2013 (Heisei 25)
Doctoral Thesis

A study on the density and sensitivity analysis concerning the maximum of SDEs

Doctoral Program in Integrated Science and Engineering
Graduate School of Science and Engineering

Tomonori Nakatsu

A study on the density and sensitivity analysis concerning the maximum of SDEs

Tomonori Nakatsu
Department of Mathematical Science
Graduate School of Science and Engineering
Ritsumeikan University

Abstract

In this thesis, we shall give some results on the existence of the density function and sensitivity analysis concerning the maximum of some stochastic differential equations (SDEs, in short). The Malliavin calculus (or stochastic calculus of variations) plays an important role to obtain the results of this thesis.

In Chapter 1, we present the introduction of this thesis and the preliminary of Malliavin calculus.

In Chapter 2, we consider an m-dimensional SDE with coefficients which depend on the maximum of the solution. First, we prove the absolute continuity of the law of the solution. Then we prove that the joint law of the maximum of the ith component of the solution and the i'th component of the solution is absolutely continuous with respect to the Lebesgue measure in a particular case.

In Chapter 3, we give a decomposition formula to calculate the vega index (the sensitivity of an option contract with respect to changes in volatility) for options depending on the extrema (maximum or minimum) of a general one-dimensional model and study its behavior. Moreover, we compare the vega index obtained in this one-dimensional model with the one in the Black-Scholes model. Our mathematical and numerical results provide mainly three interesting properties of the vega index for barrier type options in the one-dimensional model: First, the vega index can be decomposed into three components which can be called extrema sensitivity, terminal (feature) sensitivity and drift sensitivity. Second, by using an example of up-in call options, we show that there is a barrier value at which the importance of extrema and terminal sensitivity are reversed. Third, extrema sensitivity is important only for options with short maturity as far as the vega index is concerned. The comparison of the vega index in two different models clarifies that the behavior of the vega index in the one-dimensional model considered in this thesis is far away from that in the Black-Scholes model. In the case of binary barrier options, each component of the decomposition formula for the vega index involves the Dirac delta functionals. Kernel methods are used in order to estimate the vega index in this setting.

Acknowledgements

I would like to express my deepest thanks to Professor Arturo Kohatsu-Higa, for his guidance and support from the summer of 2005 when we met for the first time. He has been taking care of me in terms of about studies and researches from our first encounter (even when I was a practitioner at a company). Needless to say, his constructive ideas and thoughtful comments largely contribute to this thesis.

I am also indebted to Professor Jirô Akahori who has supported my life at Ritsumeikan University. He has organized some symposiums during my doctoral course at which I have had the opportunities to give talks.

Professor Setsuro Fujiie has given me deep kindness throughout my PhD course.

Fruitful discussions with the following people: Professor Atsushi Takeuchi, Professor Masafumi Hayashi, Professor Masaaki Fukasawa and Song Xiaoming, improved this thesis. I would like to thank them.

I am thankful to Professor Taizo Chiyonobu, Professor Kazuhiro Yasuda, Professor Takahiro Aoyama, Professor Ngô Hoàng Long and Professor Azmi Makhlouf for giving me encouragements to address my works.

I would like to thank the members the of weekly mathematical finance seminar: Professor Masatoshi Fujisaki, Professor Kenji Yasutomi, Professor Takuya Watanabe, Yuri Imamura, Nienlin Liu, Gô Yûki, Zhong Jie and Libo Li, for valuable talks.

I have received encouragements through talking with several practitioners and I must give my thanks to them.

My colleagues of Ristumeikan University: Hideyuki Tanaka, Hidemi Aihara and Takafumi Amaba, have given interesting and inspiring discussions. I greatly appreciate them.

Finally, I would like to thank my parents for their supports.

Contents

Chapter 1

Introduction and preliminary

1.1 Problem of the existence of the density function

In probability theory, we often consider an infinite-dimensional probability space, called the Wiener space. On the Wiener space, computing the expectation of a random variable implies integrating the random variable with respect to a probability measure defined on this infinite-dimensional space, called the Wiener measure. If we can prove the existence of the density function of the random variable, this integral with respect to the Wiener measure can be transformed to the integral with respect to the Lebesgue measure, namely, a measure on a finite-dimensional space.

Meanwhile, in mathematical finance, we often deal with options with non-smooth payoff functions (e.g. European call option or European put option). The price of options is defined by the expectation of random variables and the risks involved in options are defined by the sensitivities of the price of options with respect to market parameters. Thus, in order to compute these sensitivities, we are required to differentiate non-smooth payoff functions. The existence of the density function of the random variable guarantees that we can differentiate non-smooth payoff functions, as long as the Lebesgue measure of the set of all non-smooth points of the payoff functions is zero.

Therefore, to study the existence of the density function of random variables is one of the most important subject from a theoretical and a practical point of view.

Chapter 2 of this thesis is concerned with the problem of the existence of the density functions of an SDE whose coefficients are dependent on the maximum of the solution. One may interpret a result obtained in this chapter as an extension of a result in [7]. However, we shall give some results on the joint laws which are not considered in [7]. The results of Chapter 2 are taken from the published paper [15].

1.2 Problem of the computation of Greeks

In mathematical finance, the computation of the risks involved in options, called Greeks, is one of the most important problem since practitioners begin the hedging procedures for options based on the values of Greeks. There are some kinds of Greeks. For example, the sensitivity of option prices with respect to the current underlying asset's price price is called the delta, and the sensitivity of the delta with respect to the current asset price is called the gamma. A market parameter which describes the variance of asset prices is called the volatility, and the sensitivity with respect to the volatility is called the vega or vega index.

In the Black-Scholes model, the simplest financial model, the Greeks can be computed explicitly. However, in the other models which may perform better than the Black-Scholes model, the Greeks do not have the explicit formulas, therefore we are required to use some numerical techniques to compute the Greeks, such as the Monte Carlo simulation. Hence, the problem how we can express the Greeks is an interesting and important problem, mathematically and practically.

In practice, various types of options are traded by practitioners. A European option may be exercised only at the expiration date of the option (e.g. European call option or European put option). An option whose payoff is determined by the average underlying price over some pre-set period of time is called an Asian option. A lookback option is an option with the payoff depends on the maximum (or minimum) underlying asset's price occurring over the life of the option. A barrier option is an option on the underlying asset whose price breaching the pre-set barrier level either springs the option into existence or extinguishes an already existing option.

In [6], the authors used the Malliavin calculus to calculate the Greeks for the first time. They obtained some expressions to compute the Greeks of some European options and Asian options, and showed that these expressions provide the better numerical results than ones obtained by a classical method, called the finite difference method. One can find a formula to compute the vega index for Asian options, in [1]. In [9], a method to compute the delta and gamma of lookback and barrier options is discussed and numerical results are also given.

In Chapter 3, we focus on the problem of the computations of the vega index for lookback and barrier options. We shall give an expression of the vega index, numerical results and a method to simulate the vega index for some specific options. The results of Chapter 3 are taken from the submitted paper [16].

1.3 Preliminary of Malliavin calculus

Recent advances of a differential calculus on the Wiener space, called the Malliavin calculus (or stochastic calculus of variations) provides many useful tools to us in order to try the problems mentioned in the previous subsections.

We introduce some basic tools of Malliavin calculus that will be used throughout the thesis. We refer to [17] to introduce Malliavin calculus. Let $(\Omega, \mathcal{F}, P)$ be the canonical Wiener space which supports a d-dimensional Brownian morion W.

The class of real random variables of the form $F = f(W_{t_1}, \cdots, W_{t_n})$, $f \in C_b^\infty(\mathbb{R}^{nd}; \mathbb{R}), 0 \leq t_1, \cdots, t_n \leq t$ is denoted by $\mathcal{S}$. $\mathbb{D}^{1,p}$ denotes a Banach space which is the completion of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} = E[|F|^p]^{\frac{1}{p}} + \left(E\left[\left(\int_0^t \sum_{j=1}^d |D_r^j F|^2 dr \right)^{\frac{p}{2}} \right] \right)^{\frac{1}{p}},$$

where

$$D_r^j F = \sum_{i=1}^n \frac{\partial f}{\partial x_{ji}} (W_{t_1}, \cdots, W_{t_n}) \mathbf{1}_{[0,t_i]}(r).$$

$\mathbb{D}^{k,p}$ is defined analogously, and its associated norm is denoted by $\|\cdot\|_{k,p}$. Also, we define $\mathbb{D}^{k,\infty} = \cap_{p\geq 1}\mathbb{D}^{k,p}$ and $\mathbb{D}^{\infty} = \cap_{p\geq 1}\cap_{k\geq 1}\mathbb{D}^{k,p}$. For $F, G \in \mathbb{D}^{1,2}$ we define $\langle DF, DG\rangle_H := \int_0^t \sum_{j=1}^d D_r^j F D_r^j G dr$ and $\|DF\|_H^2 := \int_0^t \sum_{j=1}^d |D_r^j F|^2 dr$.

Now let us introduce a localization of $\mathbb{D}^{k,p}$. $\mathbb{D}_{loc}^{k,p}$ denotes the set of random variables F such that there exists a sequence $\{(\Omega_n, F_n), n \geq 1\} \subset \mathcal{F} \times \mathbb{D}^{k,p}$ with the following properties:

(i) $\Omega_n \uparrow \Omega$, a.s.

(ii) $F = F_n$, a.s. on Ω_n.

The following theorem is well-known and we shall use this theorem of obtain the results in Chapter 1.

Theorem 1. *(Theorem 2.1.2 of [17])* Let $F = (F^1, \cdots, F^m)$ be a random vector satisfying the following conditions.

(i) F^i belongs to the space $\mathbb{D}_{loc}^{1,p}, p > 1$, for all $i = 1, \cdots, m$.

(ii) The matrix $\gamma_F := (\langle DF^i, DF^j\rangle_H)_{1\leq i,j\leq m}$ is invertible a.s.

Then the law of F is absolute continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.

Chapter 2

Absolute continuity of the laws of a multi-dimensional stochastic differential equation with coefficients depending on the maximum

2.1 Introduction

In this chapter, we deal with the following m-dimensional stochastic differential equation (SDE):

$$X_t^i = x_0^i + \sum_{l=1}^{d} \int_0^t A_l^i(s, X_s, M_s) dW_s^l + \int_0^t B^i(s, X_s, M_s) ds, \ 1 \leq i \leq m \tag{2.1}$$

where W denotes a d-dimensional Brownian motion, $A_l^i, B^i : [0, \infty) \times \mathbb{R}^{2m} \to \mathbb{R}\,, 1 \leq i \leq m, 1 \leq l \leq d$ are measurable functions and $M_s = (M_s^1, \cdots, M_s^m) := (\max_{u \leq s} X_s^1, \cdots, \max_{u \leq s} X_s^m)$. The purpose of this chapter is to prove the absolute continuity of the joint law concerning the solution to (2.1) with Lipschitz continuous coefficients using Malliavin calculus. In [7], the authors proved that if $m = d = 1$, A and B are Hölder continuous, for $t > 0$ the law of X_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, where X_t is a weak or strong solution to (2.1). The authors used the method to analyze the characteristic function of X_t to prove the absolute continuity of the law of X_t in [7].

In this chapter, first we prove the absolute continuity of the law of $X_t = (X_t^1, \cdots, X_t^m)$ with respect to the Lebesgue measure on $\mathbb{R}^m$. Then we prove the absolute continuity of the law of $(M_t^i, X_t^{i'})$, $1 \leq i, i' \leq m$, with respect to the Lebesgue measure on $\mathbb{R}^2$ when A_l^i does not depend on the second space variable. To analyze the law of $(M_t^i, X_t^{i'})$ may be important in the field of applications such as finance. Throughout this chapter, we use C or C_i, $i \in \mathbb{N}$ to denote a positive constant which may depend on constants K, L, d, p, t and x_0.

2.2 The existence, uniqueness and differentiability of the solution to (2.1) and the absolute continuity of the probability law of X_t

In this section, firstly we prove the existence, uniqueness and differentiability of the solution to (2.1). Secondly we prove for $t > 0$, the absolute continuity of the probability law of X_t where X is the solution to (2.1).

We assume the following:

(A1) There exist $K, M, c > 0$ such that

$$\begin{aligned}|A(t,x_1,x_2) - A(t,x_1',x_2')| + |B(t,x_1,x_2) - B(t,x_1',x_2')| &\leq K(|x_1 - x_1'| + |x_2 - x_2'|)\\ |A(t,x_1,x_2)| + |B(t,x_1,x_2)| &\leq L,\end{aligned}$$

for any $x_1, x_2, x_1', x_2' \in \mathbb{R}^m$ and $t \geq 0$,

(A2) $A(t,x_1,x_2)$ is continuous with respect to (t,x_1,x_2),

(A3) there exists $c > 0$ such that

$$|v^T A(t,x_1,x_2)|^2 \geq c|v|^2,$$

for any $v \in \mathbb{R}^m$ and $x_1, x_2 \in \mathbb{R}^m$ and $t \geq 0$.

First, let us state a lemma on the existence of a unique solution to (2.1).

Lemma 1. *Assume* **(A1)***, then (2.1) has a unique strong solution for any initial value $x_0 \in \mathbb{R}^m$. Moreover we have $E[|M_t^i|^p] \leq C$ for any $t \geq 0, 1 \leq i \leq m$ and $p \geq 2$.*

Proof. For $s \in [0,t]$ we define

$$\begin{aligned}X_s^{(0),i} &:= x^i\\ X_s^{(n+1),i} &:= x^i + \sum_{l=1}^d \int_0^s A_l^i(u, X_u^{(n)}, M_u^{(n)})dW_u^l + \int_0^s B^i(u, X_u^{(n)}, M_u^{(n)})du,\ 1 \leq i \leq m,\ n \geq 0,\end{aligned} \tag{2.2}$$

where $X_u^{(n)} := (X_u^{(n),1}, \cdots, X_u^{(n),m})$ and $M_u^{(n)} := (\max_{v\leq u} X_v^{(n),1}, \cdots, \max_{v\leq u} X_v^{(n),m})$. From Hölder's inequality and Burkholder-Davis-Gundy's inequality, it is easy to see that

$$\begin{aligned}&E\left[\max_{u\leq s}\left|X_u^{(n+1),i} - X_u^{(n),i}\right|^2\right]\\ &\leq C\left(\sum_{l=1}^d E\left[\int_0^s \left(A_l^i(u, X_u^{(n)}, M_u^{(n)}) - A_l^i(u, X_u^{(n-1)}, M_u^{(n-1)})\right)^2 du\right]\right.\\ &\left.+E\left[\int_0^s \left(B^i(u, X_u^{(n)}, M_u^{(n)}) - B^i(u, X_u^{(n-1)}, M_u^{(n-1)})\right)^2 du\right]\right)\end{aligned}$$

holds for $s \in [0,t]$ and $1 \le i \le m$. By **(A1)** and a trivial inequality $|\max_{v\le u} X_v^{(n),i} - \max_{v\le u} X_v^{(n-1),i}| \le \max_{v\le u}|X_v^{(n),i} - X_v^{(n-1),i}|$, we get

$$\begin{aligned} &E\left[\max_{u\le s}|X_u^{(n+1),i} - X_u^{(n),i}|^2\right] \\ &\le CE\left[\int_0^s \left(|X_u^{(n)} - X_u^{(n-1)}|^2 + |M_u^{(n)} - M_u^{(n-1)}|^2\right) du\right] \\ &\le C\int_0^s E\left[\sum_{i=1}^m \max_{v\le u}|X_v^{(n),i} - X_v^{(n-1),i}|^2\right] du, \end{aligned}$$

therefore,

$$E\left[\sum_{i=1}^m \max_{u\le s}|X_u^{(n+1),i} - X_u^{(n+1),i}|^2\right] \le C_1\int_0^s E\left[\sum_{i=1}^m \max_{v\le u}|X_v^{(n),i} - X_v^{(n-1),i}|^2\right] du, \tag{2.3}$$

for $s \in [0,t]$ and $n \in \mathbb{N}$.

We define, for $s \in [0,t]$ and $n \in \mathbb{N}$, $f^{(n)}(s) := E\left[\sum_{i=1}^m \max_{u\le s}|X_u^{(n+1),i} - X_u^{(n+1),i}|^2\right]$, then we have

$$f^{(n)}(s) \le C_1^n \int_0^s \int_0^{u_1} \cdots \int_0^{u_{n-1}} f^{(0)}(u_n) du_n \cdots du_1,$$

by (2.3). Now due to **(A1)**, we obtain

$$f^{(0)}(s) = E\left[\sum_{i=1}^m \max_{u\le s}|X_u^{(1),i} - x_0|^2\right] \le C_2,$$

for $s \in [0,t]$, thus we have

$$E\left[\sum_{i=1}^m \max_{u\le s}|X_u^{(n+1),i} - X_u^{(n),i}|^2\right] = f^{(n)}(s) \le \frac{(C_1 s)^n}{n!} C_2. \tag{2.4}$$

Relation (2.4) and the Čebyšev's inequality give

$$P\left[\sum_{i=1}^m \max_{s\le t}|X_s^{(n+1),i} - X_s^{(n),i}| \ge \frac{1}{2^{n+1}}\right] \le 2C_2\frac{(2C_1 t)^n}{n!}, \tag{2.5}$$

for $n \in \mathbb{N}$ and the right hand side of (2.5) is a convergent series. From the Borel-Cantelli's lemma, there exists $\tilde{\Omega} \in \mathcal{F}$ with $P(\tilde{\Omega}) = 1$ such that for every $\omega \in \tilde{\Omega}$ there exists $N(\omega) \in \mathbb{N}$ with $\sum_{i=1}^m \max_{s\le t}|X_s^{(k+1),i} - X_s^{(k),i}| < 2^{-(k+1)}$ for $k \ge N(\omega)$. Moreover, this implies that

$$\sum_{i=1}^m \max_{s\le t}|X_s^{(k+m'),i} - X_s^{(k),i}| \le 2^{-k}, \tag{2.6}$$

for every $m' \in \mathbb{N}, k \ge N(\omega)$. We see then that the sequence of sample paths $\{X_s^{(n)}, s \in [0,t]\}$ is convergent in the supremum norm on continuous functions, which concludes the existence of a continuous limit $\{X_s, s \in [0,t]\}$ for all $\omega \in \tilde{\Omega}$.

Now let us prove that $\{X_s, s \in [0,t]\}$ satisfies (2.1). Firstly, we shall consider the Lebesgue integral part. Due to **(A1)**, we have for $s \in [0,t]$

$$\left| \int_0^s B^i(u, X_u^{(n)}, M_u^{(n)})du - \int_0^s B^i(u, X_u, M_u)du \right|^2 \le C \int_0^s |M_u^{(n)} - M_u|^2 du,$$

and (2.6) gives that $\max_{u \le s} |X_u^i - X_u^{(n),i}| \le 2^{-n}$ for $n \ge N(\omega)$. Thus, we get that

$$\left| \int_0^s B^i(u, X_u^{(n)}, M_u^{(n)})du - \int_0^s B^i(u, X_u, M_u)du \right|^2 \to 0,$$

holds as $n \to \infty$, a.s.

Next, we shall consider the stochastic integral part. We observe from (2.4) that for fixed $u \in [0,t]$, the sequence of random variables $\{M_u^{(n),i}\}_{n \in \mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}, P)$. Indeed, from (2.4) we get

$$\begin{aligned} E\left[|M_u^{(n),i} - M_u^{(n'),i}|^2\right] &\le E\left[\max_{v \le u} |X_v^{(n),i} - X_v^{(n'),i}|\right] \\ &\le \sum_{j=n'}^{n-1} E\left[\max_{v \le u} |X_v^{(j+1),i} - X_v^{(j),i}|^2\right] \to 0, \end{aligned}$$

as $n, n' \to \infty$. Therefore, there exists $\tilde{M}_u^i$ such that $M_u^{(n),i} \to \tilde{M}_u^i$ in $L^2(\Omega, \mathcal{F}, P)$. Since $M_u^{(n),i} \to M_u^i$, a.s., we have $E[|M_u^i - \tilde{M}_u^i|] \le \liminf_{n\to\infty} E[|M_u^{(n),i} - \tilde{M}_u|] = 0$ and this implies that $E[|M_u^{(n),i} - M_u^i|^2] \to 0$ as $n \to \infty$. From **(A1)**, we have for $s \in [0,t]$

$$E\left[\left| \sum_{l=1}^d \int_0^s \left(A_l^i(u, X_u^{(n)}, M_u^{(n)}) - A_l^i(u, X_u, M_u)\right) dW_u^l \right|^2\right] \le C \int_0^s E\left[|M_u^{(n)} - M_u|^2\right] du. \tag{2.7}$$

By **(A1)**, we have $E[|M_u^{(n),i}|^2] \le C_3$ and Fatou's lemma gives $E[|M_u^i|^2] \le \liminf_{n\to\infty} E[|M_u^{(n),i}|^2] \le C_3$. From (2.7) and the bounded convergence theorem, we have

$$E\left[\left| \sum_{l=1}^d \int_0^s \left(A_l^i(u, X_u^{(n)}, M_u^{(n)}) - A_l^i(u, X_u, M_u)\right) dW_u^l \right|^2\right] \to 0,$$

as $n \to \infty$, thus, by taking a subsequence one has

$$\left| \sum_{l=1}^d \int_0^s \left(A_l^i(u, X_u^{(n)}, M_u^{(n)}) - A_l^i(u, X_u, M_u)\right) dW_u^l \right|^2 \to 0, \tag{2.8}$$

as $n \to \infty$, a.s. Therefore, for $t > 0$, $\{X_s, s \in [0,t]\}$ satisfies (2.1).

Next, we shall prove the pathwise uniqueness of the equation (2.1). We assume that for fixed $t > 0$, $\{X_s, s \in [0,t]\}$ and $\{\check{X}_s, s \in [0,t]\}$ satisfy (2.1). From **(A1)**, it is easy to see that

$$E\left[\max_{u \le s} |X_u^i - \check{X}_u^i|^2\right] \le C \int_0^s E\left[\sum_{i=1}^m \max_{u \le v} |X_u^i - \check{X}_v^i|^2\right] dv,$$

thus,

$$E\left[\sum_{i=1}^{m}\max_{u\leq s}|X_u^i-\check{X}_u^i|^2\right]\leq C\int_0^s E\left[\sum_{i=1}^{m}\max_{u\leq v}|X_u^i-\check{X}_v^i|^2\right]dv$$

holds, then by defining $g(s):=E\left[\sum_{i=1}^{m}\max_{u\leq s}|X_u^i-\check{X}_u^i|^2\right]$ and applying the Gronwall's lemma to $g(s)$, we have $E\left[\sum_{i=1}^{m}\max_{u\leq s}|X_u^i-\check{X}_u^i|^2\right]=0$ for $s\in[0,t]$. Therefore, one has the pathwise uniqueness of the solution $\{X_s, s\in[0,t]\}$ to (2.1) for fixed $t>0$ by the continuity of $\{X_s, s\in[0,t]\}$.

Since $t>0$ is arbitrary, we have the existence of a unique strong solution $\{X_s, s\in[0,\infty)\}$ to (2.1).

Moreover, $E[|M_t^i|^p]\leq C$ for $p\geq 2$ is a consequence of **(A1)**. □

Now, let us prove the property of the time when one-dimensional process $\{X_s^i, s\in[0,t]\}$ attains its maximum on $[0,t]$. This property plays an important role to prove the absolute continuity of the joint law of $(M_t^i, X_t^{i'})$.

Lemma 2. *Under* **(A1)**-**(A3)**, *for any $t\geq 0$ and $1\leq i\leq m$, $\{X_s^i, s\in[0,t]\}$ attains its maximum on $[0,t]$ on a unique point τ_t^i and $0<\tau_t^i<t$, a.s.*

Proof. We define a new probability measure $\tilde{P}$ by

$$\frac{d\tilde{P}}{dP}:=\exp\left[-\sum_{l=1}^{d}\int_0^t C^l(s,X_s,M_s)dW_s^l\right],$$

where for $x_1,x_2\in\mathbb{R}^m$ and $s>0$, d-dimensional vector $C(s,x_1,x_2)$ is defined by $C(s,x_s,x_2):=[A^T(AA^T)^{-1}B](s,x_1,x_2)$. Define a d-dimensional process $\tilde{W}$ by

$$\tilde{W}_s:=W_s^l+\int_0^s C^l(u,X_uM_u), 1\leq l\leq d.$$

Then by the Girsanov's theorem, $\{\tilde{W}_s, s\in[0,t]\}$ is a d-dimensional Brownian motion under $\tilde{P}$, therefore $\{X_s, s\in[0,t]\}$ can be expressed as

$$X_s^i=x_0^i+\sum_{l=1}^{d}\int_0^s A_l^i(u,X_u,M_u)d\tilde{W}_u^l,\ 1\leq i\leq m, \tag{2.9}$$

and for each $1\leq i\leq m$, $\{X_s^i, s\in[0,t]\}$ is a martingale by **(A1)**. Let $\{\mathcal{F}_s, s\in[0,t]\}$ be the augmentation of the Brownian filtration generated by $\{\tilde{W}_s, s\in[0,t]\}$. For $s\in[0,t]$, we define $T(s):=\inf\{t>0:\langle X^i\rangle_t>s\}$ then the time-changed process

$$B_s:=X_{T(s)}^i, \mathcal{G}_s:=\mathcal{F}_{T(s)}, s\in[0,t]$$

is a standard one-dimensional Brownian motion. Moreover, by **(A3)**, $\{X_s^i, s\in[0,t]\}$ can be written as

$$X_s^i=x_0^i+B_{\langle X^i\rangle_s}$$

for $s\in[0,t]$.

