Optimal stopping and sensitivity analysis in regime switching models

Liu, Yue

2015

Liu, Y. (2015). Optimal stopping and sensitivity analysis in regime switching models.
Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/65834>

<https://doi.org/10.32657/10356/65834>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**Optimal Stopping and Sensitivity
Analysis in Regime Switching
Models**

YUE LIU

**DIVISION OF MATHEMATICAL SCIENCES
SCHOOL OF PHYSICAL AND MATHEMATICAL
SCIENCES**

2015

b

Optimal Stopping and Sensitivity Analysis in Regime Switching Models

YUE LIU

Division of Mathematical Sciences
School of Physical and Mathematical Sciences

A thesis submitted to Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy in Mathematical Sciences

2015

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Nicolas Privault for the guidance and providing me the privilege of working in this interesting research area. Without his invaluable supervision, encouragement, and wisdom over these years, this thesis could not have reached its present form.

I am grateful to Nanyang Technological University for awarding me the Research Scholarship. I am also benefited from the young, energetic and productive environment of the new School of Physical and Mathematical Sciences, where I could grow and develop. In particular, I appreciate the management and leadership of Prof. San Ling, Prof. Yeow Meng Chee, Prof. Huaxiong Wang and Prof. Guohua Wu. I am appreciative of the help from the IT and Admin. staff.

I would like to express my appreciation to Prof. Guangming Pan, Prof. Liming Xiang, Prof. Xin Chen and Prof. Heng Lian for the lectures and seminars. I am also indebted to other members in our group for their friendship and help.

Last but not least, I wish to thank my parents for their love and support.

Contents

Acknowledgements	i
1 Introduction	1
1.1 Summary	1
1.2 Optimal stopping in regime switching models	4
1.2.1 Regime switching models	4
1.2.2 Optimal stopping problems	6
1.3 A recursive algorithm for optimal selling in regime switching models	11
1.4 Sensitivity analysis by Wiener-Malliavin calculus in regime switching models	13
1.5 Integration by parts for Markov chains and its applications to sensitivity analysis in regime switching models	17
2 Optimal stopping in regime switching models	21
2.1 Problem formulation	22
2.2 Stopping set and boundary functions	27
2.3 Continuity lemmas	34
2.4 Solution of the free boundary problem	45

3	A recursive algorithm for optimal selling	57
3.1	Problem formulation and solutions	58
3.2	Proof of Theorem 3.1.1	65
3.3	Proof of Theorem 3.1.2	70
4	Sensitivity analysis by Wiener-Malliavin calculus	79
4.1	Formulation	80
4.2	Computation of the Greeks based on the partial Malliavin calculus	81
4.2.1	Variations in the initial price	82
4.2.2	Variations in the initial price for second order . . .	84
4.2.3	Variations in the diffusion coefficient	85
4.2.4	Variations in the drift coefficient	87
4.2.5	Optimization of convergence	88
4.3	Extension to non-differentiable payoff function	89
5	Sensitivity analysis by integration by parts for Markov chains	95
5.1	Poisson process setting	96
5.2	Birth process setting	97
5.3	Markov chain setting	102
5.4	Sensitivity analysis in a regime switching model	104
5.5	Extension to non-differentiable payoff functions	109
5.6	Another proof of integration by parts formula for birth processes	113
6	Conclusion	119
	Bibliography	121

List of Figures

1.1	Optimal stopping according boundary curves for different regime states.	9
1.2	Computation of <i>Delta</i> with respect to the number of samples.	16
2.1	Boundary curves for different regime states.	56
3.1	Boundary function computed from Theorems 3.1.1-3.1.2 <i>vs</i> (2.57).	64
3.2	Convergence graph from Theorems 3.1.1-3.1.2 <i>vs</i> the approach by Volterra equation.	64
3.3	Non-monotone boundary function under drifts of mixed signs.	65
4.1	Computation of <i>Delta</i> with respect to the number of samples.	90
5.1	Computation of <i>Delta</i> with respect to the number of samples.	109

Chapter 1

Introduction

1.1 Summary

This thesis focuses on topics related to regime switching models. Two main problems are concerned, one is optimal stopping in regime switching models, and the other is sensitivity analysis with its application to calculate Greeks based on regime switching models. In particular, the regime switching model is characterized by a geometric Brownian motion with its nondeterministic drift and volatility driven by a time-continuous Markov chain which is independent of the Brownian motion, regime switching simulated by the shifts of Markov states reflects the underlying changes in the state of the economy.

On one hand, we consider the optimal stopping problem in this framework, specifically, the ultimate maximum of an optimal value function over the optimal stopping times is predicted, and the optimal stopping time to achieve the maximum value is right time for selling, hence an optimal stopping strategy is derived depending on the current regime states. This work extends existing results for geometric Brownian motion with

constant drift and volatility, by applying the method of free-boundary problems and the techniques of local time-space calculus, we characterize the optimal stopping boundaries as solutions of a system of integral equations, and we address uniqueness of solutions under positive drifts, while immediate exercise is obtained under negative drifts. Our approach replaces the use of closed form expressions, which are not available in our setting, with PDE arguments which also simplify the approach of Du Toit and Peskir (2009) in the classical geometric Brownian case.

In addition, for the same problem of optimal stopping in regime switching models, we propose a recursive algorithm for the numerical computation of exercise prices. By applying a discretization of the time interval, a setting of backward inductions is designed. Based on the computational results of optimal value function for all states, we can determine the boundary of the optimal stopping set, and the optimal stopping strategy is thereby derived. This method avoids to solve the Volterra integral equations. Comparing with the first approach to solve it as free boundary problem, this one applies in particular when regime switching drifts have mixed signs, in which case the boundary functions may not be monotone and the Volterra equations are not available.

On the other hand, we present a Malliavin calculus approach to sensitivity analysis based on regime switching models. As is well known, risk management in portfolio hedging is often achieved by estimating the Greeks, which are price sensitivities relative to variations in the model parameters. Following a method initiated on the Wiener space in Fournié et al. (1999) and applying the integration by parts formula of Malliavin calculus, Greeks are calculated based on the regime switching model driven by

geometric Brownian motion and Markov chain. We show by Monte Carlo simulation that this approach outperforms that by finite differences in terms of computational efficiency. By developing this method for sensitivity analysis, we have multiple versions of Greeks expression, optimization by minimizing the variance of weight is available among those alternatives. Although the classical Malliavin calculus approach requires the differentiability of the payoff function, we extend the results for models with non-differentiable payoff function.

Paralleling to the approach by Malliavin calculus with respect to the randomness of Gaussian part, similar investigations regarding Markov chains are developed, the core technique is to establish the integration by parts formula with a well defined gradient with respect to the randomness from the Markovian part. Indeed, we set up this version of integration by parts formula for the random functionals of a continuous-time Markov chain, based on partial differentiation with respect to jump times, while leaving the discrete-time embedded chain unchanged. As an application we also compute sensitivities (Greeks) in asset price models with Markovian regime switching. In comparison with classical approaches based on the partial Malliavin calculus for Brownian motion, our method has a wider scope of application by taking both sources of Markovian and Gaussian noises into account. We also improve on some recent results in the literature on integration by parts for Markov chains.

1.2 Optimal stopping in regime switching models

1.2.1 Regime switching models

Geometric Brownian motion model as follows is applied in the classical Black-Scholes formula to capture the dynamics of the asset price.

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad t \in \mathbb{R}_+, \quad (1.1)$$

where $(B_t)_{t \in \mathbb{R}_+}$ is a standard Brownian motion, $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$. In this model, the drift factor μ and volatility σ are assumed to be constants. However, shown by practices and researches, the Black-Scholes model is not able to reflect the stochastic variability in real parameters. Besides, the widely admitted "volatility smile" structure cannot be captured by this model. Indicated by the "volatility smile", the volatility implied by the option price calculated by the Black-Scholes formula is not reasonable to be assumed as a constant. To make for these shortfalls and meet the need of more realistic models that reflect and match the market more closely, regime switching models show their advantages and draw more interests in both research and practice.

Regime switching models was first introduced by Hamilton (1989) in discrete time and are among the most popular and effective risky asset models. The regime switching property is reflected in the changes of states of a jump process, which stands for the influence of external market factors. Regime switching models are widely applied and studied in recent decades. In a diffusion model of Di Masi et al. (1994), drift and volatility coefficients are driven by a Markov chain, which indicates the regime switching. To present a valuation technique for both European

and American-style option, Bollen (1998) applies a lattice-based method in regime switching models. Parameters fitting for some classical regime switching models are discussed by Hardy (2001). For pricing the volatility derivatives such as variance swaps and volatility swaps, Buffington and Elliott (2002) state their partial differential equation approach under a continuous-time Markov-modulated version of the stochastic volatility model developed by Heston (1993). With consideration of the feedback effects from returns to volatility, a regime switching process is used to capture the stock price dynamic and efficient algorithm for American option pricing based on this model is developed by Duan et al. (2002). By Liu et al. (2006), a fast Fourier transform approach is developed for the model of asset price governed by a regime switching geometric Brownian motion. European options have been priced in continuous-time regime switching models by Yao et al. (2006) via a recursive algorithm. By Papanicolaou and Sircar (2014), the sharp regime shifts are applied to capture VIX implied volatilities, which cannot be captured by the basic Heston model. For other studies and applications of regime switching models, we also refer to Guo (1999), Zhang (2001), Shepp (2002), Yin et al. (2003), Zhang and Yin (2004).

In this thesis, we consider the typical regime switching model (1.2) of geometric Brownian motion and Markov chains as that by Yao et al. (2006) and Le and Wang (2010).

$$dY_t = \mu(\beta_t)Y_t dt + \sigma(\beta_t)Y_t dB_t, \quad 0 \leq t \leq T, \quad (1.2)$$

with regime switching driven by a finite-state continuous-time Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ on the state space $\mathcal{M} := \{1, \dots, m\}$.

1.2.2 Optimal stopping problems

Optimal stopping problems have been studied for half century, which are structured for Brownian motion by Bather (1970), for Markov chains by Engelbert (1973), and also considered for selling derivatives such as bonds by Boyce (1970). Theoretical studies about optimal stopping problem in early years can be found by Chow et al. (1971). During the last decade, theory and applications of optimal stopping have been greatly developed again, c.f. books by Peskir and Shiryaev (2006). Investigating the optimal stopping problem modelled by adjoining a hidden Markov process to the classical Black-Scholes geometric Brownian motion, Guo (2001) prices the Russian option (i.e. perpetual look-back options), another work by Guo and Zhang (2005) is based on a similar model while the optimal stopping rule is of a threshold type for each state. Based on a regime switching model of geometric Brownian motion and two state Markov chains, Ly Vath and Pham (2007) deal with an optimal switching problem. Du Toit and Peskir (2007) consider the optimal prediction problems based on the value function that

$$V = \inf_{0 \leq \tau \leq T} E[B_\tau^\mu - S_T^\mu]^2 \quad \mu \in \mathbf{R},$$

where $(B_t^\mu)_{t \in [0, T]}$ is a Brownian with drift μ and $S_T^\mu = \max_{0 \leq t \leq T} B_t^\mu$. This work is followed by Du Toit et al. (2008), which aims to predict the last zero of Brownian motion with drift in an optimal stopping setting. Optimal selling under threshold rules has been dealt with in Eloë et al. (2008) in an exponential Gaussian diffusion model with regime switching, where a stopping time is defined as the hitting time of either take-profit level or stop-loss level. By Shiryaev et al. (2008) optimal selling problems based on different models and object functions are summarized and concluded. The problem of selling a stock at the optimal maximum has been consid-

ered in Du Toit and Peskir (2009) when the underlying asset price is given by a geometric Brownian motion, we refer to Shiryaev (1978) and Peskir and Shiryaev (2006) for background on the characterization of stopping times and optimal rewards based on a regime switching model.

We consider a similar optimal prediction problem as that by Du Toit and Peskir (2009):

$$V_t = \inf_{t \leq \tau \leq T} E \left[\sup_{0 \leq s \leq T} \frac{Y_s}{Y_\tau} \middle| \mathcal{F}_t \right], \quad (1.3)$$

over all $\mathcal{F}_t$ stopping times τ , where the asset price $(Y_t)_{t \in \mathbb{R}_+}$ is modeled by (1.2), with the drift and volatility driven by Markov chains to reflect the regime switching, while the model by Du Toit and Peskir (2009) is built on a geometric Brownian motion (1.1).

The optimal stopping model (1.3) is chosen based on the following considerations. Letting the stopping time τ denote a strategy of selling and take values in $[0, T]$, we find several ways to measure its performance. Consider the expected ratio

$$V_1(\tau) := E \left[\sup_{0 \leq s \leq T} \frac{X_\tau}{X_s} \right],$$

where $(X_s)_{s \in [0, T]}$ is a geometric Brownian motion (1.1) with constant drift μ and volatility σ , we deem τ to be the optimal strategy if V_1 is relatively large, since $\sup_{0 \leq s \leq T} X_s$ denotes the maximum of profit one may take by selling the derivative. Or we consider

$$V_2(\tau) := E \left[\sup_{0 \leq s \leq T} \frac{X_s}{X_\tau} \right]$$

to conclude it is good when V_2 is small, by which logic the optimal value function (1.3) is designed. The optimal problems based on $V_1(\tau)$ and $V_2(\tau)$ are first recorded in Graversen and Shiryaev (2000), although they are in

similar form of expression, conclusions for them obtained by Du Toit and Peskir (2009) are different for some value of parameters. Regarding $V_1(\tau)$, Yam et al. (2012) provide independent and purely probabilistic proofs for the case $\mu = \frac{1}{2}\sigma^2$ and $\mu > \frac{1}{2}\sigma^2$ respectively. It is also reasonable to evaluate on the basis of

$$V_3(\tau) := E[(\sup_{0 \leq s \leq T} X_s - X_\tau)^p], \quad p > 0,$$

which is similar to the model in Du Toit and Peskir (2007). Or we have more complicated form:

$$V_4(\tau) := E[|\inf\{t \in [0, T] | X_t = \sup_{0 \leq s \leq T} X_s\} - \tau|],$$

which is applied by Du Toit and Peskir (2008). Other alternatives like

$$V_5(\tau) := E[X_\tau], \quad V_6(\tau) := E\left[\log\left(\frac{\sup_{0 \leq s \leq T} X_s}{X_\tau}\right)\right]$$

can be simplified into a trivial case and even the volatility σ vanishes, as shown for $V_6(\tau)$ by Shiryaev et al. (2008), the solution is a simple bang-bang structure (stop either at the beginning or at the end), c.f. Yam et al. (2009) and Yam et al. (2013). The advantage in the model (1.3) is that it has the effect of stripping away the monetary value of the derivative and concentrate on the underlying randomness.

