r_{t_k})' \left(\Delta a^d_{t_k}\right)^2 \theta \\
&- 2 \sum_{t_k \leq T} \Phi^g_{t_k} \widetilde{W}^g_{t_k} \Delta\widetilde{X}^g_{t_k} \Delta a^g_{t_k} + 2 \sum_{t_k < T} \Phi^g_{t_k} \widetilde{W}^g_{t_k} (\Phi^g_{t_k})' \left(\Delta a^g_{t_k}\right)^2 \theta \\
&- 2 \sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} \Delta\widetilde{X}^c_{t_k} \Delta a^c_{t_k} \left(a^c_{t_k}\right)^{\oplus} + 2 \sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} (\Phi^r_{t_k})' \Delta a^c_{t_k} \theta
\end{aligned}
$$

where ∇_θ is the gradient with respect to θ. The equation $\nabla_\theta L_\delta(\theta) = 0$ yields the estimator

$$
\begin{aligned}
\hat{\theta}_\delta = & \Bigg[\sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} (\Phi^r_{t_k})' \left(\Delta a^d_{t_k}\right)^2 + \sum_{t_k < T} \Phi^g_{t_k} \widetilde{W}^g_{t_k} (\Phi^g_{t_k})' \left(\Delta a^g_{t_k}\right)^2 \\
& + \sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} (\Phi^r_{t_k})' \Delta a^c_{t_k} \Bigg]^{-1} \\
& \times \Bigg[\sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} \Delta\widetilde{X}^d_{t_k} \Delta a^d_{t_k} + \sum_{t_k \leq T} \Phi^g_{t_k} \widetilde{W}^g_{t_k} \Delta\widetilde{X}^g_{t_k} \Delta a^g_{t_k} + \sum_{t_k < T} \Phi^r_{t_k} \widetilde{W}^r_{t_k} \Delta\widetilde{X}^c_{t_k} \Bigg].
\end{aligned}
$$

where we make use of the equality $\Delta\widetilde{X}^c_{t_k} \Delta a^c_{t_k} \left(a^c_{t_k}\right)^{\oplus} = \Delta\widetilde{X}^c_{t_k}$, which is true, because the process X^c does not change on constancy intervals of the process a^c (see [87]).

Taking a sequence of partitions $\delta_n = \{t^n_0, t^n_1, \ldots\}$, $0 = t^n_0 < t^n_1 < \ldots < t^n_n = T$, such that $\max_k \left(t^n_{k+1} - t^n_k\right) \to 0$ as $n \to \infty$ we obtain the following result:

Theorem 7.2.1 *Let, for all $t > 0$,*

$$\int_{]0,t]} tr\left[\Phi_s^r \widetilde{W}_s^r (\Phi_s^r)'\right]\left(\Delta a_s + \mathbf{1}_{\{\Delta a_s = 0\}}\right) da_s^r < \infty \qquad (a.s.),$$

$$\int_{]0,t]} tr\left[\Phi_s^r \widetilde{W}_s^r B_s^r \widetilde{W}_s^r (\Phi_s^r)'\right]\left((\Delta a_s)^2 + \mathbf{1}_{\{\Delta a_s = 0\}}\right) da_s^r < \infty \qquad (a.s.),$$

$$\int_{[0,t[} tr\left[\Phi_s^g \widetilde{W}_s^g (\Phi_s^g)'\right]\Delta^+ a_s da_{s+}^g < \infty \qquad (a.s.),$$

$$\int_{[0,t[} tr\left[\Phi_s^g \widetilde{W}_s^g B_s^g \widetilde{W}_s^g (\Phi_s^g)'\right](\Delta^+ a_s)^2 da_{s+}^g < \infty \qquad (a.s.),$$

and the matrix

$$\int_{]0,T]} \left[\Phi_s^r \widetilde{W}_s^r (\Phi_s^r)'\right]\left(\Delta a_s + \mathbf{1}_{\{\Delta a_s = 0\}}\right) da_s^r + \int_{[0,T[} \left[\Phi_s^g \widetilde{W}_s^g (\Phi_s^g)'\right]\Delta^+ a_s da_{s+}^g$$

be invertible for sufficiently large T a.s.

Then $\hat{\theta}_{\delta_n} \to \hat{\theta}_T$ in probability as $n \to \infty$ and $\max_k(t_{k+1}^n - t_k^n) \to 0$, where

$$\begin{aligned}\hat{\theta}_T = &\left[\int_{]0,T]} \Phi_s^r \widetilde{W}_s^r (\Phi_s^r)' \Delta a_s da_t^r + \int_{[0,T[} \Phi_s^g \widetilde{W}_s^g (\Phi_s^g)' \Delta^+ a_s da_{s+}^g\right.\\ &\left.+ \int_{]0,T]} \Phi_s^r \widetilde{W}_s^r (\Phi_s^r)' da_s^c\right]^{-1}\\ &\times\left[\int_{]0,T]} \Phi_s^r \widetilde{W}_s^r \Delta a_s d\widetilde{X}_s^d + \int_{[0,T[} \Phi_s^g \widetilde{W}_s^g \Delta^+ a_s d\widetilde{X}_{s+}^g + \int_{]0,T]} \Phi_s^r \widetilde{W}_s^r d\widetilde{X}_s^c\right].\end{aligned} \tag{7.4}$$

By making use of the equalities

$$\begin{aligned}\Delta a_s dX_s^r &= \Delta a_s (d\widetilde{X}_s^c + d\widetilde{X}_s^d)\\ &= \Delta a_s \mathbf{1}_{\{\Delta a_s = 0\}} dX_s^r + \Delta a_s d\widetilde{X}_s^d\\ &= \Delta a_s d\widetilde{X}_s^d;\\ \Delta^+ a_s dX_{s+}^g &= \Delta^+ a_s \mathbf{1}_{\{\Delta^+ a_s = 0\}} dX_{s+}^g + \Delta^+ a_s d\widetilde{X}_{s+}^g\\ &= \Delta^+ a_s d\widetilde{X}_{s+}^g;\end{aligned}$$

we can rewrite $\hat{\theta}_T$ as follows:

$$
\begin{aligned}
\hat{\theta}_T &= \left[\int_{]0,T]} \Phi^r_s W^r_s (\Phi^r_s)' da^r_t + \int_{[0,T[} \Phi^g_s W^g_s (\Phi^g_s)' da^g_{s+}\right]^{-1} \\
&\quad \times \left[\int_{]0,T]} \Phi^r_s W^r_s dX^r_s + \int_{[0,T[} \Phi^g_s W^g_s dX^g_{s+}\right] \\
&= \left[\Phi W (\Phi)' \circ a_T\right]^{-1} \left[\Phi W \circ X_T\right]
\end{aligned}
\tag{7.5}
$$

where

$$
W^r_s = \widetilde{W}^r_s(\Delta a_s + \mathbf{1}_{\{\Delta a_s = 0\}}) \quad \text{and} \quad W^g_s = \widetilde{W}^g_s \Delta^+ a_s.
$$

The estimator $\hat{\theta}_T$ is called the *weighted least-squares estimator (LSE)*.

7.3 Strong consistency of weighted LSE

The main tool in the proof of strong consistency of weighted LSE in (7.5) is SLLN for multidimensional optional martingales. In order to prove it we need the following lemmas.

For $N = (N^1, \ldots, N^p)' \in \mathcal{M}^2_{loc}(\mathbb{R}^p)$ assume

$$
Q^r_t = \left(\frac{d\langle N^{ir}, N^{jr}\rangle_t}{d\langle N^r\rangle_t}\right)_{1 \le i,j \le p} \quad \text{and} \quad Q^g_t = \left(\frac{d\langle N^{ig}, N^{jg}\rangle_t}{d\langle N^g\rangle_t}\right)_{1 \le i,j \le p},
$$

where $Q^r \in \mathcal{P}, Q^g \in \mathcal{P}_s$ and $N^i = N^{ir} + N^{ig}$.

Let $A = (A^{i,j}_t)_{i,j \le p} \in \mathcal{V}^+ \cap \mathcal{P}_s$ and $A_+ = (A^{i,j}_t)_{i,j \le p} \in \mathcal{V}^+$ such that

$$
\int_{]0,t]} tr A^{-1}_s Q^r_s (A^{-1}_s)' d\langle N^r\rangle_s < \infty,
$$

$$
\int_{[0,t[} tr A^{-1}_{s+} Q^r_{s+} (A^{-1}_{s+})' d\langle N^r\rangle_{s+} < \infty
$$

Then the process

$$
Y_t = \int_{]0,t]} A^{-1}_s dN^r_s + \int_{[0,t[} A^{-1}_{s+} dN^g_{s+} \in \mathcal{M}^2_{loc}(\mathbb{R}^p)
$$

is well-defined.

Consider a p-dimensional version of Kronecker's Lemma for matrices.

Lemma 7.3.1 *Let $N \in \mathcal{M}^2_{loc}(R^p), A \in \mathcal{V}^+ \cap \mathcal{P}_s$ and $A_+ \in \mathcal{V}^+$. Then*

$$(\lambda_{min}(A_t) \to \infty) \cap (\limsup_{t\to\infty} \lambda^{-1}_{min}(A_t)\lambda_{max}(A^r_t) < \infty) \cap$$
$$(\limsup_{t\to\infty} \lambda^{-1}_{min}(A_t)\lambda_{max}(A^g_t) < \infty) \cap (Y_t \to) \subseteq (\left\|A^{-1}_t N_t\right\| \to 0) \ (a.s.)$$

where $\lambda_{min}(A_t)$ and $\lambda_{max}(A_t)$ are smallest and largest eigenvalues of matrix A_t.

Proof. It is easy to see that $N_t = \int_{]0,t]} A_s dY^r_s + \int_{[0,t[} A_{s+} dY^g_{s+}$ (a.s.).

Next, using integration by parts formula we get

$$A_t Y_t = \int_{]0,t]} A_s dY^r_s + \int_{[0,t[} A_{s+} dY^g_{s+} + \int_{]0,t]} dA^r_s Y_{s-} + \int_{[0,t[} dA^g_{s+} Y_s.$$

From this we have

$$A^{-1}_t N_t = A^{-1}_t \left[\int_{]0,t]} (Y_t - Y_{s-}) dA^r_s + \int_{[0,t[} (Y_t - Y_s) dA^g_{s+} \right].$$

Taking into account the conditions of the lemma, we get for $u < t, v < t$

$$
\begin{aligned}
\left\|A_t^{-1}N_t\right\| \leq & \lambda_{min}^{-1}(A_t)\left[\left\|\int_{]0,t]}(Y_t - Y_{s-})dA_s^r\right\| + \left\|\int_{[0,t[}(Y_t - Y_s)dA_{s+}^g\right\|\right] \\
\leq & \lambda_{min}^{-1}(A_t)\Bigg[\left\|\int_{]0,u]}(Y_t - Y_{s-})dA_s^r\right\| + \left\|\int_{]u,t]}(Y_t - Y_{s-})dA_s^r\right\| \\
& + \left\|\int_{[0,v[}(Y_t - Y_s)dA_{s+}^g\right\| + \left\|\int_{[v,t[}(Y_t - Y_s)dA_{s+}^g\right\|\Bigg] \\
\leq & \lambda_{min}^{-1}(A_t)\Bigg[\left\|\int_{]0,u]}(Y_t - Y_{s-})dA_s^r\right\| + \left\|\int_{]u,t]}(Y_\infty - Y_t)dA_s^r\right\| \\
& + \left\|\int_{]u,t]}(Y_\infty - Y_{s-})dA_s^r\right\| + \left\|\int_{[0,v[}(Y_t - Y_s)dA_{s+}^g\right\| \\
& + \left\|\int_{[v,t[}(Y_\infty - Y_t)dA_{s+}^g\right\| + \left\|\int_{[v,t[}(Y_\infty - Y_s)dA_{s+}^g\right\|\Bigg] \\
\leq & \lambda_{min}^{-1}(A_t)\Bigg[2n\sup_{s\geq 0}\|Y_s\|\left[\lambda_{max}(A_u^r) + \lambda_{max}(A_v^g)\right] \\
& + n\left\|Y_\infty - Y_t\right\|\left[\lambda_{max}(A_t^r) + \lambda_{max}(A_t^g)\right] \\
& + \left\|\int_{]u,t]}(Y_\infty - Y_{s-})dA_s^r\right\| + \left\|\int_{[v,t[}(Y_\infty - Y_s)dA_{s+}^g\right\|\Bigg]
\end{aligned} \tag{7.6}
$$

For sufficiently large t by choosing appropriate values u and v the right hand-side in (7.6) can be made arbitrary small. ∎

From here, we obtain the following version of SLLN for p-dimensional optional martingales.

Lemma 7.3.2 *Suppose $N \in \mathcal{M}_{loc}^2(\mathbb{R}^p), A \in \mathcal{V}^+ \cap \mathcal{P}_s, A_+ \in \mathcal{V}^+$ and the following conditions hold (a.s.):*

(a) $\lambda_{min}(A_t) \to \infty$ as $t \to \infty$;

(b)

$$\limsup_{t\to\infty} \lambda_{min}^{-1}(A_t)\lambda_{max}(A_t^r) < \infty,$$
$$\limsup_{t\to\infty} \lambda_{min}^{-1}(A_t)\lambda_{max}(A_t^g) < \infty;$$

(c)

$$\int_{]0,\infty]} trA_s^{-1}Q_s^r(A_s^{-1})'d\langle N^r\rangle_s < \infty,$$
$$\int_{[0,\infty[} trA_{s+}^{-1}Q_{s+}^r(A_{s+}^{-1})'d\langle N^r\rangle_{s+} < \infty.$$

Then $\left\|A_t^{-1}N_t\right\| \to 0$ *a.s. as* $t \to \infty$.