First, let us prove that

$$\tilde{P}\left(\max_{s\le t} X_s^i = x_0^i\right) = 0.$$

By the law of iterated logarithm for Brownian motion, we have

$$\limsup_{s\downarrow 0} \frac{X_s^i - x_0^i}{\sqrt{2\langle X^i\rangle_s \log\log\left(\frac{1}{\langle X^i\rangle_s}\right)}} = \limsup_{s\downarrow 0} \frac{B_{\langle X^i\rangle_s}}{\sqrt{2\langle X^i\rangle_s \log\log\left(\frac{1}{\langle X^i\rangle_s}\right)}} = 1,$$

$\tilde{P}$-a.s., thus

$$\tilde{P}\left(\max_{s\le t} X_s^i = x_0^i\right) \le \tilde{P}\left(X_s^i \le x_0^i, \forall s \in [0,t]\right) = 0.$$

Then, we shall prove that

$$\tilde{P}\left(\max_{s\le t} X_s^i = X_t^i\right) = 0.$$

We note that for $t > 0$, $\langle X^i\rangle_t$ is a stopping time for the filtration $\mathcal{G}_s$, since

$$\{\langle X^i\rangle_t \le s\} = \{T(s) \ge t\} \in \mathcal{F}_{T(s)} = \mathcal{G}_s.$$

We define

$$\check{B}_s := \begin{cases} B_s, s \in [0,t] \\ B_t + \hat{W}_s - \hat{W}_t, s \in (t,\infty) \end{cases} \tag{2.10}$$

and $\check{\mathcal{F}}_s := \sigma(B_u, u \le s) \vee \sigma(\tilde{W}_u, u \le s)$, where $\{\hat{W}_s, s \in [0,\infty)\}$ is a one-dimensional Brownian motion independent of B, then we find that $\{\check{B}_s, s \in [0,\infty)\}$ is a one-dimensional Brownian motion, since we can easily check that $\{\check{B}_s, s \in [0,\infty)\}$ is a $\check{\mathcal{F}}_s$-martingale and its quadratic variation is given by $\{s, s \in [0,\infty)\}$. Let $\{B'_s, s \in [0,\infty)\}$ be a one-dimensional Brownian motion independent of $\{\check{B}_s, s \in [0,\infty)\}$. Define

$$\hat{B}_s := \begin{cases} \check{B}_s, s \in [0,\infty) \\ B'_{-s}, s \in (-\infty,0) \end{cases} \tag{2.11}$$

then $\{\hat{B}_s, s \in (-\infty,\infty)\}$ is a two sided Brownian motion. By the definition of $\check{\mathcal{F}}_s$, $\langle X^i\rangle_t$ is a stopping time for the filtration $\check{\mathcal{F}}_s$, therefore from Exercise 2.4 of [14], $\{\hat{B}_{\langle X^i\rangle_t - s} - \hat{B}_{\langle X^i\rangle_t}, s \in [0,\infty)\}$ is a one-dimensional Brownian motion. Again, by the law of iterated logarithm for Brownian motion, we have

$$\begin{aligned} -1 &= \liminf_{s\downarrow 0} \frac{\hat{B}_{\langle X^i\rangle_t} - \hat{B}_{\langle X^i\rangle_t - s}}{\sqrt{2s\log\log(\frac{1}{s})}} = \liminf_{s\downarrow 0} \frac{\hat{B}_{\langle X^i\rangle_t} - \hat{B}_{\langle X^i\rangle_{t-s}}}{\sqrt{2(\langle X^i\rangle_t - \langle X^i\rangle_{t-s})\log\log\left(\frac{1}{\langle X^i\rangle_t - \langle X^i\rangle_{t-s}}\right)}} \\ &= \liminf_{s\downarrow 0} \frac{B_{\langle X^i\rangle_t} - B_{\langle X^i\rangle_{t-s}}}{\sqrt{2(\langle X^i\rangle_t - \langle X^i\rangle_{t-s})\log\log\left(\frac{1}{\langle X^i\rangle_t - \langle X^i\rangle_{t-s}}\right)}} \\ &= \liminf_{s\downarrow 0} \frac{X_t^i - X_{t-s}^i}{\sqrt{2(\langle X^i\rangle_t - \langle X^i\rangle_{t-s})\log\log\left(\frac{1}{\langle X^i\rangle_t - \langle X^i\rangle_{t-s}}\right)}}, \end{aligned}$$

$\tilde{P}$-a.s. Therefore,

$$\tilde{P}\left(\max_{s\le t} X_s^i = X_t^i\right) \le \tilde{P}\left(X_s^i \le X_t^i, \forall s\in[0,t]\right) = 0.$$

Finally, let us prove the uniqueness of τ_t^i on $[0,t]$. As mentioned before, we can write

$$X_s^i = x_0^i + B_{\langle X^i\rangle_s}$$

for $s\in[0,t]$. Define

$$\begin{aligned}
\bar{\theta}_t &:= \sup\left\{s\le t : B_{\langle X^i\rangle_s} = \sup_{0\le u\le t} B_{\langle X^i\rangle_u}\right\},\\
\underline{\theta}_t &:= \inf\left\{s\le t : B_{\langle X^i\rangle_s} = \sup_{0\le u\le t} B_{\langle X^i\rangle_u}\right\},
\end{aligned}$$

and

$$\begin{aligned}
\bar{\tau}_t &:= \sup\left\{s\le t : B_s = \sup_{0\le u\le t} B_u\right\},\\
\underline{\tau}_t &:= \inf\left\{s\le t : B_s = \sup_{0\le u\le t} B_u\right\}.
\end{aligned}$$

Then by the definitions, we have

$$\bar{\theta}_t = \langle X^i\rangle^{-1}_{\bar{\tau}_{\langle X^i\rangle_t}},\ \underline{\theta}_t = \langle X^i\rangle^{-1}_{\underline{\tau}_{\langle X^i\rangle_t}}.$$

Thus, one has

$$\begin{aligned}
\tilde{P}\left(\underline{\theta}_t < \bar{\theta}_t < t\right) &= \tilde{P}\left(\underline{\tau}_{\langle X^i\rangle_t} < \bar{\tau}_{\langle X^i\rangle_t} < \langle X^i\rangle_t\right)\\
&= \tilde{P}\left(\bigcup_{\substack{r_1,r_2\in\mathbb{Q}\\ r_1<r_2}} \left\{\underline{\tau}_{\langle X^i\rangle_t} < r_1 < \bar{\tau}_{\langle X^i\rangle_t} < r_2 < \langle X^i\rangle_t\right\}\right)\\
&= \sum_{\substack{r_1,r_2\in\mathbb{Q}\\ r_1<r_2}} \tilde{P}\left(\underline{\tau}_{\langle X^i\rangle_t} < r_1 < \bar{\tau}_{\langle X^i\rangle_t} < r_2 < \langle X^i\rangle_t\right),
\end{aligned}$$

where $\mathbb{Q}$ denotes the set of all rational numbers. On $\{\underline{\tau}_{\langle X^i\rangle_t} < r_1 < \bar{\tau}_{\langle X^i\rangle_t} < r_2 < \langle X^i\rangle_t\}$, the definition of $\bar{\tau}$ shows

$$\bar{\tau}_{\langle X^i\rangle_t} = \sup\left\{s\le r_2 : B_s = \sup_{0\le u\le r_2} B_u\right\}, \tag{2.12}$$

and the definition of $\underline{\tau}$ shows

$$\underline{\tau}_{\langle X^i\rangle_t} = \inf\left\{s\le r_1 : B_s = \sup_{0\le u\le r_1} B_u\right\}. \tag{2.13}$$

Since $B_{\bar{\tau}_{\langle X^i\rangle_t}} = B_{\underline{\tau}_{\langle X^i\rangle_t}}$ holds $\tilde{P}$-a.s., (2.12) and (2.13) imply that

$$\sup_{0\le u\le r_1} B_u = \sup_{0\le u\le r_2} B_u$$

holds $\tilde{P}$- a.s. Therefore, we have

$$\begin{aligned}
&\tilde{P}\left(\underline{\tau}_{\langle X^i\rangle_t} < r_1 < \bar{\tau}_{\langle X^i\rangle_t} < r_2 < \langle X^i\rangle_t\right)\\
\le\ & \tilde{P}\left(\bar{\tau}_{\langle X^i\rangle_t} \in (r_1,r_2),\ \sup_{0\le u\le r_1} B_u = \sup_{0\le u\le r_2} B_u\right)\\
=\ & \tilde{P}\left(\sup\left\{s\le r_2 : B_s = \sup_{0\le u\le r_2} B_u\right\} \in (r_1,r_2),\ \sup_{0\le u\le r_1} B_u = \sup_{0\le u\le r_2} B_u\right) = 0,
\end{aligned}$$

where the last equality follows from Proposition 4 in Section VI of [2]. This finishes the proof. □

Let us prove a lemma on the differentiability of the maximum of a continuous process which is similar to Proposition 2.1.10 of [17]

Lemma 3. *For $t\ge 0$, let $\{\hat{X}_s, s\in[0,t]\}$ be a one-dimensional continuous process. Suppose that*

(i) $E[\sup_{s\le t}|\hat{X}_s|^2]<\infty$,

(ii) *for any $s\in[0,t]$, $\hat{X}_s\in\mathbb{D}^{1,2}$ and $E[\sup_{s\le t}\|D\hat{X}_t\|_H^2]<\infty$.*

Then $\hat{M}_t=\sup_{s\le t}\hat{X}_s\in\mathbb{D}^{1,2}$ and we have

$$E\left[\|D\hat{M}_t\|_H^2\right] \le E\left[\sup_{s\le t}\|D\hat{X}_s\|_H^2\right]. \tag{2.14}$$

Moreover, if we assume that

(iii) *$\{\hat{X}_s, s\in[0,t]\}$ attains its maximum on a unique point $\hat{\tau}_t^*$,*

(iv) *for $1\le j\le d$, and almost every r, $\{D_r^j\hat{X}_s, s\in[0,t]\}$ is continuous except for $s=r$, and*

(v) *for $1\le j\le d$, $E[\int_0^t \sup_{r\le s\le t}|D_r^j\hat{X}_s|^2dr]<\infty$,*

then we have

$$D_r^j\hat{M}_t = D_r\hat{X}_{\hat{\tau}_t^*},\ a.e.r, \tag{2.15}$$

where we have defined $D_r\hat{X}_{\hat{\tau}_t^*} := D_r\hat{X}_s|_{s=\hat{\tau}_t^*}$.

Proof. Let $\{t_k\}_{k\ge 0}$ be a dense subset of $[0,t]$ and define

$$\hat{M}_t^n := \max\{\hat{X}_{t_1},\cdots,\hat{X}_{t_n}\}.$$

Define

$$A_1 := \{\hat{X}_{t_1}=\hat{M}_t^n\},\ A_k := \{\hat{X}_{t_1}\ne\hat{M}_t^n,\cdots,\hat{X}_{t_{k-1}}\ne\hat{M}_t^n, \hat{X}_{t_k}=\hat{M}_t^n\},\ 2\le k\le n.$$

Then, by the local property of operator D we have

$$D\hat{M}_t^n = \sum_{k=1}^n \mathbf{1}_{A_k} D\hat{X}_{t_k}.$$

By Proposition 2.1.10 of [17], $\hat{M}_t = \sup_{0\le s\le t} \hat{X}_s$ belongs to $\mathbb{D}^{1,2}$ and $D\hat{M}_t^n \to D\hat{M}_t$ $(n\to\infty)$ in the weak topology of $L^2(\Omega; L^2([0,t];\mathbb{R}^d))$ under (i) and (ii). We obtain (2.14) from $E[\|D\hat{M}_t\|_H^2] \le \liminf_{n\to\infty} E[\|D\hat{M}_t^n\|_H^2]$.

Let us prove (2.15). For $\omega \in A_k$ we define $\hat{\tau}_n^* := t_k$. Then $\hat{\tau}_n^* \to \hat{\tau}_t^*$, a.s. due to (iii), and we have

$$D\hat{M}_t^n = \sum_{k=1}^n \mathbf{1}_{A_k} D\hat{X}_{\hat{\tau}_n^*} = D\hat{X}_{\hat{\tau}_n^*},$$

where we have defined $D\hat{X}_{\hat{\tau}_n^*} := D\hat{X}_s|_{s=\hat{\tau}_n^*}$. Note that, if $r = \hat{\tau}_n^*$, then $D\hat{X}_{\hat{\tau}_n^*}$ is not well defined, due to the discontinuity; thus the rigorous meaning of the above equality is that $D_r\hat{M}_t^n = D_r\hat{X}_{\hat{\tau}_n^*}$ for almost every r with probability 1.

Now let us prove

$$E\left[\int_0^t \sum_{j=1}^d D_r^j \hat{X}_{\hat{\tau}_n^*} u_r^j dr\right] \to E\left[\int_0^t \sum_{j=1}^d D_r^j \hat{X}_{\hat{\tau}_t^*} u_r^j dr\right], \tag{2.16}$$

for any $u \in L^2(\Omega; L^2([0,t];\mathbb{R}^d))$. We have

$$E\left[\int_0^t \sum_{j=1}^d D_r^j \hat{X}_{\hat{\tau}_n^*} u_r^j dr\right] - E\left[\int_0^t \sum_{j=1}^d D_r^j \hat{X}_{\hat{\tau}_t^*} u_r^j dr\right] = E\left[\int_0^t \sum_{j=1}^d (D_r^j \hat{X}_{\hat{\tau}_n^*} - D_r^j \hat{X}_{\hat{\tau}_t^*}) u_r^j dr\right]. \tag{2.17}$$

From (iv), we have $D_r^j \hat{X}_{\hat{\tau}_n^*} \to D_r^j \hat{X}_{\hat{\tau}_t^*}$ for $r \ne \hat{\tau}_t^*$ then $D_r^j \hat{X}_{\hat{\tau}_n^*} \to D_r^j \hat{X}_{\hat{\tau}_t^*}$, for almost every r with probability 1. As $|D_r^j \hat{X}_{\hat{\tau}_n^*} - D_r^j \hat{X}_{\hat{\tau}_t^*}|^2 \le 2\sup_{0\le s\le t} |D_r^j \hat{X}_s|^2$ and (v), we have

$$\int_0^t \sum_{j=1}^d |D_r^j \hat{X}_{\hat{\tau}_n^*} - D_r^j \hat{X}_{\hat{\tau}_t^*}|^2 dr \to 0\ (n\to\infty), a.s.$$

Due to $E[|D_r^j \hat{X}_{\hat{\tau}_n^*} - D_r^j \hat{X}_{\hat{\tau}_t^*}|^2] \le 2E[\sup_{0\le s\le t} |D_r^j \hat{X}_s|^2]$ and (v), we have

$$\lim_{n\to\infty} E\left[\int_0^t \sum_{j=1}^d |D_r^j \hat{X}_{\hat{\tau}_n^*} - D_r^j \hat{X}_{\hat{\tau}_t^*}|^2\right] dr = 0.$$

Then we obtain (2.16). Since $D\hat{M}_t^n$ converges to $D\hat{M}_t$ weakly in $L^2(\Omega; L^2([0,t];\mathbb{R}^d))$ and (2.16) holds, we have

$$E\left[\int_0^t \sum_{j=1}^d \left(D_r^j \hat{X}_{\hat{\tau}_t^*} - D_r^j \hat{M}_t\right) u_r^j dr\right] = 0,$$

for any $u \in L^2(\Omega; L^2([0,t];\mathbb{R}^d))$. By the fact that $\hat{M}_t$ belongs to $\mathbb{D}^{1,2}$ and (v), we have $\{D_r\hat{X}_{\hat{\tau}_t^*} - D_r\hat{M}_t, r \in [0,t]\} \in L^2(\Omega; L^2([0,t];\mathbb{R}^d))$. Therefore we have (2.15) with taking $u_r = D_r\hat{X}_{\hat{\tau}_t^*} - D_r\hat{M}_t$ and this finishes the proof. □

Remark 1. *In Lemma 3, if we assume that $\{\hat{X}_s, s \in [0,t]\}$ is adapted, then we have $D_r\hat{X}_{\hat{\tau}_t^*} = 0$ for almost every r such that $r > \hat{\tau}_t^*$. Thus, in this case, we can write $D_r\hat{M}_t = \mathbf{1}_{[0,\hat{\tau}_t^*)}(r)D_r\hat{X}_{\hat{\tau}_t^*}$, for almost every r.*

Next, let us prove the differentiability of the solution to (2.1) in Malliavin sense.

Lemma 4. *Assume* **(A1)-(A3)**. *Then, for $s \in [0,t]$ and $1 \leq i \leq m$, X_s^i, M_s^i belong to $\mathbb{D}^{1,2}$. Moreover, $\{D_r^jX_s^i, s \in [r,t]\}$ satisfies the following equation:*

$$\begin{aligned} D_r^jX_s^i &= A_j^i(r, X_r, M_r) + \int_r^s (\bar{A}_{k,l}^i(u)D_r^jX_u^k + \tilde{A}_{k,l}^i(u)D_r^jM_u^k)dW_u^l \\ &\quad + \int_r^t (\bar{B}_k^i(u)D_r^jX_u^k + \tilde{B}_k^i(u)D_r^jM_u^k)du \end{aligned} \tag{2.18}$$

for $r \leq s$, a.e., and

$$D_r^iX_s^i = 0, \tag{2.19}$$

for $r > s$, a.e., where $\bar{A}_{k,l}(u), \tilde{A}_{k,l}(u), \bar{B}_k(u)$ and $\tilde{B}_k(u)$ are uniformly bounded and adapted m-dimensional processes.

Proof. We will use the Picard approximation from Lemma 1, so $X_s^{(n)}$, $M_s^{(n)}$ are the processes constructed by recurrence there. The proof of this lemma uses the proof of Theorem 2.2.1 of [17]. We need to extend the proof to equation with coefficients which depend on the maximum process. We start by proving $X_s^{(n),i} \in \mathbb{D}^{1,2}$ for $s \in [0,t]$, $1 \leq i \leq m$ and $n \geq 0$. If we assume $X_s^{(n),i} \in \mathbb{D}^{1,2}$ and $E[\int_0^s \sup_{u\leq v} \|DX_u^{(n),i}\|_H^2 dv] < \infty$ for $s \in [0,t]$ then we have $M_s^{(n),i} \in \mathbb{D}^{1,2}$ for $s \in [0,t]$ by Lemma 3 and

$$\begin{aligned} & E\left[\int_0^t\int_0^t |D_r^j(A_l^i(u, X_u^{(n)}, M_u^{(n)}))|^2 drdu\right] \\ \leq\ & C\left(\sum_{k=1}^m E\left[\int_0^t\int_0^t |D_r^jX_u^{(n),k}|^2 drdu\right] + \sum_{k=m+1}^{2m} E\left[\int_0^t\int_0^t |D_r^jM_u^{(n),k}|^2 drdu\right]\right) \\ =\ & C\left(\sum_{k=1}^m \int_0^t E\left[\int_0^u |D_r^jX_u^{(n),k}|^2 dr\right] du + \sum_{k=m+1}^{2m} \int_0^t E\left[\int_0^u |D_r^jM_u^{(n),k}|^2 dr\right] du\right) \\ \leq\ & C\left(\sum_{k=1}^m \int_0^t E\left[\int_0^u \sum_{j=1}^d |D_r^jX_u^{(n),k}|^2 dr\right] du + \sum_{k=m+1}^{2m} \int_0^t E\left[\int_0^u \sum_{j=1}^d |D_r^jM_u^{(n),k}|^2 dr\right] du\right) \\ =\ & C\left(\sum_{k=1}^m E\left[\int_0^t \sup_{u\leq s} \|DX_u^{(n),k}\|_H^2 ds\right] + \sum_{k=1}^m E\left[\int_0^t \sup_{u\leq s} \|DM_u^{(n),k}\|_H^2 ds\right]\right) \\ \leq\ & C\sum_{k=1}^m E\left[\int_0^t \sup_{u\leq s} \|DX_u^{(n),k}\|_H^2 ds\right] < \infty \end{aligned} \tag{2.20}$$

by (2.14). Therefore, from Proposition 1.3.8 of [17], we have for $s \in [0,t]$, $X_s^{(n+1),i} \in \mathbb{D}^{1,2}$ and

$$\begin{aligned} & D_r^j X_s^{(n+1),i} \\ = \ & A_j^i(r, X_r^{(n)}, M_r^{(n)}) + \int_r^s \left[\bar{A}_{k,l}^{(n),i}(u) D_r^j X_u^{(n),k} + \tilde{A}_{k,l}^{(n),i}(u) D_r^j M_u^{(n),k} \right] dW_u^l \\ & + \int_r^t \left[\bar{B}_k^{(n),i}(u) D_r^j X_u^{(n),k} + \tilde{B}_k^{(n),i}(u) D_r^j M_u^{(n),k} \right] du, \end{aligned}$$

where $\bar{A}_{k,l}^{(n)}, \tilde{A}_{k,l}^{(n)}, \bar{B}_k^{(n)}$ and $\tilde{B}_k^{(n)}$ are uniformly bounded and adapted m-dimensional processes. Now, by **(A1)** and (2.14), one has

$$\begin{aligned} & E\left[\sup_{u \le s} \sum_{j=1}^d \int_0^u \left| \sum_{l=1}^d \sum_{k=1}^m \int_r^u \left[\bar{A}_{k,l}^{(n),i}(v) D_r^j X_v^{(n),k} + \tilde{A}_{k,l}^{(n),i}(v) D_r^j M_v^{(n),k} \right] dW_v^l \right|^2 dr \right] \\ \le \ & E\left[\sup_{u \le s} \sum_{j=1}^d \int_0^u \sup_{u \le s} \left| \sum_{l=1}^d \sum_{k=1}^m \int_r^u \left[\bar{A}_{k,l}^{(n),i}(v) D_r^j X_v^{(n),k} + \tilde{A}_{k,l}^{(n),i}(v) D_r^j M_v^{(n),k} \right] dW_v^l \right|^2 dr \right] \\ = \ & \sum_{j=1}^d \int_0^s E\left[\sup_{u \le s} \left| \sum_{l=1}^d \sum_{k=1}^m \int_r^u \left[\bar{A}_{k,l}^{(n),i}(v) D_r^j X_v^{(n),k} + \tilde{A}_{k,l}^{(n),i}(v) D_r^j M_v^{(n),k} \right] dW_v^l \right|^2 \right] dr \\ \le \ & C \sum_{j=1}^d \int_0^s E\left[\sum_{l=1}^d \int_r^s \left| \sum_{k=1}^m \left[\bar{A}_{k,l}^{(n),i}(v) D_r^j X_v^{(n),k} + \tilde{A}_{k,l}^{(n),i}(v) D_r^j M_v^{(n),k} \right] dv \right|^2 \right] dr \\ \le \ & C \sum_{j=1}^d \int_0^s E\left[\sum_{l=1}^d \int_r^s \sum_{k=1}^m \left(|D_r^j X_v^{(n),k}|^2 + |D_r^j M_v^{(n),k}|^2 \right) dv \right] dr \\ = \ & C \sum_{j=1}^d \int_0^s E\left[\sum_{l=1}^d \int_0^s \sum_{k=1}^m \left(|D_r^j X_v^{(n),k}|^2 + |D_r^j M_v^{(n),k}|^2 \right) dv \right] dr \\ = \ & C \sum_{j=1}^d \int_0^s E\left[\sum_{l=1}^d \int_0^s \sum_{k=1}^m \left(|D_r^j X_v^{(n),k}|^2 + |D_r^j M_v^{(n),k}|^2 \right) dr \right] dv \\ = \ & C \sum_{j=1}^d \int_0^s E\left[\sum_{l=1}^d \int_0^v \sum_{k=1}^m \left(|D_r^j X_v^{(n),k}|^2 + |D_r^j M_v^{(n),k}|^2 \right) dr \right] dv \\ = \ & C \sum_{k=1}^m \int_0^s E\left[\|DX_v^{(n),k}\|_H^2 + \|DM_v^{(n),k}\|_H^2 \right] \le C \sum_{k=1}^m \int_0^s E\left[\sup_{u \le v} \|DX_u^{(n),k}\|^2 \right] dv, \end{aligned}$$

and the same computation as the above gives

$$E\left[\sup_{u \le s} \int_0^u \left| \sum_{k=1}^m \int_r^s \left[\bar{B}_k^{(n),i}(v) D_r^j X_v^{(n),k} + \tilde{B}_k^{(n),i}(v) D_r^j M_v^{(n),k} \right] dv \right|^2 dr \right] \le C \sum_{k=1}^m \int_0^s E\left[\sup_{u \le v} \|DX_u^{(n),k}\|^2 \right] dv.$$

Thus, we have

$$\sum_{i=1}^m E\left[\sup_{u \le s} \|DX_u^{(n+1),i}\|_H^2 \right] \le C_1 + C_2 \int_0^s \sum_{i=1}^m E\left[\sup_{u \le v} \|DX_u^{(n),i}\|_H^2 \right] dv, \tag{2.21}$$

and this implies $M_s^{(n+1),i} \in \mathbb{D}^{1,2}$ and $E[\int_0^s \sup_{u\le v} \|DX_u^{(n+1),i}\|_H^2 dv] < \infty$ for $s \in [0,t]$ by Lemma 3.