Main contribution

In the first part of this thesis, we extend the result of Du Toit and Peskir (2009) to the framework of Markovian regime switching. Firstly, we write the optimal value of this problem as a function of time, the regime state, and the relative maximum of the underlying asset. We also identify the optimal stopping time τ_D and determine the structure of the optimal stopping set D via its boundary functions

$$b_D(t, j) := \inf\{x \in [1, \infty) : (t, j, x) \in D\}$$

for j in the state space $\mathcal{M} := \{1, \dots, m\}$ of the regime switching chain. Thereby a stopping rule is characterised by the boundary functions, as illuminated by Figure 1.1 below, in which case we have two regime states, so we have two boundary lines in this graph, plot the price curve of $\hat{Y}_{0,t}/Y_t$ with solid blue line when $\beta_t = 1$ and with dash red line when $\beta_t = 2$. Observe and wait till the curve hits the boundary curve of the same color. Figure 1.1 records a practice of this strategy by this stopping rule, as indicated in the graph, at the point S , marked by a red circle, the price curve of $\hat{Y}_{0,t}/Y_t$ hits the boundary curve with the same state for the first time, indicating that selling at this moment is optimal.

Figure 1.1: Optimal stopping according boundary curves for different regime states.

Some of our results are natural extensions of those of Du Toit and Peskir (2009) by averaging over the regime switching component, however the

regime switching case presents notable differences and additional difficulties compared with the classical Brownian case. For instance, the optimal boundary functions depend on the regime states of the system, and they may not be monotone if the drift coefficients have switching signs, one numerical example is provided at the end of Subsection 3.1 with Figure 3.3. Without the monotonicity of the boundary functions, we cannot apply the local time change of variable formula of Peskir (2005), hence fail to obtain a solution in the form of a Volterra equation. So we treat respectively for different ranges of drift value. When the drift parameters of the regime switching chain are all nonnegative we prove the continuity and monotonicity of boundary functions $b_D(t, j)$ of $t \in [0, T]$ for each $j \in \mathcal{M}$. Under such conditions, we derive the Volterra integral equation as follows:

$$G(t, j, b_D(t, j)) = J(t, j, b_D(t, j)) - \int_t^T K(t, r, j, b_D(t, j)) dr,$$

$0 \leq t \leq T$, with the boundary condition $b_D(T, j) = 1$, $j \in \mathcal{M}$, where

$$G(t, j, x) := E \left[\max \left(x, \frac{\hat{Y}_{t,T}}{Y_t} \right) \mid \beta_t = j \right],$$

$$J(t, j, x) := E[X_T^{t,x} \mid \beta_t = j], \quad X_r^{t,x} := \frac{1}{Y_r} \max \left(x Y_t, \hat{Y}_{t,r} \right),$$

and

$$K(t, r, j, x) := E \left[\mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > b_D(r, \beta_r)\}} \mid \beta_t = j \right],$$

for $0 \leq t \leq r \leq T$ and $x \geq 1$. This equation characterizes the boundary function $b_D(t, j)$ of the stopping set.

We also show the uniqueness of solutions to the boundary integral equation within a class of continuous functions on $[0, T]$ under the condition $\mu(j) \geq 0$ for all $j \in \mathcal{M}$. In addition, we show that immediate exercise is optimal when all drift parameters $\mu(j)$ are negative, $j \in \mathcal{M}$, while

exercise at maturity becomes optimal when $\mu(j) \geq \sigma^2(j)$ for all $j \in \mathcal{M}$, where $\sigma(j)$ are the volatility parameters.

In addition we can no longer rely on closed form expressions as in Du Toit and Peskir (2009), which are not available in our setting, instead we use PDE arguments, which also simplify the original approach.

1.3 A recursive algorithm for optimal selling in regime switching models

In the second part, we develop a recursive algorithm for the same optimal selling problem in regime switching models as in the first part. However it concludes to all those cases the approaches in the first part fail to cover and provides evidences for the facts that boundary functions may not be monotonous if the drift coefficients have switching signs. This method computes the optimal value function through backward inductions and allows us to determine the boundary functions of the optimal stopping set without having to solve their associated Volterra integral equations.

Main contribution

Our algorithm is inspired by and also extends the method of Yao et al. (2006) by solving an optimal problem instead of only computing an expectation. However it is different even when restricted to expectations $E[\phi(Y_T)]$ of payoff functions $\phi(Y_T)$, where $(Y_t)_{t \in [0, T]}$ follows (1.2). In particular the recursion of Yao et al. (2006) is based on the jump times of the Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ whereas we apply a discretization of the time interval $[0, T]$, and our algorithm involves the Monte Carlo method only in (1.4) below.

In particular, we construct a recursive algorithm for the numerical solution of the optimal stopping problem (1.3), by determining the stopping set D from the values of $V(t, j, a)$ and $G(t, j, a)$ computed as follows,

(i) For all $t \in [0, T]$, $j \in \mathcal{M} := \{1, \dots, m\}$ and $a \geq 1$ we have

$$V(t, j, a) = \lim_{n \rightarrow \infty} V_n(t, j, a),$$

where the value of $V_n(t, j, a)$ can be computed by the backward induction

$$V_n(t - \delta_n, j, a) = E \left[G \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \wedge V_n \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \mid \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right], \quad (1.4)$$

$t \in [\delta_n, T]$, where $\delta_n = T/n$, under the terminal condition $V_n(T, j, a) = G(T, j, a) = a \geq 1$, where $G(t, j, a)$ is computed through the relation $G(t, j, a) = \Psi(t, t, j, a)$ for any $t \in [0, T]$, $a \geq 1$ and $j \in \mathcal{M}$, and for $0 \leq t \leq h \leq T$, $\Psi(t, h, j, a)$ is calculated by the following induction.

(ii) For any $0 \leq t \leq h \leq T$, $j \in \mathcal{M}$ and $a \geq 1$, we have

$$\Psi(t, h, j, a) = \lim_{n \rightarrow \infty} \Psi_n(t, h, j, a),$$

where $\Psi_n(t, h, j, a)$ satisfies the backward induction

$$\begin{aligned} \Psi_n(t, h, j, a) &:= e^{q_{j,j}\delta_n} \int_0^\infty dy \int_{-\infty}^y e^{(u(j)+\mathbf{1}_{\{t=h\}}\sigma(j))x - u^2(j)T/(2n)} \\ &\quad \times \Psi_n(t + \delta_n, h + \mathbf{1}_{\{t=h\}}\delta_n, j, a \vee (\sigma(j)y) - \sigma(j)x) \varphi_{\delta_n}(x, y) dx \\ &\quad + \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} dr \int_0^\infty dy \int_{-\infty}^y e^{(u(j)+\mathbf{1}_{\{t=h\}}\sigma(j))x - u^2(j)r/2} \\ &\quad \times \Psi_n(t + \delta_n, h + \mathbf{1}_{\{t=h\}}\delta_n, l, a \vee (\sigma(j)y) - \sigma(j)x) e^{q_{j,j}r} \varphi_r(x, y) dx, \end{aligned}$$

$t \leq T - \delta_n$, $h \geq t$, $j \in \mathcal{M}$ and $a \geq 1$, with the terminal condition $\Psi_n(t, h, j, a) = a$, $n \geq 1$ for all t, h that $T - \delta_n < t \leq h \leq T$.

As this approach does not rely on the Volterra equation, it allows us in particular to determine the boundary function $b_D(t, j)$ without requiring the condition $\mu(j) \geq 0$ for all $j \in \mathcal{M}$ as in Liu and Privault (2015b). In comparison with Du Toit and Peskir (2009) our method does not rely on closed form expressions of value functions as in Du Toit and Peskir (2009) as they are no longer available in the regime switching setting. In addition we do not require Monte Carlo simulations in the absence of regime switching, since in this case (1.4) can be computed in deterministic form, cf. (2.7) in Du Toit and Peskir (2009).

1.4 Sensitivity analysis by Wiener-Malliavin calculus in regime switching models

In the third part of this thesis, we consider the sensitivity analysis in regime switching models. Emerging interests focus on sensitivity analysis with its wide application in risk management, especially for hedging strategy. In particular, Greeks are the sensitivities with respect to parameters in the generalized Black–Scholes’ models.

Starting from Fournié et al. (1999), Malliavin calculus is applied for computation of Greeks such like Δ , ρ , respectively representing the sensitivities of option value to spot price, risk-free rate. Via this approaches, fast algorithms for Greeks computations are designed. Following the method initiated on the Wiener space in Fournié et al. (1999), El-Khatib and Privault (2004) calculate the Greeks in a model of jump process driven by Poisson jump times. By Debelley and Privault (2004b), sensitivities in a jump diffusion model are computed by using the Malliavin calculus. By Davis and Johansson (2006), Malliavin calculus is ap-

plied for Lévy processes and integration by parts formula is developed for Greeks calculation, where the computational efficiency is assessed by comparison through Monte Carlo simulation. Using the Malliavin calculus on Poisson space, Bavouzet-Morel and Messaoud (2006) compute the sensitivities for European and Asian options simulated by a jump type diffusion. Applying the Malliavin calculus in time inhomogeneous jump-diffusion models, David and Privault (2009) obtain an expression for the sensitivity Theta of an option price as the expectation of the option payoff multiplied by a stochastic weight. Also, Bayazit and Nolder (2013) present sensitivities for options when the underlying dynamic follows an exponential Lévy process. For other recent works on computation of Greeks in asset price models, also c.f. Kawai and Kohatsu-Higa (2010) and Kawai and Takeuchi (2013).

In this chapter, applying the Malliavin calculus, we proceed the sensitivity analysis in the framework of the regime switching model (1.2) of geometric Brownian motion $(W_t)_{t \in \mathbb{R}_+}$ with finite-state continuous-time Markov chains $(\alpha_t)_{t \in \mathbb{R}_+}$ on the state space $\mathcal{M} := \{1, \dots, m\}$. The value function $V(X_0, \alpha_0)$ we consider is:

$$V(x_0, \alpha_0) = e^{-rT} E[\phi(X_T) \mid X_0 = x_0, \alpha_0],$$

where $r > 0$ denotes the risk-free rate and $(X_t)_{t \in \mathbb{R}_+}$ follows the SDE below.

$$dX_t = X_t[\mu(\alpha_t)dt + \sigma(\alpha_t)dW_t], \quad 0 \leq t \leq T$$

with $X_0 > 0, \alpha_0 \in \mathcal{M}$, and $\mu : \mathcal{M} \rightarrow \mathbb{R}$, and $\sigma : \mathcal{M} \rightarrow (0, \infty)$ are deterministic functions. We aim to calculate the sensitivities regarding model coefficients listed as follows,

$$\Delta = \frac{\partial V(x_0, \alpha_0)}{\partial x_0}$$

which is the sensitivity of value function V to the initial price X_0 ,

$$\Gamma = \frac{\partial^2 V(x_0, \alpha_0)}{\partial x_0^2}$$

which is the second-order variations in the initial price X_0 . Although usually the Greeks Vega is defined as

$$\text{Vega} = \frac{\partial V(x_0, \alpha_0)}{\partial \sigma},$$

it is not computable since the volatility σ in our model is not a constant. Therefore we introduce a perturbed process by

$$dX_t^{\sigma, \varepsilon} = X_t^{\sigma, \varepsilon} [\mu(\alpha_t) dt + (\sigma(\alpha_t) + \varepsilon) dW_t], \quad \varepsilon \in \mathbf{R},$$

and accordingly,

$$V^{\sigma, \varepsilon}(X_0, \alpha_0) := e^{-rT} E[\phi(X_T^{\sigma, \varepsilon}) \mid X_0 = x_0, \alpha_0],$$

hence the variation in the diffusion coefficient is defined by

$$\text{Vega} = \left. \frac{\partial V^{\sigma, \varepsilon}(x_0, \alpha_0)}{\partial \varepsilon} \right|_{\varepsilon=0}.$$

Variation in drift coefficient

$$\Lambda = \frac{\partial V(x_0, \alpha_0)}{\partial \mu}$$

is similarly defined by introducing the perturbed process regarding the drift part.

Main contribution

In this chapter, we follow the method initiated on the Wiener space in Fournié et al. (1999) to calculate the above mentioned Greeks by applying the integration by parts formula of Malliavin calculus. Verified by

Monte Carlo simulation, this approach outperforms the method of finite differences in terms of computational efficiency, indicated by the following Figure 1.2, computed according Proposition 4.2.1 with $g(\cdot) = \sigma(\cdot)$.

Figure 1.2: *Computation of Delta with respect to the number of samples.*

By developing this method for sensitivity analysis, we have multiple versions of Greeks expressions, comparison in terms of variance among those alternatives enable us to further improve the methods of computation. Although the classical Malliavin calculus approach requires the differentiability of the payoff function, we extend the results to the class $\Lambda(\mathbb{R}; \mathbb{R})$ of non-differentiable payoff function, defined as follows,

$$\Lambda(\mathbb{R}; \mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \mid f = \sum_{i=1}^n f_i \mathbf{1}_{\{A_i\}}, n \geq 1, f_i \in C_L(\mathbb{R}; \mathbb{R}), \right.$$

A_i are intervals of $\mathbb{R}$,

where

$$C_L(\mathbb{R}; \mathbb{R}) := \{f \in C(\mathbb{R}; \mathbb{R}) \mid |f(x) - f(y)| \leq k|x - y| \text{ for some } k \geq 0\}.$$

1.5 Integration by parts for Markov chains and its applications to sensitivity anal- ysis in regime switching models

As mentioned in Subsection 1.4, the use of Malliavin calculus and integration by parts formulas for the fast sensitivity analysis of random systems has been introduced in Fournié et al. (1999) for the calculation of Greeks in option hedging.

Integration by parts formulas can be based on multiple stochastic integral expansions as on the Wiener space, or on the absolute continuity of jump times, see e.g. Chapters 6 and 7 of Privault (2009). In the case of discrete systems driven by Markov chains, discrete multiple stochastic integral expansions for random variables have been constructed in e.g. Kroecker (1980), Biane (1989), Privault and Schoutens (2002).

More recently, an integration by parts formula for the functionals of a Markov process $(X_t)_{t \in \mathbb{R}_+}$ based on functions of numbers of jumps has been proposed in Siu (2014). However this construction relies on an erroneous application of the characterization of pure jump martingales, cf. Theorem 4 therein. In addition such a formula, if correct, would imply the absolute continuity of X_t itself, which is impossible since it has a discrete state space.

Main contribution

In this chapter we propose an integration by parts formula for continuous-time Markov processes, based on the jump times of its embedded discrete-time chain as in e.g. El-Khatib and Privault (2004) or Privault and Wei (2004), cf. also § 8.3 of Malliavin and Thalmaier (2006). Note that here we do not require the existence of absolutely continuous jump sizes as in e.g. Bally et al. (2007).

In this way we establish a “partial” Malliavin calculus with respect to the absolutely continuous jump times while the spatial discrete jump component remains unchanged. This approach has some similarities to the partial Malliavin calculus of Davis and Johansson (2006) or Debelley and Privault (2004a), in which only the Brownian component of a jump diffusion is subjected to a random perturbation.

Note that a similar Markov chain model has been considered in El-Khatib and Privault (2004) for continuous Markov chains whose matrix generator has a constant diagonal, i.e. for state-independent exponential distribution of interjump times, with application to Asian options. On the other hand, in this chapter we consider general continuous-time Markov chains with application to the sensitivity analysis of regime switching models.