Proof. From condition (c), it follows that $\langle Y\rangle_\infty < \infty$ a.s.. Then from Theorem 2.2 in [36] we have $\|Y_t - Y_\infty\| \to 0$ a.s. and by Lemma 7.3.1 we obtain $\left\|A_t^{-1}N_t\right\| \to 0$ a.s. ■

Suppose

$$N_t = \int_{]0,t]} \Psi_s^r(W_s^r)^{1/2}dm_s^r + \int_{[0,t[} \Psi_s^g(W_s^g)^{1/2}dm_{s+}^g \ \in \mathcal{M}_{loc}^2(R^p),$$
$$A_t = \int_{]0,t]} \Psi_s^r(\Psi_s^r)'da_s^r + \int_{[0,t[} \Psi_s^g(\Psi_s^g)'da_{s+}^g, \tag{7.7}$$

where

$$\Psi_t^j = \Phi_t^j(W_t^j)^{1/2},\ j = r, g, A_t \in \mathcal{V}^+ \cap \mathcal{P}_s, A_{t+} \in \mathcal{V}^+.$$

Theorem 7.3.1 *(cf. [73]). Suppose the following conditions hold for the regressors in model* (7.1) *and the weight matrix* $W = (W^r, W^g)$ *(a.s.)*

(C1)

$$\int_{]0,t]} tr[\Psi_s^r(\Psi_s^r)']\max(1, \Delta a_s + \mathbf{1}_{\{\Delta a_s=0\}})da_s^r$$
$$+ \int_{[0,t[} tr[\Psi_s^g(\Psi_s^g)']\max(1, \Delta^+ a_s)da_{s+}^g < \infty$$

(C2)

$$W^{1/2}BW^{1/2} \leq I, \quad dP \times da - a.e.$$

where I is the identity matrix of size $n \times n$.

(C3) $\lambda_{min}(A_t) \to \infty$ *as* $t \to \infty$*;*

(C4)

$$\limsup_{t\to\infty} \lambda_{min}^{-1}(A_t)\lambda_{max}(A_t^r) < \infty,$$
$$\limsup_{t\to\infty} \lambda_{min}^{-1}(A_t)\lambda_{max}(A_t^g) < \infty;$$

(C5)

$$\int_{]0,\infty]} trA_s^{-1}\Psi_s^r(\Psi_s^r)'(A_s^{-1})'da_s^r <\infty \quad (a.s.),$$
$$\int_{[0,\infty[} trA_{s+}^{-1}\Psi_{s+}^g(\Psi_{s+}^g)'(A_{s+}^{-1})'da_{s+}^g <\infty \quad (a.s.).$$

Then weighted LSE in (7.5) *is strongly consistent* $(T > 0)$.

Proof. By (C1) and (C2) the weighted LSE in (7.5) is well-defined. It is easy to see that $\hat{\theta}_T = \theta + A_T^{-1}N_T$. Then from Theorem 7.3.2 we have $\left\|\hat{\theta}_T - \theta\right\| \to 0$ a.s. as $T \to \infty$. ■

Example 7.3.1 *Let us investigate strong consistency of the weighted LSE estimate* (7.5) *for the model introduced in Example 7.1.1. If we take $W_t = (B_t)^{-1}$ as the weight matrix, then conditions (a) and (b) of Theorem 7.3.1 obviously hold. The process A_t in* (7.7) *has a simple form $A_t = \beta kt$. Thus, $\lambda_{min}(A_t) = \lambda_{max}(A_t) = \beta kt \to \infty$ as $t \to \infty$.*

Further,

$$\int_{]0,\infty]} trA_s^{-1}\Psi_s^r(\Psi_s^r)'(A_s^{-1})'da_s^r = \int_{]0,\infty]} \left[\frac{1}{\beta kt}\beta\frac{1}{\beta kt}\right] dt < \infty. \quad (a.s.)$$

As a result, all assumptions of Theorem 7.3.1 are satisfied, and the weighted LSE estimate in (7.5) *is strongly consistent.*

Example 7.3.2 *Here we consider the model introduced in Example 7.1.2, and we will show that for this model it is actually sufficient to assume only condition (c) of Theorem 7.3.1 and take* $(W^r, W^g) = ((B^r)^{-1}, (B^g)^{-1})$ *for the weighted LSE* (7.5) *to be strongly consistent.*

First we show convergence of θ_t *to* θ *in probability as* $t \to \infty$*. To accomplish this, we compute*

$$\mathbf{E}\theta_t = \theta + A_t^{-1}\mathbf{E}N_t = \theta,$$

where

$$\begin{aligned} N_t =& N_t^r + N_t^g := \int_{]0,t]} \Phi_s^r (B_s^r)^{-1} dm_s^r + \int_{[0,t[} \Phi_s^g (B_s^g)^{-1} dm_{s+}^g, \\ A_t :=& \int_{]0,t]} \Phi_s^r (B_s^r)^{-1} (\Phi_s^r)' da_s^r + \int_{[0,t[} \Phi_s^g (B_s^g)^{-1} (\Phi_s^g)' da_{s+}^g. \end{aligned}$$

Further, using orthogonality of N^r *and* N^g *we have*

$$\begin{aligned} \left(\langle N, N\rangle_t^{ij}\right)_{1\le i,j\le n} =& \left(\langle N^r, N^r\rangle_t^{ij}\right)_{1\le i,j\le n} + \left(\langle N^g, N^g\rangle_t^{ij}\right)_{1\le i,j\le n} \\ =& \int_{]0,t]} \Phi_s^r (B_s^r)^{-1} (\Phi_s^r)' da_s^r + \int_{[0,t[} \Phi_s^g (B_s^g)^{-1} (\Phi_s^g)' da_{s+}^g \\ =& A_t; \\ \mathbf{Cov}\left(A_t^{-1}N_t, A_t^{-1}N_t\right) =& \mathbf{E}\left[A_t^{-1}N_t\left(N_t\right)'\left(A_t^{-1}\right)'\right] \\ =& A_t^{-1}\mathbf{E}\left[N_t\left(N_t\right)'\right]\left(A_t^{-1}\right)' \\ =& A_t^{-1}. \end{aligned}$$

Next,

$$tr\ \mathbf{Cov}\left(A_t^{-1}N_t, A_t^{-1}N_t\right) \le n\lambda_{min}^{-1}(A_t) \to 0,$$

as $t \to \infty$*, and, therefore, from Chebyshev's inequality the consistency of* θ_t *follows.*

To show the convergence (a.s.) we consider the random variable

$$t_c = \inf\{t : \lambda_{min}(A_t) \geq c\}, \quad c > 0,$$

and the stochastic process defined by

$$Y_t = A_t^{-1} N_t - A_{t_c}^{-1} N_{t_c}.$$

Let us compute the covariance of the martingale Y_t

$$\begin{aligned}\mathbf{Cov}(Y_t, Y_t) = \mathbf{E}(Y_t Y_t') =& \mathbf{E}\Big[A_t^{-1} N_t (N_t)' \left(A_t^{-1}\right)' - A_t^{-1} N_t (N_{t_c})' \left(A_{t_c}^{-1}\right)' \\ & - A_{t_c}^{-1} N_{t_c} (N_t)' \left(A_t^{-1}\right)' + A_{t_c}^{-1} N_{t_c} (N_{t_c})' \left(A_{t_c}^{-1}\right)'\Big] \\ =& A_{t_c}^{-1} + A_t^{-1} - A_t^{-1} - A_t^{-1} = A_{t_c}^{-1} - A_t^{-1}.\end{aligned}$$

As a result, $tr\mathbf{Cov}(Y_t, Y_t) \leq n\lambda_{min}^{-1}(A_{t_c}) < \infty$. By Theorem 2.2 in [36] Y_t converges (a.s.) to a finite limit. Consequently, $A_t^{-1} N_t$ also converges (a.s.) to a finite limit. Since we have shown its convergence in probability to zero, this limit is zero.

7.4 Unbiased prescribed precision estimation for multidimensional processes

Let us consider the model (7.1) under the condition that both the process X_t and the vector θ are multidimensional but the dimension of unknown parameter vector θ does not exceed the dimension of X_t:

$$\dim \theta = p \leq n = \dim X_t \quad \forall t \geq 0.$$

In this case one can construct unbiased sequential estimators for θ with prescribed mean-square error by using the special weight matrix W and stopping rules.

Assume that the matrices B^r and $\Phi^r B^r (\Phi^r)'$ are positive definite $dP \times da^r$-a.e., and the matrices B^g and $\Phi^g B^g (\Phi^g)'$ are positive definite $dP \times da^g$-a.e.

We begin with the weighted LSE

$$\hat{\theta}_T = \left[\int_{]0,T]} \Phi^r_s W^r_s (\Phi^r_s)' da^r_t + \int_{[0,T[} \Phi^g_s W^g_s (\Phi^g_s)' da^g_{s+}\right]^{-1} \times \left[\int_{]0,T]} \Phi^r_s W^r_s dX^r_s + \int_{[0,T[} \Phi^g_s W^g_s dX^g_{s+}\right]. \quad (7.8)$$

Let the weight matrix $W_t = (W^r_t, W^g_t)$ be such that

$$\Phi^r_t W^r_t (\Phi^r_t)' = c^r_t I, \qquad \Phi^g_t W^g_t (\Phi^g_t)' = c^g_t I \quad (7.9)$$

$$tr\,[\Phi^r_t W^r_t B^r_t W^r_t (\Phi^r_t)'] \leq c^r_t, \qquad tr\,[\Phi^g_t W^g_t B^g_t W^g_t (\Phi^g_t)'] \leq c^g_t \quad (7.10)$$

where $c^r_t \in \mathcal{P}$ and $c^g_t \in \mathcal{O}$ are positive functions. Equation (7.9) is satisfied for

$$\begin{aligned} W^r_t &= c^r_t (B^r_t)^{-1} (\Phi^r_t)' \left(\Phi^r_t (B^r_t)^{-1} (\Phi^r_t)'\right)^{-2} \Phi^r_t (B^r_t)^{-1}, \\ W^g_t &= c^g_t (B^g_t)^{-1} (\Phi^g_t)' \left(\Phi^g_t (B^g_t)^{-1} (\Phi^g_t)'\right)^{-2} \Phi^g_t (B^g_t)^{-1}. \end{aligned} \quad (7.11)$$

Substituting these functions in inequalities (7.10) yields

$$\begin{aligned} (c^r_t)^2 tr\left[\left(\Phi^r_t (B^r_t)^{-1} (\Phi^r_t)'\right)^{-1}\right] &\leq c^r_t, \\ (c^g_t)^2 tr\left[\left(\Phi^g_t (B^g_t)^{-1} (\Phi^g_t)'\right)^{-1}\right] &\leq c^g_t. \end{aligned}$$

Let

$$\begin{aligned} c^r_t &= \left\{tr\left[\left(\Phi^r_t (B^r_t)^{-1} (\Phi^r_t)'\right)^{-1}\right]\right\}^{-1}, \\ c^g_t &= \left\{tr\left[\left(\Phi^g_t (B^g_t)^{-1} (\Phi^g_t)'\right)^{-1}\right]\right\}^{-1}. \end{aligned} \quad (7.12)$$

Conditions (7.9) and (7.10) enable us to invert the matrix in (7.8) and reduce the problem of constructing a sequential estimator for the vector θ to the scalar case.

For each $h > 0$ we introduce a stopping time τ_h as

$$\tau_h = \inf\left\{t \geq 0 : \int_{]0,t]} \frac{da^r_s}{tr\left[\left(\Phi^r_s (B^r_s)^{-1} (\Phi^r_s)'\right)^{-1}\right]} + \int_{[0,t]} \frac{da^g_{s+}}{tr\left[\left(\Phi^g_s (B^g_s)^{-1} (\Phi^g_s)'\right)^{-1}\right]} > h\right\}, \quad (7.13)$$

and a random variable β_h, with values in $[0,1]$, uniquely determined from the equation

$$\int_{]0,\tau_h[} \frac{da_s^r}{tr\left[(\Phi_s^r(B_s^r)^{-1}(\Phi_s^r)')^{-1}\right]} + \int_{[0,\tau_h[} \frac{da_{s+}^g}{tr\left[(\Phi_s^g(B_s^g)^{-1}(\Phi_s^g)')^{-1}\right]}$$
$$+ \beta_h \left[\frac{\Delta a_{\tau_h}^r}{tr\left[\left(\Phi_{\tau_h}^r(B_{\tau_h}^r)^{-1}(\Phi_{\tau_h}^r)'\right)^{-1}\right]} + \frac{\Delta^+ a_{\tau_h}^g}{tr\left[(\Phi_{\tau_h}^g(B_{\tau_h}^g)^{-1}(\Phi_{\tau_h}^g)')^{-1}\right]} \right] = h. \quad (7.14)$$

The random variable β_h is $\mathcal{F}_{\tau_h}$-measurable.