Due to (2.21) and Lemma 3, we have $\sup_n E[\|DX_s^{(n),i}\|_H^2] < \infty$ and $\sup_n E[\|DM_s^{(n),i}\|_H^2] \le \sup_n E[\sup_{u\le s} \|DX_u^{(n),i}\|] < \infty$. By the fact that $X_s^{(n),i} \to X_s^i$, $M_s^{(n),i} \to M_s^i$ in $L^2(\Omega)$ and Lemma 1.2.3 of [17], X_s^i and M_s^i belong to $\mathbb{D}^{1,2}$ for $s \in [0,t]$. Moreover $DX_s^{(n),i}$ and $DM_s^{(n),i}$ converge to DX_s^i and DM_s^i in the weak topology of $L^2(\Omega; L^2([0,t];\mathbb{R}^d))$.

Let us prove (2.18). We have

$$E\left[\int_0^t\int_0^t |D_r^j(A_l^i(u,X_u,M_u))|^2 dr du\right] \le C\sum_{k=1}^m \int_0^t E\left[\|DX_u^k\|_H^2 + \|DM_u^k\|_H^2\right] du < \infty$$

by the same calculation as (2.20), the fact that $E[\|DM_u^k\|_H^2] \le \liminf_{n\to\infty} E[\|DM_u^{(n),k}\|_H^2] \le \liminf_{n\to\infty} E[\sup_{v\le u} \|DX_v^{(n),k}\|_H^2]$ holds and (2.21). Therefore, we have (2.18) and the proof is completed. $\square$

Lemma 5. *Assume* **(A1)**-**(A3)**. *Then, for* $\{X_s^i, s\in[0,t]\}$ *and* $p \ge 2$ *we have*

$$E\left[\int_0^s \sup_{r\le u\le s} |D^j X_u^i|^p dr\right] < \infty, \tag{2.22}$$

and assumptions (i)-(v) of Lemma 3 hold. Moreover, for $s \in [0,t]$ *and* $p \ge 2$, $X_s^i, M_s^i \in \mathbb{D}^{1,p}$.

Proof. First, let us prove (2.22) for $p = 2$. We have

$$\begin{aligned}
&\sum_{j=1}^d E\left[\int_0^s \sup_{r\le u\le s} |D_r^j X_u^i|^2 dr\right] \\
\le\ & C_1 + C_2\left(\int_0^s E\left[\sup_{r\le u\le s}\left|\int_r^u \sum_{l=1}^d \left(\bar{A}_{k,l}^i(v) D_r^j X_v^k + \tilde{A}_{k,l}^i(v) D_r^j M_v^k\right) dW_v^l\right|^2\right] dr\right. \\
&\left. + \int_0^s E\left[\sup_{r\le u\le s}\left|\int_r^u \left(\bar{B}_k^i(v) D_r^j X_v^k + \tilde{B}_k^i(v) D_r^j M_v^k\right) dv\right|^2\right] dr\right) \\
\le\ & C_1 + C_2\left(\int_0^s E\left[\sum_{l=1}^d \int_r^s \left(\bar{A}_{k,l}^i(v) D_r^j X_v^k + \tilde{A}_{k,l}^i(v) D_r^j M_v^k\right)^2 dv\right] dr\right. \\
&\left. + \int_0^s E\left[\int_r^s \left(\bar{B}_k^i(v) D_r^j X_v^k + \tilde{B}_k^i(v) D_r^j M_v^k\right)^2 dv\right] dr\right) \\
\le\ & C_1 + C_2 \int_0^s E\left[\int_r^s \sum_{k=1}^m \left(\|D_r^j X_v^k\|^2 + \|D_r^j M_v^k\|^2\right) dv\right] dr \\
=\ & C_1 + C_2 E\left[\int_0^s\int_0^s \sum_{k=1}^m \left(\|D_r^j X_v^k\|^2 + \|D_r^j M_v^k\|^2\right) dr dv\right] \\
=\ & C_1 + C_2 \int_0^s E\left[\int_0^v \sum_{k=1}^m \left(\|D_r^j X_v^k\|^2 + \|D_r^j M_v^k\|^2\right) dr\right] dv \\
=\ & C_1 + C_2 \sum_{k=1}^m \int_0^s E\left[\|DX_v^k\|_H^2 + \|DM_v^k\|_H^2\right] dv < \infty,
\end{aligned} \tag{2.23}$$

by the fact that $E[\|DM_v^k\|_H^2] \leq \liminf_{n\to\infty} E[\|DM_v^{(n),k}\|_H^2] \leq \liminf_{n\to\infty} E[\sup_{u\leq v}\|DX_u^{(n),k}\|_H^2]$ is true, (2.18) and (2.21). This implies that (v) holds for X^i. (i) follows from Lemma 1 and we have (ii) by (2.22) for $p=2$. (iii) holds due to Lemma 2 and we have (iv) by (2.18) and (2.19).

Let us prove (2.22) for $p>2$. It suffices to prove

$$E\left[\int_0^t \sum_{i=1}^m \sum_{j=1}^d \sup_{r\leq s\leq t} |D_r^j X_s^i|^p dr\right] \leq C_1 + C_2 \int_0^t E\left[\int_0^u \sum_{i=1}^m \sum_{j=1}^d \sup_{r\leq s\leq u} |D_r^j X_s^i|^p dr\right] du. \tag{2.24}$$

However, we get (2.24) from the same computation as (2.23) and an inequality

$$\|DM_u^k\|_H^p \leq C \sum_{j=1}^d \int_0^u \sup_{r\leq s\leq u} |D_r^j X_u^k|^p dr,$$

which follows from (2.18), (2.19) and (2.15). From (2.22) we have $X_s^i, M_s^i \in \mathbb{D}^{1,p}$ for $s\in[0,t]$ and $p\geq 2$. $\square$

Now we consider two $m\times m$ matrix-valued process defined by

$$Y_j^i(s) = \delta_j^i + \int_0^s \bar{A}_{k,l}^i(u) Y_j^k(u) dW_u^l + \int_0^s \bar{B}_k^i(u) Y_j^k(u) du \,, 1\leq i,j\leq m \tag{2.25}$$

and

$$Z_j^i(s) = \delta_j^i - \int_0^s Z_k^i(u) \bar{A}_{j,l}^k(u) dW_u^l - \int_0^s Z_k^i(u)[\bar{B}_j^k(u) - \bar{A}_{\alpha,l}^k(u)\bar{A}_{j,l}^\alpha(u)]du, \ 1\leq i,j\leq m. \tag{2.26}$$

By the argument in section 2.3 of [17], we have $Y^{-1}(s) = Z(s)$. Let us express $D_r^j X_s^i$ by using $Y(s)$ and $Z(s)$.

Lemma 6. *For $s\in[r,t]$ and $1\leq i\leq m, 1\leq j\leq d$, $D^j X_s^i$ satisfies*

$$\begin{aligned} D_r^j X_s^i &= Y_k^i(s) Z_{k'}^k(r) A_j^{k'}(r) + Y_k^i(s) \int_r^s Z_{k'}^k(u) \tilde{A}_{l',l}^{k'}(u) D_r^j M_u^{l'} dW_u^l \\ &\quad + Y_k^i(s) \int_r^s Z_{k'}^k(u)[\tilde{B}_{l'}^{k'}(u) - \bar{A}_{\alpha,l}^{k'} \tilde{A}_{l',l}^\alpha(u)] D_r^j M_u^{l'} du. \end{aligned} \tag{2.27}$$

Proof. From (2.18), (2.25), (2.26) and Itô's formula, one has for $1\leq i\leq m$ and $1\leq j\leq d$,

$$\begin{aligned} \sum_{k'=1}^m Z_{k'}^i(s) D_r^j X_s^{k'} &= \sum_{k'=1}^m Z_{k'}^i(r) A_j^{k'}(r, X_r, M_r) + \sum_{k'=1}^m \int_r^s Z_{k'}^i(u) d(D_r^j X_u^{k'}) + \sum_{k'=1}^m \int_r^s D_r^j X_u^{k'} dZ_{k'}^i(u) \\ &\quad + \sum_{k'=1}^m \int_r^s d\langle Z_{k'}^i(\cdot), D_r^j X_\cdot^{k'}\rangle_u \\ &= Z_{k'}^i(r) A_j^{k'}(r, X_r, M_r) + \int_r^s Z_{k'}^i(u) \bar{A}_{l',l}^{k'}(u) D_r^j M^{l'} dW_u^l \\ &\quad + \int_r^s Z_{k'}^i(u)[\tilde{B}_{l'}^{k'}(u) - \bar{A}_{\alpha,l}^{k'}(u) \tilde{A}_{l',l}^\alpha(u)] D_r^j M_u^{l'} du. \end{aligned}$$

By the definition of (2.25) and (2.26), we have

$$D_r^j X_s^i = \sum_{k,k'=1}^m Y_k^i(s) Z_{k'}^k(s) D^j X_s^{k'},$$

therefore, the result follows. $\square$

Now we prove the absolute continuity of the law of X_t which is the main theorem of this section.

Theorem 2. *Assume* **(A1)-(A3)***, then for $t > 0$, X_t has the absolutely continuous probability law with respect to the Lebesgue measure on $\mathbb{R}^m$.*

Proof. Let us prove $\int_0^t |v^T D_r X_t|^2 dr > 0$ for nonzero vector $v \in \mathbb{R}^m$. By (2.27) and a trivial inequality $(a+b)^2 \geq \frac{a^2}{2} - b^2, a, b \in \mathbb{R}$, we have

$$\begin{aligned}
&|v^T D_r X_t|^2 \\
\geq\ & \frac{1}{2}\sum_{j=1}^d \left| \sum_{i=1}^m v_i Y_k^i(t) Z_{k'}^k(r) A_j^{k'}(r) \right|^2 \\
& - \sum_{j=1}^d \left| \sum_{i=1}^m v_i Y_k^i(t) \left(\int_r^t Z_{k'}^k(s) \tilde{A}_{l',l}^{k'}(s) D_r^j M_s^{l'} dW_s^l + \int_r^t Z_{k'}^k(s) [\tilde{B}_{l'}^{k'}(s) - \bar{A}_{\alpha,l}^{k'} \tilde{A}_{l',l}^{\alpha}(s)] D_r^j M_s^{l'} ds \right) \right|^2 \\
=:\ & \frac{1}{2}\sum_{j=1}^d \left| \sum_{i=1}^m v_i Y_k^i(t) Z_{k'}^k(r) A_j^{k'}(r) \right|^2 + A_{r,t}.
\end{aligned}$$

Then we have

$$\begin{aligned}
& E\left[\frac{1}{\varepsilon}\int_{t-\varepsilon}^t A_{r,t} dr\right] \\
\leq\ & \frac{1}{\varepsilon} E\left[\int_{t-\varepsilon}^t \sum_{j=1}^d \left| \sum_{i=1}^m v_i Y_k^i(t) \left(\int_r^t Z_{k'}^k(s) \tilde{A}_{l',l}^{k'}(s) D_r^j M_s^{l'} dW_s^l \right.\right.\right. \\
& \left.\left.\left. + \int_r^t Z_{k'}^k(s) [\tilde{B}_{l'}^{k'}(s) - \bar{A}_{\alpha,l}^{k'} \tilde{A}_{l',l}^{\alpha}(s)] D_r^j M_s^{l'} ds \right) \right|^2 dr \right] \\
\leq\ & \frac{C}{\varepsilon} \sum_{j=1}^d \sum_{i=1}^m |v_i|^2 E\left[\int_{t-\varepsilon}^t \left| Y_k^i(t) \left(\int_r^t Z_{k'}^k(s) \tilde{A}_{l',l}^{k'}(s) D_r^j M_s^{l'} dW_s^l \right.\right.\right. \\
& \left.\left.\left. + \int_r^t Z_{k'}^k(s) [\tilde{B}_{l'}^{k'}(s) - \bar{A}_{\alpha,l}^{k'} \tilde{A}_{l',l}^{\alpha}(s)] D_r^j M_s^{l'} ds \right) \right|^2 dr \right] \\
\leq\ & \frac{C}{\varepsilon} \sum_{j=1}^d \sum_{i=1}^m |v_i|^2 \left(E\left[\int_{t-\varepsilon}^t \left| Y_k^i(t) \int_r^t Z_{k'}^k(s) \tilde{A}_{l',l}^{k'}(s) D_r^j M_s^{l'} dW_s^l \right|^2 dr \right] \right. \\
& \left. + E\left[\int_{t-\varepsilon}^t \left| Y_k^i(t) \int_r^t Z_{k'}^k(s) [\tilde{B}_{l'}^{k'}(s) - \bar{A}_{\alpha,l}^{k'} \tilde{A}_{l',l}^{\alpha}(s)] D_r^j M_s^{l'} ds \right|^2 dr \right] \right).
\end{aligned}$$

Now, one has

$$
\begin{aligned}
&E\left[\int_{t-\varepsilon}^{t}\left|Y_k^i(t)\int_r^t Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}dW_s^l\right|^2 dr\right] \\
&= E\left[\int_{t-\varepsilon}^{t}\sum_{k,k',l'=1}^{m}\left|Y_k^i(t)\int_r^t Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}dW_s^l\right|^2 dr\right] \\
&\leq C\sum_{k=1}^{m}E\left[|Y_k^i(t)|^2\int_{t-\varepsilon}^{t}\left|\int_r^t Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}dW_s^l\right|^2 dr\right] \\
&\leq C\sum_{k,k'=1}^{m}E\left[|Y_k^i(t)|^2\int_{t-\varepsilon}^{t}\left|\int_r^t Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}dW_s^l\right|^2 dr\right] \\
&\leq C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^2\int_{t-\varepsilon}^{t}\left|\int_r^t Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}dW_s^l\right|^2 dr\right] \\
&\leq C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\int_{t-\varepsilon}^{t}E\left[\left|\sum_{l=1}^{d}\int_r^t Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}dW_s^l\right|^4\right]^{\frac{1}{2}} dr \\
&\leq C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\int_{t-\varepsilon}^{t}E\left[\left(\int_r^t\sum_{l=1}^{d}|Z_{k'}^k(s)\tilde{A}_{l',l}^{k'}(s)D_r^jM_s^{l'}|^2ds\right)^2\right]^{\frac{1}{2}} dr \\
&\leq C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\int_{t-\varepsilon}^{t}E\left[\sum_{l=1}^{d}\left(\int_r^t|Z_{k'}^k(s)D_r^jM_s^{l'}|^2ds\right)^2\right]^{\frac{1}{2}} dr \\
&\leq C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\int_{t-\varepsilon}^{t}E\left[\int_r^t|Z_{k'}^k(s)D_r^jM_s^{l'}|^4ds\right]^{\frac{1}{2}}(t-r)^{\frac{1}{2}}dr \\
&\leq \varepsilon C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\left(\int_{t-\varepsilon}^{t}E\left[\int_r^t|Z_{k'}^k(s)D_r^jM_s^{l'}|^4ds\right]dr\right)^{\frac{1}{2}} \\
&\leq \varepsilon C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\left(\int_{t-\varepsilon}^{t}\int_0^t E\left[|Z_{k'}^k(s)D_r^jM_s^{l'}|^4\right]dsdr\right)^{\frac{1}{2}} \\
&\leq \varepsilon C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}\left(\int_{t-\varepsilon}^{t}E\left[|Z_{k'}^k(s)|^4\int_{t-\varepsilon}^{t}\sup_{r\leq s\leq t}|D_r^jX_s^{l'}|^4dr\right]ds\right)^{\frac{1}{2}} \\
&\leq \varepsilon^{\frac{3}{2}}C\sum_{k,k',l'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^4\int_0^t\sup_{r\leq s\leq t}|D_r^jX_s^{l'}|^4dr\right]^{\frac{1}{2}} \\
&\leq \varepsilon^{\frac{3}{2}}C\sum_{k,k'=1}^{m}E\left[|Y_k^i(t)|^4\right]^{\frac{1}{2}}E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}}\sum_{l'=1}^{m}\left(E\left[\int_0^t\sup_{r\leq s\leq t}|D_r^jX_s^{l'}|^8dr\right]\right)^{\frac{1}{4}},
\end{aligned}
$$

and

$$
\begin{aligned}
& E\left[\int_{t-\varepsilon}^{t}\left|Y_k^i(t)\int_r^t Z_{k'}^k(s)[\tilde{B}_{l'}^{k'}(s)-\bar{A}_{\alpha,l}^{k'}(s)\tilde{A}_{l',l}^{\alpha}(s)]D_r^j M_s^{l'}ds\right|^2 dr\right] \\
\leq\ & C\sum_{k,k',l'=1}^{m} E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} E\left[\left(\int_{t-\varepsilon}^{t}\left(\int_r^t |D_r^j M_s^{l'}|ds\right)^2 dr\right)^2\right]^{\frac{1}{2}} \\
\leq\ & C\sum_{k,k',l'=1}^{m} E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} E\left[\left(\int_{t-\varepsilon}^{t}\int_{t-\varepsilon}^{t} |D_r^j M_s^{l'}|^2(t-r)drds\right)^2\right]^{\frac{1}{2}} \\
\leq\ & C\sum_{k,k',l'=1}^{m} E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} \\
& \times E\left[\left(\int_{t-\varepsilon}^{t}\left(\int_{t-\varepsilon}^{t}|D_r^j M_s^{l'}|^4 dr\right)^{\frac{1}{2}}\left(\int_{t-\varepsilon}^{t}|t-r|^2 dr\right)^{\frac{1}{2}} ds\right)^2\right]^{\frac{1}{2}} \\
=\ & \varepsilon^{\frac{3}{2}} C\sum_{k,k',l'=1}^{m} E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} E\left[\left(\int_{t-\varepsilon}^{t}\left(\int_{t-\varepsilon}^{t}|D_r^j M_s^{l'}|^4 dr\right)^{\frac{1}{2}} ds\right)^2\right]^{\frac{1}{2}} \\
\leq\ & \varepsilon^{\frac{3}{2}} C\sum_{k,k',l'=1}^{m} E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} E\left[\int_0^t\int_0^t |D_r^j M_s^{l'}|^4 drds\right]^{\frac{1}{2}} \\
\leq\ & \varepsilon^{\frac{3}{2}} C\sum_{k,k'=1}^{m} E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t}|Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} \sum_{l'=1}^{m}\left(E\left[\int_0^t \sup_{r\leq s\leq t}|D_r^j X_s^{l'}|^4 dr\right]\right)^{\frac{1}{2}}.
\end{aligned}
$$

This shows that $\frac{1}{\varepsilon}\int_{t-\varepsilon}^{t} A_{r,t}dr \to 0$ in $L^1(\Omega)$ as ε tends to 0. Note that we must choose $\varepsilon > 0$ such that $t-\varepsilon > 0$ holds. Therefore, there exists $\{\varepsilon_n\}_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty}\frac{1}{\varepsilon_n}\int_{t-\varepsilon_n}^{t} A_{r,t}dr = 0, a.s.$$

On the other hand, by the continuity of A_j^i, we have

$$\lim_{n\to\infty}\frac{1}{\varepsilon_n}\int_{t-\varepsilon_n}^{t}\sum_{j=1}^{d}\left|\sum_{i=1}^{m} v_i Y_k^i(t) Z_{k'}^k(r) A_j^{k'}(r, X_r, M_r)\right|^2 dr = \sum_{j=1}^{d}\left|\sum_{i=1}^{m} v_i A_j^i(t, X_t, M_t)\right|^2 > 0,$$

for any nonzero vector $v\in\mathbb{R}^m$ by **(A3)**. By Lemma 5 and Theorem 1 the proof is completed. □

2.3 The absolute continuity of the probability law of $(X_t^i, M_t^{i'})$

In this section, we prove the absolute continuity of the law of $(X_t^i, M_t^{i'}), 1 \leq i, i' \leq m$, in a special case. That is:

(A4) $A^i_l, 1 \le i \le m, 1 \le l \le d$, do not depend on the second space variable,

in addition to **(A1)**-**(A3)**.

Remark 2. *Under* **(A4)**, $\tilde{A}^{k'}_{l',l} = 0$ *in (2.27).*

The following theorem is the main theorem of this section.

Theorem 3. *Assume* **(A1)**-**(A4)**. *Then, for $t > 0$ and $1 \le i, i' \le m$, the law of $(X^i_t, M^{i'}_t)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^2$.*

Proof. Let $v_1, v_2 \in \mathbb{R}\setminus\{0\}$. Note that, by Lemma 3 and 5, for $t > 0$, we have $D^j M^{i'}_t = \mathbf{1}_{[0,\tau^{i'}_t)}(r) D^j_r X^{i'}_{\tau^{i'}_t}$.