Precisely, we consider a continuous-time Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ with transition rate matrix $Q = (q_{i,j})_{1 \leq i,j \leq m}$. Let T_n^β denote the n -th jump time of $(\beta_t)_{t \in [0,T]}$ with $T_0^\beta = 0$, and let

$$N_t^\beta := \min \left\{ n \in \mathbb{N} : t < T_{n+1}^\beta \right\}$$

denote the number of jumps of $(\beta_t)_{t \in \mathbb{R}_+}$ up to time $t \in [0, T]$. The Markov

chain $(\beta_t)_{t \in \mathbb{R}_+}$ can be represented as

$$\beta_t = Z_{N_t^\beta}, \quad t \in \mathbb{R}_+, \quad \text{where} \quad Z_n := \beta_{T_n^\beta}, \quad n \in \mathbb{N},$$

denotes the embedded chain of $(\beta_t)_{t \in \mathbb{R}_+}$, with $Z_0 = \beta_0$. Define a partial gradient

$$D_t^\beta F := -q_{\beta_t, \beta_t} D_{\lambda(t)} F, \quad t \in \mathbb{R}_+,$$

on the space $\mathcal{S}_T^\beta$ of functionals of the form

$$F = f_0 \mathbf{1}_{\{N_T^\beta=0\}} + \sum_{k=1}^n \mathbf{1}_{\{N_T^\beta=k\}} f_k(T_1^\beta, \dots, T_k^\beta, Z_1, \dots, Z_k),$$

where $f_0 \in \mathbb{R}$ and $f_k(\cdot, z_1, \dots, z_k) \in \mathcal{C}_b^1(\mathbb{R}_+^k)$, $z_1, \dots, z_k \in \mathcal{M}$, $k = 1, \dots, n$, $n \geq 1$. Given $F \in \text{Dom}(D^\beta)$ and u a square-integrable $(\mathcal{F}_t^\beta)_{t \in \mathbb{R}_+}$ -adapted process we have the integration by parts formula

$$E \left[\int_0^T D_t^\beta F u_t dt \right] = E \left[F \int_0^T u_t (dN_t^\beta + q_{\beta_t, \beta_t} dt) \right].$$

Besides, another proof of integration by parts formula for birth processes is provided by Subsection 5.6, in this proof we mainly applies the basic techniques of fundamental calculus and achieve the same formula of integration by parts, in particular, we calculate the gradient by differentiating the joint density function of interjump times. Also, we apply the integration by parts formula to sensitivity analysis regarding the initial price in a regime switching model, and finally a similar approach as that in Chapter 4 extends the results for non-differentiable payoff functions.

Chapter 2

Optimal stopping in regime switching models

This chapter deals with optimal prediction in a regime switching model driven by a continuous-time Markov chain. We extend existing results for geometric Brownian motion by deriving an optimal stopping strategy depending on the current regime state. In particular we characterize the optimal stopping boundaries as solutions of a system of integral equations, and we address uniqueness of solutions under positive drifts, while immediate exercise is obtained under negative drifts. Our approach replaces the use of closed form expressions, which are not available in our setting, with PDE arguments which also simplify the approach of Du Toit and Peskir (2009) in the classical Brownian case.

In Lemma 2.1.1 we write the optimal value of the problem as a function of time, the regime state, and the relative maximum of the underlying asset. We also identify the optimal stopping time τ_D in Proposition 2.2.1, and in Proposition 2.2.2 we determine the structure of the optimal stopping set via its boundary functions $b_D(t, j)$ for j in the state space $\mathcal{M}$ of the

regime switching chain.

The free boundary problem is solved by constructing a Volterra integral equation when the drift parameters of the regime switching chain are all nonnegative. In this context, we prove the continuity and monotonicity of boundary functions $b_D(t, j)$ in Proposition 2.4.1, by extending arguments of Du Toit and Peskir (2009) to the regime switching setting. Under such conditions, in Proposition 2.4.2 we derive the Volterra integral equation (2.57) that characterizes the boundary function $b_D(t, j)$ of the stopping set. In Proposition 2.4.3 we show the uniqueness of solutions to the boundary integral equation (2.57) within a class of continuous functions on $[0, T]$ under the condition $\mu(j) \geq 0$ for all $j \in \mathcal{M}$. In Proposition 2.4.4 we show that immediate exercise is optimal when all drift parameters $\mu(j)$ are negative, $j \in \mathcal{M}$, while exercise at maturity becomes optimal when $\mu(j) \geq \sigma^2(j)$ for all $j \in \mathcal{M}$, where $\sigma(j)$ are the volatility parameters.

While most of analysis in Du Toit and Peskir (2009) relies on some propositions revealed by the closed form expressions which are not available in our setting, we develop PDE arguments, cf. e.g. Lemma 2.3.3, based on which we analyse in another way.

2.1 Problem formulation

Given a standard Brownian motion $(B_t)_{t \in \mathbb{R}_+}$ independent of the Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$, and the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ generated by $(B_t)_{t \in \mathbb{R}_+}$ and $(\beta_t)_{t \in \mathbb{R}_+}$, we consider the optimal prediction problem

$$V_t = \inf_{t \leq \tau \leq T} E \left[\sup_{0 \leq s \leq T} \frac{Y_s}{Y_\tau} \middle| \mathcal{F}_t \right], \quad (2.1)$$

over all $\mathcal{F}_t$ stopping times τ , where the asset price $(Y_t)_{t \in \mathbb{R}_+}$ is modeled by a geometric Brownian motion

$$dY_t = \mu(\beta_t)Y_t dt + \sigma(\beta_t)Y_t dB_t, \quad 0 \leq t \leq T, \quad (2.2)$$

with regime switching driven by a finite-state continuous-time Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ on $\mathcal{M} = \{1, 2, \dots, m\}$. Here, $\mu : \mathcal{M} \rightarrow \mathbb{R}$, and $\sigma : \mathcal{M} \rightarrow (0, \infty)$ are deterministic functions.

The next Lemma 2.1.1 shows that the optimal value function V_t in (2.1) can be written as a function of $(t, \beta_t, \hat{Y}_{0,t}/Y_t)$, where $\hat{Y}_{0,t}$ is defined by

$$\hat{Y}_{t,s} := \max_{t \leq r \leq s} Y_r, \quad 0 \leq t \leq s \leq T. \quad (2.3)$$

Lemma 2.1.1. The optimal value function V_t in (2.1) takes the form

$$V_t = V(t, \beta_t, \hat{Y}_{0,t}/Y_t),$$

where the function $V : [0, T] \times \mathcal{M} \times [1, +\infty) \rightarrow \mathbb{R}_+$ is given by

$$V(t, j, x) = \inf_{t \leq \tau \leq T} E \left[\frac{1}{Y_\tau} \max \left(xY_t, \hat{Y}_{t,T} \right) \mid \beta_t = j \right], \quad (2.4)$$

$0 \leq t \leq T$, $x \geq 1$, $j \in \mathcal{M}$.

Proof. Given $t \in [0, T]$, using the drifted Brownian motion

$$\hat{B}_u^t := B_{u+t} - B_t + \int_t^{t+u} \left(\frac{\mu(\beta_s)}{\sigma(\beta_s)} - \frac{\sigma(\beta_s)}{2} \right) ds, \quad u \in [0, T-t], \quad (2.5)$$

we rewrite the solution of (2.2) as

$$Y_s = Y_t \exp \left(\int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right), \quad s \in [t, T], \quad (2.6)$$

and define

$$\hat{S}_s^t := \sup_{0 \leq r \leq s} \int_0^r \sigma(\beta_{u+t}) d\hat{B}_u^t, \quad s \in [0, T-t]. \quad (2.7)$$

By the definition of $\hat{Y}_{t,s}$ in (2.3) and expression (2.6), and from the conditional independence of

$$\left((\hat{B}_{r-t}^t)_{r \geq t}, (\hat{S}_{r-t}^t)_{r \geq t} \right)$$

with $\mathcal{F}_t$ given β_t we have, for any stopping time τ with values in $[t, T]$, letting $a \vee b = \max(a, b)$,

$$\begin{aligned} E \left[\sup_{0 \leq s \leq T} \frac{Y_s}{Y_\tau} \middle| \mathcal{F}_t \right] &= E \left[\frac{\hat{Y}_{0,t}}{Y_\tau} \vee \frac{\hat{Y}_{t,T}}{Y_\tau} \middle| \mathcal{F}_t \right] \\ &= E \left[\left(\frac{\hat{Y}_{0,t}}{Y_t} e^{-\int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \right) \vee e^{\hat{S}_{T-t}^t - \int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \middle| \mathcal{F}_t \right] \\ &= E \left[\left(\frac{\hat{Y}_{0,t}}{Y_t} e^{-\int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \right) \vee e^{\hat{S}_{T-t}^t - \int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \middle| \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] \\ &= E \left[\frac{\hat{Y}_{0,t} \vee \hat{Y}_{t,T}}{Y_\tau} \middle| \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] \\ &= E \left[\frac{1}{Y_\tau} \max(Y_t \cdot \frac{\hat{Y}_{0,t}}{Y_t}, \hat{Y}_{t,T}) \middle| \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] \\ &= E \left[\frac{1}{Y_\tau} \max(xY_t, \hat{Y}_{t,T}) \middle| \beta_t \right]_{x=\hat{Y}_{0,t}/Y_t} \end{aligned}$$

where the last line follows from the conditional independence between $\hat{Y}_{0,t}/Y_t$ and

$$\left(\frac{Y_\tau}{Y_t}, \frac{\hat{Y}_{t,T}}{Y_\tau} \right) = \left(\exp \left(\int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right), \exp \left(\hat{S}_{T-t}^t - \int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \right)$$

given β_t . Therefore by definition (2.1) and expression (2.4), we obtain

$$\begin{aligned} V_t &= \inf_{t \leq \tau \leq T} E \left[\sup_{0 \leq s \leq T} \frac{Y_s}{Y_\tau} \middle| \mathcal{F}_t \right] \\ &= \inf_{t \leq \tau \leq T} E \left[\frac{1}{Y_\tau} \max(xY_t, \hat{Y}_{t,T}) \middle| \beta_t \right]_{x=\hat{Y}_{0,t}/Y_t} \\ &= V \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right). \end{aligned}$$

□

In the next lemma we rewrite the optimal stopping problem (2.1) in the standard form (2.9) below, using the function

$$G(t, i, x) := E \left[\max \left(x, \frac{\hat{Y}_{t,T}}{Y_t} \right) \mid \beta_t = i \right], \quad t \in [0, T], \quad i \in \mathcal{M}, \quad x \geq 1, \quad (2.8)$$

with $G(T, i, x) = x, x \geq 1$.

Lemma 2.1.2. The function $V : [0, T] \times \mathcal{M} \times [1, +\infty) \rightarrow \mathbf{R}_+$ defined by (2.1.1) admits the expression

$$V(t, j, x) = \inf_{t \leq \tau \leq T} E \left[G(\tau, \beta_\tau, X_\tau^{t,x}) \mid \beta_t = j \right], \quad (2.9)$$

for $t \in [0, T], j \in \mathcal{M}, x \geq 1$, where

$$X_r^{t,x} := \frac{1}{Y_r} \max \left(x Y_t, \hat{Y}_{t,r} \right), \quad r \in [t, T], \quad x \geq 1. \quad (2.10)$$

Proof. Similarly to the conditional independence argument in the proof of Lemma 2.1.1, for any $s \in [t, T]$, we find the independence between

$$\frac{\hat{Y}_{0,s}}{Y_s} \quad \text{and} \quad \frac{\hat{Y}_{s,T}}{Y_s}$$

given β_s , hence we have

$$\begin{aligned} E \left[\frac{\hat{Y}_{0,T}}{Y_s} \mid \mathcal{F}_s \right] &= E \left[\frac{\hat{Y}_{0,s} \vee \hat{Y}_{s,T}}{Y_s} \mid \mathcal{F}_s \right] \\ &= E \left[\frac{\hat{Y}_{0,s} \vee \hat{Y}_{s,T}}{Y_s} \mid \frac{\hat{Y}_{0,s}}{Y_s}, \beta_s \right] \\ &= E \left[y \vee \frac{\hat{Y}_{s,T}}{Y_s} \mid \beta_s \right]_{y=\hat{Y}_{0,s}/Y_s} \\ &= G \left(s, \beta_s, \frac{\hat{Y}_{0,s}}{Y_s} \right). \end{aligned} \quad (2.11)$$

Next we extend the above relation (2.11) to stopping times τ written as the limit of a decreasing sequence of discrete stopping times by checking

that for any discrete stopping time

$$\tau = \sum_{i=1}^n \mathbf{1}_{\{\tau=s_i\}} s_i, \quad s_1, \dots, s_n \in [t, T], \quad n \geq 1.$$

By (2.11) we have

$$\begin{aligned} E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \mathcal{F}_\tau \right] &= \sum_{i=1}^n E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mathbf{1}_{\{\tau=s_i\}} \mid \mathcal{F}_\tau \right] \\ &= \sum_{i=1}^n E \left[\frac{\hat{Y}_{0,T}}{Y_{s_i}} \mathbf{1}_{\{\tau=s_i\}} \mid \mathcal{F}_{s_i} \right] \\ &= \sum_{i=1}^n E \left[\frac{\hat{Y}_{0,T}}{Y_{s_i}} \mid \mathcal{F}_{s_i} \right] \mathbf{1}_{\{\tau=s_i\}} \\ &= \sum_{i=1}^n G \left(s_i, \beta_{s_i}, \frac{\hat{Y}_{0,s_i}}{Y_{s_i}} \right) \mathbf{1}_{\{\tau=s_i\}} \\ &= G \left(\tau, \beta_\tau, \frac{\hat{Y}_{0,\tau}}{Y_\tau} \right). \end{aligned}$$

Taking the conditional expectation

$$E \left[\cdot \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right]$$

on both sides of the above equality, we obtain

$$E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] = E \left[G \left(\tau, \beta_\tau, \frac{\hat{Y}_{0,\tau}}{Y_\tau} \right) \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right]. \quad (2.12)$$

By (2.10) and the conditional independence between

$$\frac{\hat{Y}_{0,t}}{Y_t} \quad \text{and} \quad \left(\frac{Y_t}{Y_\tau}, \frac{\hat{Y}_{t,\tau}}{Y_\tau} \right)$$

given $\beta_t = j$, we find

$$\begin{aligned} &E \left[\frac{1}{Y_\tau} \max \left(x Y_t, \hat{Y}_{t,T} \right) \mid \beta_t = j \right] \\ &= E \left[\frac{1}{Y_\tau} \max \left(\frac{\hat{Y}_{0,t}}{Y_t} \cdot Y_t, \hat{Y}_{t,T} \right) \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \end{aligned}$$

$$\begin{aligned}
&= E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \\
&= E \left[G \left(\tau, \beta_\tau, \frac{\hat{Y}_{0,\tau}}{Y_\tau} \right) \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \\
&= E \left[G \left(\tau, \beta_\tau, \frac{(xY_t) \vee \hat{Y}_{t,\tau}}{Y_\tau} \right) \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \\
&= E \left[G \left(\tau, \beta_\tau, X_\tau^{t,x} \right) \mid \beta_t = j \right], \tag{2.13}
\end{aligned}$$

which completes the proof by (2.4). $\square$

2.2 Stopping set and boundary functions

In this section we apply Corollary 2.9 in Peskir and Shiryaev (2006) in the framework of the regime switching model (2.2) in order to specify the stopping set and optimal stopping time associated to the optimal stopping problem (2.1), cf. Proposition 2.2.1 below. In order to deal with the existence of an optimal stopping time for (2.1) rewritten as (2.9), we define the set

$$D := \{(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : V(t, j, x) = G(t, j, x)\}. \tag{2.14}$$

From the relation

$$V(T, j, x) = G(T, j, x) = x, \quad j \in \mathcal{M}, x \geq 1,$$

we check that

$$\{T\} \times \mathcal{M} \times [1, \infty) \subset D,$$

which is consistent with the fact that the infimum in (2.1) is over stopping times $\tau \in [t, T]$.