On the basis of the estimator (7.8) with the weight matrix given by (7.11) and (7.12), we define the sequential estimator for the vector θ as

$$\begin{aligned}
\theta^*(h) =& h^{-1} \left[\int_{]0,\tau_h[} \Phi_s^r W_s^r dX_s^r + \int_{[0,\tau_h[} \Phi_s^g W_s^g dX_{s+}^g \right. \\
& \left. + \beta_h \left(\Phi_{\tau_h}^r W_{\tau_h}^r \Delta X_{\tau_h} + \Phi_{\tau_h}^g W_{\tau_h}^g \Delta^+ X_{\tau_h} \right) \right] \\
=& h^{-1} \left[\int_{]0,\tau_h]} \Phi_s^r W_s^r \left(\mathbf{1}_{]0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s) \right) dX_s^r \right. \\
& \left. + \int_{[0,\tau_h]} \Phi_s^g W_s^g \left(\mathbf{1}_{[0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s) \right) dX_{s+}^g \right]. \qquad (7.15)
\end{aligned}$$

where

$$\begin{aligned}
W_t^r &= \left\{ tr\left[\left(\Phi_t^r (B_t^r)^{-1} (\Phi_t^r)' \right)^{-1} \right] \right\}^{-1} (B_t^r)^{-1} (\Phi_t^r)' \left(\Phi_t^r (B_t^r)^{-1} (\Phi_t^r)' \right)^{-2} \Phi_t^r (B_t^r)^{-1}, \\
W_t^g &= \left\{ tr\left[\left(\Phi_t^g (B_t^g)^{-1} (\Phi_t^g)' \right)^{-1} \right] \right\}^{-1} (B_t^g)^{-1} (\Phi_t^g)' \left(\Phi_t^g (B_t^g)^{-1} (\Phi_t^g)' \right)^{-2} \Phi_t^g (B_t^g)^{-1}.
\end{aligned} \qquad (7.16)$$

This estimator has the following properties.

Theorem 7.4.1 *Let the matrices B^r and $\Phi^r(B^r)^{-1}(\Phi^r)'$ be not degenerate $dP \times da^r$-a.e., and $B^g, \Phi^g(B^g)^{-1}(\Phi^g)'$ be not degenerate $dP \times da^g$-a.e.; the integral*

$$\int_{]0,t]} \frac{da_s^r}{tr\left[(\Phi_s^r(B_s^r)^{-1}(\Phi_s^r)')^{-1}\right]} + \int_{[0,t]} \frac{da_{s+}^g}{tr\left[(\Phi_s^g(B_s^g)^{-1}(\Phi_s^g)')^{-1}\right]} \qquad (7.17)$$

be finite for $0 < t < \infty$ a.s. and converging to $+\infty$ as $t \to +\infty$ a.s.

Then, for each $h > 0$,

$$
\begin{aligned}
\tau_h <& \infty \quad a.s., \\
\mathbf{E}_\theta \theta^*(h) =& \theta, \\
\mathbf{E}_\theta \left\| \theta^*(h) - \theta \right\|^2 \leq & h^{-1},
\end{aligned}
$$

where $\mathbf{E}_\theta$ denotes the average by the distribution P_θ of the process X with given parameter θ.

Proof. Since the integral (7.17) tends to $+\infty$, as $t \to \infty$, the stopping time τ_h is finite a.s. for all $h > 0$.

from (7.1) and (7.15) we have

$$
\begin{aligned}
\theta^*(h) = h^{-1} \Bigg[& \int_{]0,\tau_h]} \Phi_s^r W_s^r \left(\mathbf{1}_{]0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s) \right) \theta da_s^r \\
& + \int_{[0,\tau_h]} \Phi_s^g W_s^g \left(\mathbf{1}_{[0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s) \right) \theta da_{s+}^g + M_{\tau_h +} \Bigg]
\end{aligned}
$$

where

$$
\begin{aligned}
M_{\tau_h +} = & \int_{]0,\tau_h]} \Phi_s^r W_s^r \left(\mathbf{1}_{]0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s) \right) dm_s^r \\
& + \int_{[0,\tau_h]} \Phi_s^g W_s^g \left(\mathbf{1}_{[0,\tau_h]}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s) \right) dm_{s+}^g .
\end{aligned}
$$

By (7.11)-(7.14), we get

$$
\theta^*(h) = \theta + h^{-1} M_{\tau_h}. \tag{7.18}
$$

Since $M \in \mathcal{M}^2_{loc}(\mathbb{R}^p)$, then in view of (7.16),

$$
\begin{aligned}
& tr\left[(\langle M^i_{\tau_h}, M^j_{\tau_h}\rangle)_{1\le i,j\le n}\right] \\
&= \int_{]0,\tau_h]} tr\left[\Phi^r_s W^r_s B^r_s W^r_s (\Phi^r_s)'\right]\left(\mathbf{1}_{]0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s)\right)^2 da^r_s \\
&\quad + \int_{[0,\tau_h]} tr\left[\Phi^g_s W^g_s B^g_s W^g_s (\Phi^g_s)\right]\left(\mathbf{1}_{[0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s)\right)^2 da^g_{s+} \\
&= \int_{]0,\tau_h]} \frac{\left(\mathbf{1}_{]0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s)\right)^2 da^r_s}{tr\left[(\Phi^r_s (B^r_s)^{-1}(\Phi^r_s)')^{-1}\right]} \\
&\quad + \int_{[0,\tau_h]} \frac{\left(\mathbf{1}_{[0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s)\right)^2 da^g_{s+}}{tr\left[(\Phi^g_s (B^g_s)^{-1}(\Phi^g_s)')^{-1}\right]} \\
&\le \int_{]0,\tau_h]} \frac{\left(\mathbf{1}_{]0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s)\right) da^r_s}{tr\left[(\Phi^r_s (B^r_s)^{-1}(\Phi^r_s)')^{-1}\right]} \\
&\quad + \int_{[0,\tau_h]} \frac{\left(\mathbf{1}_{[0,\tau_h[}(s) + \beta_h \mathbf{1}_{\{\tau_h\}}(s)\right) da^g_{s+}}{tr\left[(\Phi^g_s (B^g_s)^{-1}(\Phi^g_s)')^{-1}\right]} = h.
\end{aligned}
\tag{7.19}
$$

Hence, $\mathbf{E}_\theta M_{\tau_h} = 0$ and from (7.18) we obtain $\mathbf{E}_\theta \theta^*(h) = \theta$.

Further from (7.18) and (7.19), it follows that

$$
\mathbf{E}_\theta \left\|\theta^*(h) - \theta\right\|^2 = h^{-2}\mathbf{E}_\theta tr\left[(\langle M^i_{\tau_h}, M^j_{\tau_h}\rangle)_{1\le i,j\le n}\right] \le h^{-1}.
$$

■

Example 7.4.1 *Consider the model from Example 7.1.1. Then condition* (7.17) *becomes*

$$
\int_{]0,t]} \frac{da^r_s}{tr\left[(\Phi^r_s (B^r_s)^{-1}(\Phi^r_s)')^{-1}\right]} + \int_{[0,t[} \frac{da^g_{s+}}{tr\left[(\Phi^g_s (B^g_s)^{-1}(\Phi^g_s)')^{-1}\right]} = \int_{]0,t]} \beta k ds = \beta k t,
\tag{7.20}
$$

and it is finite for $0 < t < \infty$ *a.s. and converging to* $+\infty$ *as* $t \to +\infty$ *a.s. Thus, Theorem 7.4.1 holds.*

7.5 Construction of the two-step sequential procedure in general case

In the case when the number of unknown parameters in model is arbitrary, the guaranteed estimator for θ can also be constructed on the basis of the weighted LSE, defined by (7.5). It is convenient to rewrite this estimate as

$$\theta_t = A_t^{-1}\left[\int_{]0,t]} \Psi_s^r (W_s^r)^{1/2} dX_s^r + \int_{[0,t[} \Psi_s^g (W_s^g)^{1/2} dX_{s+}^g\right], \qquad (7.21)$$

where the information matrix A_t is as in (7.7) and its inverse A_t^{-1} is assumed to exist.

In the sequel the following conditions are imposed on the regressors and on the weight matrix W:

(A_1) The regressors matrix-valued functions $\Phi^r \in \mathcal{P}$ and $\Phi^g \in \mathcal{O}$, and such that for all $t \geq 0$

$$\int_{]0,t]} \|\Phi_s^r\|\, da_s^r + \int_{[0,t[} \|\Phi_s^g\|\, da_{s+}^g < \infty \quad (a.s.)$$

(A_2) The weight matrix W is such that

$$W^{1/2} B W^{1/2} \leq I, \quad dP \times da - a.e.$$

where I is the identity matrix of size $n \times n$.

(A_3) Both integrals in (7.21) are well defined if for all $t \geq 0$

$$\int_{]0,t]} tr[\Psi_s^r (\Psi_s^r)'] \max(1, \Delta a_s + \mathbf{1}_{\{\Delta a_s = 0\}}) da_s^r + \int_{[0,t[} tr[\Psi_s^g (\Psi_s^g)'] \max(1, \Delta^+ a_s) da_{s+}^g < \infty \quad (a.s.)$$

(A_4) $\lim_{t\to\infty} \lambda_{min}(A_{t+}) = +\infty, \quad (a.s.)$

(A_5) There exists $\delta,\ 0 < \delta < 1$, such that

$$\liminf_{t\to\infty} \lambda_{min}^{\delta}(A_{t+}) / \ln \lambda_{max}(A_{t+}) > 0 \quad (a.s.)$$

The procedure is constructed in two steps.

Step 1. Let $(C_j)_{j\geq 1},\ (\beta_j)_{j\geq 1}$ be two sequences of positive numbers such that

$$C_j \uparrow \infty, \quad \sum_{j\geq 1} \beta_j < \infty, \quad \sum_{j\geq 1} \beta_j C_j^{\frac{1-\delta}{\delta}} = \infty.$$

Here $\delta,\ 0 < \delta < 1$, is the same as in condition (A_5).

By virtue of condition (A_4), for any given positive constant C_0 we can define the a.s. finite wide sense stopping time T as

$$T = \inf\{t \geq 0 : \lambda_{min}(A_{t+}) > C_0\}, \quad \inf\{\emptyset\} = +\infty. \tag{7.22}$$

Next we introduce the sequence of wide sense stopping times $\tau_j,\ j \geq 1$, as

$$\begin{aligned}\tau_j = \inf \Bigg\{ t \geq T : C_0^{-1} \left(\int_{]0,T]} \|\Psi_s^r(\Psi_s^r)'\| \, da_s^r + \int_{[0,T[} \|\Psi_s^g(\Psi_s^g)'\| \, da_{s+}^g \right) \\ + tr \int_{]T,t]} \Psi_s^r A_s^{-1} (\Psi_s^r)' da_s^r + tr \int_{[T,t]} \Psi_s^g A_{s+}^{-1} (\Psi_s^g)' da_{s+}^g > C_j \Bigg\},\end{aligned} \tag{7.23}$$

and the sequence of estimators

$$\theta_j = \theta(\tau_j) = A_{\tau_j}^{-1} \left[\int_{]0,\tau_j]} \Psi_s^r (W_s^r)^{1/2} dX_s^r + \int_{[0,\tau_j]} \Psi_s^g (W_s^g)^{1/2} dX_{s+}^g \right].$$

On the basis of these estimators we define the desired sequential estimators of the unknown vector θ by applying a special smoothing procedure.

Step 2. Let us define the estimator θ_h^* as a weighted average of estimators θ_j :

$$\theta_h^* = \left[\sum_{j=1}^{\sigma(h)} b_j \right]^{-1} \sum_{j=1}^{\sigma(h)} b_j \theta_j, \tag{7.24}$$

where h is a positive parameter; $\sigma(h)$ is the wide sense stopping time given by

$$\sigma(h) = \inf \left(n \geq 1 : \sum_{j=1}^{n} b_j > h \right),$$

$$b_j = \beta_j/[C_j tr[A^{-1}_{\tau_j+}]].$$

Denote

$$N(h) = \tau_{\sigma(h)}.$$

The main result of this section is the following.