First, we assume $v_1 \neq 0, v_2 \neq 0$. By Schwarz's inequality and a trivial inequality $a^2 + b^2 \ge 2ab, a, b \in \mathbb{R}$, we have

$$
\begin{aligned}
&\int_0^t \left| (v_1, v_2) \begin{pmatrix} D^1_r X^i_t \cdots D^d_r X^i_t \\ D^1_r M^{i'}_t \cdots D^d_r M^{i'}_t \end{pmatrix} \right|^2 dr \\
= {} & \int_0^t \sum_{j=1}^d |v_1 D^j_r X^i_t|^2 dr + 2 \int_0^{\tau^{i'}_t} \sum_{j=1}^d v_1 D^j_r X^i_t v_2 D^j_r X^{i'}_{\tau^{i'}_t} dr + \int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_2 D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \\
\ge {} & \int_0^t \sum_{j=1}^d |v_1 D^j_r X^i_t|^2 dr - 2 \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_1 D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_2 D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \right)^{\frac{1}{2}} \\
& + \int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_2 D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \\
\ge {} & 2 \left(\int_0^t \sum_{j=1}^d |v_1 D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_2 D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \right)^{\frac{1}{2}} \\
& -2 \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_1 D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |v_2 D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \right)^{\frac{1}{2}} \\
= {} & 2|v_1||v_2| \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \right)^{\frac{1}{2}} \left[\left(\int_0^t \sum_{j=1}^d |D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} - \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} \right]. \\
= {} & 2|v_1||v_2| \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |D^j_r X^{i'}_{\tau^{i'}_t}|^2 dr \right)^{\frac{1}{2}} \left[\left(\int_{\tau^{i'}_t}^t \sum_{j=1}^d |D^j_r X^i_t|^2 dr + \int_0^{\tau^{i'}_t} \sum_{j=1}^d |D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} \right. \\
& \left. - \left(\int_0^{\tau^{i'}_t} \sum_{j=1}^d |D^j_r X^i_t|^2 dr \right)^{\frac{1}{2}} \right]. \qquad (2.28)
\end{aligned}
$$

Let us prove that

$$\int_0^{\tau_t^{i'}} \sum_{j=1}^d |D_r^j X_{\tau_t^{i'}}^{i'}|^2 dr > 0, \int_{\tau_t^{i'}}^t \sum_{j=1}^d |D_r^j X_t^i|^2 dr > 0, a.s. \tag{2.29}$$

From the same computation as the proof of Theorem 2, we get

$$\begin{aligned}
& E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} |D_r^j X_{\tau_t^{i'}}^{i'}|^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right] \\
\geq \ & \frac{1}{2} E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} |Y_k^i(t) Z_{k'}^k(r) A_j^{k'}(r, X_r)|^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right] \\
& -E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} \left|Y_k^i(t) \int_r^{\tau_t^{i'}} Z_{k'}^k(s) \tilde{B}_{l'}^{k'}(s) D_r^j M_s^{l'} ds\right|^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right],
\end{aligned}$$

and

$$\begin{aligned}
& E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} \left|Y_k^i(t) \int_r^{\tau_t^{i'}} Z_{k'}^k(s) \tilde{B}_{l'}^{k'}(s) D_r^j M_s^{l'} ds\right|^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right] \\
\leq \ & C \sum_{k,k',l'=1}^m E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t} |Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} E\left[\left(\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} \left|\int_r^{\tau_t^{i'}} \tilde{B}_{l'}^{k'}(s) D_r^j M_s^{l'} ds\right|^2 dr\right)^2 \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right]^{\frac{1}{2}} \\
\leq \ & C \sum_{k,k',l'=1}^m E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t} |Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} \\
& \times E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} \left(\int_r^{\tau_t^{i'}} |D_r^j M_s^{l'}|^2 ds\right)^2 dr \int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} (\tau_t^{i'} - r)^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right]^{\frac{1}{2}} \\
= \ & \varepsilon^{\frac{3}{2}} C \sum_{k,k',l'=1}^m E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t} |Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} \left(\int_r^{\tau_t^{i'}} |D_r^j M_s^{l'}|^2 ds\right)^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right]^{\frac{1}{2}} \\
\leq \ & \varepsilon^{\frac{3}{2}} C \sum_{k,k'=1}^m E\left[|Y_k^i(t)|^8\right]^{\frac{1}{4}} E\left[\sup_{0\leq s\leq t} |Z_{k'}^k(s)|^8\right]^{\frac{1}{4}} \sum_{l'=1}^m E\left[\int_0^t \sup_{0\leq s\leq t} |D_r^j X_s^{l'}|^4 dr\right]^{\frac{1}{2}}.
\end{aligned}$$

Therefore, one has

$$\frac{1}{\varepsilon} E\left[\int_{\tau_t^{i'}-\varepsilon}^{\tau_t^{i'}} \left|Y_k^i(t) \int_r^{\tau_t^{i'}} Z_{k'}^k(s) \tilde{B}_{l'}^{k'}(s) D_r^j M_s^{l'} ds\right|^2 dr \mathbf{1}_{\{\tau_t^{i'}-\varepsilon>0\}}\right] \to 0, \tag{2.30}$$

as ε tends to 0. By (2.30) and the proof of Theorem 2, there exists $\{\varepsilon_n^i\}_{n\in\mathbb{N}}$ and $\{\varepsilon_n^{i'}\}_{n\in\mathbb{N}}$ such that $\varepsilon_n^i \downarrow 0$, $\varepsilon_n^{i'} \downarrow 0$ $(n \to \infty)$ and

$$\lim_{n\to\infty} \frac{1}{\varepsilon_n^i} \int_{t-\varepsilon_n^i}^t \sum_{j=1}^d |D_r^j X_t^i|^2 dr \geq \frac{1}{2} \sum_{j=1}^d |A_j^i(t, X_t)|^2 \tag{2.31}$$

and

$$\lim_{n\to\infty}\frac{1}{\varepsilon_n^{i'}}\int_{\tau_t^{i'}-\varepsilon_n^{i'}}^{\tau_t^{i'}}\sum_{j=1}^{d}|D_r^jX_{\tau_t^{i'}}^{i'}|^2dr\mathbf{1}_{\{\tau_t^{i'}-\varepsilon_n^{i'}>0\}}\geq\frac{1}{2}\sum_{j=1}^{d}|A_j^{i'}(\tau_t^{i'},X_{\tau_t^{i'}})|^2, \tag{2.32}$$

almost surely, where we have used the fact that $\lim_{n\to\infty}\mathbf{1}_{\{\tau_t^{i'}-\varepsilon_n^{i'}>0\}}=1$, a.s., which is a consequence of Lemma 2. By Lemma 2, we have $t>\tau_t^i$, thus, there exists $N'\in\mathbb{N}$ such that

$$\begin{aligned}\int_{t-\varepsilon_n^i}^{t}\sum_{j=1}^{d}|D_r^jX_t^i|^2dr &\geq \frac{\varepsilon_n^i}{2}\sum_{j=1}^{d}|A_j^i(t,X_t)|^2,\\ \int_{\tau_t^{i'}-\varepsilon_t^{i'}}^{\tau_t^{i'}}\sum_{j=1}^{d}|D_r^jX_{\tau_t^{i'}}^{i'}|^2dr &\geq \frac{\varepsilon_n^{i'}}{2}\sum_{j=1}^{d}|A_j^{i'}(\tau_t^{i'},X_{\tau_t^{i'}})|^2,\\ t-\varepsilon_n^i &> \tau_t^{i'},\end{aligned}$$

for any $n\geq N'$, almost surely. This implies that (2.29) is true. Therefore, the right hand side of (2.28) is strictly positive for any $v\in\mathbb{R}^2$ such that $v_1\neq 0, v_2\neq 0$.

Second, in the case $v_1=0$ or $v_2=0$ we have

$$\int_0^t\left|(v_1,v_2)\begin{pmatrix}D_r^1X_t^i\cdots D_r^dX_t^i\\ D_r^1M_t^{i'}\cdots D_r^dM_t^{i'}\end{pmatrix}\right|^2dr=\begin{cases}\int_0^t\sum_{j=1}^d|v_1D_r^jX_t^i|^2dr,\text{if } v_2\text{=}0,\\ \int_0^{\tau_t^{i'}}\sum_{j=1}^d|v_2D_r^jM_t^{i'}|^2dr,\text{if } v_1\text{=}0.\end{cases}$$

Therefore, we obtain

$$\int_0^t\left|(v_1,v_2)\begin{pmatrix}D_r^1X_t^i\cdots D_r^dX_t^i\\ D_r^1M_t^{i'}\cdots D_r^dM_t^{i'}\end{pmatrix}\right|^2dr>0, a.s.$$

by (2.31) and (2.32). This finishes the proof. □

Remark 3. *The general $2m$-dimensional study of the law of $(X_t^1,\cdots,X_t^m,M_t^1,\cdots,M_t^m)$ does not follow with the arguments presented here, due to the particular structure used in the calculation of (2.28). Indeed, in the proof of Theorem 3, we have used an inequality $a^2+b^2\geq 2ab, a,b\in\mathbb{R}$.*

Corollary 1. *Under* **(A1)**-**(A4)**, *by the same calculation as that in Theorem 3, for $t>0$ and $1\leq i\neq i'\leq m$, we can prove the absolute continuity of the law of $(M_t^i,M_t^{i'})$ conditioned by the set $\{\tau_t^i\neq\tau_t^{i'}\}$.*

Now we give an example for A_l^i and $A_l^{i'}$ that $\{\tau_t^i\neq\tau_t^{i'}\}$ holds, a.s.

Example 1. *For each $k=i,i'$, let $\{X_s^k, s\in[0,t]\}$ satisfies*

$$X_s^k=x^k+\int_0^s B^k(u,X_u,M_u)du+A_k^kW_s^k,$$

where A_k^k is a nonzero constant then $\tau_t^i\neq\tau_t^{i'}$, a.s.

Proof. By Girsanov theorem, the independence of Brownian motions, and the explicit density function for $\tau_t^k, k=i,i'$ (Problem 8.17 in Chapter 2 of [11]), we obtain the existence of the density function for $\tau_t^i-\tau_t^{i'}$. Then we have $P(\tau_t^i=\tau_t^{i'})=0$. □

2.4 A concluding remark

In this chapter, we proved the absolute continuity of the law of X_t and $(X_t^i, M_t^{i'})$ with Lipschitz coefficients under some additional assumptions. We end this chapter with some remarks on the law of the maximum of processes. There are some theoretical and applicable results about the law of the maximum of continuous processes. In [17] the smoothness of the density function of the maximum of the Wiener sheet is proven. In [9], authors derived some integration by parts formulae involving the maximum and minimum of a one dimensional diffusion to compute the sensitivities of the price of financial products with respect to market parameters called Greeks. Recently, the smoothness of density function of the joint law of a multi-dimensional diffusion at the time when a component attains its maximum time was proven in [10]. In these articles, Garsia-Rodemich-Rumsey's lemma (Lemma A.3.1 of [17]) plays an important role to obtain the results.

Chapter 3

Volatility risk for options depending on extrema and its estimation using kernel methods

3.1 Introduction

The Black-Scholes model has been widely used by practitioners due to its simplicity and the existence of some explicit probability density functions concerning the model. This model assumes a constant volatility. However, in the option market data, we observe that the volatility can not be a constant. This phenomenon is often called "volatility smile" after the shape of observed data-implied volatilities (see [5] or [8]). For this reason, it is natural to consider a general model which may perform better than the Black-Scholes model. On the other hand, in a general model, usually one knows neither the associated explicit density functions nor explicit formulas for option prices. Therefore, the risks involved in options, called Greeks, can only be computed through numerical approximations.

In this chapter, we consider the sensitivity of the model to changes in the volatility parameter for options depending on the extrema (maximum or minimum). We call this sensitivity the vega index and we focus our discussion on the calculation of the vega index. In a general model, the volatility is not a constant and this makes the discussion complicated mathematically. We introduce a perturbation parameter to consider the directional derivatives for the diffusion coefficients to calculate the vega index. In particular, this problem has been discussed by some authors. In [6], the authors obtained a formula to calculate the vega index for options whose payoffs depend on the prices of underlying at fixed times through Malliavin calculus. Other Greeks, such as delta and gamma, which are defined by the sensitivities with respect to the current price of the underlying, for options depending on the extrema are discussed in [9]. In [1], a formula to compute the vega index was obtained in the case of options with payoffs depending on the underlying smoothly (e.g. Asian type option) by using Malliavin calculus.

However, the vega index for options depending on the extrema has not been considered yet, since the extrema of a diffusion process is not sufficiently smooth and therefore difficult to treat from the mathematical point of view. In mathematical finance various credit linked and barrier type products have this kind of feature.

There are mainly two goals in this chapter: One is to consider various options which may depend on the extrema of the underlying and obtain some financial conclusions about the properties of the vega index in a one-dimensional model. The other is to give a methodology to compute the vega index for a specific option by using so-called kernel methods.

To study the structure of vega index, we draw the vega risk profiles in the one-dimensional model and compare the vega index obtained in this one-dimensional model with the one in the Black-Scholes model (see Table 3.1 and Figure 3.4). According to Table 3.1 and Figure 3.4, these different models give different values of the vega index, even if the payoff functions are the same, and this difference is crucial for practitioners, since in practice hedging procedures are done based on the value of vega index obtained in each model.

Technically, in this chapter, we consider a one-dimensional stochastic differential equation (SDE) with time-independent coefficients as the dynamics of an asset price under the pricing measure P. The results obtained in this chapter may be a breakthrough to study the Greeks in so-called stochastic volatility models which are often used by practitioners (see [4], for a relationship between one-dimensional models and stochastic volatility models). To deal with the extrema of diffusion process, we use the Lamperti method (see Exercise 5.2.20 of [11], for example). That is, first we transform the SDE using Girsanov's theorem to a Stratonovich type SDE without drift coefficient which can then be expressed as a monotone transformation of a Wiener process. This method is different from the one considered in [9] where the Garsia-Rodemich-Rumsey's lemma (see Lemma A.3.1 of [17]) plays an important role. Although techniques used in [9] are quite interesting, the formulas obtained there have high computational complexity. However, the formula obtained in this chapter is much simpler.

By working under a new measure, we can express the extrema of diffusion process in a simple fashion and calculate the directional derivatives. In addition, we use the duality formula of Malliavin calculus as it appears in [17, Page 37] to obtain a formula that may give a better expression to the vega index for some numerical methods such as Monte Carlo simulation.

The formula of the vega index obtained in this chapter allows one to decompose it into three components: the extrema and maturity feature of options, and a by-product of the Girsanov transformation. The intention of the current research is to try to reveal some properties of the structure of these three components for realistic options. Through simulation studies in Section 3 of this chapter, one can see that the decomposition of the vega index for barrier type options has some interesting properties. For example, when we consider an up-in call option, our Monte Carlo analysis shows that for the option with lower barrier, the vega index is mostly conveyed by the maturity feature of the payoff, while for the option with higher barrier, the extrema feature controls most of the vega index. We can see the existence of a barrier that determines which component in the decomposition is of most importance (see Figure 3.1). Moreover, we observe that for the options with short maturity, we have to pay more attention to the change of the value of vega index with respect to the maturity (see Figure 3.2 and 3.3). These results seem to be valid among several types of options, according to our numerical experiments.

Unfortunately, each component of the decomposition formula obtained here for binary barrier options involves the Dirac delta functionals, therefore, we give a method to approximate the delta functionals called kernel methods. The kernel method is quite effective to some numerical problems appearing in various fields such as finance. A basic kernel method to estimate probability density functions is given in [18, Chapter 2-4], and it is applied in [13] to compute the Greeks for options with discontinuous payoffs. To address this method, we shall define an estimator for the delta functional by using a so-called kernel function and bandwidth parameter. The bandwidth parameter controls the bias and variance of the estimator, therefore, its choice is quite important for using the kernel method. In order to choose the best bandwidth, we define an asymptotic mean squared error (AMSE) as an error for the estimator, then we look for a bandwidth so that the AMSE is as small as possible. If there exists a bandwidth which minimizes the AMSE, then we call it the optimal bandwidth. A theorem to express the optimal bandwidth is stated as the main theorem of Section 4 and some numerical results obtained by the kernel method are also given.

This chapter is organized as follows. In Section 2, we provide the mathematical result on the decomposition of vega index. In Section 3, we carry out Monte Carlo simulations and obtain some results on the structure of vega index as mentioned in the previous paragraph. In Section 4, we consider a binary barrier option and discuss the kernel method. Then we apply it to the computation of the vega index for this option. Some numerical results obtained with the kernel method are given in Section 5. In the Appendices, we give some lemmas and proofs of our results.

Throughout the chapter, we use $C_b^k(A,B)$ to denote the space of B-valued k times continuously differentiable functions defined on A with bounded derivatives. For a differentiable function F from $\mathbb{R}^m$ to $\mathbb{R}$ where $m \in \mathbb{N}$, we define $\partial_i F(x) := \frac{\partial F}{\partial x_i}(x)$ for $x \in \mathbb{R}^m$ and $1 \le i \le m$. The letters C and $C_i, i \in \mathbb{N}$ denote positive constants which may depend on f, p, x and T that will appear in this chapter, and the values of C and C_i may change from line to line. We define $\mathbb{R}_+ := (0,\infty)$ and E^P as the expectation under a probability measure P.

3.2 Main result: Vega index for options depending on the extrema

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space which supports a one-dimensional Wiener process $\{W_t, t \in [0,T]\}$. For $\sigma, \hat{\sigma}, b : \mathbb{R} \to \mathbb{R}$ and $x > 0$, we consider the following stochastic differential equation (SDE in short),

$$\begin{cases} dS_t^\varepsilon = b(S_t^\varepsilon)dt + \sigma^\varepsilon(S_t^\varepsilon)dW_t \\ S_0^\varepsilon = x, \end{cases} \tag{3.1}$$

where σ^ε is of the form $\sigma^\varepsilon(z) = \sigma(z) + \varepsilon\hat{\sigma}(z)$, $\varepsilon \in [0,1]$.

For $f : \mathbb{R}^2 \to \mathbb{R}$, we consider the quantity $\Pi^\varepsilon := E^P[f(\max_{0\le t\le T} S_t^\varepsilon, S_T^\varepsilon)]$. We assume the following hypotheses.

(H1) $\sigma,\ \hat{\sigma} \in C_b^2(\mathbb{R}_+, \mathbb{R}_+)$ and $b \in C_b^2(\mathbb{R}_+, \mathbb{R}_+)$.

(H2) There exists $\sigma_0 > 0$ such that $\sigma^\varepsilon(y) \ge \sigma_0 y$ for all $y \in \mathbb{R}_+$ and $\varepsilon \in [0,1]$.

(H3) There exists $r_0 \in (0,1]$ and $\sigma_1 > 0$ such that $\sigma_0 y \le \sigma^\varepsilon(y) \le \sigma_1 y$, for all $y \in \mathbb{R}_+$ such that $|y| < r_0$.

(H4) The function $\frac{b}{\sigma^\varepsilon}$ is bounded.

(H5) $f \in C_b^1(\mathbb{R}_+^2, \mathbb{R}_+)$.

Note that by (**H1**), for all $\varepsilon \in [0,1]$, (3.1) has a unique strong solution, and let $S^\varepsilon = \{S_t^\varepsilon, t \in [0,T]\}$ be the solution to (3.1). In finance, Π^ε defines a perturbed option price with a payoff function f. We consider the quantity $\frac{\partial \Pi^\varepsilon}{\partial \varepsilon}\big|_{\varepsilon=0}$ and we call this the vega index of this option.

Our main result is the following theorem. It gives the decomposition formula for vega index.

Theorem 4. *Assume the above hypotheses* **(H1)-(H5)**. *Then the following expression for vega index is valid.*

$$\begin{aligned}\left.\frac{\partial \Pi^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} &= E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)\sigma(\max_{0\le t\le T} S_t)\int_x^{\max_{0\le t\le T} S_t}\frac{\hat\sigma}{(\sigma)^2}(y)dy\right]\\ &+E^P\left[\partial_2 f(\max_{0\le t\le T} S_t, S_T)\sigma(S_T)\int_x^{S_T}\frac{\hat\sigma}{(\sigma)^2}(y)dy\right]\\ &+E^P\left[\int_0^T\left(\partial_1 f(\max_{0\le t\le T} S_t, S_T)Y_\eta I_{[0,\eta]}(t)+\partial_2 f(\max_{0\le t\le T} S_t, S_T)Y_T\right)\right.\\ &\left.\times\left\{\frac{(b'\sigma-b\sigma'-\frac{(\sigma)^2\sigma''}{2})(S_t)}{Y_t}\left(\int_x^{S_t}\frac{\hat\sigma}{(\sigma)^2}(y)dy\right)-\frac{(\frac{b\hat\sigma}{\sigma}+\frac{\sigma\hat\sigma'}{2})(S_t)}{Y_t}\right\}dt\right]\end{aligned} \tag{3.2}$$

where $S_t := S_t^0$, $Y_t := \frac{\partial S_t}{\partial x}$ *and* $\eta := \arg\max_{0\le t\le T} S_t$.

Before giving the proof of theorem, let us state some remarks and prove some preparatory lemmas.

Remark 4. A. *One can prove that* η *in Theorem 4 is almost surely unique by using Theorem 5 of this chapter and the fact that the time at which the maximum of a one-dimensional Wiener process over* $[0,T]$ *is attained is almost surely unique (see Remark 2.8.16 of [11]).*

B. *We have given the above theorem a general mathematical form. In finance, we should assume that* P *is the equivalent martingale measure, the interest rate is zero for convenience and that* S^ε *is a martingale. Then, in that particular case,* Π^ε *has the interpretation of a perturbed option price. Furthermore, as our main goal is to describe the structure of vega index, we do not discuss the possible mathematical extensions to payoff functions* f *to avoid cumbersome technicalities and long arguments. In Section 3, we will obtain some numerical results for irregular* f *with special* b, σ *and* $\hat\sigma$. *This extension can be done by using the explicit density functions for* $\max_{0\le t\le T} S_t$ *and* $(\max_{0\le t\le T} S_t, S_T)$. *The proof of this extension can be found in Appendix B.*

C. *The first term of (3.2) comes from the differentiation with respect to the maximum of asset price, the second term is due to the maturity price of the asset and the third term is a result of a change of measure. We call these three terms "extrema sensitivity", "terminal sensitivity" and "drift sensitivity", respectively.*

D. *When we consider the following measure change*

$$\frac{dQ(\varepsilon)}{dP} := \exp\left\{\int_0^T\left(\frac{\sigma^{\varepsilon\prime}}{2}-\frac{b}{\sigma^\varepsilon}\right)(S_t^\varepsilon)dW_t-\frac{1}{2}\int_0^T\left(\frac{\sigma^{\varepsilon\prime}}{2}-\frac{b}{\sigma^\varepsilon}\right)^2(S_t^\varepsilon)dt\right\},$$

where we have defined $\sigma^{\varepsilon\prime} := (\sigma^\varepsilon)'$, *then under* $Q(\varepsilon)$

$$\hat W_t^\varepsilon := W_t - \int_0^t\left(\frac{\sigma^{\varepsilon\prime}}{2}-\frac{b}{\sigma^\varepsilon}\right)(S_u^\varepsilon)du$$

is a one-dimensional Wiener process. Note that due to the boundedness for $\frac{b}{\sigma^\varepsilon}$ *and* $\sigma^{\varepsilon\prime}$ *the Novikov condition is clearly satisfied. Then under* $Q(\varepsilon)$, S^ε *can be written as*

$$\begin{cases} dS_t^\varepsilon = \frac{1}{2}\sigma^\varepsilon\sigma^{\varepsilon\prime}(S_t^\varepsilon)dt+\sigma^\varepsilon(S_t^\varepsilon)d\hat W_t^\varepsilon = \sigma^\varepsilon(S_t^\varepsilon)\circ d\hat W_t^\varepsilon,\\ S_0^\varepsilon = x,\end{cases}$$

where $\circ\, d\hat{W}_t^\varepsilon$ *denotes Stratonovich integral. Finally, we can write* Π^ε *as follows,*

$$\Pi^\varepsilon = E^P[f(\max_{0\le t\le T} S_t^\varepsilon, S_T^\varepsilon)] = E^{Q(\varepsilon)}\left[f(\max_{0\le t\le T} S_t^\varepsilon, S_T^\varepsilon)\frac{dP}{dQ(\varepsilon)}\right].$$

E. *Under* $Q(\varepsilon)$*,* S^ε *is driven by* $\hat{W}^\varepsilon$ *and the distribution of* $\hat{W}^\varepsilon$ *does not depend on* ε *under* $Q(\varepsilon)$*. Let* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{Q})$ *be another complete probability space and* $\tilde{W}$ *be a one-dimensional Wiener process under* $\tilde{Q}$*. Then* Π^ε *is written as*

$$\Pi^\varepsilon = E^{\tilde{Q}}[f(\max_{0\le t\le T} S_t^\varepsilon, S_T^\varepsilon)\exp(X_T^\varepsilon)],$$

where S_t^ε *satisfies*

$$\begin{cases} dS_t^\varepsilon = \sigma^\varepsilon(S_t^\varepsilon)\circ d\tilde{W}_t \\ S_0^\varepsilon = x, \end{cases} \tag{3.3}$$

and

$$X_T^\varepsilon = -\int_0^T \left(\frac{\sigma^{\varepsilon\prime}}{2} - \frac{b}{\sigma^\varepsilon}\right)(S_t^\varepsilon)d\tilde{W}_t - \frac{1}{2}\int_0^T \left(\frac{\sigma^{\varepsilon\prime}}{2} - \frac{b}{\sigma^\varepsilon}\right)^2 (S_t^\varepsilon)dt.$$

We use the SDE of the form (3.3) to write down S^ε *with only* $\tilde{W}$ *so that we can express* $\max_{0\le t\le T} S_t^\varepsilon$ *by* $\max_{0\le t\le T}\tilde{W}_t$*. From now on, we use the notation* $X_T := X_T^0$*.*

We introduce a function F_ε which is used to express the solution to (3.3) in an explicit form.

Definition 1. *(Lamperti transform) For* $\varepsilon \in [0,1]$*, define* $F_\varepsilon : \mathbb{R}_+ \to \mathbb{R}$ *as*

$$F_\varepsilon(z) := \int_1^z \frac{1}{\sigma^\varepsilon(y)}dy.$$

Note that the inverse function F_ε^{-1} exists, since F_ε is a continuous monotone increasing function due to the assumption **(H2)**. Furthermore, it is clear that F_ε and F_ε^{-1} are differentiable with respect to z and we have

$$\begin{aligned} \frac{\partial F_\varepsilon}{\partial z}(z) &= \frac{1}{\sigma^\varepsilon(z)}, \\ \frac{\partial F_\varepsilon^{-1}}{\partial z}(z) &= \frac{1}{\frac{\partial F_\varepsilon}{\partial z}(F_\varepsilon^{-1}(z))} = \sigma^\varepsilon(F_\varepsilon^{-1}(z)). \end{aligned} \tag{3.4}$$

In this set-up one has the following result.

Theorem 5. *Under hypothesis* **(H1)-(H2)***, there exists a unique strong solution to (3.3). Furthermore, under* $\tilde{Q}$*, the solution to the SDE (3.3) can be written as follows,*

$$S_t^\varepsilon = F_\varepsilon^{-1}(F_\varepsilon(x) + \tilde{W}_t). \tag{3.5}$$

Therefore, one has $\max_{0\le t\le T} S_t^\varepsilon = F_\varepsilon^{-1}(F_\varepsilon(x) + \max_{0\le t\le T}\tilde{W}_t)$*.*

Proof. By (3.4) and applying Itô's formula for Stratonovich integral to $F_\varepsilon^{-1}(F_\varepsilon(x)+z)$, it is easy to see that (3.5) is a solution to (3.3). On the other hand, if there exists a solution to (3.3), then, again by applying Itô's formula for Stratonovich integral to $F_\varepsilon(z)$, the solution can be expressed by (3.5). Thus, one obtains the uniqueness.