Proposition 2.2.1. Let $t \in [0, T]$. Given $\beta_t = j \in \mathcal{M}$ and $\hat{Y}_{0,t}/Y_t = x \in [1, \infty)$, the stopping time

$$\tau_D(t, j, x) := \inf \left\{ r \geq t : \left(r, \beta_r, \frac{\hat{Y}_{0,r}}{Y_r} \right) \in D \right\} \quad (2.15)$$

is an optimal stopping time for (2.1), or equivalently for (2.9), provided it is a.s. finite.

Proof. By Corollary 2.9 in Peskir and Shiryaev (2006) the optimal stopping time for problem (2.9) exists and is equal to $\tau_D(t, j, x)$ in (2.15) provided we check that for all $t \in [0, T]$ we have:

- a) $G(t, j, x)$ is lower semicontinuous with respect to x , which follows directly from the definition (2.8) of $G(t, j, x)$.
- b) $V(t, j, x)$ is upper semicontinuous with respect to x , which follows from the continuity Lemma 2.3.6 below.
- c) We have $E \left[\sup_{t \leq s \leq T} |G(s, \beta_s, X_s^{t,x})| \right] < \infty$. Indeed, from (2.10) and (2.6), for $s \in [t, T]$ we have

$$\begin{aligned} X_s^{t,x} &= \frac{1}{Y_s} \max \left(x Y_t, \hat{Y}_{t,s} \right) \\ &= Y_t^{-1} e^{-\int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \max \left(x Y_t, Y_t e^{\hat{S}_{s-t}^t} \right) \\ &= \exp \left(\max \left(\log x, \hat{S}_{s-t}^t \right) - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right). \end{aligned} \quad (2.16)$$

Hence by (2.8) and the conditional independence between

$$X_s^{t,x} = \max \left(x \frac{Y_t}{Y_s}, \frac{\hat{Y}_{t,s}}{Y_s} \right) \quad s \in [t, T]$$

and $\hat{Y}_{s,T}/Y_s$ given β_s , we find that

$$G(s, \beta_s, X_s^{t,x}) = E \left[y \vee \frac{\hat{Y}_{s,T}}{Y_s} \mid \beta_s \right]_{y=X_s^{t,x}}$$

$$\begin{aligned}
&= E \left[X_s^{t,x} \vee \frac{\hat{Y}_{s,T}}{Y_s} \mid \beta_s, X_s^{t,x} \right] \\
&= E \left[\exp \left(\max \left(\log x, \hat{S}_{s-t}^t \right) - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \vee \frac{\hat{Y}_{s,T}}{Y_s} \mid \beta_s, X_s^{t,x} \right] \\
&= E \left[\exp \left(- \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \right. \\
&\quad \cdot \max \left(\exp \left(\max \left(\log x, \sup_{t \leq r \leq s} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \right), \right. \\
&\quad \left. \left. \exp \left(\sup_{s \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \right) \mid \beta_s, X_s^{t,x} \right] \\
&= E \left[\exp \left(\max \left(\log x, \hat{S}_{T-t}^t \right) - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \mid \beta_s, X_s^{t,x} \right].
\end{aligned} \tag{2.17}$$

Letting

$$\begin{aligned}
\check{S}_{T-t}^t &:= \inf_{t \leq s \leq T} \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \\
&= \inf_{t \leq s \leq T} \left[\int_0^{s-t} \sigma(\beta_{u+t}) dB_{t+u} + \int_0^{s-t} (\mu(\beta_{u+t}) - \sigma^2(\beta_{u+t})/2) du \right],
\end{aligned} \tag{2.18}$$

we conclude that

$$\begin{aligned}
&E \left[\sup_{t \leq s \leq T} |G(s, \beta_s, X_s^{t,x})| \right] \\
&= E \left[\sup_{t \leq s \leq T} E \left[e^{\max(\log x, \hat{S}_{T-t}^t) - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_s, X_s^{t,x} \right] \right] \\
&\leq x E \left[\sup_{t \leq s \leq T} E \left[e^{\hat{S}_{T-t}^t - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_s, X_s^{t,x} \right] \right] \\
&\leq x E \left[\sup_{t \leq s \leq T} E \left[e^{\hat{S}_{T-t}^t - \inf_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_s, X_s^{t,x} \right] \right] \\
&= x E \left[e^{\hat{S}_{T-t}^t - \inf_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \right] \\
&= x E \left[e^{\hat{S}_{T-t}^t - \check{S}_{T-t}^t} \right] \\
&\leq x E \left[e^{2\hat{S}_{T-t}^t} \right]^{1/2} E \left[e^{-2\check{S}_{T-t}^t} \right]^{1/2} \\
&\leq x E \left[e^{2\hat{S}_{T-t}^t} \right]^{1/2} E \left[e^{2\check{S}_{T-t}^t} \right]^{1/2} e^{\max_{i \in \mathcal{M}} |\sigma^2(i) - 2\mu(i)|(T-t)} \\
&\leq x E \left[e^{2\hat{S}_{T-t}^t} \right] e^{\max_{i \in \mathcal{M}} |\sigma^2(i) - 2\mu(i)|(T-t)}
\end{aligned}$$

$$< \infty, \quad (2.19)$$

where we applied the Cauchy–Schwarz inequality.

□

In Proposition 2.2.2 we characterize the shape of the stopping set D defined in (2.14) in terms of the boundary function $b_D(t, j)$ defined by

$$b_D(t, j) := \inf\{x \in [1, \infty) : (t, j, x) \in D\}. \quad (2.20)$$

From the relation $\{T\} \times \mathcal{M} \times [1, \infty) \subset D$ we deduce the terminal condition

$$b_D(T, j) = 1, \quad j \in \mathcal{M}.$$

Define

$$F(t, j, x) := V(t, j, x) - G(t, j, x) \leq 0, \quad (2.21)$$

which is negative by (2.9), $t \in [0, T]$, $j \in \mathcal{M}$, $x \geq 1$, so that we have

$$D = \{(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : F(t, j, x) = 0\}, \quad (2.22)$$

hence D is closed from the continuity of $(t, x) \mapsto F(t, j, x)$ is continuous on $[0, T] \times [1, \infty)$, cf. Lemma 2.3.6 below. The continuation set $C = D^c$ is an open set that can be written as

$$C = \{(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : F(t, j, x) < 0\}.$$

Proposition 2.2.2. For any $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$ such that $(t, j, x) \in D$ we have

$$\{t\} \times \{j\} \times [x, \infty) \subset D. \quad (2.23)$$

and

$$D = \{(t, j, y) \in [0, T] \times \mathcal{M} \times [1, \infty) : y \geq b_D(t, j)\}. \quad (2.24)$$

Proof. Let

$$y := \sup\{z \in [x, \infty) : \{t\} \times \{j\} \times [x, z] \subset D\}.$$

If $y < \infty$ then we have $(t, j, y) \in D$ by the closedness of D , and from the monotonicity property of $F(t, j, x)$ stated in Lemma 2.2.1, $(t, j, y) \in D$ admits a right neighborhood of the form

$$\{t\} \times \{j\} \times [x, x + \eta] \subset D$$

for some $\eta > 0$, which leads to a contradiction. Hence $y = +\infty$ and (2.23) holds. Relation (2.24) follows from the equivalence

$$(t, j, x) \in D \iff \{t\} \times \{j\} \times [x, \infty) \subset D \iff x \geq b_D(t, j)$$

that follows from (2.20). $\square$

The following lemma has been used in the proof of Proposition 2.2.2.

Lemma 2.2.1. For any $(t, j, x) \in D$, we have

$$\liminf_{\varepsilon \searrow 0} \frac{F(t, j, x + \varepsilon) - F(t, j, x)}{\varepsilon} \geq 0. \quad (2.25)$$

Proof. We split the proof into two parts.

(i) From (2.17) we have

$$\begin{aligned} G(s, \beta_s, X_s^{t,x}) &= E \left[X_s^{t,x} \vee \frac{\hat{Y}_{s,T}}{Y_s} \middle| \beta_s, X_s^{t,x} \right] = E \left[X_s^{t,x} \vee \frac{\hat{Y}_{s,T}}{Y_s} \middle| \mathcal{F}_s \right] \\ &= E \left[e^{\max(\log x, \hat{S}_{T-t}^t) - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \middle| \mathcal{F}_s \right], \quad s \in [t, T], \end{aligned}$$

which extends to any stopping time $\tau \in [t, T]$ as

$$G(\tau, \beta_\tau, X_\tau^{t,x}) = E \left[e^{\max(\log x, \hat{S}_{T-t}^t) - \int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \middle| \mathcal{F}_\tau \right], \quad (2.26)$$

as in (2.11)-(2.12) above. For all $x \geq 1$ and $\varepsilon > 0$, consider the stopping time

$$\tau_\varepsilon^+ := \tau_D(t, j, x + \varepsilon) \in [t, T]$$

defined by (2.15), which solves the optimal stopping problem

$$\begin{aligned} V(t, j, x + \varepsilon) &= \inf_{t \leq \tau \leq T} E \left[G(\tau, \beta_\tau, X_\tau^{t, x + \varepsilon}) \mid \beta_t = j \right] \\ &= E \left[G(\tau_\varepsilon^+, \beta_{\tau_\varepsilon^+}, X_{\tau_\varepsilon^+}^{t, x + \varepsilon}) \mid \beta_t = j \right], \end{aligned}$$

cf. (2.9). The following argument relies on the fact that for any $(t, j, x) \in D$ we have

$$\lim_{\varepsilon \rightarrow 0} \tau_D(t, j, x + \varepsilon) = t, \quad (2.27)$$

as will be shown in part (ii) below. Relations (2.8), (2.9), (2.26) and (2.27) imply

$$\begin{aligned} & \liminf_{\varepsilon \searrow 0} \frac{V(t, j, x + \varepsilon) - V(t, j, x)}{\varepsilon} \\ & \geq \liminf_{\varepsilon \searrow 0} \frac{1}{\varepsilon} E \left[G(\tau_\varepsilon^+, \beta_{\tau_\varepsilon^+}, X_{\tau_\varepsilon^+}^{t, x + \varepsilon}) - G(\tau_\varepsilon^+, \beta_{\tau_\varepsilon^+}, X_{\tau_\varepsilon^+}^{t, x}) \mid \beta_t = j \right] \\ & = \liminf_{\varepsilon \searrow 0} \frac{1}{\varepsilon} E \left[E \left[e^{\log(x + \varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_\varepsilon^+ - t} \sigma(\beta_{u+t}) d\hat{B}_u^t} - e^{\log x \vee \hat{S}_{T-t}^t - \int_0^{\tau_\varepsilon^+ - t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \right. \right. \\ & \quad \left. \left. \mid \mathcal{F}_{\tau_\varepsilon^+} \right] \mid \beta_t = j \right] \\ & = \liminf_{\varepsilon \searrow 0} \frac{1}{\varepsilon} E \left[e^{\log(x + \varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_\varepsilon^+ - t} \sigma(\beta_{u+t}) d\hat{B}_u^t} - e^{\log x \vee \hat{S}_{T-t}^t - \int_0^{\tau_\varepsilon^+ - t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_t = j \right] \\ & = \liminf_{\varepsilon \searrow 0} \frac{1}{\varepsilon} E \left[e^{\log(x + \varepsilon) \vee \hat{S}_{T-t}^t} - e^{\log x \vee \hat{S}_{T-t}^t} \mid \beta_t = j \right] \\ & = \frac{\partial}{\partial x} E \left[e^{\max(\log x, \hat{S}_{T-t}^t)} \mid \beta_t = j \right] \\ & = \frac{\partial G}{\partial x}(t, j, x), \end{aligned} \quad (2.28)$$

hence we conclude to (2.25). Here we used the dominated convergence theorem with the bound

$$\begin{aligned} & \frac{1}{\varepsilon} \left| e^{\log(x + \varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_\varepsilon^+ - t} \sigma(\beta_{u+t}) d\hat{B}_u^t} - e^{\log x \vee \hat{S}_{T-t}^t - \int_0^{\tau_\varepsilon^+ - t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \right| \\ & \leq \left| \frac{e^{\log(x + \varepsilon)} - e^{\log x}}{\varepsilon} \right| e^{-\inf_{0 \leq s \leq T-t} \int_0^s \sigma(\beta_{t+u}) d\hat{B}_u^t} \\ & = e^{-\hat{S}_{T-t}^t}, \end{aligned} \quad (2.29)$$

where $\check{S}_{T-t}^t$ is defined in (2.18) and the right-hand side is integrable as in the derivation of (2.19).