Theorem 7.5.1 *Let the regressor matrix-valued functions* (Φ^r, Φ^g), $\Phi^r \in \mathcal{P}$ *and* $\Phi^g \in \mathcal{O}$ *in model* (7.1) *and the corresponding weight matrices* (W^r, W^g), $W^r \in \mathcal{P}$, *and* $W^g \in \mathcal{O}$ *satisfy conditions* $(A_1) - (A_5)$. *Then the sequential design* $(N(h), \theta^*_h)$ *has the following properties: for any* $h > 0$,

$$N(h) < \infty \quad a.s.$$

$$\mathbf{E}_\theta \left\|\theta^*_h - \theta\right\|^2 \leq h^{-1} \sum_{j=1}^{\infty} \beta_j (1 + pC_j^{-1}).$$

Proof. By condition (A_4) we have $T < \infty$ a.s. From the definition of τ_j and Lemma 7.5.4 we have $\tau_j < \infty$ a.s. and $\tau_j \uparrow +\infty$, as $j \to \infty$. Therefore, the inequality $N(h) < \infty$ is true provided that

$$\sum_{j \geq 1} b_j = +\infty \quad a.s. \tag{7.25}$$

Let us verify this equality. From the definitions of b_j and τ_j it follows that

$$\begin{aligned}
b_j =& \frac{\beta_j}{C_j tr[A^{-1}_{\tau_j+}]} \geq \frac{\beta_j}{C_j p \lambda_{max}(A^{-1}_{\tau_j+})} \\
=& \frac{\beta_j}{C_j p} \lambda_{min}(A_{\tau_j+}) \\
=& \beta_j p^{-1} C_j^{1/\delta - 1} (\lambda^\delta_{min}(A_{\tau_j+})/C_j)^{1/\delta} \\
\geq& \beta_j p^{-1} C_j^{1/\delta - 1} \Bigg[\lambda^\delta_{min}(A_{\tau_j+}) \\
& \times \left(g(T) + tr \int_{]T, \tau_j]} \Psi^r_s A^{-1}_s (\Psi^r_s)' da^r_s + tr \int_{[T, \tau_j]} \Psi^g_s A^{-1}_{s+} (\Psi^g_s)' da^g_{s+} \right)^{-1} \Bigg]^{1/\delta},
\end{aligned}$$

where

$$g(T) = C_0^{-1}\left(\int_{]0,T]} \|\Psi_s^r\|^2\, da_s^r + \int_{[0,T[} \|\Psi_s^g\|^2\, da_{s+}^g\right).$$

By making use of Lemma 7.5.4 we obtain

$$b_j \geq \beta_j p^{-1} C_j^{1/\delta - 1}\left[\lambda_{min}^{\delta}(A_{\tau_j+}) \Big/ \left(g(T) + C \ln \lambda_{max}(A_{\tau_j+})\right)\right]^{1/\delta},$$

where $C > 0$ is some constant. From this, the properties of β_j, C_j and condition (A_5) we obtain (7.25).

Further, we have

$$\theta_j - \theta = A_{\tau_j+}^{-1} N_{\tau_j+},$$

where

$$N_{t+} = \int_{]0,t]} \Psi_s^r (W_s^r)^{1/2} dm_s^r + \int_{[0,t]} \Psi_s^g (W_s^g)^{1/2} dm_{s+}^g.$$

From this it follows that

$$\begin{aligned}
\|\theta_j - \theta\|^2 &= \left\| A_{\tau_j+}^{-1/2} A_{\tau_j+}^{-1/2} N_{\tau_j+} \right\|^2 \\
&\leq \left\| A_{\tau_j+}^{-1/2} \right\|^2 \left\| A_{\tau_j+}^{-1/2} N_{\tau_j+} \right\|^2 = Q_{\tau_j+} tr A_{\tau_j+}^{-1},
\end{aligned}$$

where

$$Q_{\tau_j+} = N'_{\tau_j+} A_{\tau_j+}^{-1} N_{\tau_j+}. \tag{7.26}$$

Taking into account the definition of b_j and applying the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
\|\theta^*(h) - \theta\|^2 \leq & \sum_{i=1}^{\sigma(h)} b_i \left(\sum_{j=1}^{\sigma(h)} b_j \|\theta_j - \theta\|^2\right)\left(\sum_{j=1}^{\sigma(h)} b_j\right)^{-2} \\
\leq & h^{-1} \sum_{j\geq 1} b_j \|\theta_j - \theta\|^2 \leq h^{-1} \sum_{j\geq 1} b_j Q_{\tau_j+} tr A_{\tau_j+}^{-1} \\
= & h^{-1} \sum_{j\geq 1} \beta_j Q_{\tau_j+} / C_j.
\end{aligned}$$

Hence,

$$\mathbf{E}\left\|\theta^*(h)-\theta\right\|^2 \le h^{-1}\sum_{j\ge 1}\beta_j \mathbf{E}Q_{\tau_j+}/C_j.$$

From this and Lemma 7.5.5 we obtain the desired result. ■

Example 7.5.1 *Consider the model*

$$X_t = \theta^1 kt + \theta^2 k\int_{]0,t]} f(X_{s-})ds + m_t,$$

where m_t is the same as in (7.3). *This is a non-homogeneous Ornstein –Uhlenbeck process (mean-reverting process in mathematical finance). In this example, we can not apply results from Section 7.4 because the number of parameters is greater than number of the observations X_t. However, we can use the technique described in Section 7.5. We assume that the function f satisfies Lipschitz continuity condition for the existence and uniqueness of the solution X_t (see [13]).*

Suppose that

$$\liminf_{t\to\infty}\frac{\int_{]0,t]} f^2(X_{s-})ds}{t} > 0, \tag{7.27}$$

$$\int_{]0,t]} f^2(X_{s-})ds = O(t). \tag{7.28}$$

Now, we check the assumptions of Theorem .

Conditions (A_1) and (A_3) are obviously satisfied, and (A_2) holds with $W_t^r = 1$.

The process A_t in this example has the following form

$$A_t = \begin{bmatrix} kt & \int_{]0,t]} f(X_{s-})ds \\ \int_{]0,t]} f(X_{s-})ds & \int_{]0,t]} f^2(X_{s-})ds \end{bmatrix}.$$

Thus,

$$\lambda_{max,min}(A_t) = k\Bigg[\int_{]0,t]} f^2(X_{s-})ds + t \pm \sqrt{\left(\int_{]0,t]} f^2(X_{s-})ds - t\right)^2 + 4\left(\int_{]0,t]} f(X_{s-})ds\right)^2}\Bigg]$$

Further, using (7.27)

$$\lambda_{min}(A_t) = 2k\left[t\int_{]0,t]} f^2(X_{s-})ds - \left(\int_{]0,t]} f(X_{s-})ds\right)^2\right] \Bigg/ \Bigg[\int_{]0,t]} f^2(X_{s-})ds$$
$$+ t + \sqrt{\left(\int_{]0,t]} f^2(X_{s-})ds - t\right)^2 + 4\left(\int_{]0,t]} f(X_{s-})ds\right)^2}\,\Bigg]$$
$$\geq \frac{2k(1-1/t)}{1/t + \frac{1}{\int_{]0,t]} f^2(X_{s-})ds} + \sqrt{\left(1/t - \frac{1}{\int_{]0,t]} f^2(X_{s-})ds}\right)^2 + 4/t}} \to \infty$$

as $t \to \infty$.

Again, using (7.27)

$$\liminf_{t\to\infty} \frac{\lambda_{min}(A_t)}{t} \geq \frac{2k(1-1/t)}{1 + \frac{t}{\int_{]0,t]} f^2(X_{s-})ds} + \sqrt{\left(1 - \frac{t}{\int_{]0,t]} f^2(X_{s-})ds}\right)^2 + \frac{4}{\int_{]0,t]} f^2(X_{s-})ds}}}$$
$$> 0,$$

and

$$\lambda_{max}(A_t) = \int_{]0,t]} f^2(X_{s-})ds + t$$
$$+ \sqrt{\left(\int_{]0,t]} f^2(X_{s-})ds - t\right)^2 + 4\left(\int_{]0,t]} f(X_{s-})ds\right)^2}$$
$$\leq \int_{]0,t]} f^2(X_{s-})ds + t + \int_{]0,t]} f^2(X_{s-})ds + t$$
$$+ 2\sqrt{\int_{]0,t]} f^2(X_{s-})ds}.$$

Thus, by (7.28) $\lambda_{max}(A_t) = O(t)$.

Therefore,

$$\liminf_{t\to\infty} \lambda^{\delta}_{min}(A_t)/\ln\lambda_{max}(A_t) \geq \liminf_{t\to\infty} t^{\delta}/\ln(t) > 0 \quad (a.s.)$$

for all $\delta,\ 0 < \delta < 1$. *Consequently, the implications of Theorem 7.5.1 hold.*

Note that the conditions (7.27) and (7.28) can be obtained, for instance, if the function f is bounded, i.e. $c_1 \le f(x) \le c_2$, which is enough for most real-world applications like modelling of financial assets or capital processes of firms.

7.5.1 Auxiliary results

Lemma 7.5.1 *Let D be a $p \times p$ matrix and F be a $p \times p$ symmetric non-negative definite non-zero matrix of real numbers. If the matrix $C = D + F$ is non-singular and $rank F = r$, then*

$$tr[C^{-1}F] = \sum_{i=1}^{r} [|C| - |C - \lambda_i e_i e_i'|]/|C|,$$

where (λ_i) and (e_i) are the eigenvalues and eigenvectors of matrix F respectively. If, besides, D is symmetric non-negative definite, then

$$tr[C^{-1}F] \le r[|C| - |D|]/|C| \le r.$$

Lemma 7.5.2 *Under assumptions (A_1) and (A_3) the following inequalities are satisfied:*

$$\ln \frac{|A_t|}{|A_T|} \le \int_{]T,t]} \frac{d|A_s^r|}{|A_{s-}|} + \int_{[T,t[} \frac{d|A_{s+}^g|}{|A_s|} \quad a.s.$$

If A_t is a continuous matrix-valued process, then

$$\ln \frac{|A_t|}{|A_T|} = \int_{]T,t]} \frac{d|A_s|}{|A_{s-}|} \quad a.s.$$

Proof. By applying change of variables formula to the process $\ln|A_t|,\ t \ge T$, we obtain

$$\begin{aligned}\ln|A_t| = {}& \ln|A_T| + \int_{]T,t]} \frac{d|A_s^r|}{|A_{s-}|} + \int_{[T,t[} \frac{d|A_{s+}^g|}{|A_s|} \\ & + \sum_{T<s\le t} \left(\ln \frac{|A_s|}{|A_{s-}|} - \frac{\Delta|A_s|}{|A_{s-}|} \right) + \sum_{T\le s<t} \left(\ln \frac{|A_{s+}|}{|A_s|} - \frac{\Delta^+|A_s|}{|A_s|} \right) \quad a.s.\end{aligned}$$

Thus,

$$
\begin{aligned}
\int_{]T,t]} \frac{d|A_s^r|}{|A_{s-}|} + \int_{[T,t[} \frac{d|A_{s+}^g|}{|A_s|} &= \ln \frac{|A_t|}{|A_T|} \\
&\quad + \sum_{T<s\le t} \left[\frac{\Delta |A_s|}{|A_{s-}|} - \ln \left(1 + \frac{\Delta |A_s|}{|A_{s-}|} \right) \right] \\
&\quad + \sum_{T\le s<t} \left[\frac{\Delta^+ |A_s|}{|A_s|} - \ln \left(1 + \frac{\Delta^+ |A_s|}{|A_s|} \right) \right] \\
&\ge \ln \frac{|A_t|}{|A_T|} \quad a.s.
\end{aligned}
$$

From here using the inequality $\ln(1+x) \le x, \ x \ge 0$, we come to the desired result.

■

Lemma 7.5.3 *Let the regressor matrices Φ^r, Φ^g in* (7.1) *and the weight matrices W^r, W^g satisfy conditions (A_1), (A_3) and T be defined as in* (7.22).

Then for any $t \ge T$

$$
tr \left(\int_{]T,t]} (\Psi_s^r)' \widetilde{A}_s^{-1} \Psi_s^r da_s^c \right) = \int_{]T,t]} \frac{d|\widetilde{A}_s|}{|\widetilde{A}_s|} \quad a.s., \tag{7.29}
$$

where

$$
\widetilde{A}_t = A_T + \int_{]T,t]} \Psi_s^r (\Psi_s^r)' da_s^c. \tag{7.30}
$$

Proof. First we verify that the integral in the left-hand side of equality is well-defined, that is, for all $t \ge T$,

$$
\int_{]T,t]} \left\| (\Psi_s^r)' \widetilde{A}_s^{-1} \Psi_s^r \right\| da_s^c < \infty \quad a.s. \tag{7.31}
$$

By the inequality

$$
\left\| (\Psi_s^r)' \widetilde{A}_s^{-1} \Psi_s^r \right\| \le \|(\Psi_s^r)' \Psi_s^r\| \, tr \widetilde{A}_s^{-1}
$$

we have

$$\begin{aligned}\int_{]T,t]} \left\|(\Psi_s^r)'\widetilde{A}_s^{-1}\Psi_s^r\right\| da_s^c \le & \int_{]T,t]} \|\Psi_s^r\|^2 \, tr\widetilde{A}_s^{-1} da_s^c \\ \le & p\int_{]T,t]} \lambda_{max}(\widetilde{A}_s^{-1}) tr(\Psi_s^r)'\Psi_s^r da_s^c \\ \le & p\lambda_{min}^{-1}(A_T)\int_{]T,t]} tr(\Psi_s^r)'\Psi_s^r da_s^r.\end{aligned}$$

In view, of condition (A_3), we obtain (7.31).