The equality $\max_{0\leq t\leq T} S_t^\varepsilon = F_\varepsilon^{-1}(F_\varepsilon(x)+\max_{0\leq t\leq T}\tilde{W}_t)$ is a conclusion of the monotonicity for $F_\varepsilon^{-1}(z)$. □

Remark 5. *Although the representation of $F_\varepsilon^{-1}(F_\varepsilon(x)+\tilde{W}_t)$ is clearly continuous in (t,ε), this does not imply the continuity of the solution to (3.3) in ε, since the exceptional set such that Itô's formula does not hold may depend on ε.*

To overcome this problem, we modify the solution to (3.3) to be continuous in (t,ε). This procedure will be done in Appendix A.2.

The above representation is the key formula which allows us to obtain Theorem 4. Our next step is to state some results on the regularity of S_t^ε and X_T^ε with respect to ε and the exchange between $E^{\tilde{Q}}[\cdot]$ and $\frac{\partial}{\partial\varepsilon}(\cdot)|_{\varepsilon=0}$. The following four lemmas will be listed and their proofs can be found in Appendix A.

Lemma 7. *Let* **(H1)-(H2)** *be satisfied. Let S^ε be the solution to (3.3). Then, $\max\limits_{0\leq t\leq T} S_t^\varepsilon$ is differentiable with respect to $\varepsilon\in[0,1]$ and the following equation holds,*

$$\frac{\partial}{\partial\varepsilon}\left(\max_{0\leq t\leq T} S_t^\varepsilon\right)\Bigg|_{\varepsilon=0} = \sigma(\max_{0\leq t\leq T} S_t)\int_x^{\max\limits_{0\leq t\leq T} S_t}\frac{\hat{\sigma}}{(\sigma)^2}(y)dy, a.s.$$

Lemma 8. *Let* **(H1)-(H2)** *be satisfied. Let S^ε be the solution to (3.3). Then, S^ε is differentiable with respect to $\varepsilon\in[0,1]$ and we have*

$$Z_t := \frac{\partial S_t^\varepsilon}{\partial\varepsilon}\bigg|_{\varepsilon=0} = \sigma(S_t)\int_x^{S_t}\frac{\hat{\sigma}}{(\sigma)^2}(y)dy,\ \forall t\in[0,T],\ a.s. \tag{3.6}$$

Lemma 9. *Let* **(H1)-(H3)** *be satisfied. Then, X_T^ε is differentiable with respect to $\varepsilon\in[0,1]$ and it holds that*

$$\begin{aligned}
\frac{\partial X_T^\varepsilon}{\partial\varepsilon}\bigg|_{\varepsilon=0} &= -\frac{1}{2}\int_0^T\left(\sigma''(S_t)Z_t+\hat{\sigma}'(S_t)\right)d\tilde{W}_t \\
&\quad+\int_0^T\frac{(b'\sigma)(S_t)Z_t-b(S_t)\left(\sigma'(S_t)Z_t+\hat{\sigma}(S_t)\right)}{(\sigma)^2(S_t)}d\tilde{W}_t \\
&\quad-\frac{1}{2}\int_0^T\left(\frac{\sigma'}{2}-\frac{b}{\sigma}\right)(S_t)\left(\sigma''(S_t)Z_t+\hat{\sigma}'(S_t)\right)dt \\
&\quad+\int_0^T\left(\frac{\sigma'}{2}-\frac{b}{\sigma}\right)(S_t)\frac{(b'\sigma)(S_t)Z_t-b(S_t)\left(\sigma'(S_t)Z_t+\hat{\sigma}(S_t)\right)}{(\sigma)^2(S_t)}dt, a.s.
\end{aligned}$$

From the above three lemmas one can see the correspondence of the derivatives to (3.2). In fact, the three lemmas correspond to the derivative of the maximum, the underlying at maturity and the change of measure with respect to ε, respectively.

In addition to the above lemmas, we need the following lemma about the exchange between $E^{\tilde{Q}}[\cdot]$ and $\frac{\partial}{\partial\varepsilon}(\cdot)|_{\varepsilon=0}$.

Lemma 10. *Let* **(H1)-(H5)** *be satisfied. Then, we have the following equation,*

$$\frac{\partial}{\partial\varepsilon}\left(E^{\tilde{Q}}[f(\max_{0\le t\le T}S_t^\varepsilon, S_T^\varepsilon)\exp(X_T^\varepsilon)]\right)\bigg|_{\varepsilon=0} = E^{\tilde{Q}}\left[\frac{\partial}{\partial\varepsilon}\left(f(\max_{0\le t\le T}S_t^\varepsilon, S_T^\varepsilon)\exp(X_T^\varepsilon)\right)\bigg|_{\varepsilon=0}\right].$$

Now, let us prove the main theorem.

Proof of Theorem 4. As mentioned in Remark 4.**E**, Π^ε is expressed using $\tilde{Q}$ which does not depend on ε, therefore, we have

$$\begin{aligned}\frac{\partial\Pi^\varepsilon}{\partial\varepsilon}\bigg|_{\varepsilon=0} &= E^{\tilde{Q}}\left[\partial_1 f(\max_{0\le t\le T}S_t^\varepsilon, S_T^\varepsilon)\frac{\partial}{\partial\varepsilon}\left(\max_{0\le t\le T}S_t^\varepsilon\right)\exp(X_T^\varepsilon)\bigg|_{\varepsilon=0}\right]\cdots\text{(i)}\\ &+ E^{\tilde{Q}}\left[\partial_2 f(\max_{0\le t\le T}S_t^\varepsilon, S_T^\varepsilon)\frac{\partial S^\varepsilon}{\partial\varepsilon}(T)\exp(X_T^\varepsilon)\bigg|_{\varepsilon=0}\right]\cdots\text{(ii)}\\ &+ E^{\tilde{Q}}\left[f(\max_{0\le t\le T}S_t^\varepsilon, S_T^\varepsilon)\frac{\partial}{\partial\varepsilon}\left(\exp(X_T^\varepsilon)\right)\bigg|_{\varepsilon=0}\right]\cdots\text{(iii)}.\end{aligned}$$

(i).By Theorem 5 and Lemma 7 we obtain that (i) equals the first term on the right of (3.2) after the application of Girsanov's theorem.
(ii).By Lemma 8 as in the proof of (i), the result is trivial.
(iii).By Lemma 9, $\frac{\partial X_T^\varepsilon}{\partial\varepsilon}\big|_{\varepsilon=0}$ is

$$\begin{aligned}\frac{\partial X_T^\varepsilon}{\partial\varepsilon}\bigg|_{\varepsilon=0} &= -\frac{1}{2}\int_0^T(\sigma''(S_t)Z_t+\hat{\sigma}'(S_t))d\tilde{W}_t + \int_0^T\frac{(b'\sigma)(S_t)Z_t - b(S_t)(\sigma'(S_t)Z_t+\hat{\sigma}(S_t))}{(\sigma)^2(S_t)}d\tilde{W}_t\\ &-\int_0^T\left(\frac{1}{2}\sigma'-\frac{b}{\sigma}\right)(S_t)\left\{\frac{1}{2}(\sigma''(S_t)Z_t+\hat{\sigma}'(S_t)) - \frac{(b'\sigma)(S_t)Z_t - b(S_t)(\sigma'(S_t)Z_t+\hat{\sigma}(S_t))}{(\sigma)^2(S_t)}\right\}dt\end{aligned}\tag{3.7}$$

where Z_t is defined by (3.6). By Lemma 8, we can express Z as follows;

$$Z_t = (\sigma G)(S_t).\tag{3.8}$$

where G is defined as

$$G(z) := \int_x^z \frac{\hat{\sigma}}{(\sigma)^2}(y)dy.$$

Using (3.7), (3.8) and the duality formula in Malliavin calculus we obtain

$$\begin{aligned}\text{(iii)} &= E^{\tilde{Q}}\Bigg[f(\max_{0\le t\le T}S_t, S_T)\exp(X_T)\Bigg\{-\int_0^T\left(\frac{\sigma\sigma''G+\hat{\sigma}'}{2} - \frac{(\sigma)^2 b'G - b(\sigma\sigma'G+\hat{\sigma})}{(\sigma)^2}\right)(S_t)d\tilde{W}_t\\ &\quad -\int_0^T\left(\frac{\sigma'}{2}-\frac{b}{\sigma}\right)\left(\frac{\sigma\sigma''G+\hat{\sigma}'}{2} - \frac{(\sigma)^2 b'G - b(\sigma\sigma'G+\hat{\sigma})}{(\sigma)^2}\right)(S_t)dt\Bigg\}\Bigg]\\ &= E^{\hat{P}}\left[f(\max_{0\le t\le T}S_t, S_T)\left\{-\int_0^T\left(\frac{\sigma\sigma''G+\hat{\sigma}'}{2} - \frac{(\sigma)^2 b'G - b(\sigma\sigma'G+\hat{\sigma})}{(\sigma)^2}\right)(S_t)dB_t\right\}\right]\\ &= E^{\hat{P}}\left[\int_0^T D_t\left(f(\max_{0\le t\le T}S_t, S_T)\right)\left\{\left(b'-\frac{b\sigma'}{\sigma}-\frac{\sigma\sigma''}{2}\right)G - \left(\frac{b\hat{\sigma}}{(\sigma)^2}+\frac{\hat{\sigma}'}{2}\right)\right\}(S_t)dt\right],\end{aligned}$$

where we have used a change of measure $\frac{d\hat{P}}{d\hat{Q}} := \exp(X_T)$ and B is a $\hat{P}$-Wiener process. Moreover, in the above equality, D_t denotes Malliavin derivative operator with respect to B.

Furthermore, we have

$$D_t\left(f(\max_{0\le t\le T} S_t, S_T)\right) = \partial_1 f(\max_{0\le t\le T} S_t, S_T) D_t\left(\max_{0\le t\le T} S_t\right) + \partial_2 f(\max_{0\le t\le T} S_t, S_T) D_t(S_T).$$

Due to Theorem 2.2.1 of [17] and Lemma 3, we obtain

$$D_t\left(\max_{0\le t\le T} S_t\right) = \frac{Y_\eta}{Y_t}\sigma(S_t) I_{[0,\eta]}(t),$$

where $Y_t := \frac{\partial S_t}{\partial x}$ and $\eta := \arg\max_{0\le t\le T} S_t^\varepsilon$, and

$$D_t(S_T) = \frac{Y_T}{Y_t}\sigma(S_t).$$

Finally we obtain (3.2). □

Remark 6. *In the calculation of (iii), it is clear that we can avoid the appearance of the derivative of f without the duality formula and obtain*

$$\text{(iii)} = E^{\hat{P}}\left[f(\max_{0\le t\le T} S_t, S_T)\left\{-\int_0^T \left(\frac{\sigma\sigma''G + \hat{\sigma}'}{2} - \frac{(\sigma)^2 b'G - b(\sigma\sigma'G + \hat{\sigma})}{(\sigma)^2}\right)(S_t)dB_t\right\}\right]. \tag{3.9}$$

However, we still prefer to avoid the stochastic integrals in (3.9) in order to obtain the stability of Monte Carlo estimates.

Remark 7. *We can apply the above technique to obtain the representation of the vega index for options whose payoffs depend on the minimum of the asset. We define $\tilde{\Pi}^\varepsilon := E^P[f(\min_{0\le t\le T} S_t^\varepsilon, S_T^\varepsilon)]$, then we have the following formula for the vega index,*

$$\begin{aligned}
\left.\frac{\partial\tilde{\Pi}^\varepsilon}{\partial\varepsilon}\right|_{\varepsilon=0} &= E^P\left[\partial_1 f(\min_{0\le t\le T} S_t, S_T)\sigma(\min_{0\le t\le T} S_t)\int_x^{\min_{0\le t\le T} S_t} \frac{\hat{\sigma}}{(\sigma)^2}(y)dy\right] \\
&+ E^P\left[\partial_2 f(\min_{0\le t\le T} S_t, S_T)\sigma(S_T)\int_x^{S_T}\frac{\hat{\sigma}}{(\sigma)^2}(y)dy\right] \\
&+ E^P\left[\int_0^T\left(\partial_1 f(\min_{0\le t\le T} S_t, S_T)Y_{\tilde{\eta}} I_{[0,\tilde{\eta}]}(t) + \partial_2 f(\min_{0\le t\le T} S_t, S_T)Y_T\right)\right. \\
&\left.\times\left\{\frac{(b'\sigma - b\sigma' - \frac{(\sigma)^2\sigma''}{2})(S_t)}{Y_t}\left(\int_x^{S_t}\frac{\hat{\sigma}}{(\sigma)^2}(y)dy\right) - \frac{(\frac{b\hat{\sigma}}{\sigma} + \frac{\sigma\hat{\sigma}'}{2})(S_t)}{Y_t}\right\}dt\right],
\end{aligned}$$

where $\tilde{\eta} := \arg\min_{0\le t\le T} S_t$. This formula provides possibilities to applications to credit linked products.

3.3 Numerical experiment 1: Structure of the vega index

In this section, we obtain some numerical results by using (3.2). There are mainly two aims to carry out the numerical experiments: One is to study the structure of the vega index in the model of form (3.1), and the other is to compare the values of vega index obtained with (3.2) with the values of vega obtained in the Black-Scholes model.

In this section, for convenience, we call the vega index obtained in (3.2) "LV vega" (LV stands for "local volatility") and the vega in the Black-Scholes model "BS vega". BS vega will be defined in the next subsection.

3.3.1 Preliminary: assumptions of the one-dimensional model and the definition of the vega in the Black-Scholes model

The one-dimensional model we will consider is specified as follows. Let $b(z) = 0$, $\sigma(z) = \tilde{\sigma} z$ and

$$\hat{\sigma}(z) = \begin{cases} az & (z \leq c), \\ \frac{ac^{1-\beta}}{\beta} z^{\beta} + (1 - \frac{1}{\beta})ac & (z > c), \end{cases} \tag{3.10}$$

where $a > 0$, $0 < \beta \leq 1$, $c > 1$ and $\tilde{\sigma} > 0$ are constants. This setting characterizes the Black-Scholes model perturbed using a CEV like model. Note that the parameter "a" is the gradient and "β" is the convexity of the volatility surface. We introduce the parameter "c" so that $\hat{\sigma}$ satisfies hypothesis **(H3)**.

The function $\hat{\sigma}(z)$ defined by (3.10) does not satisfy **(H1)**, since $\hat{\sigma}''(z)$ is not continuous at $z = c$. However, we have the following lemma. The proof can be found in Appendix C.

Lemma 11. *Let $b(z) = 0$, $\sigma(z) = \tilde{\sigma} z$. Assume that $\hat{\sigma}$ is defined by (3.10). Then, (3.2) holds for $f \in C_b^1(\mathbb{R}_+^2, \mathbb{R}_+)$.*

Now, let us define the vega in the Black-Scholes model ("BS vega"). Let S^{BS} be the solution to

$$\begin{cases} dS_t^{BS} = \tilde{\sigma} S_t^{BS} dW_t, \\ S_0^{BS} = x, \end{cases} \tag{3.11}$$

where $\tilde{\sigma} > 0$ is a positive constant called the volatility parameter of Black-Scholes model. For a payoff function f we define $\Pi^{BS}(\tilde{\sigma})$ by

$$\Pi^{BS}(\tilde{\sigma}) := E^P[f(\max_{0 \leq t \leq T} S_t^{BS}, S_T^{BS})],$$

and we define "BS vega" by

$$\frac{\partial \Pi^{BS}}{\partial \tilde{\sigma}}(\tilde{\sigma}). \tag{3.12}$$

Note that we can obtain the explicit formula for $\frac{\partial \Pi^{BS}}{\partial \tilde{\sigma}}(\tilde{\sigma})$ using the explicit density function for $(\max_{0 \leq t \leq T} S_t^{BS}, S_T^{BS})$.

Remark 8. *Under the above assumptions, we have $\Pi^0 = \Pi^{BS}(\tilde{\sigma})$ for all $\tilde{\sigma} > 0$ and payoff functions f, therefore, the prices of options are the same in two models (3.1) and (3.11).*

To compute LV vega, we set the initial price of a stock $x = 80$ and the volatility $\tilde{\sigma} = 0.3$. For the parameters of the perturbation function we set $a = 3$, $\beta = 0.9$ and $c = 50$. We first consider options with the strike $K = 100$ and the maturity $T = 1$, then we shall change the value of maturity T. We set the number of partitions of the interval $[0, T]$ $n = 10^3$ and the number of simulations $N = 10^6$, since under this setting our preparatory simulations converged well.

We consider first the case of payoff function which depends only on the maximum in Subsection 3.3.2 and then another that depends on the maximum and the terminal value of the stock, in Subsection 3.3.3, so that we can study the three components of LV vega as mentioned in Remark 4.**C**.

3.3.2 The case of payoff functions depending on only one component

We assume f is of the form

$$f(y, z) = f(y) = (y - K)_+,$$

where $K > x$. This option is called a lookback call option with strike K. In this case we can show that (3.2) is valid with

$$f'(y) := I_{(K,\infty)}(y),$$

although f does not belong to $C_b^1(\mathbb{R}_+^2, \mathbb{R}_+)$. Namely, the following equation holds (recall that $\eta = \arg\max_{0 \le t \le T} S_t$),

$$\begin{aligned} \left.\frac{\partial \Pi^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} &= \frac{1}{\tilde{\sigma}} E^P \left[I_{(K,\infty)}(\max_{0 \le t \le T} S_t) \max_{0 \le t \le T} S_t \int_x^{\max_{0 \le t \le T} S_t} \frac{\hat{\sigma}(y)}{y^2} dy \right] \\ &\quad - \frac{\tilde{\sigma}}{2} E^P \left[I_{(K,\infty)}(\max_{0 \le t \le T} S_t) \max_{0 \le t \le T} S_t \int_0^\eta \hat{\sigma}'(S_t) dt \right]. \end{aligned} \tag{3.13}$$

The proof of (3.13) can be found in Appendix B. We have the following numerical results,

	LV vega	BS vega
Lookback call option	171.523	60.145
European call option	79.791	26.757

Table 3.1: Vega index in two models.

In Table 3.1, LV vega for a European call option is obtained by replacing the payoff function $f(y, z) = (y - K)_+$ with $f(y, z) = (z - K)_+$ and $\max_{0 \le t \le T} S_t$ with S_T in (3.13).

We observe from the above table that there is a large difference between LV vega and BS vega while these two different models provide the same option price for an arbitrary payoff function (see Remark 8). This difference may be crucial for traders of financial institutions, since once they trade an option, they start hedging procedures using the risks calculated at the same time with the option price. Therefore, if they use only Black-Scholes model, they have much hedging (model) error in the case that the volatility surface changes in the direction previously indicated in (3.10). At this point, the difference of the values of vega index between the lookback call option and the European call option is large.

We will study the extrema and maturity features of LV vega in details with an example of a barrier type option.

3.3.3 The case of payoff functions depending on the extrema and the terminal value of the underlying

We assume f is of the form

$$f(y,z) = I_{(U,\infty)}(y)(z-K)_+,$$

where $x < K < U$. This option is called an up-in call option with the strike K and barrier U. In this case we have

$$\begin{aligned}
\left.\frac{\partial \Pi^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} &= \frac{1}{\tilde{\sigma}} E^P \left[\delta_U(\max_{0\le t\le T} S_t)(S_T-K)_+ \max_{0\le t\le T} S_t \int_x^{\max_{0\le t\le T} S_t} \frac{\hat{\sigma}(y)}{y^2} dy \right] \\
&\quad + \frac{1}{\tilde{\sigma}} E^P \left[I_{(U,\infty)}(\max_{0\le t\le T} S_t) I_{(K,\infty)}(S_T) S_T \int_x^{S_T} \frac{\hat{\sigma}(y)}{y^2} dy \right] \\
&\quad - \frac{1}{2} E^P \left[I_{(U,\infty)}(\max_{0\le t\le T} S_t)(S_T-K)_+ \int_0^T \hat{\sigma}'(S_t) dB_t \right] \\
&=: E(U,K) + T(U,K) + D(U,K), \qquad (3.14)
\end{aligned}$$

where δ_U denotes the Dirac delta functional at U. For drift sensitivity $D(U,K)$, we have used the expression of the form (3.9) to avoid the appearance of the delta functional as we mentioned in Remark 6. Note that we can compute extrema sensitivity $E(U,K)$ and terminal sensitivity $T(U,K)$ explicitly from the explicit density function for $(\max_{0\le t\le T} S_t, S_T)$.

As we mentioned in the previous subsection, in order to obtain the above equation we have to extend f to an irregular function. This extension can be done using the same method in the proof of (3.13). Thus, we omit the proof.

Using (3.14), we have the following numerical results,

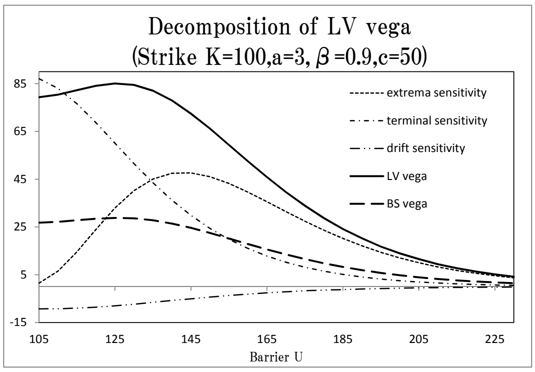

Figure 3.1: Decomposition of LV vega.

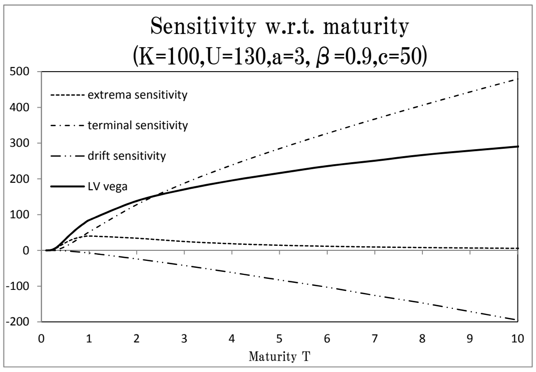

Figure 3.2: LV vega in terms of the maturity.

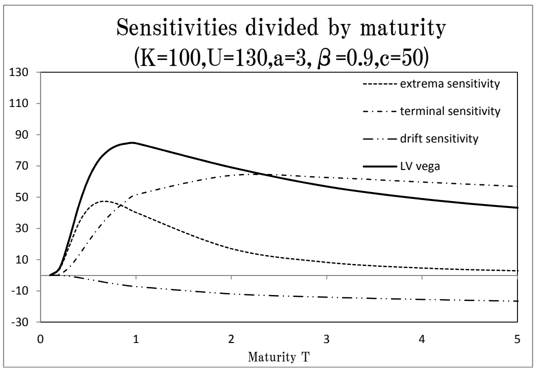

Figure 3.3: Standardized LV vega.

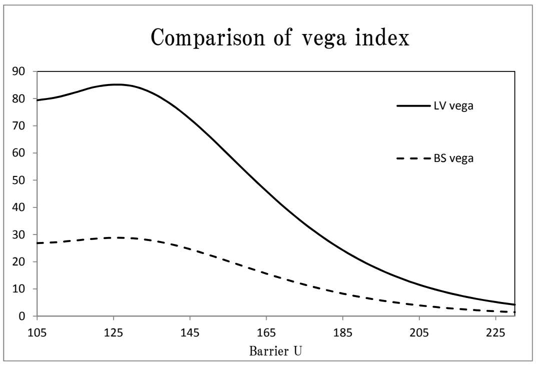

Figure 3.4: Vega index in two different models.

In Figure 3.1, we plot the values of each sensitivity in LV vega against the barrier U. We can observe that in the range of $U \leq 130$, extrema sensitivities are smaller than terminal sensitivities, while the results are inverted if $U > 130$. The existence of this critical barrier (say U^*) is important, since in the case $U \leq U^*$ we are required to pay more attention to LV vega caused by the terminal feature than the maximum feature, and the opposite occurs in the case $U > U^*$. We will give a result on the existence of U^* later (Theorem 6).

In Figure 3.2, we plot LV vega against the maturity T. The value of barrier is fixed with $U = 130$. We observe that for large T, extrema sensitivities are small. The mathematical reasoning is that, for this option, extrema sensitivity becomes small as the probability of $\{\omega \in \Omega : \max_{0\leq t\leq T} S_t < U < \max_{0\leq t\leq T} S_t^\varepsilon\}$ (S_t^ε denotes the solution to (3.1) with its perturbation parameter ε) becomes small. Note that this probability is small for large T. From this result we can conclude that extrema sensitivity is less important than terminal sensitivity and drift sensitivity for the option with long maturity. Thus, for barrier options with long maturity, we may ignore LV vega due to the extrema of the underlying when constructing a hedging strategy by using LV vega.