(ii) We turn to the proof of (2.27). From the expression (2.4) in Lemma 2.1.1, we have

$$\begin{aligned} V(t, j, x) &= \inf_{t \leq \tau \leq T} E \left[\frac{xY_t \vee \hat{Y}_{t,T}}{Y_\tau} \mid \beta_t = j, \mathcal{F}_t \right] \\ &= \inf_{t \leq \tau \leq T} E \left[e^{-\int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} (x \vee e^{\hat{S}_{T-t}^t}) \mid \beta_t = j, \mathcal{F}_t \right]. \end{aligned} \quad (2.30)$$

From (2.30) and

$$X_r^{t,x+\varepsilon} = e^{-\int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} (x + \varepsilon \vee e^{\hat{S}_{r-t}^t})$$

cf. (2.16), we obtain

$$\begin{aligned} V(r, \beta_r, X_r^{t,x+\varepsilon}) &= \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (y \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right]_{y=X_r^{t,x+\varepsilon}} \\ &= \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x+\varepsilon} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right]. \end{aligned} \quad (2.31)$$

Next, from the definition (2.15) of $\tau_D(t, j, x + \varepsilon)$ and (2.31) we have, on the event $\{\beta_t = j\}$,

$$\tau_D(t, j, x + \varepsilon) = \inf\{r \geq t : (r, \beta_r, X_r^{t,x+\varepsilon}) \in D\} \quad (2.32)$$

$$\begin{aligned} &= \inf \left\{ r \geq t : \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x+\varepsilon} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right] \right. \\ &\quad \left. = E \left[X_r^{t,x+\varepsilon} \vee e^{\hat{S}_{T-r}^r} \mid \mathcal{F}_r \right] \right\} \\ &\leq \inf \left\{ r \geq t : \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right] \right. \\ &\quad \left. \geq E \left[X_r^{t,x+\varepsilon} \vee e^{\hat{S}_{T-r}^r} \mid \mathcal{F}_r \right] \right\} \\ &\leq \inf \left\{ r \geq t : \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right] \right. \\ &\quad \left. \geq e^\varepsilon E \left[X_r^{t,x} \vee e^{\hat{S}_{T-r}^r} \mid \mathcal{F}_r \right] \right\}, \end{aligned} \quad (2.33)$$

where we applied the inequality

$$X_r^{t,x+\varepsilon} = e^{-\int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} (e^{\log(x+\varepsilon)} \vee e^{\hat{S}_{r-t}^t}) \leq e^{-\int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t + \varepsilon} (e^{\log(x)} \vee e^{\hat{S}_{r-t}^t})$$

$$= e^\varepsilon X_r^{t,x},$$

$x \geq 1$, $\varepsilon \geq 0$, $r \in [t, T]$. This implies

$$\lim_{\varepsilon \rightarrow 0} \tau_D(t, j, x + \varepsilon) \tag{2.34}$$

$$\begin{aligned} &\leq \liminf_{\varepsilon \rightarrow 0} \left\{ r \geq t : \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right] \right. \\ &\quad \left. \geq e^\varepsilon E \left[X_r^{t,x} \vee e^{\hat{S}_{T-r}^r} \mid \mathcal{F}_r \right] \right\} \\ &= \inf \left\{ r \geq t : \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right] \right. \\ &\quad \left. \geq E \left[X_r^{t,x} \vee e^{\hat{S}_{T-r}^r} \mid \mathcal{F}_r \right] \right\} \end{aligned} \tag{2.35}$$

$$\begin{aligned} &= \inf \left\{ r \geq t : \inf_{r \leq \tau \leq T} E \left[e^{-\int_0^{\tau-r} \sigma(\beta_{u+r}) d\hat{B}_u^r} (X_r^{t,x} \vee e^{\hat{S}_{T-r}^r}) \mid \mathcal{F}_r \right] \right. \\ &\quad \left. = E \left[X_r^{t,x} \vee e^{\hat{S}_{T-r}^r} \mid \mathcal{F}_r \right] \right\} \end{aligned} \tag{2.36}$$

$$\begin{aligned} &= \inf \{ r \geq t : (r, \beta_r, X_r^{t,x}) \in D \} \\ &= t, \end{aligned} \tag{2.37}$$

since $(t, j, x) \in D$, $\beta_t = j$ and $X_t^{t,x} = x$. Since $\tau_D(t, j, x + \varepsilon) \geq t$ we conclude to (2.27). $\square$

2.3 Continuity lemmas

The following property of smooth fit, namely the continuity of the function $\frac{\partial V}{\partial y}$ over the optimal stopping boundary ∂C , will be needed in the proof of Proposition 2.4.2 below.

Lemma 2.3.1. For any $(t, j, y) \in \partial C$, $y > 1$, we have

$$\frac{\partial V}{\partial y}(t, j, y+) = \frac{\partial V}{\partial y}(t, j, y-).$$

Proof. For any $\varepsilon \in (0, y - 1)$, let

$$\tau_\varepsilon^- = \tau_D(t, j, y - \varepsilon) \in [t, T],$$

cf. (2.15). Since $(t, j, y) \in \partial C$ and D is closed we have $(t, j, y) \in D$. Similarly to (2.32) to (2.34), τ_ε^- converges to t a.s. when ε tends to 0. By the same approach as in (2.28), replacing $y + \varepsilon$ with $y - \varepsilon$ shows that

$$\frac{\partial G}{\partial y}(t, j, y) \leq \liminf_{\varepsilon \searrow 0} \frac{V(t, j, y - \varepsilon) - V(t, j, y)}{\varepsilon}. \quad (2.38)$$

On the other hand, since $(t, j, y) \in \partial C \subset D$, we have

$$\limsup_{\varepsilon \searrow 0} \frac{V(t, j, y - \varepsilon) - V(t, j, y)}{\varepsilon} \leq \lim_{\varepsilon \searrow 0} \frac{G(t, j, y - \varepsilon) - G(t, j, y)}{\varepsilon} = \frac{\partial G}{\partial y}(t, j, y),$$

hence

$$\frac{\partial V}{\partial y}(t, j, y-) = \frac{\partial G}{\partial y}(t, j, y).$$

Finally the fact that $V = G$ on the closed set D implies

$$\frac{\partial V}{\partial y}(t, j, y-) = \frac{\partial V}{\partial y}(t, j, y+) = \frac{\partial G}{\partial y}(t, j, y).$$

□

In the next proposition, which will be used in the proof of Proposition 2.4.2, we show the normal reflection of the free boundary problem by proving that the right derivative of the value function $V(t, j, y)$ vanishes at $y = 1$, cf. also page 264 of Peskir and Shiryaev (2006) without regime switching.

Lemma 2.3.2. For any $t \in [0, T]$ and $j \in \mathcal{M}$ we have

$$\frac{\partial V}{\partial y}(t, j, 1+) = 0.$$

Proof. For convenience of notation we set $\tau_0 = \tau_D(t, j, 1)$, and note that

$$\begin{aligned} & \limsup_{\varepsilon \searrow 0} \frac{V(t, j, 1 + \varepsilon) - V(t, j, 1)}{\varepsilon} \\ & \leq \limsup_{\varepsilon \searrow 0} \frac{1}{\varepsilon} E[G(\tau_0, \beta_{\tau_0}, X_{\tau_0}^{t, 1+\varepsilon}) - G(\tau_0, \beta_{\tau_0}, X_{\tau_0}^{t, 1}) \mid \beta_t = j] \\ & = \limsup_{\varepsilon \searrow 0} \frac{1}{\varepsilon} E \left[\exp \left(\log(1 + \varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+r}) d\hat{B}_r^t \right) \right] \end{aligned}$$

$$\begin{aligned}
 & - \exp \left(\hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+r}) d\hat{B}_r^t \right) \Big| \beta_t = j \Big] \\
 = & \limsup_{\varepsilon \searrow 0} E \left[\frac{1}{\varepsilon} \left(\exp \left(\log(1+\varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+r}) d\hat{B}_r^t \right) \right. \right. \\
 & \left. \left. - \exp \left(\hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+r}) d\hat{B}_r^t \right) \right) \mathbf{1}_{\{\hat{S}_{T-t}^t < \log(1+\varepsilon)\}} \Big| \beta_t = j \right] \\
 = & E \left[\limsup_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \left(\exp \left(\log(1+\varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+r}) d\hat{B}_r^t \right) \right. \right. \\
 & \left. \left. - \exp \left(\hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+r}) d\hat{B}_r^t \right) \right) \mathbf{1}_{\{\hat{S}_{T-t}^t < \log(1+\varepsilon)\}} \Big| \beta_t = j \right] \\
 = & 0,
 \end{aligned}$$

since $\lim_{\varepsilon \searrow 0} \mathbf{1}_{\{\hat{S}_{T-t}^t < \log(1+\varepsilon)\}} = 0$, where we applied the dominated convergence theorem as in the proof of Lemma 2.2.1 and the following inequality.

$$\begin{aligned}
 & \frac{1}{\varepsilon} \left| e^{\log(1+\varepsilon) \vee \hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+u}) d\hat{B}_u^t} - e^{\hat{S}_{T-t}^t - \int_0^{\tau_0-t} \sigma(\beta_{t+u}) d\hat{B}_u^t} \right| \\
 & \leq \left| \frac{e^{\log(1+\varepsilon)} - 1}{\varepsilon} \right| e^{-\inf_{0 \leq s \leq T-t} \int_0^s \sigma(\beta_{t+u}) d\hat{B}_u^t} \\
 & = e^{-\check{S}_{T-t}^t}.
 \end{aligned}$$

where $\check{S}_{T-t}^t$ is defined in (2.18) and the right-hand side is integrable as in the derivation of (2.19). Since $V(t, j, y)$ is nondecreasing in $y \in [1, \infty)$, we have

$$\liminf_{\varepsilon \searrow 0} \frac{V(t, j, 1+\varepsilon) - V(t, j, 1)}{\varepsilon} \geq 0,$$

which shows that

$$\frac{\partial V}{\partial y}(t, j, 1+) = \lim_{\varepsilon \searrow 0} \frac{V(t, j, 1+\varepsilon) - V(t, j, 1)}{\varepsilon} = 0.$$

□

Next we consider the infinitesimal generator

$$\mathbb{L}f(s, j, x) = \left(\frac{\partial}{\partial s} + x(\sigma^2(j) - \mu(j)) \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2(j) x^2 \frac{\partial^2}{\partial x^2} \right) f(s, j, x)$$

$$+ \sum_{i=1}^m q_{j,i} f(s, i, x), \quad (2.39)$$

of the Markov process $(s, \beta_s, X_s^{t,x})_{s \in [t, T]}$, where $Q = (q_{ij})_{i,j=1,\dots,m}$ is the infinitesimal matrix generator of the Markov process $(\beta_t)_{t \in [0, T]}$, for any sufficiently differentiable function f of $(s, j, y) \in [0, T] \times \mathcal{M} \times [1, \infty)$, cf. Lemma 2.3.4 below.

The following lemmas will be used in the proof of Proposition 2.4.4 below. In Lemma 2.3.3 we replace the use of closed form expressions for $\mathbb{L}G(t, j, x)$, which are no longer available in our setting, with the differential expression (2.41).

Lemma 2.3.3. We have

$$\frac{\partial G}{\partial x}(t, j, 1+) = 0, \quad t \in [0, T], \quad (2.40)$$

and

$$\mathbb{L}G(t, j, x) = x\sigma^2(j) \frac{\partial G}{\partial x}(t, j, x) - \mu(j)G(t, j, x), \quad t \in [0, T], \quad (2.41)$$

with $\mathbb{L}G(T, j, x) = -\mu(j)x$, $j \in \mathcal{M}$, $x \in [1, \infty)$. In particular, for any $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$ we have

$$\begin{cases} \mathbb{L}G(t, j, x) > 0, & \text{when } \mu(j) \leq 0, \\ \mathbb{L}G(t, j, x) < 0, & \text{when } \mu(j) \geq \sigma^2(j). \end{cases} \quad (2.42)$$

In addition, $\mathbb{L}G(t, j, x)$ is nondecreasing and continuous in t for all $x \geq 1$ when $\mu(j) \geq 0$.

Proof. For all $j \in \mathcal{M}$ we let

$$f(t, j, x, y) := yG\left(t, j, \frac{y}{x}\right) \quad (2.43)$$

$$=E \left[\max \left(y, x \frac{\hat{Y}_{t,T}}{Y_t} \right) \mid \beta_t = j \right], \quad t \in [0, T], \quad x, y > 0.$$

By (2.2) and the Itô formula we have

$$\begin{aligned} & df(t, \beta_t, Y_t, \hat{Y}_{0,t}) \\ &= \frac{\partial f}{\partial t}(t, \beta_t, Y_t, \hat{Y}_{0,t})dt + \frac{\partial f}{\partial x}(t, \beta_t, Y_t, \hat{Y}_{0,t})dY_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, \beta_t, Y_t, \hat{Y}_{0,t})(dY_t)^2 \\ &\quad + \frac{\partial f}{\partial y}(t, \beta_t, Y_t, \hat{Y}_{0,t})d\hat{Y}_{0,t} + f(t, \beta_t, Y_t, \hat{Y}_{0,t}) - f(t, \beta_{t-}, Y_t, \hat{Y}_{0,t}) \\ &= \frac{\partial f}{\partial t}(t, \beta_t, Y_t, \hat{Y}_{0,t})dt + \frac{\partial f}{\partial x}(t, \beta_t, Y_t, \hat{Y}_{0,t})(\mu(\beta_t)Y_tdt + \sigma(\beta_t)Y_tdB_t) \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, \beta_t, Y_t, \hat{Y}_{0,t})(\sigma^2(\beta_t)Y_t^2dt) + \frac{\partial f}{\partial y}(t, \beta_t, Y_t, \hat{Y}_{0,t})d\hat{Y}_{0,t} \\ &\quad + f(t, \beta_t, Y_t, \hat{Y}_{0,t}) - f(t, \beta_{t-}, Y_t, \hat{Y}_{0,t}) \\ &= \frac{\partial f}{\partial t}(t, \beta_t, Y_t, \hat{Y}_{0,t})dt + \mu(\beta_t)Y_t \frac{\partial f}{\partial x}(t, \beta_t, Y_t, \hat{Y}_{0,t})dt \\ &\quad + \sigma(\beta_t)Y_t \frac{\partial f}{\partial x}(t, \beta_t, Y_t, \hat{Y}_{0,t})dB_t + \frac{1}{2} \sigma^2(\beta_t)Y_t^2 \frac{\partial^2 f}{\partial x^2}(t, \beta_t, Y_t, \hat{Y}_{0,t})dt \\ &\quad + \frac{\partial f}{\partial y}(t, \beta_t, Y_t, \hat{Y}_{0,t})d\hat{Y}_{0,t} + f(t, \beta_t, Y_t, \hat{Y}_{0,t}) - f(t, \beta_{t-}, Y_t, \hat{Y}_{0,t}), \quad (2.44) \end{aligned}$$

and given that

$$\begin{aligned} f(t, \beta_t, Y_t, \hat{Y}_{0,t}) &= E[\hat{Y}_{0,T} \mid \beta_t, Y_t, \hat{Y}_{0,t}] \\ &= E[\hat{Y}_{0,T} \mid \mathcal{F}_t], \quad t \in [0, T], \end{aligned}$$

is a martingale and $(\hat{Y}_{0,t})_{t \in [0, T]}$ has finite variation, we find

$$\begin{aligned} & \frac{\partial f}{\partial t}(t, j, y, z) + \mu(j)y \frac{\partial f}{\partial x}(t, j, y, z) + \frac{1}{2} \sigma^2(j)y^2 \frac{\partial^2 f}{\partial x^2}(t, j, y, z) \\ & + \sum_{i=1}^m q_{j,i} f(t, i, y, z) = 0, \end{aligned} \quad (2.45)$$

and

$$\frac{\partial f}{\partial y}(t, j, x, y)_{x=y} = 0.$$

Substituting (2.43) into (2.45) shows that

$$y \frac{\partial G}{\partial t} \left(t, j, \frac{z}{y} \right) + \mu(j)y \left(G \left(t, j, \frac{z}{y} \right) + y \frac{\partial G}{\partial x} \left(t, j, \frac{z}{y} \right) \left(-\frac{z}{y^2} \right) \right)$$