Equality (7.29) is equivalent to the one for the differentials:

$$\frac{d|\widetilde{A}_s|}{|\widetilde{A}_s|} = tr\left((\Psi_s^r)'\widetilde{A}_s^{-1}\Psi_s^r\right) da_s^c.$$

Let us find $d|\widetilde{A}_s|$. By the definition of a determinant we have

$$|\widetilde{A}_t| = \sum_{(i_1,\ldots,i_p)} (-1)^{[i_1,\ldots,i_p]}[\widetilde{A}_t]_{i_1,1}\cdots[\widetilde{A}_t]_{i_p,p}$$

where $[\widetilde{A}_t]_{ik}$ is the (i,k)-th element of the matrix $\widetilde{A}_t$ and the summation is taken over all permutations $(i_1,\ldots,i_p)$ of numbers $1,\ldots,p$, and $[i_1,\ldots,i_p]$ denotes the number of inversions in a permutation $(i_1,\ldots,i_p)$. Since the matrix-valued process $\widetilde{A}_t$ is continuous with bounded variation then by the Ito formula we obtain

$$\prod_{i=1}^{p}[\widetilde{A}_t]_{i_l,l} = \sum_{k=1}^{p}\left(\prod_{l=1,l\neq k}^{p}[\widetilde{A}_t]_{i_l,l}\right) d[\widetilde{A}_t]_{i_k,k}$$

and, hence,

$$d|[\widetilde{A}_t]_{i_l,l}| = \sum_{k=1}^{p}\sum_{(i_1,\ldots,i_p)}(-1)^{[i_1,\ldots,i_p]}\left(\prod_{l=1,l\neq k}^{p}[\widetilde{A}_t]_{i_l,l}\right) d[\widetilde{A}_t]_{i_k,k}$$

By (7.30)

$$d[\widetilde{A}_t]_{ik} = [\Psi_t^r(\Psi_t^r)']_{ik}\, da_t^c$$

and therefore

$$d|\widetilde{A}_t| = \sum_{k=1}^{p} |\widetilde{A}_t^{(k)}| da_t^c,$$

where $\widetilde{A}^{(k)}$ is the determinant which is obtained from $\widetilde{A}_t$ by replacing the k-th column by the column-vector $\Big([\Psi_t^r(\Psi_t^r)']_{1k}, \ldots, [\Psi_t^r(\Psi_t^r)']_{pk}\Big)$. Decomposing the determinant $|\widetilde{A}_t^{(k)}|$ by the elements of k-th column yields

$$|\widetilde{A}_t^{(k)}| = \sum_{i=1}^{p} (\widetilde{A}_t)_{ik} \, [\Psi_t^r(\Psi_t^r)']_{ik}$$

where $(\widetilde{A}_t)_{ik}$ is the algebraic adjoint for the element $[\widetilde{A}_t]_{ik}$ of the matrix $\widetilde{A}_t$. Thus

$$d|\widetilde{A}_t| \sum_{k=1}^{p} \sum_{i=1}^{p} (\widetilde{A}_t)_{ik} \, [\Psi_t^r(\Psi_t^r)']_{ik} \, da_t^c.$$

From here it follows that

$$\frac{d|\widetilde{A}_t|}{|\widetilde{A}_t|} = tr \left((\Psi_t^r)' \widetilde{A}_t^{-1} \Psi_t^r \right) da_t^c.$$

■

Lemma 7.5.4 *Under assumptions* (A_1), (A_3) *and* (A_4),

$$tr \int_{]T,t]} (\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^r + tr \int_{[T,t[} (\Psi_s^g)' A_{s+}^{-1} \Psi_s^g da_{s+}^g = O(\ln \lambda_{max}(A_t)),$$
$$t \to \infty \quad a.s., \qquad (7.32)$$

$$\lim_{t\to\infty} \left[tr \int_{]T,t]} (\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^r + tr \int_{[T,t[} (\Psi_s^g)' A_{s+}^{-1} \Psi_s^g da_{s+}^g \right] = +\infty \quad a.s., \qquad (7.33)$$

where $f(t) = O(g(t))$, $t \to \infty$, *means that there exist* $t_0 > T$ *and* $0 < C < \infty$ *such that* $|f(t)| \le C|g(t)|$ *for all* $t \ge t_0$.

Proof. We have

$$\begin{aligned} tr \int_{]T,t]} (\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^r + tr \int_{[T,t[} (\Psi_s^g)' A_s^{-1} \Psi_s^g da_{s+}^g =& tr \int_{]T,t]} (\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^c \\ &+ tr \sum_{T<s\le t} (\Psi_s^r)' A_s^{-1} \Psi_s^r \Delta a_s \qquad (7.34) \\ &+ \sum_{T\le s<t} (\Psi_s^g)' A_{s+}^{-1} \Psi_s^g \Delta^+ a_s. \end{aligned}$$

Let us introduce the process

$$\widetilde{A}_t = A_T + \int_{]T,t]} \Psi^r_s (\Psi^r_s)' da^c_s.$$

This process is continuous and satisfies the inequality $\widetilde{A}_t \leq A_t$ which implies $A_t^{-1} \leq \widetilde{A}^{-1}$. From this and Lemma 7.5.3 and 7.5.2 it follows that for all $t \geq T$,

$$\begin{aligned} tr \int_{]T,t]} (\Psi^r_s)' A_s^{-1} \Psi^r_s da^c_s \leq & tr \int_{]T,t]} (\Psi^r_s)' \widetilde{A}_s^{-1} \Psi^r_s da^c_s \\ & = \int_{]T,t]} \frac{d|\widetilde{A}_s|}{|\widetilde{A}_s|} \\ & = \ln \frac{|\widetilde{A}_t|}{|\widetilde{A}_T|} \leq \ln \frac{|A_t|}{|A_T|}. \end{aligned} \tag{7.35}$$

Now we find the upper bound for the second term in the right-hand side of (7.34). Denoting

$$\hat{A}_t = A_T + \sum_{T<d\leq t} \Psi^r_s (\Psi^r_s)' \Delta a_s$$

and applying Lemma 7.5.1, we obtain

$$\begin{aligned} tr \sum_{T<d\leq t} (\Psi^r_s)' A_s^{-1} \Psi^r_s \Delta a_s \leq & tr \sum_{T<d\leq t} (\Psi^r_s)' \hat{A}_s^{-1} \Psi^r_s \Delta a_s \\ \leq & p \sum_{T<s\leq t} [|\hat{A}_s| - |\hat{A}_{s-}|]/|\hat{A}_s| \\ \leq & p \sum_{T<s\leq t} \int_{|\hat{A}_{s-}|}^{|\hat{A}_s|} \frac{dx}{x} \\ \leq & p \int_{|\hat{A}_T|}^{|\hat{A}_t|} \frac{dx}{x} = p \ln \frac{|\hat{A}_t|}{|\hat{A}_T|} \leq p \ln \frac{|A_t|}{|A_T|}. \end{aligned} \tag{7.36}$$

Similarly, denoting

$$\bar{A}_t = A_T + \sum_{T\leq d<t} \Psi^g_s (\Psi^g_s)' \Delta^+ a_s$$

and applying Lemma 7.5.1, we get

$$
\begin{aligned}
tr \sum_{T\le d<t} (\Psi_s^g)' A_{s+}^{-1} \Psi_s^g \Delta a_s \le & tr \sum_{T\le d<t} (\Psi_s^g)' \bar{A}_{s+}^{-1} \Psi_s^g \Delta a_s \\
\le & p \sum_{T\le d<t} [|\bar{A}_{s+}| - |\bar{A}_s|]/|\bar{A}_{s+}| \\
\le & p \sum_{T\le d<t} \int_{|\bar{A}_s|}^{|\bar{A}_{s+}|} \frac{dx}{x} \\
\le & p \int_{|\bar{A}_T|}^{|\bar{A}_t|} \frac{dx}{x} = p \ln \frac{|\bar{A}_t|}{|\bar{A}_T|} \le p \ln \frac{|A_t|}{|A_T|}.
\end{aligned} \tag{7.37}
$$

Substituting the estimates (7.35)-(7.37) in (7.34) yields

$$
\begin{aligned}
tr \int_{]T,t]} (\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^r + tr \int_{[T,t[} (\Psi_s^g)' A_{s+}^{-1} \Psi_s^g da_{s+}^g \le & (2p+1) \ln \frac{|A_t|}{|A_T|} \\
\le & p(2p+1) \ln \frac{\lambda_{max}(A_t)}{|A_T|}.
\end{aligned}
$$

From this in view of condition (A_4) we obtain (7.32).

Now we verify (7.33). The integrand in (7.33) can be estimated from below by

$$
\begin{aligned}
tr(\Psi_s^r)' A_s^{-1} \Psi_s^r \ge & \lambda_{max}((\Psi_s^r)' A_s^{-1} \Psi_s^r) \\
\ge & \lambda_{min}(A_s^{-1}) \sup_{z\ne 0} \frac{\|\Psi_s^r z\|^2}{\|z\|^2} \\
\ge & C \lambda_{max}^{-1}(A_s) \|\Psi_s^r\|^2 \\
\ge & C \|\Psi_s^r\|^2 / tr A_s = C \|\Psi_s^r\|^2 / V_s
\end{aligned}
$$

where $V_s = \int_{]0,t]} \|\Psi_s^r\|^2 da_s^r + \int_{[0,t[} \|\Psi_s^g\|^2 da_{s+}^g$ and C is some positive constant.

$$
\begin{aligned}
tr(\Psi_s^g)' A_{s+}^{-1} \Psi_s^g \ge & \lambda_{max}((\Psi_s^g)' A_{s+}^{-1} \Psi_s^g) \\
\ge & \lambda_{min}(A_{s+}^{-1}) \sup_{z\ne 0} \frac{\|\Psi_s^g z\|^2}{\|z\|^2} \\
\ge & C \lambda_{max}^{-1}(A_{s+}) \|\Psi_s^g\|^2 \\
\ge & C \|\Psi_s^g\|^2 / tr A_{s+} = C \|\Psi_s^g\|^2 / V_{s+}.
\end{aligned}
$$

Hence,

$$\begin{aligned}
&\lim_{t\to\infty}\left[tr\int_{]T,t]}(\Psi^r_s)'A_s^{-1}\Psi^r_s da^r_s + tr\int_{[T,t[}(\Psi^g_s)'A_{s+}^{-1}\Psi^g_s da^g_{s+}\right]\\
&\qquad \geq C\lim_{t\to\infty}\left[\int_{]T,t]}\frac{\|\Psi^r_s\|^2 da^r_s}{V_s} + \int_{[T,t[}\frac{\|\Psi^g_s\|^2 da^g_{s+}}{V_{s+}}\right]\\
&\qquad \geq C\lim_{t\to\infty}\left[\int_{]T,t]}\frac{dV^r_s}{V_s} + \int_{[T,t[}\frac{dV^g_{s+}}{V_{s+}}\right]
\end{aligned}$$

Assume that (7.33) is not true. Then with positive probability

$$\int_{]T,\infty]}\frac{dV^r_s}{V_s} < \infty, \quad \int_{[T,\infty[}\frac{dV^g_{s+}}{V_{s+}} < \infty.$$

From here it follows that

$$\lim_{t\to\infty}\frac{V^r_{t-}}{V_t} = 1, \quad \lim_{t\to\infty}\frac{V^g_t}{V_{t+}} = 1,$$

and there exists $T_1 > T$ that for all $t \geq T_1$

$$\frac{V^r_{t-}}{V_t} \geq 1/2, \quad \frac{V^g_t}{V_{t+}} \geq 1/2.$$

Thus,

$$\frac{V_{t-}}{V_t} \geq 1/2, \quad \frac{V_t}{V_{t+}} \geq 1/2.$$

By making use of these inequalities and Lemma 7.5.2 we obtain

$$\begin{aligned}
+\infty > \lim_{t\to\infty}\left[\int_{]T,t]}\frac{dV^r_s}{V_s} + \int_{[T,t[}\frac{dV^g_{s+}}{V_{s+}}\right] &\geq \lim_{t\to\infty}\left[\int_{]T_1,t]}\frac{dV^r_s}{V_{s-}}\frac{V_{s-}}{V_s} + \int_{[T,t[}\frac{dV^g_{s+}}{V_s}\frac{V_s}{V_{s+}}\right]\\
\geq 2^{-1}\lim_{t\to\infty}\left[\int_{]T_1,t]}\frac{dV^r_s}{V_{s-}} + \int_{[T,t[}\frac{dV^g_{s+}}{V_s}\right] &\geq 2^{-1}\lim_{t\to\infty}\ln\frac{V_t}{V_{T_1}}
\end{aligned}$$

Thus, with positive probability,

$$\lim_{t\to\infty}\ln\lambda_{min}(A_t) \leq \lim_{t\to\infty}\ln tr(A_t) = \lim_{t\to\infty}\ln V_t < +\infty.$$

This contradicts to the condition (A_4). ■

Lemma 7.5.5 *Under the assumptions* $(A_1)-(A_5)$ *the function* $Q(\tau_j)$ *in* (7.26) *satisfies the inequality*

$$\mathbf{E}Q_{\tau_j} \le C_j + p, \quad j \ge 1,$$

where the sequence $(C_j)_{j\ge1}$ *is the same as in* (7.23).