Next, let us observe the standardized LV vega which is defined by LV vega divided by the maturity T. We set $U = 130$. In Figure 3.3, we plot the values of standardized LV vega against the maturity in order to understand LV vega per unit of time. We observe that the growth of each sensitivity is sharp for small T and almost linear for large T. This numerical result shows that we must be more careful about LV vega of the options with short maturity than that of the options with long maturity. Moreover, from Figure 3.3, we can observe that, for small T, the behavior of extrema sensitivity is the sharpest one of three sensitivities.

Finally, we observe the values of vega index obtained in two different models. We compare LV vega with BS vega which is defined by (3.12). The value of maturity is fixed with $T = 1$. In Figure 3.4, we plot the values of LV vega and BS vega against the barrier U. The difference between LV vega and BS vega in Figure 3.4 (and Table 3.1) represents the importance of the selection of pricing models from the point of view of vega index, since the prices of options are the same in these two different models, as we have mentioned in Remark 8. This figure implies that, as far as the vega index is concerned, the property of one-dimensional model dealt in this chapter is far away from the Black-Scholes model.

Now we give a theorem that guarantees the existence of U^*.

Theorem 6. *Assume* **(H1)-(H4)** *and* $\sigma(z) = \tilde{\sigma} z, f(y,z) = I_{(U,\infty)}(y)(z-K)_+, x < K < U$. *Let* $E(U,K)$ *and* $T(U,K)$ *be defined in (3.14). Then for any* $K > 0$ *there exists* $U^*(> K)$ *such that* $E(U^*,K) = T(U^*,K)$.

Proof. By **(H1)-(H4)**, we have

$$\begin{aligned} E(U,K) &\geq \frac{\sigma_0}{\tilde{\sigma}} E^P\left[\delta_U(\max_{0\leq t\leq T} S_t)(S_T - K)_+ \max_{0\leq t\leq T} S_t \log\left(\frac{\max\limits_{0\leq t\leq T} S_t}{x}\right)\right], \\ T(U,K) &\leq \frac{K_1}{\tilde{\sigma}} E^P\left[I_{(U,\infty)}(\max_{0\leq t\leq T} S_t) I_{(K,\infty)}(S_T) S_T \left(\log S_T + \log x + \frac{1}{S_T} + \frac{1}{x}\right)\right]. \end{aligned}$$

It is easy to obtain

$$\lim_{U\to\infty} \frac{T(U,K)}{E(U,K)} = 0,$$

by using the joint density function

$$P\left(\max_{0\le t\le T}(-\frac{\tilde{\sigma}}{2}t+W_t)\in dy, -\frac{\tilde{\sigma}}{2}T+W_T\in dz\right)=\frac{2(2y-z)}{T\sqrt{2\pi T}}e^{-\frac{\tilde{\sigma}}{2}z-\frac{\tilde{\sigma}^2}{8}T-\frac{1}{2T}(2y-z)^2}dydz, y\ge z$$

which is obtained from Formula 1.1.8 of [3]. □

3.4 Kernel method

In this section, we assume that the payoff function f is of the form $f(y,z)=I_{(U,\infty)}(y)I_{(K,\infty)}(z)$, where $x<K<U$. This option is called a binary barrier option with strike K and barrier U. In this case, each component of (3.2) involves delta functionals.

The goal of this section is to introduce a method to compute (3.2) for this option by using the so-called kernel method. We shall provide the results under the simplest assumption for the coefficients of (3.1). Still, we may extend the results obtained in this section to more general cases with some additional assumptions which are remarked as Remark 11.

We consider the method to compute three components (extrema, terminal and drift sensitivity) in (3.2), altogether. In the case of an up-in call option, we have seen the existence of a barrier value at which the importance of extrema and terminal sensitivity are reversed in Figure 3.1 of Subsection 3.3.3. This interesting property still holds under the assumptions in this section, when we focus on the behavior of each individual component of (3.2). However, we omit giving the figures on it, since the goal of this section is to compute the components of (3.2) all at once. In the case that we consider a separate kernel method for each component in (3.2), the results in this section are not applicable.

In this section we assume $b(z)=0, \sigma(z)=\tilde{\sigma}_1 z$ and $\hat{\sigma}(z)=\tilde{\sigma}_2 z$, where $\tilde{\sigma}_1,\tilde{\sigma}_2$ are positive constants, moreover, we define $M_T^\varepsilon := \max_{0\le t\le T} S_t^\varepsilon$ for simplicity of notations.

Under these assumptions, (3.2) is represented as

$$\begin{aligned}
\left.\frac{\partial \Pi^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} &= \frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}E^P\left[\delta_U(M_T)I_{(K,\infty)}(S_T)M_T\log\left(\frac{M_T}{x}\right)\right] \\
&\quad +\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}E^P\left[\delta_K(S_T)I_{(U,\infty)}(M_T)S_T\log\left(\frac{S_T}{x}\right)\right] \\
&\quad -\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{2}E^P\left[\delta_U(M_T)I_{(K,\infty)}(S_T)M_T\cdot\eta+\delta_K(S_T)I_{(U,\infty)}(M_T)S_T\cdot T\right] \\
&= E(U,K)+T(U,K)+D(U,K)=: V(U,K). \qquad (3.15)
\end{aligned}$$

Remark 9. *Under our assumptions in this section, we can compute* $\Pi^\varepsilon = E^P[I_{(U,\infty)}(M_T^\varepsilon)I_{(K,\infty)}(S_T^\varepsilon)]$ *explicitly, thus, we obtain the explicit formula for* $\left.\frac{\partial\Pi^\varepsilon}{\partial\varepsilon}\right|_{\varepsilon=0}$. *Nevertheless, in general, neither* Π^ε *nor* $\left.\frac{\partial\Pi^\varepsilon}{\partial\varepsilon}\right|_{\varepsilon=0}$ *has the explicit formula.*

Now let us consider a problem to compute (3.15) by using Monte Carlo simulation. Equation (3.15) involves the delta functionals, therefore, we will define estimators for $E(U,K), T(U,K)$ and $D(U,K)$ by approximating the delta functionals using so-called kernel functions.

Let us give the definition of a kernel function.

Definition 2. *(Kernel function) We call a symmetric function* $\mathcal{K}:\mathbb{R}\to\mathbb{R}_+$ *a kernel function if* $\mathcal{K}$ *satisfies the following three properties,*

$$\int_{\mathbb{R}}\mathcal{K}(y)dy=1,\ \int_{\mathbb{R}}y\mathcal{K}(y)dy=0,\ \int_{\mathbb{R}}y^2\mathcal{K}(y)dy<\infty. \qquad (3.16)$$

Now let us define the estimators for $E(U,K), T(U,K)$ and $D(U,K)$ by using kernel functions. Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be two kernel functions and $h_1>0, h_2>0$. We define

$$\begin{aligned}
\hat{E}(U,K) &:= \frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_1}\mathcal{K}_1\left(\frac{U-M_T^{(i)}}{h_1}\right)I_{(K,\infty)}(S_T^{(i)})M_T^{(i)}\log\left(\frac{M_T^{(i)}}{x}\right),\\
\hat{T}(U,K) &:= \frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_2}\mathcal{K}_2\left(\frac{K-S_T^{(i)}}{h_2}\right)I_{(U,\infty)}(M_T^{(i)})S_T^{(i)}\log\left(\frac{S_T^{(i)}}{x}\right),\\
\hat{D}(U,K) &:= -\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{2}\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{h_1}\mathcal{K}_1\left(\frac{U-M_T^{(i)}}{h_1}\right)I_{(K,\infty)}(S_T^{(i)})M_T^{(i)}\eta^{(i)}\right.\\
&\qquad \left.+\frac{1}{h_2}\mathcal{K}_2\left(\frac{K-S_T^{(i)}}{h_2}\right)I_{(U,\infty)}(M_T^{(i)})S_T^{(i)}T\right]\\
\hat{V}(U,K) &:= \hat{E}(U,K)+\hat{T}(U,K)+\hat{D}(U,K),
\end{aligned}\tag{3.17}$$

where $\{M_T^{(i)}\}_{1\le i\le N}$, $\{S_T^{(i)}\}_{1\le i\le N}$ and $\{\eta^{(i)}\}_{1\le i\le N}$ are sequences of independent random variables defined on $(\Omega,\mathcal{F},P)$ with the same laws as M_T, S_T and η, respectively. We note that (h_1,h_2) is called the bandwidth parameter and (h_1,h_2) depends on N, although we omit the dependence on N for simplicity of notation.

For $\mathcal{K}_i$ and $h_i, i=1,2$, we assume the following conditions.

(K1) There exist $c_1>0$ and $c_2>3$ such that for each $i=1,2$, $\mathcal{K}_i(y)\le c_1|y|^{-c_2}$ holds for all $y\in\mathbb{R}$.

(K2) For each $i=1,2$, $\mathcal{K}_i$ is bounded and $\int_{\mathbb{R}}\mathcal{K}_i(y)|y|^3dy<\infty$ holds.

(K3) For each $i=1,2$, $h_i\to 0$ and $Nh_i\to+\infty$, as N tends to infinity.

(K4) There exist $c_3>0$ and $c_4>0$ such that $c_3\le\frac{h_1}{h_2}\le c_4$ holds for all $N\in\mathbb{N}$.

Note that (**K2**) implies $\int_{\mathbb{R}}|\mathcal{K}_i(y)|^2dy<\infty, i=1,2$.

One can show that $\hat{E}(U,K),\hat{T}(U,K)$ and $\hat{D}(U,K)$ converge in probability to $E(U,K),T(U,K)$ and $D(U,K)$ respectively as $N\to\infty$. Thus $\hat{V}(U,K)$ converges to $V(U,K)$ in probability, as N tends to infinity. We can compute $\hat{E}(U,K),\hat{T}(U,K)$ and $\hat{D}(U,K)$ with Monte Carlo simulation, since they do not involve the delta functionals any more. To compute $\hat{E}(U,K),\hat{T}(U,K)$ and $\hat{D}(U,K)$, we are required to choose (h_1,h_2) appropriately for fixed N. For this, we define the following so-called mean squared error.

Definition 3. *(Mean squared error) We define the mean squared error of the estimator $\hat{V}(U,K)$ by*

$$MSE(U,K):=E^P[|\hat{V}(U,K)-V(U,K)|^2].\tag{3.18}$$

The following lemma decomposes $MSE(U,K)$ into the bias part and the variance part of $\hat{V}(U,K)$.

Lemma 12. *We have the following decomposition for $MSE(U,K)$,*

$$MSE(U,K)=E^P[\hat{V}(U,K)-V(U,K)]^2+Var[\hat{V}(U,K)].\tag{3.19}$$

Proof. The result follows from the definition of $\hat{V}(U,K)$ easily. □

Now our aim is to find (h_1, h_2) so that $MSE(U,K)$ is as small as possible. Unfortunately, it seems difficult to express $MSE(U,K)$ by (h_1, h_2) explicitly. Therefore, we define an asymptotic mean squared error of the estimator $\hat{V}(U,K)$.

Definition 4. *(Asymptotic mean squared error) We define an asymptotic mean squared error of the estimator* $\hat{V}(U,K)$ *by*

$$\begin{aligned}
AMSE(U,K) \quad := \quad & \left(\left[\frac{1}{2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}(\psi_1 p_M)''(U) - \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{4}(\phi_1 p_M)''(U) \right] \mu_2(\mathcal{K}_1)h_1^2 \right. \\
& \left. + \left[\frac{1}{2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}(\psi_2 p_S)''(K) - \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{4}(\phi_2 p_S)''(K) \right] \mu_2(\mathcal{K}_2)h_2^2 \right)^2 \\
& + \left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2 \Psi_1(U) - \tilde{\sigma}_2^2\Psi_2(U) + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4}\Psi_3(U) \right] p_M(U)R_2(\mathcal{K}_1)\frac{1}{Nh_1} \\
& + \left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2 \Phi_1(K) - \tilde{\sigma}_2^2\Phi_2(K) + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4}\Phi_3(K) \right] p_S(K)R_2(\mathcal{K}_2)\frac{1}{Nh_2},
\end{aligned} \tag{3.20}$$

where $R(\mathcal{K}_i) := \int_{\mathbb{R}} \mathcal{K}_i(y)^2 dy$, $\mu_2(\mathcal{K}_i) := \int_{\mathbb{R}} y^2\mathcal{K}_i(y)dy, i = 1,2$ *and* $p_M(y), p_S(z)$ *denote the probability density functions of* M_T *and* S_T*, respectively, and*

$$\left.\begin{aligned}
\psi_1(y) &:= y\log\left(\tfrac{y}{x}\right) E^P[I_{(K,\infty)}(S_T)|M_T = y], \\
\psi_2(y) &:= y\log\left(\tfrac{y}{x}\right) E^P[I_{(U,\infty)}(M_T)|S_T = y], \\
\phi_1(y) &:= yE^P[I_{(K,\infty)}(S_T)\eta|M_T = y], \\
\phi_2(y) &:= yTE^P[I_{(U,\infty)}(M_T)|S_T = y], \\
\Psi_1(y) &:= y^2\left(\log\left(\tfrac{y}{x}\right)\right)^2 E^P[I_{(K,\infty)}(S_T)|M_T = y], \\
\Psi_2(y) &:= y^2\log\left(\tfrac{y}{x}\right) E^P[I_{(K,\infty)}(S_T)\eta|M_T = y], \\
\Psi_3(y) &:= y^2E^P[I_{(K,\infty)}(S_T)\eta^2|M_T = y], \\
\Phi_1(y) &:= y^2\left(\log\left(\tfrac{y}{x}\right)\right)^2 E^P[I_{(U,\infty)}(M_T)|S_T = y], \\
\Phi_2(y) &:= y^2\log\left(\tfrac{y}{x}\right) TE^P[I_{(U,\infty)}(M_T)|S_T = y], \\
\Phi_3(y) &:= y^2T^2E^P[I_{(U,\infty)}(M_T)|S_T = y].
\end{aligned}\right\} \tag{3.21}$$

The following lemma provides the justification to use $AMSE(U,K)$. The proof can be found in Appendix D.

Lemma 13. *Assume* **(K1)-(K4)**. *Let* $MSE(U,K)$ *and* $AMSE(U,K)$ *be defined by (3.18) and (3.20), respectively. Then, we have*

$$AMSE(U,K) = MSE(U,K) + o(h_1^4 + h_2^4 + \frac{1}{Nh_1} + \frac{1}{Nh_2}).$$

The main result in this section is the following theorem on the optimal selection of (h_1, h_2).

Theorem 7. *Let $AMSE(U,K)$ be defined by (3.20). We define*

$$\left.\begin{aligned} a_1 &:= \left[\frac{1}{2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}(\psi_1 p_M)''(U) - \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{4}(\phi_1 p_M)''(U)\right]\mu_2(\mathcal{K}_1) \\ a_2 &:= \left[\frac{1}{2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}(\psi_2 p_S)''(K) - \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{4}(\phi_2 p_S)''(K)\right]\mu_2(\mathcal{K}_2) \\ b_1 &:= \left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2\Psi_1(U) - \tilde{\sigma}_2^2\Psi_2(U) + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4}\Psi_3(U)\right]p_M(U)R_2(\mathcal{K}_1)\frac{1}{N} \\ b_2 &:= \left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2\Phi_1(K) - \tilde{\sigma}_2^2\Phi_2(K) + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4}\Phi_3(K)\right]p_S(K)R_2(\mathcal{K}_2)\frac{1}{N} \end{aligned}\right\} \tag{3.22}$$

and assume that $a_1, a_2, b_1, b_2 > 0$. Then, (h_1^*, h_2^*) given by

$$(h_1^*, h_2^*) = \left(\left[\frac{b_1}{4a_1^2 + 4\left(\frac{a_1^5 a_2 b_2^2}{b_1^2}\right)^{\frac{1}{3}}}\right]^{\frac{1}{5}}, \left[\frac{b_2}{4a_2^2 + 4\left(\frac{a_1 a_2^5 b_1^2}{b_2^2}\right)^{\frac{1}{3}}}\right]^{\frac{1}{5}}\right) \tag{3.23}$$

minimizes $AMSE(U,K)$.

Proof. Under the condition $a_1, a_2, b_1, b_2 > 0$, $AMSE(U,K)$ is a strictly convex function with respect to (h_1, h_2), then the result follows easily. □

Remark 10. *p_S, p_M and many of (3.21) can be computed explicitly under assumptions in this section. Indeed, the conditional expectations in (3.21) which do not involve η can be calculated explicitly from the joint density function for (M_T, S_T). Moreover, the conditional expectations in (3.21) involving η can be expressed by using the joint density function for (M_T, S_T, η), therefore, we can compute them with numerical integrations. The joint density function for (M_T, S_T, η) can be obtained by using Formula 1.13.8 of [3].*

Remark 11. *Lemma 13 can be extended to more general cases by assuming the smoothness and bounded derivatives of some density functions (the smoothness and bounded derivatives of the density function of M_T and S_T) and conditional expectations (e.g. the smoothness and bounded derivatives of $y \mapsto E^P[I_{(K,\infty)}\sigma(S_T)\int_x^{M_T}\frac{\hat{\sigma}}{(\sigma)^2}(z)dz|M_T = y]$).*

Moreover, if we obtain the values of these derivatives, then we can compute the optimal bandwidth for the general cases by the same method as Theorem 7.

Remark 12. *In general, when we use the kernel method to estimate a d-dimensional probability density function, to find the optimal bandwidth is a quite difficult problem, since in this case the bandwidth is a $d \times d$-symmetric matrix. Meanwhile, in our case, two delta functionals appear independently in (3.15), thus the problem is much easier and we can successfully find out the optimal bandwidth.*

3.5 Numerical experiment 2: Estimation of the vega index using the kernel method

In this section, we obtain some numerical results by using (3.17), (3.22) and (3.23). We use the optimal bandwidth given by (3.23). We use parameters $x = 80$, $\tilde{\sigma}_1 = 0.3$, $\tilde{\sigma}_2 = 1$ for the model and $K = 85$, $U =$

90, $T = 1$ for the option. Under this setting, a_1, a_2, b_1 and b_2 in (3.22) are calculated as follows,

a_1	a_2	b_1	b_2
0.003110	0.000428	1.89×10^{-7}	3.95×10^{-8}

Table 3.2: Values of (3.22).

Therefore, we can apply Theorem 7 to obtain the optimal bandwidth and it is computed as

h_1^*	h_2^*
0.3337	0.3834

Table 3.3: Optimal bandwidth.

By using the above optimal bandwidth, we have the following numerical result for the vega index,

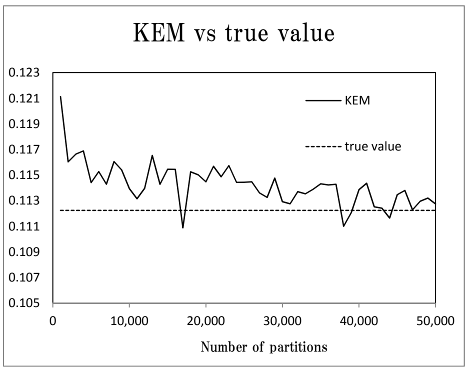

Figure 3.5: KEM vs True value.

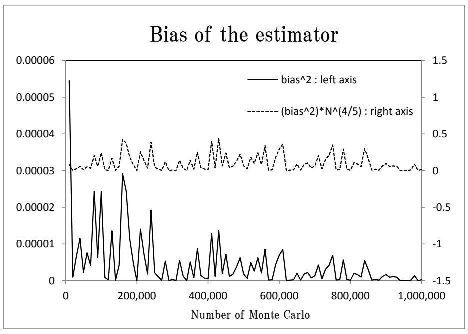

Figure 3.6: Bias of $\hat{V}(K,U)$.

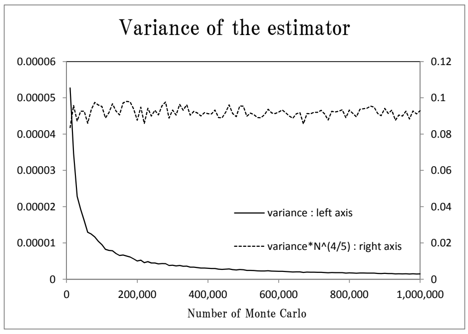

Figure 3.7: Variance of $\hat{V}(K,U)$.

In Figure 3.5, we plot the vega index obtained by using (3.17) (denoted by KEM in the figure) and the explicit formula (denoted by true value in the figure) against the number of partitions of time to maturity. The number of Monte Carlo simulations is fixed with $N = 10^6$. We have used the density function of the normal distribution with the mean 0 and variance 1 for the kernel functions (i.e. $\mathcal{K}_i(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}, i = 1, 2$). Figure 3.5 shows that (3.17) converges to the true value when we increase the number of partitions.

Next, let us focus on the bias and variance of the estimator $\hat{V}(K,U)$. We fix the number of partitions of the interval $[0,T]$ with $n = 5 \times 10^4$. In Figure 3.6 and 3.7, we plot the squared bias and variance of the estimator $\hat{V}(K,U)$ against the number of simulations N, respectively. The squared bias of $\hat{V}(K,U)$ is defined by $E^P[\hat{V}(U,K) - V(U,K)]^2$ and the variance is defined by $E^P[\hat{V}(U,K)^2] - E^P[\hat{V}(U,K)]^2$. In Figure 3.6, the solid line (with left axis) represents the values of squared bias and the broken line (with right axis) expresses the values of "squared bias$\times N^{\frac{4}{5}}$". In Figure 3.7, the values of variance and "variance$\times N^{\frac{4}{5}}$" are represented by the solid line (with left axis) and the broken line (with right axis), respectively.

From the above figures, we can observe that the variance decreases with the order $N^{-\frac{4}{5}}$ as the number of simulations increases. Moreover, we may see that, for large N, the squared bias converges with the order $N^{-\frac{4}{5}}$. Therefore, Figure 3.6 and 3.7 indicate that the order of $MSE(U,K)$ defined by (3.18) with the optimal bandwidth (3.23) is $N^{-\frac{4}{5}}$, numerically.

3.6 Conclusion and final remarks

In this chapter, we have obtained a formula to calculate the vega index for options whose payoff functions may depend on the maximum or minimum of a one-dimensional SDE. The key technique is the Lamperti transform which enables us to calculate the directional derivatives with respect to the diffusion coefficient.

This formula gives a decomposition of the vega index into three sensitivities: extrema sensitivity, terminal sensitivity and drift sensitivity. Numerical tests illustrate that there are some important relationships between extrema sensitivity and terminal sensitivity in realistic options.

In [10], formulas of this type in the multi-dimensional case under some commutativity conditions on the diffusion coefficient are used to prove the smoothness of the density function concerning the supremum of a multi-dimensional diffusion. The authors obtain the formulas by means of Garsia-Rodemich-Rumsey's lemma.

The numerical result on the comparison of the vega index in two different models tells us that the traditional Black-Scholes model is far away from the one-dimensional model dealt in this chapter as far as the vega index is concerned. Today, some practitioners are using so-called stochastic volatility models which deal with stochastic diffusion coefficients to express the dynamics of economy (see [8], for example). There are many difficulties to deal with stochastic volatility models, however to compute the vega index for exotic options in stochastic volatility models is a challenging problem for the future. According to [4], there is a relationship between a one-dimensional model and a stochastic volatility model, when we consider some exotic options. Thus, the results obtained in [4] may be applied to this problem.

The kernel method has been used to compute the vega index for a specific option in this chapter. The bandwidth selection problem in this chapter can be successfully solved, since it is much simpler than general bandwidth selection problems that appear in the kernel density estimation problems for multi-dimensional density functions. Our numerical result shows the optimal bandwidth works well for the problem considered in this chapter.

Appendix A

In Appendix A, we shall give the proof of Lemma 8, 7, 9 and 10 which we have used in Section 3.2. We use some preparatory lemmas. We denote by K_i, $i = 1, 2$ the appropriate Lipschitz constants that appear in the hypothesis **(H1)**.

Appendix A.1

In this subsection, we shall prove some properties which are used in order to prove Lemma 8, 7, 9 and 10.