$$\begin{aligned}
& + \frac{1}{2}\sigma^2(j)y^2 \left(\frac{\partial G}{\partial x} \left(t, j, \frac{z}{y} \right) \left(-\frac{z}{y^2} \right) + \frac{z}{y^2} \frac{\partial G}{\partial x} \left(t, j, \frac{z}{y} \right) + \frac{z^2}{y^3} \frac{\partial^2 G}{\partial x^2} \left(t, j, \frac{z}{y} \right) \right) \\
& + \sum_{i=1}^m q_{j,i} y G \left(t, i, \frac{z}{y} \right) = 0,
\end{aligned} \tag{2.46}$$

which shows that the function $G(t, j, x)$ satisfies the PDE

$$\begin{aligned}
& \mu(j)G(t, j, x) + \frac{\partial G}{\partial t}(t, j, x) - \mu(j)x \frac{\partial G}{\partial x}(t, j, x) + \frac{1}{2}\sigma^2(j)x^2 \frac{\partial^2 G}{\partial x^2}(t, j, x) \\
& + \sum_{i=1}^m q_{j,i} G(t, i, x) = 0,
\end{aligned} \tag{2.47}$$

and we conclude to (2.41) by (2.39). Next we note that (2.40) follows from

$$\begin{aligned}
\frac{\partial G}{\partial x}(t, j, x) &= P \left(\frac{\hat{Y}_{t,T}}{Y_t} < x \mid \beta_t = j \right) \\
&\leq 1, \quad (t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty),
\end{aligned} \tag{2.48}$$

cf. the definition (2.8) of G . Next, by (2.8) and Lemma 2.3.3, for any $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$, we find

$$\begin{aligned}
\mathbb{L}G(t, j, x) &= x\sigma^2(j)P \left(\frac{\hat{Y}_{t,T}}{Y_t} < x \mid \beta_t = j \right) \\
&\quad - \mu(j)E \left[\max \left(x, \hat{Y}_{t,T}/Y_t \right) \mid \beta_t = j \right] \\
&= E \left[x\sigma^2(j)\mathbf{1}_{\{\hat{Y}_{t,T}/Y_t < x\}} - \mu(j) \left(x \vee \frac{\hat{Y}_{t,T}}{Y_t} \right) \mid \beta_t = j \right] \\
&= E \left[x(\sigma^2(j) - \mu(j))\mathbf{1}_{\{\hat{Y}_{t,T}/Y_t < x\}} \mid \beta_t = j \right] \\
&\quad - E \left[\mu(j) \frac{\hat{Y}_{t,T}}{Y_t} \mathbf{1}_{\{\hat{Y}_{t,T}/Y_t \geq x\}} \mid \beta_t = j \right],
\end{aligned} \tag{2.49}$$

which shows (2.42), and implies by (2.49) that $\mathbb{L}G(t, j, x)$ is nondecreasing and continuous in $t \in [0, T]$ when $\mu(j) \geq 0$. $\square$

The following lemma gives the expression of infinitesimal generator of the Markov process $(s, \beta_s, X_s^{t,x})_{s \in [t, T]}$.

Lemma 2.3.4. The Markov process $(s, \beta_s, X_s^{t,x})_{s \in [t, T]}$ has the infinitesimal generator

$$\mathbb{L}f(s, j, y) = \left(\frac{\partial}{\partial s} + y(\sigma^2(j) - \mu(j)) \frac{\partial}{\partial y} + \frac{1}{2} \sigma^2(j) y^2 \frac{\partial^2}{\partial y^2} \right) f(s, j, y) + \sum_{i=1}^m q_{j,i} f(s, i, y),$$

$s \in [0, T]$, $j \in \mathcal{M}$, $y \in [1, \infty)$, for $f \in \text{Dom}(\mathbb{L})$ satisfying

$$\frac{\partial f}{\partial y}(s, j, 1+) = 0.$$

Proof. Letting

$$Z_s^{t,x} := \log x \vee \hat{S}_{s-t}^t - \int_0^{s-t} \sigma(\beta_{u+t}) d\hat{B}_u^t, \quad (2.50)$$

$s \in [t, T]$, $x \geq 1$, from (2.16) we have $X_s^{t,x} = \exp(Z_s^{t,x})$, $s \in [t, T]$, $x \geq 1$.

Since $(\hat{S}_r^t)_{r \in [0, T-t]}$ is nondecreasing it has finite variation, hence

$$\begin{aligned} d\langle Z_r^{t,x}, Z_r^{t,x} \rangle &= \sigma^2(\beta_r) \langle d\hat{B}_{r-t}^t, d\hat{B}_{r-t}^t \rangle \\ &= \sigma^2(\beta_r) d\langle B_r, B_r \rangle \\ &= \sigma^2(\beta_r) dr, \end{aligned}$$

which shows that

$$\begin{aligned} dX_s^{t,x} &= X_s^{t,x} dZ_s^{t,x} + \frac{1}{2} X_s^{t,x} d\langle Z_s^{t,x}, Z_s^{t,x} \rangle \\ &= X_s^{t,x} dZ_s^{t,x} + \frac{1}{2} \sigma^2(\beta_s) X_s^{t,x} ds \\ &= X_s^{t,x} d(\log x \vee \hat{S}_{s-t}^t) - \sigma(\beta_s) X_s^{t,x} d\hat{B}_{s-t}^t + \frac{1}{2} \sigma^2(\beta_s) X_s^{t,x} ds. \end{aligned} \quad (2.51)$$

Given that $\frac{\partial f}{\partial y}(s, j, 1+) = 0$ for $(s, j, y) \in [0, T] \times \mathcal{M} \times [1, \infty)$, we have

$$\begin{aligned} &\frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) d(\log x \vee \hat{S}_{s-t}^t) \\ &= \frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) \mathbf{1}_{\{X_s^{t,x} > 1\}} d(\log x \vee \hat{S}_{s-t}^t) \\ &= \frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) \mathbf{1}_{\{Z_s^{t,x} > 0\}} d(\log x \vee \hat{S}_{s-t}^t) \\ &= 0, \end{aligned}$$

since $d(\log x \vee \hat{S}_{s-t}^t) = 0$ when $Z_s^{t,x} > 0$, $s \in [t, T]$. From (2.51) this shows that

$$\frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) dX_s^{t,x} = \frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) \left(-\sigma(\beta_s) X_s^{t,x} d\hat{B}_{s-t}^t + \frac{1}{2} \sigma^2(\beta_s) X_s^{t,x} dr \right).$$

Therefore by Itô's calculus and definition of $\mathbb{L}$, for any function $f \in \text{Dom}(\mathbb{L})$ and $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$ we have

$$\begin{aligned} \mathbb{L}f(t, j, x) &= \lim_{r \searrow t} \frac{E[f(r, \beta_r, X_r^{t,x}) | \beta_t = j, X_t^{t,x} = x] - f(t, j, x)}{r - t} \\ &= \lim_{r \searrow t} \frac{1}{r - t} E \left[\int_t^r \frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) dX_s^{t,x} + \int_t^r \frac{\partial f}{\partial s}(s, \beta_s, X_s^{t,x}) ds \right. \\ &\quad \left. + \frac{1}{2} \int_t^r \frac{\partial^2 f}{\partial y^2}(s, \beta_s, X_s^{t,x}) d\langle X_s^{t,x}, X_s^{t,x} \rangle + \int_t^r \sum_{i=1}^m q_{\beta_s, i} f(s, i, X_s^{t,x}) ds \middle| \beta_t = j \right] \\ &= \lim_{r \searrow t} \frac{1}{r - t} E \left[- \int_t^r \frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) \sigma(\beta_s) X_s^{t,x} d\hat{B}_{s-t}^t \right. \\ &\quad + \frac{1}{2} \int_t^r \frac{\partial f}{\partial y}(s, \beta_s, X_s^{t,x}) \sigma^2(\beta_s) X_s^{t,x} ds \\ &\quad + \int_t^r \frac{\partial f}{\partial s}(s, \beta_s, X_s^{t,x}) ds + \frac{1}{2} \int_t^r \frac{\partial^2 f}{\partial y^2}(s, \beta_s, X_s^{t,x}) \sigma^2(\beta_s) |X_s^{t,x}|^2 ds \\ &\quad \left. + \int_t^r \sum_{i=1}^m q_{\beta_s, i} f(s, i, X_s^{t,x}) ds \middle| \beta_t = j \right] \\ &= \frac{\partial f}{\partial t}(t, j, x) + x(\sigma^2(j) - \mu(j)) \frac{\partial f}{\partial x}(t, j, x) + \frac{1}{2} \sigma^2(j) x^2 \frac{\partial^2 f}{\partial x^2}(t, j, x) \\ &\quad + \sum_{i=1}^m q_{j, i} f(t, i, x). \end{aligned}$$

□

The proof of the next lemma, which will be used in Proposition 2.4.4 below, extends the argument of Du Toit and Peskir (2009) page 993.

Lemma 2.3.5. We have

$$\{(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : \mathbb{L}G(t, j, x) < 0\} \subset C,$$

where $C = D^c$ is the continuation set.

Proof. Assume that $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$ is such that $\mathbb{L}G(t, j, x) < 0$. By the continuity of $\mathbb{L}G(t, j, x)$ with respect to t , there exists an open neighbourhood $U \subset [0, T] \times [1, \infty)$ of (t, x) such that

$$\mathbb{L}G(s, j, y) < 0 \quad \text{for all } (s, y) \in U.$$

By Lemma 2.3.4 and Lemma 1 in Yao et al. (2006) we have

$$E[G(s, \beta_s, X_s^{t,x}) \mid \beta_t = j] = G(t, j, x) + E \left[\int_t^s \mathbb{L}G(r, \beta_r, X_r^{t,x}) dr \mid \beta_t = j \right], \quad (2.52)$$

$s \in [t, T]$. Replacing s in (2.52) with the first exit time τ_U of U when $(X_s^{t,x}, \beta_s)_{s \in [t, T]}$ is started at (x, j) at time t , (2.52) above shows by optional sampling that

$$E[G(\tau_U, \beta_{\tau_U}, X_{\tau_U}^{t,x}) \mid \beta_t = j] = G(t, j, x) + E \left[\int_t^{\tau_U} \mathbb{L}G(r, \beta_r, X_r^{t,x}) dr \mid \beta_t = j \right].$$

Since $\tau_U > t$ a.s. and $\mathbb{L}G(r, \beta_r, X_r^{t,x}) < 0$ when $r \in (t, \tau_U)$, the right hand side is strictly smaller than $G(t, j, x)$, while we have

$$E[G(\tau_U, \beta_{\tau_U}, X_{\tau_U}^{t,x}) \mid \beta_t = j] \geq V(t, j, x),$$

showing that $V(t, j, x) < G(t, j, x)$, which implies that $(t, j, x) \in C$. $\square$

Next we derive the following continuity result which extends the argument of Du Toit and Peskir (2009) page 995 to the regime switching setting, and has been used in the proof of Proposition 2.2.1.

Lemma 2.3.6. For any $j \in \mathcal{M}$ the function $x \mapsto V(t, j, x)$ is continuous on $[1, \infty)$, uniformly in $t \in [0, T]$. Furthermore we have, the function $(t, x) \mapsto F(t, j, x)$ is continuous on $[0, T] \times [1, \infty)$.

Proof. We proceed the proof by the following steps,

(i). We show that $x \mapsto V(t, j, x)$ is continuous on $[1, \infty)$, uniformly in $t \in [0, T]$. By Relation (2.48) and the mean value theorem, for all

$y \in [x, \infty)$ there exists a (random) $\eta \in [X_{t+\tau}^{t,x}, X_{t+\tau}^{t,y}]$ such that for any stopping time $\tau \in [0, T - t]$ we have

$$\begin{aligned} & G(t + \tau, \beta_{t+\tau}, X_{t+\tau}^{t,y}) - G(t + \tau, \beta_{t+\tau}, X_{t+\tau}^{t,x}) \\ &= \frac{\partial G}{\partial x}(t + \tau, \beta_{t+\tau}, \eta)(X_{t+\tau}^{t,y} - X_{t+\tau}^{t,x}) \\ &\leq (y - x) \frac{Y_t}{Y_{t+\tau}}, \end{aligned} \quad (2.53)$$

since $X_{t+\tau}^{t,y} - X_{t+\tau}^{t,x} \leq (y - x)Y_t/Y_{t+\tau}$ by (2.10). Let now $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$ and consider $\tau_x := \tau(t, j, x)$ given by (2.15). By Lemma 2.1.2 we have

$$\begin{aligned} & V(t, j, y) - V(t, j, x) \\ &\leq E \left[G(t + \tau_x, \beta_{t+\tau_x}, X_{t+\tau_x}^{t,y}) - G(t + \tau_x, \beta_{t+\tau_x}, X_{t+\tau_x}^{t,x}) \mid \beta_t = j \right]. \end{aligned} \quad (2.54)$$

Since $E[Y_t/Y_{t+\tau} \mid \beta_t = j]$ is uniformly bounded as in (2.19), taking expectation on both sides of (2.53) yields

$$\lim_{y \rightarrow x} E \left[G(t + \tau, \beta_{t+\tau}, X_{t+\tau}^{t,y}) - G(t + \tau, \beta_{t+\tau}, X_{t+\tau}^{t,x}) \mid \beta_t = j \right] = 0, \quad (2.55)$$

uniformly in $t \in [0, T]$ and in the stopping times $\tau \in [t, T]$. Since $V(t, j, x)$ is increasing in $x \in [1, \infty)$, (2.54) and (2.55) yield

$$0 \leq \lim_{y \rightarrow x} (V(t, j, y) - V(t, j, x)) \leq 0,$$

which shows the continuity of $x \mapsto V(t, j, x)$, uniformly in $t \in [0, T]$, for all $j \in \mathcal{M}$.

(ii). To prove the joint continuity, we have the following proposition: For any function $f(t, x) : [0, T] \times [1, \infty) \mapsto \mathbb{R}$ which is continuous in $t \in [0, T]$ and uniformly continuous in $x \in [1, \infty)$, we have the joint continuity of f in $(t, x) \in [0, T] \times [1, \infty)$. We prove it as follows. For any $\varepsilon > 0$, by the uniform continuity of $x \mapsto f(t, x)$, there exists an $\eta_1 > 0$ such that

$|f(t, y) - f(t, x)| < \varepsilon/2$ for any pair $x, y \in [1, \infty)$ that $|y - x| < \eta_1$. By the continuity of $t \mapsto f(t, x)$, for any $(t_0, x_0) \in [0, T] \times [1, \infty)$ and the same $\varepsilon > 0$ we have an $\eta_2 > 0$ such that $|V(t, x_0) - V(t_0, x_0)| < \varepsilon/2$ for any $t \in [0, T]$ that $|t - t_0| \leq \eta_2$.

Therefore for these given $\varepsilon > 0$, and $(t_0, x_0) \in [0, T] \times [1, \infty)$, there exists an $\eta_3 = \min(\eta_1, \eta_2)$ such that, for any $(t, x) \in [0, T] \times [1, \infty)$ that $\sqrt{(t - t_0)^2 + (x - x_0)^2} < \eta_3$ we have

$$|f(t, x) - f(t_0, x_0)| \leq |f(t, x) - V(t, x_0)| + |f(t, x_0) - f(t_0, x_0)| < \varepsilon,$$

which implies the joint continuity of $(t, x) \mapsto f(t, x)$.