Proof. Let us introduce the processes

$$\begin{aligned} N_t =&(n_t^1,\dots,n_t^p)' = \int_{]0,t]} \Psi_s^r (W_s^r)^{1/2} dm_s^r + \int_{[0,t[} \Psi_s^g (W_s^g)^{1/2} dm_{s+}^g, \\ Z_t =&(N_t', (U_t^1)', \dots, (U_t^p)'), \\ F(Z_t) =&N_t' A_t^{-1} N_t, \end{aligned} \tag{7.38}$$

where U_t^i is the i-th column of the matrix A_t^{-1}. Note that Z_t is a $(p+1)p\times1$-dimensional semimartingale vector. In this notation we have

$$Q_{\tau_j} = F(Z_{\tau_j}).$$

Let us calculate the stochastic differential of the process $F(Z)$ by applying the change of variables formula. The process $F(Z_t)$ can be written as

$$\begin{aligned} F(Z_t) =&N_t'[U_t^1,\dots,U_t^p]N_t \\ =&(N_t'U_t^1,\dots,N_t'U_t^p)N_t \\ =&\sum_{i=1}^p N_t'U_t^i n_t^i. \end{aligned} \tag{7.39}$$

For the function $F:\mathbb{R}^d \to \mathbb{R}^1,\ d=p(p+1)$ and the semimartingale Z defined by

(7.38), the Ito formula has the form

$$
\begin{aligned}
F(Z_t) =& F(Z_T) + \int_{]T,t]} (\nabla_n F(Z_{s-}), dN^r_s) + \int_{[T,t[} (\nabla_n F(Z_s), dN^g_{s+}) \\
&+ \int_{]T,t]} \sum_{i=1}^{p} (\nabla_{u_i} F(Z_{s-}), dU^{ir}_s) \\
&+ \int_{[T,t[} \sum_{i=1}^{p} (\nabla_{u_i} F(Z_s), dU^{ig}_{s+}) \\
&+ \frac{1}{2} \int_{]T,t]} tr \left[\nabla_n \nabla_n F(Z_{s-}) \Psi^r_s (W^r_s)^{1/2} B^c_s (W^r_s)^{1/2} (\Psi^r_s)' \right] d\langle m^c \rangle_s \\
&+ \sum_{T<s\le t} \left[F(Z_s) - F(Z_{s-}) - (\nabla_n F(Z_{s-}), \Delta Z_s) \right] \\
&+ \sum_{T\le s<t} \left[F(Z_{s+}) - F(Z_s) - \big(\nabla_n F(Z_s), \Delta^+ Z_s\big) \right],
\end{aligned}
\tag{7.40}
$$

where $\nabla_n = (\frac{\partial}{\partial n_1}, \dots, \frac{\partial}{\partial n_p})'$, $(u,v) = v'u$ is a dot product of vectors u, v.

By (7.39) we obtain

$$
\begin{aligned}
\frac{\partial F}{\partial n_k} &= \sum_{i=1}^{p} \frac{\partial}{\partial n_k} [N'_t U^i_t] n^i_t + \sum_{i=1}^{p} N'_t U^i_t \delta^{ik} \\
&= \sum_{i=1}^{p} u^{ik}_t n^i_t + N'_t U^k_t,
\end{aligned}
$$

where $U^i_t = (u^{i1}_t, \dots, u^{ip}_t)'$. Therefore,

$$
\begin{aligned}
\nabla_n F(Z_t) &= \sum_{i=1}^{p} (u^{i1}_t, \dots, u^{ip}_t)' n^i_t + (N'_t U^1_t, \dots, N'_t U^p_t)' \\
&= \sum_{i=1}^{p} U^i_t n^i_t + A^{-1}_t N_t = 2A^{-1}_t Y_t.
\end{aligned}
\tag{7.41}
$$

Further, we have

$$
\nabla_{u^k} F(Z_t) = \sum_{i=1}^{p} [\nabla_{u^k} N'_t U^i_t] n^i_t = \sum_{i=1}^{p} N_t \delta^{ik} n^i_t = N_t n^k_t; \tag{7.42}
$$

$$
\begin{aligned}
\nabla_n\nabla_n F =& \left(\frac{\partial^2 F}{\partial n^k \partial n^j}\right)_{1\le k,j\le p};\\
\frac{\partial^2 F}{\partial n^k \partial n^j} =& \frac{\partial}{\partial n^j}\frac{\partial F(Z_t)}{\partial n^k} = \frac{\partial}{\partial n^j}\left[\sum_{i=1}^{p} u_t^{ik} n_t^i + N_t' U_t^k\right]\\
=& \sum_{i=1}^{p} u_t^{ik}\delta^{ij} + u_t^{kj} = 2u_t^{kj}.
\end{aligned}
$$

Thus,

$$
\nabla_n\nabla_n F = 2A_t^{-1}; \tag{7.43}
$$

Combining (7.40)-(7.43), it follows that

$$
F(Z_t) = F(Z_T) + 2(\mu_t^r + \mu_t^g) + I_t^1 + I_t^2 + I_t^3 + I_t^4 + I_t^5,
$$

where

$$
\begin{aligned}
\mu_t^r =& \int_{]T,t]} \left((A_{s-})^{-1}N_{s-}, \Psi_s^r (W_s^r)^{1/2} dm_s^r\right),\\
\mu_t^g =& \int_{[T,t[} \left((A_s)^{-1}N_s, \Psi_s^g (W_s^g)^{1/2} dm_{s+}^g\right)\\
I_t^1 =& \int_{]T,t]} N_{s-} d(A_s^r)^{-1} N_{s-}\\
I_t^2 =& \int_{[T,t[} N_s d(A_{s+}^g)^{-1} N_s\\
I_t^3 =& tr\int_{]T,t]} A_{s-}^{-1}\Psi_s^r (W_s^r)^{1/2} B_s^c (W_s^r)^{1/2} (\Psi_s^r)' d\langle m^c\rangle_s\\
I_t^4 =& \sum_{T<s\le t}\left[F(Z_s) - F(Z_{s-}) - 2N_{s-}' A_{s-}^{-1}\Delta N_s - N_{s-}'\Delta A_s^{-1} N_{s-}\right]\\
I_t^5 =& \sum_{T\le s<t}\left[F(Z_{s+}) - F(Z_s) - 2N_s' A_s^{-1}\Delta^+ N_s - N_s'\Delta^+ A_s^{-1} N_s\right]
\end{aligned}
$$

In order to study I_t^1 in we need to find the differential for $(A_t^r)^{-1}$. We have

$$
d\left[A_t^r (A_t^r)^{-1}\right] = [dA_t^r](A_{t-}^r)^{-1} + A_{t-}^r d(A_t^r)^{-1} = 0,
$$

Hence,

$$
d(A_t^r)^{-1} = -(A_{t-}^r)^{-1}[dA_t^r](A_{t-}^r)^{-1} = -(A_{t-}^r)^{-1}[\Psi_t^r(\Psi_t^r)'](A_{t-}^r)^{-1} da_t^r.
$$

Consequently,

$$
\begin{aligned}
I_t^1 &= \int_{]T,t]} N_{s-} d(A_s^r)^{-1} N_{s-} \\
&= -\int_{]T,t]} N_{s-} (A_{s-}^r)^{-1} [\Psi_s^r (\Psi_s^r)'] (A_{s-}^r)^{-1} N_{s-} da_t^c \\
&\quad + \sum_{T<s\leq t} N_{s-}' \Delta (A_s^r)^{-1} N_{s-}.
\end{aligned} \tag{7.44}
$$

The matrix $(A_s^r)^{-1}$ is non-increasing because the matrix A_s^r is non-decreasing. Therefore, the matrix $\Delta(A_s^r)^{-1} \leq 0$, and right-hand side of (7.44) is non-positive. Thus,

$$I_t^1 \leq 0 \quad a.s. \text{ for all } t \geq T.$$

Since A_t^g is a pure-jump process and also non-increasing, we have

$$I_t^2 = \int_{[T,t[} N_s' d(A_{s+}^g)^{-1} N_s = \sum_{[T,t[} N_s' \Delta (A_{s+}^g)^{-1} N_s \leq 0.$$

Consider the term I_t^4. We have

$$
\begin{aligned}
I_t^4 &= \sum_{T<s\leq t} \left[N_s' A_s^{-1} N_s - N_{s-}' A_{s-}^{-1} N_{s-} - 2N_{s-}' A_{s-}^{-1} \Delta N_s - N_{s-}' \Delta A_s^{-1} N_{s-} \right] \\
&= \sum_{T<s\leq t} \left[\Delta N_s' A_s^{-1} \Delta N_s \right] + 2\nu_t^r,
\end{aligned}
$$

where

$$\nu_t^r = \sum_{T<s\leq t} \left[N_{s-}' \Delta A_s^{-1} \Delta N_s \right]$$

Similarly,

$$
\begin{aligned}
I_t^5 &= \sum_{T\leq s<t} \left[\Delta^+ N_s' A_{s+}^{-1} \Delta^+ N_s \right] + 2\nu_t^g, \\
\nu_t^g &= \sum_{T\leq s<t} \left[N_s' \Delta^+ A_s^{-1} \Delta^+ N_s \right].
\end{aligned}
$$

From the obtained estimates for $I_t^1 - I_t^5$, it follows that

$$F(Z_t) \leq F(Z_T) + I_t^3 + 2(\mu_t^r + \mu_t^g) + 2(\nu_t^r + \nu_t^g) + \delta_t^r + \delta_t^g + D_t^r + D_t^g, \tag{7.45}$$

where

$$\begin{aligned} \delta_t^r &= \sum_{T<s\leq t} \left[\Delta N_s' A_s^{-1} \Delta N_s\right] - D_t^r, \\ \delta_t^g &= \sum_{T\leq s<t} \left[\Delta^+ N_s' A_{s+}^{-1} \Delta^+ N_s\right] - D_t^g, \end{aligned}$$

δ_t^r and δ_t^g are local optional martingales, and D_t^r and D_t^g are, respectively, the increasing predictable and optional processes in the Doob-Meyer decomposition of the submartingales $\sum_{T<s\leq t} [\Delta N_s' A_s^{-1} \Delta N_s]$ and $\sum_{T\leq s<t} \left[\Delta^+ N_s' A_{s+}^{-1} \Delta^+ N_s\right]$:

$$\begin{aligned} D_t^r &= \int_{]T,t]} tr[(W_s^r)^{1/2} (\Psi_s^r)' A_s^{-1} \Psi_s^r (W_s^r)^{1/2} B_s^d] d\langle m^d \rangle_s \\ &= \int_{]T,t]} tr[A_s^{-1} \Psi_s^r (W_s^r)^{1/2} B_s^d (W_s^r)^{1/2} (\Psi_s^r)'] d\langle m^d \rangle_s, \\ D_t^g &= \int_{[T,t[} tr[(W_s^g)^{1/2} (\Psi_s^g)' A_{s+}^{-1} \Psi_s^g (W_s^g)^{1/2} B_s^d] d\langle m^g \rangle_{s+} \\ &= \int_{[T,t[} tr[A_{s+}^{-1} \Psi_s^g (W_s^g)^{1/2} B_s^g (W_s^g)^{1/2} (\Psi_s^g)'] d\langle m^g \rangle_{s+}. \end{aligned}$$

The process D_t^r is well-defined, because by conditions (A_2) and (A_3) we have

$$\begin{aligned} D_t^r \leq& \int_{]T,t]} tr[A_s^{-1} \Psi_s^r (\Psi_s^r)'] d\langle m^d \rangle_s \\ \leq& \int_{]T,t]} tr[A_s^{-1} \Psi_s^r (\Psi_s^r)'] da_s^r \\ \leq& tr[A_T^{-1}] \int_{]T,t]} \|\Psi_s^r (\Psi_s^r)'\| \, da_s^r \\ \leq& tr[A_T^{-1}] \int_{]T,t]} tr \Psi_s^r (\Psi_s^r)' da_s^r < \infty \end{aligned}$$

for all $t \geq T$. The same holds for the process D_t^g.

Let us verify that the processes $\mu^j, \nu^j,\ j = r, g$ are locally square intagrable optional martingales. Their predictable quadratic variations are given by the formulae

$$\langle \mu^r \rangle_t = \int_{]T,t]} N'_{s-}(A_{s-})^{-1}\Psi^r_s(W^r_s)^{1/2}B^r_s(W^r_s)^{1/2}(\Psi^r_s)'(A_{s-})^{-1}N_{s-}da^r_s,$$

$$\langle \mu^g \rangle_t = \int_{[T,t[} N'_s(A_s)^{-1}\Psi^g_s(W^g_s)^{1/2}B^g_s(W^g_s)^{1/2}(\Psi^g_s)'(A_s)^{-1}N_s da^g_{s+},$$

$$\langle \nu^r \rangle_t = \int_{]T,t]} N'_{s-}\Delta(A_{s-})^{-1}\Psi^r_s(W^r_s)^{1/2}B^d_s(W^r_s)^{1/2}(\Psi^r_s)'\Delta(A_{s-})^{-1}N_{s-}d\langle m^d\rangle_s,$$

$$\langle \nu^g \rangle_t = \int_{[T,t[} N'_s\Delta^+(A_s)^{-1}\Psi^g_s(W^g_s)^{1/2}B^g_s(W^g_s)^{1/2}(\Psi^g_s)'\Delta^+(A_s)^{-1}N_s d\langle m^g\rangle_{s+}.$$