Lemma A1. *Assume* **(H1)** *and* **(H2)***. Then, we have the following inequalities,*

$$\begin{aligned} \frac{1}{2K_1}\left|\log\left(\frac{1+z}{2}\right)\right| \leq |F_\varepsilon(z)| &\leq \frac{1}{\sigma_0}|\log(z)|, \forall z > 0 \\ \left|\int_x^z \frac{\hat{\sigma}}{\sigma^{\varepsilon_1}\sigma^{\varepsilon_2}}(y)dy\right| &\leq \frac{K_1}{\sigma_0^2}\left(\left|\frac{1}{z} - \frac{1}{x}\right| + \left|\log\left(\frac{z}{x}\right)\right|\right), \forall x, z > 0. \end{aligned}$$

Proof. In fact, the proof follows by using the lower bound for σ^ε in the hypothesis **(H2)** and the upper bounds for σ^ε and $\hat{\sigma}$ in the hypothesis **(H1)**. □

Lemma A2. *Assume* **(H1)***,* **(H2)** *and* **(H3)***. Then, F_ε^{-1} can be evaluated as follows,*

$$\begin{aligned} &e^{\sigma_0 z} \leq F_\varepsilon^{-1}(z) \leq 2e^{2K_1 z}, \quad \textit{if } z \geq 0 \\ &r_0 e^{\sigma_1 z} \leq F_\varepsilon^{-1}(z) \leq e^{\sigma_0 z}, \quad \textit{if } z < 0. \end{aligned}$$

Proof. To prove this lemma, we consider two cases according to the sign of z. In the case that z is negative, we use hypothesis **(H3)** in addition to **(H1)** and **(H2)**. In both cases, we use the monotonicity of F_ε and F_ε^{-1}. We leave the details to the reader. □

Let us prove a lemma on the regularity of $F_\varepsilon^{-1}(z)$.

Lemma A3. *Assume* **(H1)** *and* **(H2)**. *Then,* $F_\varepsilon^{-1}(z)$ *is differentiable with respect to* ε *and*

$$\frac{\partial F_\varepsilon^{-1}}{\partial \varepsilon}(z) = \sigma^\varepsilon(F_\varepsilon^{-1}(z)) \int_1^{F_\varepsilon^{-1}(z)} \frac{\hat{\sigma}}{(\sigma^\varepsilon)^2}(y)dy.$$

Furthermore $\frac{\partial}{\partial \varepsilon}\left(F_\varepsilon^{-1}(F_\varepsilon(x)+z)\right)$ *exists and*

$$\begin{aligned}\frac{\partial}{\partial \varepsilon}\left(F_\varepsilon^{-1}(F_\varepsilon(x)+z)\right) &= \frac{\partial F_\varepsilon^{-1}}{\partial z}(F_\varepsilon(x)+z)\frac{\partial F_\varepsilon}{\partial \varepsilon}(x) + \frac{\partial F_\varepsilon^{-1}}{\partial \varepsilon}(F_\varepsilon(x)+z) \\ &= \sigma^\varepsilon(F_\varepsilon^{-1}(F_\varepsilon(x)+z)) \int_x^{F_\varepsilon^{-1}(F_\varepsilon(x)+z)} \frac{\hat{\sigma}}{(\sigma^\varepsilon)^2}(y)dy. \qquad \text{(A.1)}\end{aligned}$$

Proof. The proof follows from the implicit function theorem applied to $z = F_\varepsilon(F_\varepsilon^{-1}(z))$ and the chain rule for partial differentiation. □

Now we shall prove Lemma 7, 8 and 9. For convenience, let us prove Lemma 8 first.

Appendix A.2

In this subsection, we prove Lemma 8.

The result in Theorem 5 implies that if σ and $\hat{\sigma}$ satisfy **(H1)** and **(H2)**, and $\varepsilon \in [0,1]$ and $x > 0$ are fixed, then there exists a stochastic process S^ε such that it satisfies

$$S_t^\varepsilon = x + \int_0^t \sigma^\varepsilon(S_u^\varepsilon) \circ d\tilde{W}_u, \forall t \in [0,T], a.s. \qquad \text{(A.2)}$$

However, the exceptional set where the above equality is not satisfied may depend on ε. In order to consider the differentiation of the solution to (3.3), we modify the solution so that the exceptional set does not depend on ε. Let us prove that the solution to (3.3) is continuous with respect to ε so that we can discuss its differentiability.

We refer [12] to obtain the continuity in ε.

Lemma A4. *Assume* **(H1)-(H2)**. *Let* S_t^i *be the solution to*

$$\begin{cases} dS_t^i = \sigma^{\varepsilon_i}(S_t^i) \circ d\tilde{W}_t \left(= \sigma^{\varepsilon_i}(S_t^i)d\tilde{W}_t + \frac{1}{2}\sigma^{\varepsilon_i}(\sigma^{\varepsilon_i})'(S_t^i)dt\right) \\ S_0^i = x, \end{cases} \qquad \text{(A.3)}$$

$\varepsilon_i \in [0,1], i = 1,2$. *Then, for* $p > 2$ *there exists a constant* $C > 0$ *such that*

$$E^{\tilde{Q}}[|S_{t_2}^2 - S_{t_1}^1|^p] \le C\left\{|\varepsilon_2 - \varepsilon_1|^p + |t_2 - t_1|^{\frac{p}{2}}\right\}.$$

Proof. We have

$$\begin{aligned} E^{\tilde{Q}}[|S^2_{t_2} - S^1_{t_1}|^p] &\leq 2^{p-1}\left\{E^{\tilde{Q}}\left[|F^{-1}_{\varepsilon_2}(F_{\varepsilon_2}(x) + \tilde{W}_{t_2}) - F^{-1}_{\varepsilon_2}(F_{\varepsilon_1}(x) + \tilde{W}_{t_1})|^p\right]\right. \\ &\quad \left. + E^{\tilde{Q}}\left[|F^{-1}_{\varepsilon_2}(F_{\varepsilon_1}(x) + \tilde{W}_{t_1}) - F^{-1}_{\varepsilon_1}(F_{\varepsilon_1}(x) + \tilde{W}_{t_1})|^p\right]\right\} \\ &=: 2^{p-1}\left\{E^{\tilde{Q}}[|I_1|^p] + E^{\tilde{Q}}[|I_2|^p]\right\}. \end{aligned}$$

The arguments to deal with the above expectations are standard and similar to the proof of continuity of flows associated with the solution of stochastic differential equations. We do not give all the arguments here and only stress that to obtain the result one uses the lemmas in Appendix Appendix A.1 together with hypotheses **(H1)-(H2)** and the fact that the Wiener process has finite exponential moments. □

Lemma A4 implies that there exists a continuous modification of the solution S^ε with respect to two variables (t, ε) so that (3.3) is satisfied for all (t, ε), a.s. Then we can consider the differentiation of the solution to (3.3).

Proof of Lemma 8. Combining Lemma A3 and Lemma A4, the proof is straightforward. □

Appendix A.3

This subsection is devoted to the proof of Lemma 7.

Using the fact that $F_\varepsilon(z)$ and $F^{-1}_\varepsilon(z)$ are continuous monotone increasing functions, the term $\max_{0\leq t\leq T} S^\varepsilon_t$ can be expressed as $\max_{0\leq t\leq T} S^\varepsilon_t = F^{-1}_\varepsilon(F_\varepsilon(x) + \max_{0\leq t\leq T}\tilde{W}_t)$ under $\tilde{Q}$. We have the following continuity lemma.

Lemma A5. *Assume* **(H1)** *and* **(H2)**. *For $i = 1, 2$, let S^i_t be the solutions to (A.3). Then, for any $p > 2$, there exists a constant $C > 0$ such that*

$$E^{\tilde{Q}}\left[\left|\max_{0\leq t\leq T} S^2_t - \max_{0\leq t\leq T} S^1_t\right|^p\right] \leq C|\varepsilon_2 - \varepsilon_1|^p.$$

Proof. The proof is similar to the proof of Lemma A4. We obtain the result by replacing exponential moments of $\tilde{W}_t$ by the respective ones of $\max_{0\leq t\leq T}\tilde{W}_t$. □

Proof of Lemma 7. Using the above lemma, the explicit expression for $\max_{0\leq t\leq T} S^\varepsilon_t$ in Theorem 5 and (A.1), we obtain the result. □

Appendix A.4

The goal of this subsection is to prove Lemma 9.

First, let us prove the following lemma about L^p-boundedness of S^ε_t, $\log S^\varepsilon_t$, $\frac{\partial S^\varepsilon_t}{\partial\varepsilon}$ and $\frac{\partial^2 S^\varepsilon_t}{\partial\varepsilon^2}$ under the hypotheses **(H1)-(H3)**.

Lemma A6. *Assume* **(H1)-(H3)**. *Let S^ε be the solution to (A.2). Then, for any $p > 2$ there exists $C_1 > 0$ such that*

$$E^{\tilde{Q}}[|S^\varepsilon_t|^p] + E^{\tilde{Q}}[|\log S^\varepsilon_t|^p] \leq C_1.$$

Moreover, for any real number p there exists $C_2 > 0$ such that

$$E^{\tilde{Q}}[|S^\varepsilon_t|^p] \leq C_2.$$

Proof. The estimate for $E^{\tilde{Q}}[|S_t^\varepsilon|^p]$ follows from Lemma A1 and Lemma A2. Let us consider the finiteness of $E^{\tilde{Q}}[|\log S_t^\varepsilon|^p]$. Using the fact that $S_t^\varepsilon \geq 1 \Leftrightarrow F_\varepsilon(x) + \tilde{W}_t \geq 0$ and $S_t^\varepsilon < 1 \Leftrightarrow F_\varepsilon(x) + \tilde{W}_t < 0$ together with Lemma A2, we have

$$\begin{aligned}
&E^{\tilde{Q}}\left[|\log S_t^\varepsilon|^p\right] \\
&= E^{\tilde{Q}}\left[(\log S_t^\varepsilon)^p : S_t^\varepsilon \geq 1\right] + E^{\tilde{Q}}\left[\left(\log \frac{1}{S_t^\varepsilon}\right)^p : S_t^\varepsilon < 1\right] \\
&\leq E^{\tilde{Q}}\left[\left(\log\left(2e^{2K_1(F_\varepsilon(x)+\tilde{W}_t)}\right)\right)^p : F_\varepsilon(x) + \tilde{W}_t \geq 0\right] \\
&\quad + E^{\tilde{Q}}\left[\left(\log\left(\frac{1}{r_0}e^{-\sigma_1(F_\varepsilon(x)+\tilde{W}_t)}\right)\right)^p : F_\varepsilon(x) + \tilde{W}_t < 0\right] \\
&\leq 2^{p-1}\left\{|\log 2|^p + (2K_1)^p E^{\tilde{Q}}\left[\left|F_\varepsilon(x) + \tilde{W}_t\right|^p\right]\right\} + 2^{p-1}\left\{\left(\log\left(\frac{1}{r_0}\right)\right)^p + \sigma_1^p E^{\tilde{Q}}\left[\left|F_\varepsilon(x) + \tilde{W}_t\right|^p\right]\right\}.
\end{aligned}$$

Then by Lemma A1, we have

$$E^{\tilde{Q}}\left[\left|F_\varepsilon(x) + \tilde{W}_t\right|^p\right] \leq 2^{p-1}\left(\frac{1}{\sigma_0^p}|\log(x)|^p + E^{\tilde{Q}}\left[\left|\max_{0\leq t\leq T}\tilde{W}_t\right|^p\right]\right),$$

and the result follows. □

Now, we consider the L^p-boundedness of $\frac{\partial S_t^\varepsilon}{\partial \varepsilon}$ and $\frac{\partial^2 S_t^\varepsilon}{\partial \varepsilon^2}$. Note that as $\sigma^\varepsilon(z)$ and $\int_x^z \frac{1}{(\sigma^\varepsilon)^2}(y)dy$ are continuously differentiable with respect to ε and z, $\frac{\partial^2 S_t^\varepsilon}{\partial \varepsilon^2}$ exists for $\varepsilon \in [0,1]$ and $t \in [0,T]$, a.s. by Lemma 8.

Lemma A7. *Assume* **(H1)-(H3)**. *Let S^ε be the solution to (A.2). Then, for any $p > 2$ there exists $C > 0$ such that*

$$E^{\tilde{Q}}\left[\left|\frac{\partial S_t^\varepsilon}{\partial \varepsilon}\right|^p\right] + E^{\tilde{Q}}\left[\left|\frac{\partial^2 S_t^\varepsilon}{\partial \varepsilon^2}\right|^p\right] \leq C.$$

Proof. $E^{\tilde{Q}}\left[|\frac{\partial S_t^\varepsilon}{\partial \varepsilon}|^p\right]$ part follows from the explicit form of $\frac{\partial S_t^\varepsilon}{\partial \varepsilon}$, Lemma A1, A6 and (**H1**). Then, the upper bound for $E^{\tilde{Q}}\left[|\frac{\partial^2 S_t^\varepsilon}{\partial \varepsilon^2}|^p\right]$ is calculated as follows. By the differentiability of $\int_x^z \frac{\hat{\sigma}}{(\sigma^\varepsilon)^2}(y)dy$ with respect to ε and z, we have

$$\frac{\partial^2 S_t^\varepsilon}{\partial \varepsilon^2} = \left(\frac{\partial S_t^\varepsilon}{\partial \varepsilon}\sigma^{\varepsilon\prime}(S_t^\varepsilon) + \hat{\sigma}(S_t^\varepsilon)\right)\int_x^{S_t^\varepsilon}\frac{\hat{\sigma}}{(\sigma^\varepsilon)^2}(y)dy - 2\sigma^\varepsilon(S_t^\varepsilon)\int_x^{S_t^\varepsilon}\frac{\hat{\sigma}^2}{(\sigma^\varepsilon)^3}(y)dy + \frac{\hat{\sigma}}{\sigma^\varepsilon}(S_t^\varepsilon)\frac{\partial S_t^\varepsilon}{\partial \varepsilon}.$$

Thus, the result follows from Lemma A6 and the inequality

$$\left|\int_x^z \frac{\hat{\sigma}^2}{(\sigma^\varepsilon)^3}(y)dy\right| \leq \frac{K_1^2}{\sigma_0^3}\left(\frac{1}{2}\left|\frac{1}{x^2} - \frac{1}{z^2}\right| + 2\left|\frac{1}{x} - \frac{1}{z}\right| + \left|\log\left(\frac{z}{x}\right)\right|\right).$$

□

For the proof of Lemma 9, we use the following theorem which is proved in Theorem 10.6 of [12]. We shall use this theorem without the proof.

Theorem A7. *(Theorem 10.6 of [12]) Let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a filtration σ-field $\{\mathcal{F}_t, t \in [0,T]\}$ satisfying the usual condition. Let $f_t(\varepsilon), (t,\varepsilon) \in [0,T] \times [0,1]^d$ be a measurable random field satisfying the following properties.*

(i) For each ε, $f_t(\varepsilon)$ is predictable.

(ii) For any $p > 2$, there is a positive constant C_1 such that

$$\int_0^T E^P[|f_t(\varepsilon)|^p]dt \leq C_1,$$

for any ε.

(iii) For any $p > 2$, there is a positive constant C_2 such that

$$\int_0^T E^P[|f_t(\varepsilon_1) - f_t(\varepsilon_2)|^p]dt \leq C_2|\varepsilon_1 - \varepsilon_2|^{\alpha p},$$

where $0 < \alpha \leq 1$.

(iv) $f_t(\varepsilon)$ is m-times continuously differentiable in ε for all t, a.s. and derivatives $\frac{\partial^k f_t}{\partial \varepsilon^k}(\varepsilon), |k| \leq m$ satisfy conditions (ii) and (iii).

Let M_t be a continuous local martingale such that $\langle M \rangle_t - \langle M \rangle_s \leq t - s$ for any $t > s$, a.s. Then there is a modification of the integral which is continuous in (t,ε) and m-times continuously differentiable. Furthermore, it holds that

$$\frac{\partial^k}{\partial \varepsilon^k}\left(\int_0^t f_s(\varepsilon)dM_s\right) = \int_0^t \frac{\partial^k f_s}{\partial \varepsilon^k}(\varepsilon)dM_s$$

for any $\frac{\partial^k}{\partial \varepsilon^k}$ such that $|k| \leq m$.

Proof of Lemma 9. Firstly, we prove the stochastic integral term of X_T^ε. For the proof, we use Theorem A7 on the differentiability of the stochastic integral with ε-dependent integrand. For $f_t(\varepsilon) := -(\frac{\sigma^{\varepsilon\prime}}{2} - \frac{b}{\sigma^\varepsilon})(S_t^\varepsilon)$, it is straightforward to check that $f_t(\varepsilon)$ satisfies the sufficient conditions of Theorem A7 by (**H1**) and Lemma A6 and A7.

Secondly, let us consider the Lebesgue integral term of X_T^ε. It suffices to prove that there exists A_t such that

$$\left|\frac{\partial}{\partial \varepsilon}\left(\frac{(\sigma^\varepsilon)'}{2} - \frac{b}{\sigma^\varepsilon}\right)^2 (S_t^\varepsilon)\right| \leq A_t, \forall t \in [0,T],\ a.s.,$$

and $\int_0^T A_t dt < \infty$, *a.s.*, where A_t does not depend on ε. This follows from the fact that

$$\begin{aligned}
&\frac{\partial}{\partial \varepsilon}\left(\frac{(\sigma^\varepsilon)'}{2} - \frac{b}{\sigma^\varepsilon}\right)^2 (S_t^\varepsilon) \\
&= \left(\frac{(\sigma^\varepsilon)'}{2} - \frac{b}{\sigma^\varepsilon}\right)(S_t^\varepsilon)\left\{\sigma''(S_t^\varepsilon)\frac{\partial S_t^\varepsilon}{\partial \varepsilon} + \hat{\sigma}'(S_t^\varepsilon) + \varepsilon\hat{\sigma}''(S_t^\varepsilon)\frac{\partial S_t^\varepsilon}{\partial \varepsilon}\right\} \\
&\quad -2\left(\frac{(\sigma^\varepsilon)'}{2} - \frac{b}{\sigma^\varepsilon}\right)(S_t^\varepsilon)\frac{b'(S_t^\varepsilon)\frac{\partial S_t^\varepsilon}{\partial \varepsilon}\sigma^\varepsilon(S_t^\varepsilon) - b(S_t^\varepsilon)\left\{\sigma'(S_t^\varepsilon)\frac{\partial S_t^\varepsilon}{\partial \varepsilon} + \hat{\sigma}(S_t^\varepsilon) + \varepsilon\hat{\sigma}'(S_t^\varepsilon)\frac{\partial S_t^\varepsilon}{\partial \varepsilon}\right\}}{(\sigma^\varepsilon)^2(S_t^\varepsilon)}.
\end{aligned}$$

Then, S_t^ε, $\frac{1}{S_t^\varepsilon}$ and $\frac{\partial S_t^\varepsilon}{\partial \varepsilon}$ can be evaluated above due to Lemma A1 and A2. □

Appendix A.5

In this subsection, we shall prove Lemma 10.

We have already checked that

$$\begin{aligned} H(\varepsilon,h) &:= \frac{1}{h}\left\{f(\max_{0\le t\le T} S_t^{\varepsilon+h}, S_T^{\varepsilon+h})\exp(X_T^{\varepsilon+h}) - f(\max_{0\le t\le T} S_t^{\varepsilon}, S_T^{\varepsilon})\exp(X_T^{\varepsilon})\right\} \\ &\to \frac{\partial}{\partial\varepsilon}\left(f(\max_{0\le t\le T} S_t^{\varepsilon}, S_T^{\varepsilon})\exp(X_T^{\varepsilon})\right), \varepsilon\in[0,1], a.s. \end{aligned}$$

as h tend to 0. Then it suffices to prove the uniform integrability of $H(\varepsilon,h)$ with respect to h.

To prove the uniform integrability, we need the following lemma.

Lemma A8. *Assume* **(H1)-(H3)**. *Then, for $p>2$, there exists $C>0$ such that*

$$E^{\tilde{Q}}[|X_T^{\varepsilon_2}-X_T^{\varepsilon_1}|^p] \le C|\varepsilon_2-\varepsilon_1|^p.$$

Proof. It suffices to check that there exists $C>0$ such that

$$E^{\tilde{Q}}\left[\left|\frac{\partial X_T^{\varepsilon}}{\partial\varepsilon}\right|^p\right] \le C.$$

By using the proof of Lemma 9, we have the explicit form of $\frac{\partial X_T^{\varepsilon}}{\partial\varepsilon}$. Thus, the result follows by Burkholder-Davis-Gundy's inequality and Lemma A6 and A7. □

Now let us prove the uniform integrability.

Lemma A9. *Assume* **(H1)-(H5)**. *Let $\varepsilon\in[0,1)$ be fixed. Then, for any $p>2$, we have*

$$\sup_{h\in(0,1-\varepsilon)} E^{\tilde{Q}}[|H(\varepsilon,h)|^p] < \infty.$$

Proof. A straighforward calculation yields

$$\begin{aligned} &\left|f(\max_{0\le t\le T} S_t^{\varepsilon+h}, S_T^{\varepsilon+h})\exp(X_T^{\varepsilon+h}) - f(\max_{0\le t\le T} S_t^{\varepsilon}, S_T^{\varepsilon})\exp(X_T^{\varepsilon})\right|^p \\ &\le 3^{p-1}\left\{C_1\big|\max_{0\le t\le T} S_t^{\varepsilon+h} - \max_{0\le t\le T} S_t^{\varepsilon}\big|^p e^{pX_T^{\varepsilon+h}}\right. \\ &\quad \left. + C_2|S_T^{\varepsilon+h}-S_T^{\varepsilon}|^p e^{pX_T^{\varepsilon+h}} + |f(\max_{0\le t\le T} S_t^{\varepsilon}, S_T^{\varepsilon})|^p|X_T^{\varepsilon+h}-X_T^{\varepsilon}|^p\left|\int_0^1 e^{v((X_T^{\varepsilon+h}-X_T^{\varepsilon}))+X_T^{\varepsilon}}dv\right|^p\right\}. \end{aligned}$$

The first and second term of the right-hand side can be evaluated as

$$E^{\tilde{Q}}\left[\big|\max_{0\le t\le T} S_t^{\varepsilon+h} - \max_{0\le t\le T} S_t^{\varepsilon}\big|^p e^{pX_T^{\varepsilon+h}}\right] + E^{\tilde{Q}}[|S_T^{\varepsilon+h}-S_T^{\varepsilon}|^p e^{pX_T^{\varepsilon+h}}] \le C|h|^p.$$

due to the boundedness of $\sigma^{\varepsilon\prime}$, $\frac{b}{\sigma^{\varepsilon}}$ and Lemma A4, A5. For the third term, we have

$$\begin{aligned} &E^{\tilde{Q}}\left[|f(\max_{0\le t\le T} S_t^{\varepsilon}, S_T^{\varepsilon})|^p|X_T^{\varepsilon+h}-X_T^{\varepsilon}|^p\left|\int_0^1 e^{v((X_T^{\varepsilon+h}-X_T^{\varepsilon}))+X_T^{\varepsilon}}dv\right|^p\right] \\ &\le E^{\tilde{Q}}[|X_T^{\varepsilon+h}-X_T^{\varepsilon}|^{2p}]^{\frac{1}{2}} E^{\tilde{Q}}[|f(\max_{0\le t\le T} S_t^{\varepsilon+h}, S_T^{\varepsilon+h})|^{4p}]^{\frac{1}{4}}\left(\int_0^1 E^{\tilde{Q}}\left[e^{4pvX_T^{\varepsilon+h}+4p(1-v)X_T^{\varepsilon}}\right]dv\right)^{\frac{1}{4}}. \end{aligned}$$

Moreover, by the boundedness of $\sigma^{\varepsilon}{}'$, b and $\frac{b}{\sigma^{\varepsilon}}$ we have

$$E^{\tilde{Q}}\left[e^{4pvX_T^{\varepsilon+h}+4p(1-v)X_T^{\varepsilon}}\right] \le C < \infty.$$

Finally, by Lemma A8 we have

$$E^{\tilde{Q}}\left[|f(\max_{0\le t\le T} S_t^{\varepsilon+h}, S_T^{\varepsilon+h})|^p |X_T^{\varepsilon+h} - X_T^{\varepsilon}|^p \left|\int_0^1 e^{v((X_T^{\varepsilon+h}-X_T^{\varepsilon}))+X_T^{\varepsilon}} dv\right|^p\right] \le C|h|^p.$$

The proof is completed. □

Proof of Lemma 10. The result in Lemma 10 follows from Lemma A9. □

Appendix B

In this section, we shall prove equation (3.13).

In order to prove (3.13), we use the mollifier approximation of $f(z) = (z-K)_+$. Define

$$j(z) = \begin{cases} Ce^{\frac{1}{z^2-1}} & (|z|<1) \\ 0 & (|z|\ge 1) \end{cases} \tag{B.1}$$

where C is a constant such that $\int_{-\infty}^{\infty} j(z)dz = 1$. We consider

$$f_n(z) := (j_n * f)(z) := \int_{-\infty}^{\infty} j_n(z-y)f(y)dy, \tag{B.2}$$

where $j_n(z) = nj(nz)$. Then by the Lipschitz continuity of f, we have $f_n(z) \nearrow f(z)$, uniformly as $n \to \infty$. Moreover, f_n' exists and we have

$$f_n'(z) = \int_{-\infty}^{n(z-K)} j(y)dy,$$

therefore, we have that $\lim_{n\to\infty} f_n'(z) \to I_{(K,\infty)}(z)$, for almost every z and f_n' is bounded.