(iii). Next we prove that $(t, x) \mapsto V(t, j, x)$ is continuous on $[0, T] \times [1, \infty)$.

By (2.9) we have,

$$V(t, j, x) = \inf_{t \leq \tau \leq T} \sum_{l=1}^m [e^{(\tau-t)Q}]_{j,l} E \left[G(\tau, \beta_\tau, X_\tau^{t,x}) \mid \beta_\tau = l \right],$$

where $Q = [q_{i,j}]_{1 \leq i,j \leq m}$ is the infinitesimal generator of $(\beta_t)_{t \in [0, T]}$. It follows that $t \mapsto V(t, j, x)$ is continuous on $[0, T]$, combining this fact with above (i) and (ii) we obtain the joint continuity of $(t, x) \mapsto V(t, j, x)$.

(iv) Finally we conclude the continuity of $(t, x) \mapsto F(t, j, x)$ on $[0, T] \times [1, \infty)$ by showing the continuity of $(t, x) \mapsto G(t, j, x)$. Given a $j \in \mathcal{M}$, by Relation (2.48) and the mean value theorem, for all $y \in [x, \infty)$ there exists an $\eta \in [x, y]$ such that for any $t \in [0, T]$ we have

$$0 \leq G(t, j, y) - G(t, j, x) = \frac{\partial G}{\partial x}(t, j, \eta)(y - x) \leq y - x,$$

which indicates the uniform continuity of $x \mapsto G(t, j, x)$ on $[1, \infty)$. On the other hand, by (2.8) we see that $t \mapsto G(t, j, x)$ is continuous on $[0, T]$.

Therefore, applying the proposition proved in (ii) above, we conclude to the continuity of $(t, x) \mapsto G(t, j, x)$ on $[0, T] \times [1, \infty)$, hence complete the proof for Lemma 2.3.6. $\square$

2.4 Solution of the free boundary problem

In this section we turn to the solution of the free boundary problem (2.4). We start by providing sufficient conditions on the drift coefficients $(\mu(j))_{j \in \mathcal{M}}$ for the boundary function $b_D(t, j)$ defined by (2.20) to be non-increasing and continuous in $t \in [0, T]$. The next proposition 2.4.1 relies on Lemma 2.4.1.

Proposition 2.4.1. Assume that $\mu(j) \geq 0$, $j \in \mathcal{M}$. Then the boundary function $b_D(t, j)$ defined by (2.20) is nonincreasing in $t \in [0, T]$ and continuous in $t \in [0, T]$, for all $j \in \mathcal{M}$.

Proof. (i) Monotonicity. Let $(t, j, x) \in D$ and $s \in [t, T]$. We have $F(t, j, x) = 0$ and $F(s, j, x) = 0$ since $F(t, j, x)$ is nondecreasing in t by Lemma 2.4.1, hence

$$[t, T] \times \{j\} \times \{x\} \subset D,$$

showing that

$$(t, j, x) \in D \iff [t, T] \times \{j\} \times \{x\} \subset D.$$

Then for any $s \in (t, T]$, we have $(s, j, b_D(t, j)) \in D$ since $(t, j, b_D(t, j)) \in D$. By Proposition 2.2.2 and noting that $(s, j, b_D(s, j)) \in D$, we conclude that $b_D(s, j) \leq b_D(t, j)$.

(ii) Right continuity. Given $(t, j, b_D(t, j)) \in D$, consider a strictly decreasing sequence $(t_n)_{n \geq 1}$ such that $\lim_{n \rightarrow \infty} t_n = t$. By part (i) above we

know that

$$b_D(t_n, j) \leq b_D(t, j), \quad n \geq 1,$$

and

$$\lim_{n \rightarrow \infty} b_D(t_n, j) \leq b_D(t, j).$$

Next, by Proposition 2.2.2 we have

$$[t, T] \times \{j\} \times [b_D(t, j), \infty) \subset D,$$

and since $(t_n, j, b_D(t_n, j)) \in D$, $n \geq 1$, and D is closed, we have

$$\left(t, j, \lim_{n \rightarrow \infty} b_D(t_n, j)\right) \in D,$$

hence $\lim_{n \rightarrow \infty} b_D(t_n, j) \geq b_D(t, j)$.

(iii) Left continuity. Using Lemma 2.3.5 we can repeat the argument of Du Toit and Peskir (2009) page 998 provided we show that the function $h(t, j)$ defined by

$$h(t, j) := \inf\{x \in [1, \infty) : \mathbb{L}G(t, j, y) \geq 0, \quad \forall y \in [x, \infty)\}, \quad (2.56)$$

is continuous in $t \in [0, T]$ for all $j \in \mathcal{M}$, with $h(T, j) = 1$. By Lemma 2.3.3 the function $\mathbb{L}G(t, j, x)$ is nondecreasing in t for all $x \geq 1$ since $(j) \geq 0$ and it follows from the definition (2.56) of $h(t, j)$ that $t \mapsto h(t, j)$ is nonincreasing in $t \in [0, T]$. For any $t_0 \in [0, T)$ and decreasing sequence $(t_n)_{n \geq 1} \subset (t_0, T]$ converging t_0 from the right-hand side we have

$$\lim_{n \rightarrow \infty} h(t_n, j) \leq h(t_0, j)$$

and

$$\lim_{n \rightarrow \infty} h(t_n, j) \geq h(t_k, j)$$

for any $k \geq 1$, hence

$$\lim_{n \rightarrow \infty} h(t_n, j) \geq h(t_0, j)$$

as by the continuity of $t \rightarrow \mathbb{L}G(t, j, x)$ we have

$$\mathbb{L}G(t_0, j, \lim_{n \rightarrow \infty} h(t_n, j)) = \lim_{k \rightarrow \infty} \mathbb{L}G(t_k, j, \lim_{n \rightarrow \infty} h(t_n, j)) \geq 0,$$

and this proves that $\lim_{t \searrow t_0} h(t, j) = h(t_0, j)$. On the other hand we have

$$h(t_0-, j) := \lim_{t \uparrow t_0} h(t, j) \geq h(t_0, j)$$

for any $t_0 \in [0, T]$, $j \in \mathcal{M}$. In case $h(t_0-, j) > h(t_0, j)$ we have $\mathbb{L}G(t_0, j, x) \geq 0$ for all $x \in [h(t_0, j), \infty]$. In addition, for any $t \in [0, t_0)$ and $x \in [h(t_0, j), h(t_0-, j))$ we have $\mathbb{L}G(t, j, x) < 0$ since $h(t, j) \geq h(t_0-, j)$, hence $\mathbb{L}G(t_0, j, x) = 0$ for all $x \in [h(t_0, j), h(t_0-, j))$ by the continuity of $t \mapsto \mathbb{L}G(t, j, x)$. By Lemma 2.3.3 we would have

$$x\sigma^2(j)\frac{\partial G}{\partial x}(t_0, j, x) = \mu(j)G(t_0, j, x), \quad x \in [h(t_0, j), h(t_0-, j)),$$

which shows that $G(t_0, j, x) = C(t_0, j)x^{\mu(j)/\sigma^2(j)}$, where $C(t_0, j)$ depends only on t_0 and $j \in \mathcal{M}$. This is a contradiction since

$$\begin{aligned} \frac{\partial}{\partial x}G(t_0, j, x) &= P\left(\hat{Y}_{t_0, T}/Y_{t_0} < x \mid \beta_{t_0} = j\right) \\ &= C(t_0, j)\frac{\mu(j)}{\sigma^2(j)}x^{-1+\mu(j)/\sigma^2(j)}, \end{aligned}$$

for $x \in [h(t_0, j), h(t_0-, j))$ cannot hold when $\mu(j) < \sigma^2(j)$, and more generally $\hat{Y}_{t_0, T}/Y_{t_0}$ cannot have a power law, even locally. $\square$

Next we derive the equation (2.57) satisfied by the function $b_D(t, \beta_t)$ defined in (2.20). This will yield the boundary curves

$$\left\{ (t, x) \in [0, T] \times [1, \infty) \mid x = b_D(t, j) \right\}$$

of the optimal stopping set D in (2.14), for any $j \in \mathcal{M}$.

Proposition 2.4.2. Assume that $\mu(j) \geq 0$, $j \in \mathcal{M}$. The boundary function $b_D(t, j)$ satisfies the Volterra equation

$$G(t, j, b_D(t, j)) = J(t, j, b_D(t, j)) - \int_t^T K(t, r, j, b_D(t, j))dr, \quad (2.57)$$

$0 \leq t \leq T$, with the boundary condition $b_D(T, j) = 1$, $j \in \mathcal{M}$, where

$$J(t, j, x) := E[X_T^{t,x} \mid \beta_t = j], \quad (2.58)$$

and

$$K(t, r, j, x) := E \left[\mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > b_D(r, \beta_r)\}} \mid \beta_t = j \right], \quad (2.59)$$

for $0 \leq t \leq r \leq T$ and $x \geq 1$.

Proof. Noting that $V(t, j, x) \leq G(t, j, x)$ for all $(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty)$ by (2.9), the continuation set $C := D^c$ is given by

$$C = D^c = \{(t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : V(t, j, x) < G(t, j, x)\}. \quad (2.60)$$

According to Proposition 2.2.1, for any $(t, j, x) \in C$, we have

$$V(t, j, x) = E \left[G(\tau_D, \beta_{\tau_D}, X_{\tau_D}^{t,x}) \mid \beta_t = j \right], \quad (2.61)$$

where $\tau_D = \tau_D(t, j, x)$ is defined by (2.15). Given that $\frac{\partial V}{\partial y}(t, j, 1+) = 0$ by Lemma 2.3.2, by the application of Peskir and Shiryaev (2006), Chapter III, § 7.1.1, § 7.4.1 as in Du Toit and Peskir (2009) § 3.5, page 996, the function V in (2.61) is $\mathcal{C}^{1,2}$ in the continuation set C in (2.60) and it solves the Cauchy-Dirichlet problem

$$\begin{cases} \mathbb{L}V(t, j, y) = 0, & (t, j, y) \in C, \\ V(t, j, y) = G(t, j, y), & (t, j, y) \in \partial C, \end{cases} \quad (2.62a)$$

$$(2.62b)$$

hence $\partial C \subset D$, where ∂C denotes the boundary of the open set C . By the local time change of variable formula of Peskir (2005), and by Lemma 2.3.4 below with the property $\frac{\partial V}{\partial y}(t, j, 1+) = 0$ shown in Lemma 2.3.2 above, we have

$$E[X_T^{t,x} \mid \beta_t = j] = E[V(T, \beta_T, X_T^{t,x}) \mid \beta_t = j]$$

$$\begin{aligned}
&= V(t, j, x) + E \left[\int_t^T \mathbb{L}V(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} \neq b_D(r, \beta_r)\}} dr \mid \beta_t = j \right] \\
&\quad + \frac{1}{2} E \left[\int_t^T \left(\frac{\partial V}{\partial y}(r, \beta_r, X_r^{t,x}+) - \frac{\partial V}{\partial y}(r, \beta_r, X_r^{t,x}-) \right) \right. \\
&\quad \left. \mathbf{1}_{\{X_r^{t,x} = b_D(r, \beta_r)\}} d\ell_r^b(X^{t,x}) \mid \beta_t = j \right], \tag{2.63}
\end{aligned}$$

where we applied the equality $V(T, \beta_T, X_T^{t,x}) = X_T^{t,x}$, and $(\ell_r^b(X^{t,x}))_{r \in [t, T]}$ denotes the local time of $X^{t,x}$ on the (piecewise continuous and nonincreasing by Proposition 2.4.1) curve $r \mapsto b_D(r, \beta_r)$. By the smooth fit property shown in Lemma 2.3.1 above, the last term in (2.63) vanishes. By Proposition 2.2.2 above and the definition (2.20) of $b_D(t, j)$, Relation (2.62a) can be rewritten as

$$\mathbb{L}V(r, j, y) \mathbf{1}_{\{y < b_D(r, j)\}} = 0, \quad r \in [0, T], \quad j \in \mathcal{M}, \quad y \geq 1, \tag{2.64}$$

which implies

$$\begin{aligned}
&E \left[\int_t^T \mathbb{L}V(r, \beta_r, X_r^{t,x}) dr \mid \beta_t = j \right] \\
&= E \left[\int_t^T \mathbb{L}V(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > b_D(r, \beta_r)\}} dr \mid \beta_t = j \right]. \tag{2.65}
\end{aligned}$$

Since $V(r, j, y) = G(r, j, y)$ on D and on the open set $\{y > b_D(r, j)\}$ which is contained in D by Proposition 2.2.2, from (2.65) we get

$$\begin{aligned}
&E \left[\int_t^T \mathbb{L}V(r, \beta_r, X_r^{t,x}) dr \mid \beta_t = j \right] \\
&= E \left[\int_t^T \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > b_D(r, \beta_r)\}} dr \mid \beta_t = j \right]. \tag{2.66}
\end{aligned}$$

Combining (2.63) and (2.66), we obtain

$$E[X_T^{t,x} \mid \beta_t = j] = V(t, j, x) + \int_t^T E \left[\mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} - b_D(r, \beta_r) > 0\}} \mid \beta_t = j \right] dr. \tag{2.67}$$

Hence substituting x with $b_D(t, j)$ in (2.67) above, we find that

$$G(t, j, b_D(t, j)) = V(t, j, b_D(t, j))$$

$$\begin{aligned}
 &= E[X_T^{t, b_D(t, j)} \mid \beta_t = j] - E \left[\int_t^T \mathbb{L}G(r, \beta_r, X_r^{t, b_D(t, j)}) \mathbf{1}_{\{X_r^{t, j} \geq b_D(r, \beta_r)\}} dr \mid \beta_t = j \right] \\
 &= J(t, j, b_D(t, j)) - \int_t^T K(t, r, j, b_D(t, j)) dr,
 \end{aligned}$$

where the functions J, K are defined by (2.58)-(2.59). $\square$

In the absence of regime switching, $(Y_t)_{t \in \mathbb{R}_+}$ is modeled by a geometric Brownian motion

$$dY_t = \mu Y_t dt + \sigma Y_t dB_t, \quad 0 \leq t \leq T.$$

Relation (2.20) is replaced by

$$b_D(t) = \inf\{x \in \mathbb{R}_+ : (t, x) \in D\}, \quad t \in [0, T],$$

and the boundary equation (2.57) becomes

$$G(t, b_D(t)) = E[X_T^{t, b_D(t)}] - E \left[\int_t^T \mathbb{L}G(r, X_r^{b_D(t)}) \mathbf{1}_{\{X_r^{t, b_D(t)} > b_D(r)\}} dr \right], \quad (2.68)$$

which recovers (3.50) in Du Toit and Peskir (2009), with

$$\mathbb{L}G(r, x) = \left(\frac{\partial}{\partial r} + x(\sigma^2 - \mu) \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} \right) G(r, x), \quad r \in [0, T], \quad x \in \mathbb{R}_+.$$

Similarly to (3.22)-(3.23) in Du Toit and Peskir (2009) we show that $F(t, j, x)$ defined by (2.21) is nondecreasing in $t \in [0, T]$ for all $j \in \mathcal{M}$ and $x \in [1, \infty)$, as in the following lemma which follows Du Toit and Peskir (2009) page 994. Note that without the condition $\mu(i) \geq 0$ for all $i \in \mathcal{M}$, the function $F(t, j, x)$ in Lemma 2.4.1 may not be nondecreasing in $t \in [0, T]$, in which case the equivalence

$$(t, j, x) \in D \iff [t, T] \times \{j\} \times \{x\} \subset D$$

in the proof of the next Proposition 2.4.1 does not hold and in this situation the boundary function $b_D(t, j)$ may not be decreasing in $t \in [0, T]$.