By condition (A_2) and the Cauchy-Schwartz inequality,

$$\begin{aligned}
\langle \nu^r \rangle_t \leq & \int_{]T,t]} N'_{s-}\Delta(A_s)^{-1}\Psi^r_s(\Psi^r_s)'\Delta(A_s)^{-1}N'_{s-}da^r_s \\
= & \int_{]T,t]} (N'_{s-}\Delta(A_s)^{-1}\Psi^r_s)^2 da^r_s \\
\leq & \int_{]T,t]} \left\|\Delta(A_s)^{-1/2}N_{s-}\right\|^2 \left\|\Delta(A_s)^{-1/2}\Psi^r_s\right\|^2 da^r_s \\
\leq & (trA_T^{-1})^2 \int_{[T,t[} \|N_{s-}\|^2 \left\|\Psi^r_s(\Psi^r_s)'\right\| da^r_s < \infty \quad a.s.
\end{aligned}$$

for all $t \geq T$ due to condition (A_3) and left-continuity of N_{s-} and, similarly,

$$\begin{aligned}
\langle \nu^g \rangle_t \leq & \int_{[T,t[} N'_s\Delta^+(A_s)^{-1}\Psi^g_s(\Psi^g_s)'\Delta^+(A_s)^{-1}N'_s da^g_{s+} \\
= & \int_{[T,t[} (N'_s\Delta^+(A_s)^{-1}\Psi^g_s)^2 da^g_{s+} \\
\leq & \int_{[T,t[} \left\|\Delta^+(A_s)^{-1/2}N_s\right\|^2 \left\|\Delta^+(A_s)^{-1/2}\Psi^g_s\right\|^2 da^g_{s+} \\
\leq & (trA_T^{-1})^2 \int_{[T,t[} \|N_s\|^2 \left\|\Psi^g_s(\Psi^g_s)'\right\| da^g_{s+}
\end{aligned}$$

for all $t \geq T$. By Doob's inequality, conditions (A_2) and (A_3)

$$\begin{aligned}
\mathbf{E} \sup_{T \leq s < t} \|N_s\|^2 \leq & 4\mathbf{E}\|N_s\|^2 \\
= & 4\mathbf{E}\int_{]0,t]} tr[\Psi^r_s(\Psi^r_s)']da^r_s + \int_{[0,t[} tr[\Psi^g_s(\Psi^g_s)']da^g_{s+} \\
< & \infty.
\end{aligned}$$

Therefore, by (A_3)

$$\int_{[T,t[} \|N_s\|^2 \|\Psi^g_s(\Psi^g_s)'\| da^g_{s+} < \infty \quad a.s.$$

In a similar way one can verify that $\langle \mu^j \rangle_t < \infty,\ j = r, g$, a.s. for all $t \geq T$.

Since the process $D^r_t + \langle \mu^r \rangle_t + \langle \nu^r \rangle_t + D^g_t + \langle \mu^g \rangle_t + \langle \nu^g \rangle_t$ is strongly predictable, there exists a sequence of stopping times $\sigma^r_k, \sigma^r_k \uparrow \infty$ a.s. as $k \to \infty$, such that for any k the stopped process $D^r_{\sigma^r_k} + \langle \mu^r \rangle_{\sigma^r_k} + \langle \nu^r \rangle_{\sigma^r_k} + D^g_{\sigma^r_k} + \langle \mu^g \rangle_{\sigma^r_k} + \langle \nu^g \rangle_{\sigma^r_k}$ is bounded (see Lemma 1.7, [32]).

Then from (7.45) and condition (A_2) we obtain, for $k > 0$,

$$\begin{aligned}
\mathbf{E}F(Z_{t\wedge\tilde{\sigma}^r_k\wedge\sigma^g_k}) \leq & \mathbf{E}\Bigg[F(Z_T) + \int_{]T,t]} trA^{-1}_{s-}\Psi^r_s(W^r_s)^{1/2}B^c_s(W^r_s)^{1/2}(\Psi^r_s)'d\langle m^c\rangle_s \\
& + \int_{]T,t]} tr[A^{-1}_s\Psi^r_s(W^r_s)^{1/2}B^d_s(W^r_s)^{1/2}(\Psi^r_s)']d\langle m^d\rangle_s \\
& + \int_{[T,t[} tr[A^{-1}_{s+}\Psi^g_s(W^g_s)^{1/2}B^g_s(W^g_s)^{1/2}(\Psi^g_s)']d\langle m^g\rangle_{s+}\Bigg] \\
\leq & \mathbf{E}\Bigg[F(Z_T) + tr\int_{]T,t]} A^{-1}_{s-}\Psi^r_s(\Psi^r_s)'d\langle m^c\rangle_s \\
& + \int_{]T,t]} tr[A^{-1}_s\Psi^r_s(\Psi^r_s)']d\langle m^d\rangle_s \\
& + \int_{[T,t[} tr[A^{-1}_{s+}\Psi^g_s(\Psi^g_s)']d\langle m^g\rangle_{s+}\Bigg].
\end{aligned}$$

Letting $t = \tau_j$, taking the limit as $k \to \infty$ and applying the monotone convergence theorem, we obtain

$$\begin{aligned}
\mathbf{E}Q_{\tau_j} = & \mathbf{E}N'_{\tau_j}A^{-1}_{\tau_j}N_{\tau_j} \\
\leq & \mathbf{E}\Bigg[N'_TA^{-1}_TN_T + \int_{]T,\tau_j]} trA^{-1}_s\Psi^r_s(\Psi^r_s)'da^r_s + \int_{[T,\tau_j[} tr[A^{-1}_{s+}\Psi^g_s(\Psi^g_s)']da^g_{s+}\Bigg].
\end{aligned}$$

Now we can estimate $\mathbf{E}Q_{\tau_j}$. We have

$$\begin{aligned}\mathbf{E}Q_{\tau_j} \leq & \mathbf{E}\Bigg[N_T' A_T^{-1} N_T + \int_{]T,\tau_j[} tr(\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^r + tr(\Psi_s^r)' A_s^{-1} \Psi_s^r \Delta a_s^r \\ & + \int_{[T,\tau_j[} tr[(\Psi_s^g)' A_{s+}^{-1} \Psi_s^g] da_{s+}^g \Bigg].\end{aligned} \tag{7.46}$$

By the definition of wide sense stopping time T in (7.22), condition (A_2) and orthogonality of N_T^r and N_T^g,

$$\begin{aligned}\mathbf{E}N_T' A_T^{-1} N_T \leq & \mathbf{E}\lambda_{max}(A_T^{-1}) N_T' N_T \\ = & \mathbf{E}\lambda_{min}^{-1}(A_T^{-1}) \|N_T\|^2 \\ \leq & C_0^{-1} \mathbf{E} tr\left[(N_T^r)' N_T^r + (N_T^r)' N_T^g + (N_T^g)' N_T^r + (N_T^g)' N_T^g)\right] \\ = & C_0^{-1} \mathbf{E}\left[\int_{]0,T]} \|\Psi_s^r\|^2 da_s^r + \int_{[0,T[} \|\Psi_s^g\|^2 da_{s+}^g\right].\end{aligned}$$

Combining this inequality and (7.46) yields

$$\begin{aligned}\mathbf{E}Q_{\tau_j} \leq & \mathbf{E}\Bigg[C_0^{-1}\left(\int_{]0,T]} \|\Psi_s^r\| da_s^r + \int_{[0,T[} \|\Psi_s^g\| da_{s+}^g\right) \\ & \int_{]T,\tau_j[} tr(\Psi_s^r)' A_s^{-1} \Psi_s^r da_s^r + tr(\Psi_{\tau_j}^r)' A_{\tau_j}^{-1} \Psi_{\tau_j}^r \Delta a_{\tau_j}^r \\ & + \int_{[T,\tau_j[} tr[(\Psi_s^g)' A_{s+}^{-1} \Psi_s^g] da_{s+}^g \Bigg] \\ \leq & C_j + \mathbf{E}\left[tr(\Psi_{\tau_j}^r)' A_{\tau_j}^{-1} \Psi_{\tau_j}^r \Delta a_{\tau_j}^r\right].\end{aligned} \tag{7.47}$$

By Lemma 7.5.1,

$$\begin{aligned}tr A_{\tau_j}^{-1} \Psi_{\tau_j}^r (\Psi_{\tau_j}^r)' \Delta a_{\tau_j}^r \leq & r[|A_{\tau_j}| - |A_{\tau_j} - \Psi_{\tau_j}^r (\Psi_{\tau_j}^r)' \Delta a_{\tau_j}^r|]/[|A_{\tau_j}|] \\ \leq & r \leq p,\end{aligned}$$

where r is the rank of the matrix $\Psi_s^r (\Psi_s^r)'$. By substituting this estimate in (7.47), we come to the assertion of the Lemma. ■

Chapter 8

Conclusion

In this thesis we obtained several new results advancing the theory of optional processes and successfully applied this new theory in the areas such as mathematical finance, risk theory, and statistics. In particular, the following results are obtained:

- different versions of the comparison theorem and also a uniqueness theorem for a general class of optional stochastic differential equations were stated and proved. Furthermore, these results were applied to the pricing of financial derivatives.

- the so-called Krylov estimates for distributions of stochastic integrals using L_d-norm of a measurable function were generalized for optional semimartingales. Corresponding applications of this result were illustrated.

- a very general optional semimartingale risk model for the capital process of a company was introduced and exhaustively investigated. A general approach to the calculation of ruin probabilities of such models was shown and supported by diverse examples.

- an optional semimartingale regression model with a one-dimensional unknown parameter was introduced. The strong consistency of structural least squares estimates and the property of fixed accuracy of sequential least squares estimates were proved.

- a general optional semimartingale regression model with a multi-dimensional unknown parameter was introduced. The strong consistency of structural least squares estimates was proved. The property of fixed accuracy of sequential least squares estimates was proved for the multivariate optional regression models with the number of parameters less than or equal to the dimension of the observation process and for the general case with an arbitrary number of parameters.

References

[1] O. Aalen, "Nonparametric inference for a family of counting processes," *The Annals of Statistics*, pp. 701–726, 1978.

[2] O. Aalen, "A model for nonparametric regression analysis of counting processes," in *Mathematical statistics and probability theory*, Springer, 1980, pp. 1–25.

[3] M. Abdelghani, A. Melnikov, and A. Pak, "On krylov's estimates for optional semimartingales," *Random Operators and Stochastic Equations*, 2021.

[4] M. Abdelghani and A. V. Melnikov, *Optional Processes*. Chapman & Hall/CRC, 2020.

[5] M. Abdelghani, A. V. Melnikov, and A. A. Pak, "An optional semimartingale approach to risk theory," Submitted, 2020.

[6] M. Abdelghani, A. V. Melnikov, and A. A. Pak, "On comparison theorem for optional sdes via local times and applications," *Stochastics.*, (in press).

[7] M. N. Abdelghani and A. V. Melnikov, "Financial markets in the context of the general theory of optional processes," in *Mathematical and Computational Approaches in Advancing Modern Science and Engineering*, Springer, 2016, pp. 519–528.

[8] M. Abdelghani and A. Melnikov, "Optional decomposition of optional supermartingales and applications to filtering and finance," *Stochastics*, vol. 91, no. 6, pp. 797–816, 2019.

[9] M. Abdelghani, A. Melnikov, and A. Pak, "On statistical estimation and inferences in optional regression models," *Statistics: A Journal of Theoretical and Applied Statistics*, pp. 1–13, 2021.

[10] M. Abdelghani and A. V. Melnikov, "On linear stochastic equations of optional semimartingales and their applications," *Statistics & Probability Letters*, vol. 125, pp. 207–214, 2017.

[11] M. Abdelghani and A. V. Melnikov, "A comparison theorem for stochastic equations of optional semimartingales," *Stochastics and Dynamics*, vol. 18, no. 04, p. 1 850 029, Aug. 2018, ISSN: 0219-4937. DOI: 10.1142/S0219493718500296.

[12] M. N. Abdelghani and A. V. Melnikov, "Optional defaultable markets," *Risks*, vol. 5, no. 4, p. 56, 2017.

[13] M. N. Abdelghani and A. V. Melnikov, "Existence and uniqueness of stochastic equations of optional semimartingales under monotonicity condition," *Stochastics*, vol. 92, no. 1, pp. 67–89, 2020.

[14] S. Anulova and H. Pragarauskas, "Weak markov solutions of stochastic equations," *Lithuanian Mathematical Journal*, vol. 17, no. 2, pp. 141–155, 1977.

[15] S. Asmussen and H. Albrecher, *Ruin probabilities*. World scientific Singapore, 2010, vol. 14.

[16] M. Barlow and E. Perkins, "One-dimensional stochastic differential equations involving a singular increasing process," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 12, no. 3-4, pp. 229–249, 1984.

[17] M. Benabdallah, S. Bouhadou, and Y. Ouknine, "On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations with jumps," in *Statistical Methods and Applications in Insurance and Finance*, Springer, 2013, pp. 193–210.

[18] J. Bishwal, "Sequential maximum likelihood estimation in semimartingales," *Journal of Statistics and Applications*, vol. 1, no. 2-4, pp. 15–25, 2006.

[19] J. Bishwal, *Parameter Estimation in Stochastic Differential Equations*. Springer, Jan. 2008, vol. 1923, ISBN: 978-3-540-74447-4. DOI: `10.1007/978-3-540-74448-1`.

[20] J. P. Bishwal, "Sequential maximum likelihood estimation in nonlinear non-markov diffusion type processes," *Dynamic Systems and Applications*, vol. 27, no. 1, pp. 107–124, 2018.

[21] J. Cai and D. C. Dickson, "Upper bounds for ultimate ruin probabilities in the sparre andersen model with interest," *Insurance: Mathematics and Economics*, vol. 32, no. 1, pp. 61–71, 2003.

[22] N. Christopeit, "Quasi-least-squares estimation in semimartingale regression models," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 16, no. 3-4, pp. 255–278, 1986.