Proof *of (3.13).* For $h \in (0,1]$ we define $\Pi_n^h := E^{\tilde{Q}}[f_n(\max_{0\le t\le T} S_t^h)\exp(X_T^h)]$ and $\phi_n(h)$, $\phi(h)$ by

$$\phi_n(h) := \frac{\Pi_n^h - \Pi_n^0}{h}, \quad \phi(h) := \frac{\Pi^h - \Pi^0}{h}.$$

Then, we have

$$\lim_{n\to\infty} \sup_{0<h\le 1} |\phi_n(h) - \phi(h)| = 0. \tag{B.3}$$

Let us prove (B.3). We define $g_n(z) := f_n(z) - f(z)$, then we have

$$\phi_n(h) - \phi(h) = \frac{1}{h}E^{\tilde{Q}}[(g_n(\max_{0\le t\le T} S_t^h) - g_n(\max_{0\le t\le T} S_t))e^{X_T^h}] - \frac{1}{h}E^{\tilde{Q}}[g_n(\max_{0\le t\le T} S_t)(e^{X_T^h} - e^{X_T})]. \tag{B.4}$$

The second term on the right-hand side of (B.4) can be evaluated by Lemma A8 as

$$\frac{1}{h}E^{\tilde{Q}}[g_n(\max_{0\le t\le T} S_t)(e^{X_T^h} - e^{X_T})] \le CE^{\tilde{Q}}[|g_n(\max_{0\le t\le T} S_t)|^2]^{\frac{1}{2}},$$

and this goes to 0 as $n \to \infty$ by the uniform convergence of f_n.

Now let us evaluate the first term on the right-hand side of (B.4). By the definition of g_n we have

$$g_n(z) - g_n(z') = (z - z')\int_0^1 [f_n'(v(z - z') + z') - I_{(K,\infty)}(v(z - z') + z')]dv. \tag{B.5}$$

Using (B.5) we have

$$\begin{aligned} E^{\tilde{Q}}[|g_n(\max_{0\le t\le T} S_t^h) - g_n(\max_{0\le t\le T} S_t)|^2] &\le E^{\tilde{Q}}[|g_n(\max_{0\le t\le T} S_t^h) - g_n(\max_{0\le t\le T} S_t)|^4]^{\frac{1}{2}} \\ \times\quad E^{\tilde{Q}}\left[\left|\int_0^1 \left\{f_n'(v\max_{0\le t\le T} S_t^h + (1-v)\max_{0\le t\le T} S_t) - I_{(K,\infty)}(v\max_{0\le t\le T} S_t^h + (1-v)\max_{0\le t\le T} S_t)\right\}dv\right|^4\right]^{\frac{1}{2}}. \end{aligned}$$

The second expectation of right-hand side of the above inequality can be written as

$$\begin{aligned} \int_0^1\int_0^\infty \Big| & f_n'(vF_h^{-1}(F_h(x) + z) + (1-v)F^{-1}(F(x) + z)) \\ & -I_{(K,\infty)}(vF_h^{-1}(F_h(x) + z) + (1-v)F^{-1}(F(x) + z))\Big|^4 p_M(z)dzdv, \end{aligned} \tag{B.6}$$

where $p_M(z)$ is the density function of $\max_{0\le t\le T} \tilde{W}_t$. We define $g_{v,h}(z) := vF^{-1}(F_h(x) + z) + (1 - v)F^{-1}(F(x) + z)$ and consider a change of variable $g_{v,h}(z) = u$, then, by Lemma A2, we have $\frac{du}{dz} \ge C$, where C does not depend on v, h and z. Thus, we can show that the term in (B.6) is bounded by

$$\frac{1}{C}\int_0^1\int_x^\infty |f_n'(u) - I_{(K,\infty)}(u)|^4 p_M(g_{v,h}^{-1}(u))dudv,$$

which goes to 0 as $n \to \infty$ by the dominated convergence theorem, therefore, we have (B.3).

The limit in (B.3) asserts that

$$\left.\frac{\partial \Pi^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} = \lim_{n\to\infty} \left.\frac{\partial \Pi_n^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0}.$$

Then (3.13) follows from the dominated convergence theorem, the monotone convergence theorem and the existence of the density of $\max_{0\le t\le T} S_t$. □

Appendix C

This section is devoted to the proof of Lemma 11.

Define

$$\hat{\sigma}_n(z) := \int_{-\infty}^\infty j_n(z - y)\hat{\sigma}(y)dy, \tag{C.1}$$

where $\hat{\sigma}$ is defined by (3.10) and j_n is defined by (B.2). Furthermore, we define

$$\Pi^{\varepsilon,n} := E^P[f(\max_{0\le t\le T} S_t^{\varepsilon,n}, S_T^{\varepsilon,n})],$$

where $S^{\varepsilon,n}$ denotes the solution to (3.1) with $b(z)=0$, $\sigma(z)=\tilde{\sigma}z$ and $\hat{\sigma}$ is defined by (C.1). Then by (3.2), we have

$$\begin{aligned}
\left.\frac{\partial \Pi^{\varepsilon,n}}{\partial \varepsilon}\right|_{\varepsilon=0} &= \frac{1}{\tilde{\sigma}}E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)\max_{0\le t\le T} S_t\int_x^{\max_{0\le t\le T} S_t}\frac{\hat{\sigma}_n(y)}{y^2}dy\right] \\
&+\frac{1}{\tilde{\sigma}}E^P\left[\partial_2 f(\max_{0\le t\le T} S_t, S_T)S_T\int_x^{S_t}\frac{\hat{\sigma}_n(y)}{y^2}dy\right] \\
&-\frac{\tilde{\sigma}}{2}E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)S_\eta\int_0^\eta \hat{\sigma}_n'(S_t)dt+\partial_2 f(\max_{0\le t\le T} S_t, S_T)S_T\int_0^T \hat{\sigma}_n'(S_t)dt\right].
\end{aligned} \tag{C.2}$$

From the fact that $\hat{\sigma}\in C_b^1(\mathbb{R}_+;\mathbb{R}_+)$, we get $\hat{\sigma}_n\to\hat{\sigma}$, uniformly as $n\to\infty$. Thus, one has

$$\begin{aligned}
&\frac{1}{\tilde{\sigma}}E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)\max_{0\le t\le T} S_t\int_x^{\max_{0\le t\le T} S_t}\frac{\hat{\sigma}_n(y)}{y^2}dy\right] \\
&\quad\to\frac{1}{\tilde{\sigma}}E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)\max_{0\le t\le T} S_t\int_x^{\max_{0\le t\le T} S_t}\frac{\hat{\sigma}(y)}{y^2}dy\right], \\
&\frac{1}{\tilde{\sigma}}E^P\left[\partial_2 f(\max_{0\le t\le T} S_t, S_T)S_T\int_x^{S_t}\frac{\hat{\sigma}_n(y)}{y^2}dy\right] \\
&\quad\to\frac{1}{\tilde{\sigma}}E^P\left[\partial_2 f(\max_{0\le t\le T} S_t, S_T)S_T\int_x^{S_t}\frac{\hat{\sigma}(y)}{y^2}dy\right],
\end{aligned}$$

as $n\to\infty$.

By the definition of $\hat{\sigma}_n$, one has $\hat{\sigma}_n(z)\to\hat{\sigma}(z)$ for almost every z. Moreover, we have

$$E^P\left[\partial f(\max_{0\le t\le T} S_t, S_T)S_\eta\int_0^\eta(\hat{\sigma}_n'(S_t)-\hat{\sigma}'(S_t))dt\right]\le C\left(\int_0^T E^P\left[|\hat{\sigma}_n'(S_t)-\hat{\sigma}'(S_t)|^2\right]dt\right)^{\frac{1}{2}}.$$

Due to the boundedness of $\hat{\sigma}_n'$ and $\hat{\sigma}'$, and the existence of the density function of S_t, we get

$$\begin{aligned}
&\frac{\tilde{\sigma}}{2}E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)S_\eta\int_0^\eta \hat{\sigma}_n'(S_t)dt+\partial_2 f(\max_{0\le t\le T} S_t, S_T)S_T\int_0^T \hat{\sigma}_n'(S_t)dt\right] \\
&\quad\to-\frac{\tilde{\sigma}}{2}E^P\left[\partial_1 f(\max_{0\le t\le T} S_t, S_T)S_\eta\int_0^\eta \hat{\sigma}'(S_t)dt+\partial_2 f(\max_{0\le t\le T} S_t, S_T)S_T\int_0^T \hat{\sigma}'(S_t)dt\right],
\end{aligned}$$

as $n\to\infty$.

Finally, let us prove that

$$\lim_{n\to\infty}\lim_{\varepsilon\to 0}\frac{\Pi^{\varepsilon,n}-\Pi^{0,n}}{\varepsilon}=\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\Pi^{\varepsilon,n}-\Pi^{0,n}}{\varepsilon}. \tag{C.3}$$

We define $\varphi_n(\varepsilon):=\frac{\Pi^{\varepsilon,n}-\Pi^{0,n}}{\varepsilon}$ and $\varphi(\varepsilon):=\frac{\Pi^{\varepsilon}-\Pi^{0}}{\varepsilon}$, then

$$\varphi_n(\varepsilon)-\varphi(\varepsilon)=\frac{1}{\varepsilon}E^P\left[f(\max_{0\le t\le T}S_t^{\varepsilon,n},S_T^{\varepsilon,n})-f(\max_{0\le t\le T}S_t^{\varepsilon},S_T^{\varepsilon})\right],$$

where S^ε denotes the solution to (3.1) with $b(z)=0$, $\sigma(z)=\tilde{\sigma}z$ and $\hat{\sigma}$ is defined by (3.10). From Burkholder-Davis-Gundy's inequality, we have

$$\begin{aligned}
E^P\left[|S_T^{\varepsilon,n}-S_T^{\varepsilon}|^2\right] &\le C_1\sup_y|\hat{\sigma}_n(y)-\hat{\sigma}(y)|^2\varepsilon^2+C_2\int_0^T E^P[|S_t^{\varepsilon,n}-S_t^{\varepsilon}|^2]dt,\\
E^P\left[|\max_{0\le t\le T}S_t^{\varepsilon,n}-\max_{0\le t\le T}S_t^{\varepsilon}|^2\right] &\le C_1\sup_y|\hat{\sigma}_n(y)-\hat{\sigma}(y)|^2\varepsilon^2+C_2\int_0^T E^P\left[\max_{0\le u\le t}|S_u^{\varepsilon,n}-S_u^{\varepsilon}|^2\right]dt.
\end{aligned}$$

This implies that $\lim_{n\to\infty}\sup_{0\le\varepsilon\le 1}|\varphi_n(\varepsilon)-\varphi(\varepsilon)|=0$ is true and, therefore, (C.3) holds. The proof is completed. $\square$

Appendix D

In this section, we shall prove Lemma 13.

The proof is divided in two steps: in the first step, we deal with the bias part of (3.19), then we consider the variance part of (3.19) in the second step.

step1. Let us consider the bias part of (3.19). By the change of variables we have

$$E^P[\hat{E}(U,K)] \;=\; \frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)(\psi_1 p_M)(U-h_1y)dy. \tag{D.1}$$

From the explicit expression for $(\psi_1 p_M)$, we know that $y\in(K,\infty)\mapsto(\psi_1 p_M)(y)\in\mathbb{R}_+$ is smooth. We use the Taylor's theorem and **(K2)**, and obtain

$$\begin{aligned}
&\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)\psi_1 p_M(U-h_1y)dy\\
&=\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)\left[\psi_1 p_M(U)-h_1y(\psi_1 p_M)'(U)+\frac{1}{2}h_1^2y^2(\psi_1 p_M)''(U)\right]+o(h_1^2).
\end{aligned}$$

Due to (3.16) and **(K1)** it is easy to see

$$\begin{aligned}
\frac{1}{h_1^2}\left(\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)dy-\int_{-\infty}^{\infty}K_1(y)dy\right) &\to 0,\\
\frac{1}{h_1}\int_{-\infty}^{\frac{U-K}{h_1}}y\mathcal{K}_1(y)dy &\to 0,\\
\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)y^2dy &\to \mu_2(\mathcal{K}_1),
\end{aligned}$$

as h_1 tends to 0. This implies that

$$E^P[\hat{E}(U,K)] - E(U,K) = \frac{1}{2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}(\psi_1 p_M)''(U)\mu_2(\mathcal{K}_1)h_1^2 + o(h_1^2).$$

By the same arguments, we have

$$\begin{aligned}
E^P[\hat{T}(U,K)] - T(U,K) &= \frac{1}{2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}(\psi_2 p_S)''(K)\mu_2(\mathcal{K}_2)h_2^2 + o(h_2^2) \\
E^P[\hat{D}(U,K)] - D(U,K) &= -\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{4}\left[(\phi_1 p_M)''(U)\mu_2(\mathcal{K}_1)h_1^2 + (\phi_2 p_S)''(K)\mu_2(\mathcal{K}_2)h_2^2\right] + o(h_1^2 + h_2^2).
\end{aligned}$$

step2. Now, let us consider the variance part of (3.19). The term $Var[\hat{V}(U,K)]$ is calculated as

$$\begin{aligned}
&Var[\hat{V}(U,K)] \\
&= \frac{1}{N}E^P\Bigg[\Bigg\{\frac{1}{h_1}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)I_{(K,\infty)}(S_T)M_T\log\left(\frac{M_T}{x}\right) \\
&\quad +\frac{1}{h_2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)I_{(U,\infty)}(M_T)S_T\log\left(\frac{S_T}{x}\right) \\
&\quad -\frac{1}{h_1}\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{2}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)I_{(K,\infty)}(S_T)M_T\eta - \frac{1}{h_2}\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{2}\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)I_{(U,\infty)}(M_T)S_T T\Bigg\}^2\Bigg] \\
&\quad -\frac{1}{N}E^P\Bigg[\frac{1}{h_1}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)I_{(K,\infty)}(S_T)M_T\log\left(\frac{M_T}{x}\right) \\
&\quad +\frac{1}{h_2}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)I_{(U,\infty)}(M_T)S_T\log\left(\frac{S_T}{x}\right) \\
&\quad -\frac{1}{h_1}\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{2}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)I_{(K,\infty)}(S_T)M_T\eta - \frac{1}{h_2}\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{2}\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)I_{(U,\infty)}(M_T)S_T T\Bigg]^2 \\
&\quad =: I_1 - I_2. \qquad \text{(D.2)}
\end{aligned}$$

Let us check that I_2 in (D.2) is $o(\frac{1}{Nh_1} + \frac{1}{Nh_2})$. Using the same calculation as in (D.1), we have

$$E^P\left[\frac{1}{h_1}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)I_{(K,\infty)}(S_T)M_T\log\left(\frac{M_T}{x}\right)\right] = \frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)(\psi_1 p_M)(U-h_1 y)dy.$$

Due to the fact that $|\psi_1(y)| \le \frac{|y|^2}{|x|}, \forall y \in (K,\infty)$ and **(K1)**, we obtain

$$\lim_{h_1\to 0}\int_{-\infty}^{\frac{U-K}{h_1}}\mathcal{K}_1(y)(\psi_1 p_M)(U-h_1 y)dy < \infty.$$

We use the same calculations for the other terms of the second part of (D.2), then we can show that I_2 of (D.2) is $o(\frac{1}{Nh_1} + \frac{1}{Nh_2})$.

Then, we focus on I_1 part of (D.2). Clearly, we have

$$\begin{aligned}
I_1 &= \frac{1}{N}E^P\left[\frac{1}{h_1^2}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)^2 M_T^2 I_{(K,\infty)}(S_T)\left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\log\left(\frac{M_T}{x}\right)-\frac{\tilde\sigma_1\tilde\sigma_2}{2}\eta\right)^2\right]\\
&\quad+\frac{2}{N}E^P\left[\frac{1}{h_1h_2}\mathcal{K}_1\left(\frac{U-M_T}{h_1}\right)M_T I_{(K,\infty)}(S_T)\left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\log\left(\frac{M_T}{x}\right)-\frac{\tilde\sigma_1\tilde\sigma_2}{2}\eta\right)\right.\\
&\quad\left.\times\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)S_T I_{(U,\infty)}(M_T)\left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\log\left(\frac{S_T}{x}\right)-\frac{\tilde\sigma_1\tilde\sigma_2}{2}T\right)\right]\\
&\quad+\frac{1}{N}E^P\left[\frac{1}{h_2^2}\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)^2 S_T^2 I_{(U,\infty)}(M_T)\left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\log\left(\frac{S_T}{x}\right)-\frac{\tilde\sigma_1\tilde\sigma_2}{2}T\right)^2\right]\\
&=: I_{1,1}+I_{1,2}+I_{1,3}.
\end{aligned}\tag{D.3}$$

Let us show that $I_{1,2}=o(\frac{1}{Nh_1}+\frac{1}{Nh_2})$. We define

$$\begin{aligned}
J_1 &:= E^P\left[\frac{1}{h_1h_2}K_1\left(\frac{U-M_T}{h_1}\right)M_T I_{(K,\infty)}(S_T)\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)\right.\\
&\qquad\left.\times S_T I_{(U,\infty)}(M_T)\left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\right)^2\log\left(\frac{M_T}{x}\right)\log\left(\frac{S_T}{x}\right)\right]\\
&= \left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\right)^2\frac{1}{h_1h_2}\int_U^\infty \mathcal{K}_1\left(\frac{U-y}{h_1}\right)(\Phi p_M)(y)dy,
\end{aligned}$$

where $\Phi(y):=y\log(\frac{y}{x})E^P[I_{(K,\infty)}(S_T)\mathcal{K}_2(\frac{K-S_T}{h_2})S_T\log(\frac{S_T}{x})|M_T=y]$. By the change of variables, we obtain

$$J_1=\left(\frac{\tilde\sigma_2}{\tilde\sigma_1}\right)^2\frac{1}{h_2}\int_{-\infty}^0\mathcal{K}_1(y)(\Phi p_M)(U-h_1y)dy.$$

Now let us focus on $E^P[I_{(K,\infty)}(S_T)\mathcal{K}_2(\frac{K-S_T}{h_2})S_T\log(\frac{S_T}{x})|M_T=y]$. For $y\in(K,\infty)$, again by the change of variables, we have

$$\begin{aligned}
&E^P\left[I_{(K,\infty)}(S_T)\mathcal{K}_2\left(\frac{K-S_T}{h_2}\right)S_T\log\left(\frac{S_T}{x}\right)\middle|M_T=y\right]\\
&= \int_K^y\mathcal{K}_2\left(\frac{K-z}{h_2}\right)z\log\left(\frac{z}{x}\right)p_{S|M}(z|y)dz\\
&= \frac{1}{p_M(y)}\int_K^y\mathcal{K}_2\left(\frac{K-z}{h_2}\right)z\log\left(\frac{z}{x}\right)p_{M,S}(y,z)dz\\
&= \frac{h_2}{p_M(y)}\int_{\frac{K-y}{h_2}}^0\mathcal{K}_2(z)(K-h_2z)\log\left(\frac{K-h_2z}{x}\right)p_{M,S}(y,K-h_2z)dz,
\end{aligned}$$

where $p_{S|M}(y|z)$ denotes the conditional density function of S_T given by M_T and $p_{M,S}(y,z)$ denotes the joint density function of (M_T, S_T). Therefore, we have

$$\begin{aligned} J_1 &= \left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2 \int_{-\infty}^{0} \mathcal{K}_1(y)(U - h_1 y) \log\left(\frac{U - h_1 y}{x}\right) \\ &\quad \times \int_{\frac{K-y}{h_2}}^{0} \mathcal{K}_2(z)(K - h_2 z) \log\left(\frac{K - h_2 z}{x}\right) p_{M,S}(U - h_1 y, K - h_2 z) dz dy, \end{aligned}$$

and $J_1 = o(\frac{1}{Nh_1} + \frac{1}{Nh_2})$. The same calculation yields that the other terms of $I_{1,2}$ are also $o(\frac{1}{Nh_1} + \frac{1}{Nh_2})$.

Finally, we consider $I_{1,1}$ and $I_{1,3}$ of (D.3). By using Ψ_1, Ψ_2 and Ψ_3, we have the following expression for $I_{1,1}$,

$$\begin{aligned} I_{1,1} &= \frac{1}{Nh_1}\left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_2}\right)^2 \int_{-\infty}^{\frac{U-K}{h_1}} \mathcal{K}_1(y)^2(\Psi_1 p_M)(U - h_1 y)dy - \tilde{\sigma}_2^2 \int_{-\infty}^{\frac{U-K}{h_1}} \mathcal{K}_1(y)^2(\Psi_2 p_M)(U - h_1 y)dy \right. \\ &\quad \left. + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4} \int_{-\infty}^{\frac{U-K}{h_1}} \mathcal{K}_1(y)^2(\Psi_3 p_M)(U - h_1 y)dy \right]. \end{aligned}$$

Thus, by applying the Taylor's theorem for $(\Psi_1 p_M)(y)$, $(\Psi_2 p_M)(y)$ and $(\Psi_3 p_M)(y)$, due to **(K2)**, we get

$$I_{1,1} = \left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2 \Psi_1(U) - \tilde{\sigma}_2^2\Psi_2(U) + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4}\Psi_3(U)\right] p_M(U) R_2(\mathcal{K}_1)\frac{1}{Nh_1} + o(\frac{1}{Nh_1}).$$

For $I_{1,3}$, the same calculation applies and we obtain

$$I_{1,3} = \left[\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^2 \Phi_1(K) - \tilde{\sigma}_2^2\Phi_2(K) + \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{4}\Phi_3(K)\right] p_S(K) R_2(\mathcal{K}_2)\frac{1}{Nh_2} + o(\frac{1}{Nh_2}).$$

This completes the proof, due to **(K4)**. □

Bibliography

[1] H. P. Bermin, A. Kohatsu-Higa and M, Montero, Local vega index and variance reduction methods, *Mathematical Finance* **13** (2003) 85-97.

[2] J. Bertoin, *Lévy Processes. Cambridge Tracts in Mathematics, 121* (Cambridge University Press, 1996).

[3] A. Borodin and P. Salminen, *Handbook of Brownian motion; Facts and Formulae, Second Edition* (Birkhäuser, 2002).

[4] G. Brunick and S. Shreve, Mimicking an Itô process by a solution of a stochastic differential equation, *Annals of Applied Probability* **23** (4) (2013) 1584-1628.

[5] B. Dupire, Pricing with a smile, *Risk* **7** (1994) 18-20.

[6] E. Fournié, J.M. Lasry, J. Lebuchoux, P.L. Lions and N. Touzi, Applications of Malliavin calculus to Monte Carlo methods in finance, *Finance Stochastic* **3** (1999) 391-412.

[7] N. Fournier and J. Printems, Absolute continuity for some one-dimensional processes, *Bernoulli* **16** (2) (2010) 343-360.

[8] J. Gatheral, *The Volatility Surface: A Practitioner's Guide* (Wiley Finance, 2006).

[9] E. Gobet and A. Kohatsu-Higa, Computation of greeks for barrier and look-back options using Malliavin calculus, *Electronic Communications in Probability* **8** (2003) 51-62.

[10] M. Hayashi and A. Kohatsu-Higa, Smoothness of the distribution of the supremum of a multi-dimensional diffusion process, *Potential Analysis* **38**/1 (2013) 57-77.

[11] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus. Second Edition, Graduate Texts in Mathematics, 113* (Springer-Verlag, New York, 1991).

[12] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, *In: Ecole d'Eté de Probabilités de Saint Flour XII, 1982, Lecture Notes in Math. 1097* (1984) 144-305.

[13] G. Liu and L.J. Hong, Kernel estimation of the greeks for options with discontinuous payoffs, *Operations Research* **59** (2011) 96-108.

[14] P. Mörters and Y. Peres, *Brownian motion* (Cambridge University Press, 2010).

[15] T. Nakatsu, Absolute continuity of the laws of a multi-dimensional stochastic differential equation with coefficients depending on the maximum, *Statistics and Probability Letters* **83**/11 (2013) 2499-2506.

[16] T. Nakatsu, Volatility risk for options depending on extrema and its estimation using kernel methods.(submitted)

[17] D. Nualart, *The Malliavin Calculus and Related Topics, Second Edition* (Springer-Verlag, Berlin, 2006).

[18] M.P. Wand and M.C. Jones, *Kernel Smoothing* (Chapman & Hall, London, 1993).

[19] W. Yue and T.S. Zhang: Absolute continuity of the law of perturbed diffusion processes and perturbed reflected diffusion processes. *Journal of Theoretical Probability* (in press) DOI 10.1007/s10959-013-0499-7.