[23] J. C. Cox and S. A. Ross, "The valuation of options for alternative stochastic processes," *Journal of Financial Economics*, vol. 3, no. 1-2, pp. 145–166, Jan. 1976, ISSN: 0304-405X. DOI: `10.1016/0304-405X(76)90023-4`.

[24] A. Dassios and P. Embrechts, "Martingales and insurance risk," *Communications in Statistics. Stochastic Models*, vol. 5, no. 2, pp. 181–217, 1989.

[25] F. Delbaen and J. Haezendonck, "A martingale approach to premium calculation principles in an arbitrage free market," *Insurance: Mathematics and Economics*, vol. 8, no. 4, pp. 269–277, 1989.

[26] C. Dellacherie and P.-A. Meyer, "Un nouveau théoreme de projection et de section," in *Séminaire de Probabilités IX Université de Strasbourg*, Springer, 1975, pp. 239–245.

[27] C. Dellacherie and P.-A. Meyer, "Probabilités et potentiels 2nd ed," *Hermann, Paris)(English transl.: North-Holland, 1978)*, 1976.

[28] P. Embrechts and H. Schmidli, "Ruin estimation for a general insurance risk model," *Advances in applied probability*, vol. 26, no. 2, pp. 404–422, 1994.

[29] L. I. Gal'chuk and M. Davis, "A note on a comparison theorem for equations with different diffusions," *Stochastics*, vol. 6, no. 2, pp. 147–149, 1982.

[30] L. I. Gal'chuk, "On the Existence of Optional Modifications for Martingales," *Theory of Probability & Its Applications*, vol. 22, no. 3, pp. 572–573, Jun. 1978, ISSN: 0040-585X. DOI: 10.1137/1122068.

[31] L. I. Gal'chuk, "A Comparison Theorem for Stochastic Equations with Integrals with Respect to Martingales and Random Measures," *Theory of Probability & Its Applications*, vol. 27, no. 3, pp. 450–460, Mar. 1983, ISSN: 0040-585X. DOI: 10.1137/1127055.

[32] L. I. Gal'chuk, "Stochastic Integrals with Respect to Optional Semimartingales and Random Measures," *Theory of Probability & Its Applications*, vol. 29, no. 1, pp. 93–108, Mar. 1985. DOI: 10.1137/1129008.

[33] L. I. Galchuk, "Optional martingales," *Matematicheskii Sbornik*, vol. 154, no. 4, pp. 483–521, 1980.

[34] L. Galtchouk and V. Konev, "On sequential estimation of parameters in semimartingale regression models with continuous time parameter," *Annals of statistics*, vol. 29, pp. 1508–1536, 2001.

[35] K. Gasparyan, "On asymptotic behavior of optional martingales," *Dokl. Nats. Akad. Nauk Armen.*, vol. 93, no. 5, pp. 212–219, 1992.

[36] K. Gasparyan, "Law of the iterated logarithm for optional martingales," *Izv. Nats. Akad. Nauk Armenii Mat.*, vol. 29, no. 3, pp. 33–53, 1994.

[37] K. Gasparyan, "On change of variables formula for optional semimartingales," *Uch. Zap. Erevan. Gos. Univ.*, vol. 1, pp. 14–25, 1984.

[38] K. Gasparyan, "Stochastic equations with respect to optional semimartingales," *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, vol. 12, pp. 57–60, 1985.

[39] K. Gasparyan, "Some Questions on Theory of Optional Semimartingales. Dissertation.," *V.Kapsukas Vilnius State University Council*, 1988.

[40] H. U. Gerber, "Martingales in risk theory," *Bulletin of the Swiss Association of Actuaries*, pp. 205–216, 1973.

[41] J. Grandell, "Finite time ruin probabilities and martingales," *Informatica*, vol. 2, no. 1, pp. 3–32, 1991.

[42] J. Grandell, *Aspects of risk theory*. Springer Science & Business Media, 2012.

[43] M. Grigorova, P. Imkeller, E. Offen, Y. Ouknine, M.-C. Quenez, *et al.*, "Reflected bsdes when the obstacle is not right-continuous and optimal stopping," *Annals of Applied Probability*, vol. 27, no. 5, pp. 3153–3188, 2017.

[44] M. Grigorova, P. Imkeller, Y. Ouknine, and M.-C. Quenez, "Optimal stopping with f-expectations: The irregular case," *Stochastic Processes and their Applications*, vol. 130, no. 3, pp. 1258–1288, 2020.

[45] I. Gyöngy and N. V. Krylov, "On stochastic equations with respect to semimartingales I.," *Stochastics*, vol. 4, no. 1, pp. 1–21, Jan. 1980. DOI: `10.1080/03610918008833154`.

[46] J. Horowitz, "Optional supermartingales and the Andersen-Jessen theorem," *Probability Theory and Related Fields*, vol. 43, no. 3, pp. 263–272, 1978. DOI: `10.1007/BF00536207`.

[47] M. Jacobsen, "The time to ruin in some additive risk models with random premium rates," *Journal of Applied Probability*, vol. 49, no. 4, pp. 915–938, 2012.

[48] J. Jacod and A. Shiryaev, *Limit theorems for stochastic processes*. Springer Science & Business Media, 2013, vol. 288.

[49] I. Jarni and Y. Ouknine, "On reflection with two-sided jumps," *Journal of Theoretical Probability*, pp. 1–20, 2020.

[50] M. Kirch and A. Melnikov, "Efficient hedging and pricing of life insurance policies in a jump-diffusion model," *Stochastic analysis and applications*, vol. 23, no. 6, pp. 1213–1233, 2005.

[51] Y. Kozachenko, A. Melnikov, and Y. Mishura, "On drift parameter estimation in models with fractional brownian motion," *Statistics: A Journal of Theoretical and Applied Statistics*, vol. 49, no. 1, pp. 35–62, 2015.

[52] V. Krasin, I. Smirnov, and A. V. Melnikov, "Approximate option pricing and hedging in the CEV model via path-wise comparison of stochastic processes," *Annals of Finance*, vol. 14, no. 2, pp. 195–209, May 2018, ISSN: 1614-2446. DOI: `10.1007/s10436-017-0309-9`.

[53] V. Y. Krasin and A. V. Melnikov, "On Comparison Theorem and its Applications to Finance," *Optimality and Risk - Modern Trends in Mathematical Finance*, pp. 171–181, 2009. DOI: `10.1007/978-3-642-02608-9{\_}8`.

[54] R. Krutchenko and A. V. Mel′nikov, "Quantile hedging for a jump-diffusion finanaicl market model," in *Mathematical finance*, Springer, 2001, pp. 215–229.

[55] N. V. Krylov, "On Ito stochastic integral equations," *Teor. Veroyatnost. i Primen.*, vol. 14, no. 2, pp. 340–348, 1969.

[56] N. V. Krylov, "Some estimates for the density of the distribution of a stochastic integral," *Izv. Akad. Nauk SSSR Ser. Mat*, vol. 38, pp. 228–248, 1974.

[57] N. V. Krylov, *Controlled diffusion processes*. Springer Science & Business Media, 2008, vol. 14.

[58] C. Kühn, M. Stroh, *et al.*, "A note on stochastic integration with respect to optional semimartingales," *Electronic Communications in Probability*, vol. 14, pp. 192–201, 2009.

[59] G. S. Ladde, V. Lakshmikantham, and A. S. Vatsala, *Monotone iterative techniques for nonlinear differential equations*. Pitman, 1985, p. 236, ISBN: 027308707X.

[60] A. Le Breton and M. Musiela, "Strong consistency of least squares estimates in linear regression models driven by semimartingales," *Journal of multivariate analysis*, vol. 23, no. 1, pp. 77–92, 1987.

[61] J.-F. Le Gall, "Applications du temps local aux équations différentielles stochastiques unidimensionnelles," in *Séminaire de Probabilités XVII 1981/82*, Springer, 1983, pp. 15–31.

[62] E. Lenglart, "Tribus de meyer et théorie des processus," in *Séminaire de Probabilités XIV 1978/79*, Springer, 1980, pp. 500–546.

[63] D. Lepingle, "Sur la représentation des sauts des martingales," in *Séminaire de Probabilités XI*, Springer, 1977, pp. 418–434.

[64] J. Li, V. Todorov, and G. Tauchen, "Robust jump regressions," *Journal of the American Statistical Association*, vol. 112, no. 517, pp. 332–341, 2016.

[65] R. S. Liptser and A. N. Shiryaev, *Theory of Martingales*. Springer Netherlands, 1989.

[66] R. S. Liptser, "A strong law of large numbers for local martingales," *Stochastics*, vol. 3, no. 1-4, pp. 217–228, 1980.

[67] R. S. Liptser and A. Shiryaev, *Statistics of Random Processes, I–II. Engl. trans.* Springer-Verlag, Berlin, 1977.

[68] A. V. Mel'nikov, "On the theory of stochastic equations in components of semimartingales," *Mathematics of the USSR-Sbornik*, vol. 38, no. 3, pp. 381–394, 1981.

[69] A. V. Mel'nikov, "Stochastic equations and Krylov's estimates for semimartingales," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 10, no. 2, pp. 81–102, 1983.

[70] A. V. Mel'nikov, "On regression models with non-square integrable martingale-like errors," *Publications mathématiques et informatique de Rennes*, no. 1, pp. 97–107, 1987.

[71] A. V. Mel'nikov, "Stochastic differential equations: Singularity of coefficients, regression models, and stochastic approximation," *Russian Mathematical Surveys*, vol. 51, no. 5, p. 819, 1996.

[72] A. V. Mel'nikov and A. A. Novikov, "Sequential inferences with guaranteed accuracy for semimartingales," *Prob. Theory Appl.*, vol. 33, pp. 480–494, 1988.

[73] A. V. Melnikov, "The law of large numbers for multidimensional martingales," in *Soviet Math. Dokl.*, vol. 33, 1986, pp. 131–135.

[74] A. V. Melnikov and A. Pak, "Parameter estimation in optional semimartingale regression models," *Submitted*, 2021.

[75] A. V. Melnikov, *Risk analysis in finance and insurance*. Chapman and Hall/CRC, 2011.

[76] S. M. Nikol'skii, *Approximation of functions of several variables and embedding theorems*. Springer Science & Business Media, 2012, vol. 205.

[77] A. Novikov, "Consistency of least squares estimates in regression models with martingale errors," *Statistics and Control of Stochastic Processes (Moscow, 1984), Transl. Ser. Math. Engrg., Optimization Software, New York*, pp. 389–409, 1985.

[78] G. O'brien, "A new comparison theorem for solutions of stochastic differential equations," *Stochastics*, vol. 3, no. 1-4, pp. 245–249, Jan. 1980, ISSN: 0090-9491. DOI: 10.1080/17442508008833148.

[79] E. Perkins, "Local time and pathwise uniqueness for stochastic differential equations," in *Séminaire de Probabilités XVI 1980/81*, Springer, 1982, pp. 201–208.

[80] T. Rolski, H. Schmidli, V. Schmidt, and J. L. Teugels, *Stochastic processes for insurance and finance*. John Wiley & Sons, 2009, vol. 505.

[81] S. Rong, *Theory of stochastic differential equations with jumps and applications: mathematical and analytical techniques with applications to engineering.* Springer Science & Business Media, 2006.

[82] L. Schwartz, *Mathematical Analysis*. (Russian) Mir, Moscow, 1972, vol. 1.

[83] S. L. Sobolev, *Some applications of functional analysis in mathematical physics.* American Mathematical Soc., 2008, vol. 90.

[84] M. Sørensen, "A semimartingale approach to some problems in risk theory," *ASTIN Bulletin: The Journal of the IAA*, vol. 26, no. 1, pp. 15–23, 1996.

[85] J. Spielmann, “Classification of the Bounds on the Probability of Ruin for Lévy Processes with Light-tailed Jumps,” working paper or preprint, Feb. 2018, [Online]. Available: `https://hal.archives-ouvertes.fr/hal-01597828`.

[86] T. Sriram and R. Iaci, “Editor’s special invited paper: Sequential estimation for time series models,” *Sequential Analysis*, vol. 33, no. 2, pp. 136–157, 2014.

[87] C. Stricker, “Les intervalles de constance de $< x, x >$,” *Séminaire de probabilités de Strasbourg*, vol. 16, pp. 219–220, 1982.

[88] V. Todorov, “Estimation of continuous-time stochastic volatility models with jumps using high-frequency data,” *Journal of Econometrics*, vol. 148, no. 2, pp. 131–148, 2009.

[89] T. Yamada, “On a comparison theorem for solutions of stochastic differential equations and its applications,” *J. Math. Kyoto Univ.*, vol. 13, pp. 497–512, 1973.

[90] J. A. Yan, “A comparison theorem for semimartingales and its applications,” *Séminaire de Probabilités XX*, pp. 349–351, 1986.

[91] G. Yuan, “Semi-discretization Algorithm for Option Pricing in CEV Jump-diffusion Model,” *Rev. Téc. Ing. Univ. Zulia*, vol. 39, p. 44, 2016. DOI: `10.21311/001.39.3.7`.

[92] Q. Zang and T. Wang, “Sequential estimation for nonhomogeneous ornstein-uhlenbeck processes,” *Communications in Statistics-Theory and Methods*, vol. 48, no. 8, pp. 1955–1962, 2019.