Lemma 2.4.1. Under the condition $\mu(j) \geq 0$ for all $j \in \mathcal{M}$, the function

$$t \mapsto F(t, j, x) = V(t, j, x) - G(t, j, x)$$

is nondecreasing in $t \in [0, T]$, for any $(j, x) \in \mathcal{M} \times [1, \infty)$.

Proof. By (2.52) we have

$$E[G(s, \beta_s, X_s^{t,x}) \mid \beta_t = j] = G(t, j, x) + E \left[\int_t^s \mathbb{L}G(r, \beta_r, X_r^{t,x}) dr \mid \beta_t = j \right], \quad (2.69)$$

$s \in [t, T]$. For any $r, s \in [0, T - t]$, and $r < s$, denote

$$\tau_s := \tau_D(s, j, x) - s \in [0, T - s]$$

by the definition (2.15) of τ_D . Replacing s with τ_s , t with r in the formula (2.69) and using optional sampling we have

$$\begin{aligned} F(r, j, x) &= V(r, j, x) - G(r, j, x) \\ &\leq E[G(r + \tau_s, \beta_{r+\tau_s}, X_{r+\tau_s}^{r,x}) \mid \beta_r = j] - G(r, j, x) \\ &= E \left[\int_r^{r+\tau_s} \mathbb{L}G(v, \beta_v, X_v^{r,x}) dv \mid \beta_r = j \right] \\ &= E \left[\int_0^{\tau_s} \mathbb{L}G(v + r, \beta_{v+r}, X_{v+r}^{r,x}) dv \mid \beta_r = j \right] \\ &= E \left[\int_0^{\tau_s} \mathbb{L}G(v + r, \beta_v, X_v^{0,x}) dv \mid \beta_0 = j \right]. \end{aligned} \quad (2.70)$$

Combining (2.70) with

$$\begin{aligned} F(s, j, x) &= V(s, j, x) - G(s, j, x) \\ &= E[G(s + \tau_s, \beta_{s+\tau_s}, X_{s+\tau_s}^{s,x}) \mid \beta_s = j] - G(s, j, x) \\ &= E \left[\int_0^{\tau_s} \mathbb{L}G(v + s, \beta_v, X_v^{0,x}) dv \mid \beta_0 = j \right], \end{aligned}$$

we have

$$F(s, j, x) - F(r, j, x) \quad (2.71)$$

$$\begin{aligned}
&\geq E \left[\int_0^{\tau_s} \mathbb{L}G(v+s, \beta_v, X_v^{0,x}) dv \mid \beta_0 = j \right] \\
&\quad - E \left[\int_0^{\tau_s} \mathbb{L}G(v+r, \beta_v, X_v^{0,x}) dv \mid \beta_0 = j \right] \\
&= E \left[\int_0^{\tau_s} \mathbb{L}G(v+s, \beta_v, X_v^{0,x}) - \mathbb{L}G(v+r, \beta_v, X_v^{0,x}) dv \mid \beta_0 = j \right].
\end{aligned}$$

Since $\mathbb{L}G(t, j, x)$ is nondecreasing in t when $\mu(j) \geq 0$ by (2.49), the right hand side of (2.71) is nonnegative, thereby $F(t, j, x)$ is nondecreasing in $t \in [0, T]$. $\square$

Uniqueness of boundary functions

We now turn to the uniqueness of the solution to equation (2.57).

Proposition 2.4.3. Assume that $\mu(j) \geq 0$ for all $j \in \mathcal{M}$. Then for all $j \in \mathcal{M}$, the boundary function $b_D(t, j)$ defined by (2.20) is the unique solution to (2.57) in the class of continuous functions $t \mapsto b(t, j)$ on $[0, T]$ satisfying $b(t, j) \geq h(t, j)$ for all t in $[0, T]$, where $h(t, j)$ is defined in (2.56).

Proof. The fact that the optimal stopping boundary $b_D(t, j)$ is nonincreasing and continuous in $t \in [0, T]$ is proved in Proposition 2.4.1. In order to show the uniqueness, based on the argument of Du Toit and Peskir (2009) it suffices to check the following.

- a) We have $b_D(t, j) \geq h(t, j)$ for all $t \in [0, T]$, $j \in \mathcal{M}$. Suppose there exists a pair $(t, j) \in [0, T] \times \mathcal{M}$ that $b_D(t, j) < h(t, j)$, by the definition (2.56) of h as follows,

$$h(t, j) := \inf\{x \in [1, \infty) : \mathbb{L}G(t, j, y) \geq 0, \quad \forall y \in [x, \infty)\},$$

we have $\mathbb{L}G(t, j, b_D(t, j)) < 0$, by Lemma 2.3.5 we have $(t, j, b_D(t, j)) \in C$, which conflicts with the closedness of the stopping set D , hence $b_D(t, j) \geq h(t, j)$ for all $t \in [0, T]$, $j \in \mathcal{M}$.

- b) For any continuous function $(c(r))_{r \in [0, T]}$ the process $U_{t,x}(s)$ defined for $t \in [0, T]$ and $x \geq 1$ as

$$U_{t,x}(s) := U^c(s, \beta_s, X_s^{t,x}) - \int_t^s \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr, \quad (2.72)$$

is a martingale in $s \in [t, T]$, where we let

$$\begin{aligned} U^c(s, j, y) &:= E[G(T, \beta_T, X_T^{s,y}) \mid \beta_s = j] \\ &\quad - E \left[\int_s^T \mathbb{L}G(r, \beta_r, X_r^{s,y}) \mathbf{1}_{\{X_r^{s,y} > c(r)\}} dr \mid \beta_s = j \right], \quad (2.73) \end{aligned}$$

$s \in [0, T]$, $j \in \mathcal{M}$, $y \in [1, \infty)$. Next, by (2.73) and the Markov property of $(X_r^{t,x}, \beta_r)_{r \in [t, T]}$ we have

$$\begin{aligned} U_{t,x}(t+s) &= U^c(t+s, \beta_{t+s}, X_{t+s}^{t,x}) - \int_t^{t+s} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \\ &= E[G(T, \beta_T, X_T^{t+s,y}) \mid \beta_{t+s}, X_{t+s}^{t,x} = y] \\ &\quad - E \left[\int_{t+s}^T \mathbb{L}G(r, \beta_r, X_r^{t+s,y}) \mathbf{1}_{\{X_r^{t+s,y} > c(r)\}} dr \mid \beta_{t+s}, X_{t+s}^{t,x} = y \right] \\ &\quad - \int_t^{t+s} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr, \end{aligned}$$

hence it follows that, for $v \in (s, T - t]$,

$$\begin{aligned} &E[U^c(t+v) - U^c(t+s) \mid \mathcal{F}_{t+s}] \\ &= E[E[G(T, \beta_T, X_T^{t+v,y}) \mid \beta_{t+v}, X_{t+v}^{t,x} = y] - E[G(T, \beta_T, X_T^{t+s,y}) \\ &\quad \mid \beta_{t+s}, X_{t+s}^{t,x} = y] \mid \mathcal{F}_{t+s}] \\ &\quad - E \left[E \left[\int_{t+v}^T \mathbb{L}G(r, \beta_r, X_r^{t+v,y}) \mathbf{1}_{\{X_r^{t+v,y} > c(r)\}} dr \mid \beta_{t+v}, X_{t+v}^{t,x} = y \right] \mid \mathcal{F}_{t+s} \right] \\ &\quad - E \left[\int_t^{t+v} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\ &\quad + E \left[E \left[\int_{t+s}^T \mathbb{L}G(r, \beta_r, X_r^{t+s,y}) \mathbf{1}_{\{X_r^{t+s,y} > c(r)\}} dr \mid \beta_{t+s}, X_{t+s}^{t,x} = y \right] \mid \mathcal{F}_{t+s} \right] \\ &\quad + E \left[\int_t^{t+s} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\ &= E \left[G(T, \beta_T, X_T^{t+v, X_{t+v}^{t,x}}) - G(T, \beta_T, X_T^{t+s, X_{t+s}^{t,x}}) \mid \mathcal{F}_{t+s} \right] \end{aligned}$$

$$\begin{aligned}
& -E \left[\int_{t+v}^T \mathbb{L}G(r, \beta_r, X_r^{t+v, X_{t+v}^{t,x}}) \mathbf{1}_{\{X_r^{t+v, X_{t+v}^{t,x}} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& -E \left[\int_t^{t+v} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& +E \left[\int_{t+s}^T \mathbb{L}G(r, \beta_r, X_r^{t+s, X_{t+s}^{t,x}}) \mathbf{1}_{\{X_r^{t+s, X_{t+s}^{t,x}} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& +E \left[\int_t^{t+s} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& = E \left[G(T, \beta_T, X_T^{t,x}) - G(T, \beta_T, X_T^{t,x}) \mid \mathcal{F}_{t+s} \right] \\
& -E \left[\int_{t+v}^T \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& -E \left[\int_t^{t+v} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& +E \left[\int_{t+s}^T \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& +E \left[\int_t^{t+s} \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& = -E \left[\int_t^T \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& +E \left[\int_t^T \mathbb{L}G(r, \beta_r, X_r^{t,x}) \mathbf{1}_{\{X_r^{t,x} > c(r)\}} dr \mid \mathcal{F}_{t+s} \right] \\
& = 0.
\end{aligned}$$

- c) For any fixed $j \in \mathcal{M}$, the boundary function $b_D(t, j)$ is right-continuous in $t \in [0, T]$, as follows from Proposition 2.4.1.

□

Particular exercise strategies

Next we show that the stopping set D has a simple form in two special situations.

Proposition 2.4.4. We have the following special cases of optimal stopping sets D .

i) Immediate exercise. Under the condition

$$\mu(j) \leq 0, \quad j \in \mathcal{M},$$

we have

$$D = [0, T] \times \mathcal{M} \times [1, \infty).$$

ii) Exercise at maturity. Under the condition

$$\mu(j) \geq \sigma^2(j), \quad j \in \mathcal{M},$$

we have

$$D = \{T\} \times \mathcal{M} \times [1, \infty).$$

Proof. Replacing s in (2.52) with τ_D defined in (2.15) and using optional sampling, we find

$$V(t, i, x) = G(t, i, x) + E \left[\int_t^{\tau_D(t, i, x)} \mathbb{L}G(r, \beta_r, X_r^{t, x}) dr \right], \quad t \in [0, T]. \quad (2.74)$$

i) In case $\mu(j) \leq 0$ for all $j \in \mathcal{M}$, by Lemma 2.3.3, we have $\mathbb{L}G(t, i, x) > 0$ for all $(t, i, x) \in [0, T) \times \mathcal{M} \times [1, \infty)$, hence (2.74) implies $\tau_D(t, i, x) = 0$ a.s., otherwise it contradicts the fact that $V(t, i, x) \leq G(t, i, x)$ because of (2.74). This implies $[0, T] \times \mathcal{M} \times [1, \infty) \subset D$.

ii) In case $\mu(j) \geq \sigma^2(j)$ for all $j \in \mathcal{M}$, by Lemma 2.3.3 we have $\mathbb{L}G(t, i, x) < 0$ for all $(t, i, x) \in [0, T) \times \mathcal{M} \times [1, \infty)$, and applying Lemma 2.3.5, we see that $[0, T) \times \mathcal{M} \times [1, \infty) \subset C$, which means $D = \{T\} \times \mathcal{M} \times [1, \infty)$.

□

Numerical example

We present a numerical implementation of (2.57) in the special case of a two-state Markov chain $(\beta_t)_{t \in [0, T]}$ with $\mathcal{M} = \{1, 2\}$, $\beta_0 = 1$, $\mu(1) = 0.02$, $\mu(2) = 0.03$, $\sigma(1) = 0.25$, $\sigma(2) = 0.2$, and infinitesimal generator

$$\mathbf{Q} := \begin{bmatrix} -0.15 & 0.15 \\ 0.2 & -0.2 \end{bmatrix},$$

with $Y_0 = 100, T = 10$. Based on simulations of the path of Y_t we estimate $J(t, j, x)$ in (2.58) and $K(t, r, j, x)$ in (2.59) by Monte Carlo and we solve the equation (2.57) for $b_D(t, j)$ by dichotomy. The optimal stopping time $\tau_D(t, \beta_t, \hat{Y}_{0,t}/Y_t)$ for selling is found as soon as the condition $\hat{Y}_{0,s}/Y_s \geq b_D(s, \beta_s)$ is satisfied for some $s \in [t, T]$. The next Figure 2.1 shows the boundary curves for each state of $(\beta_t)_{t \in [0, T]}$.

Figure 2.1: Boundary curves for different regime states.

Chapter 3

A recursive algorithm for optimal selling

We propose a recursive algorithm for the numerical computation of exercise prices in a stochastic optimal stopping problem with regime switching. This method allows us to determine the boundary functions of the optimal stopping set without having to solve their associated Volterra integral equations. It applies in particular when regime switching drifts have mixed signs, in which case the boundary functions may not be monotone and the Volterra equations are not available.

Our algorithm applies a discretization of the time interval $[0, T]$ whereas the recursion of Yao et al. (2006) is based on the jump times of the Markov chain $(\beta_t)_{t \in \mathbf{R}_+}$, besides solving the optimal stopping problem, our method is able to compute a large class of expectation of Markovian process as the recursion of Yao et al. (2006) does. As this approach does not rely on the Volterra equation, it allows us in particular to determine the boundary function $b_D(t, j)$ without requiring the condition $\mu(j) \geq 0$ for all $j \in \mathcal{M}$ as in Liu and Privault (2015b). In comparison with Du Toit

and Peskir (2009) our method does not rely on closed form expressions of value functions as in Du Toit and Peskir (2009) as they are no longer available in the regime switching setting. In addition we do not require Monte Carlo simulations in the absence of regime switching, since in this case (1.4) can be computed in deterministic form, cf. (2.7) in Du Toit and Peskir (2009). This Chapter is written according to our research paper Liu and Privault (2015a).

3.1 Problem formulation and solutions

In this chapter we continue to study the optimal stopping problem

$$V_t = \inf_{t \leq \tau \leq T} E \left[\sup_{0 \leq s \leq T} \frac{Y_s}{Y_\tau} \middle| \mathcal{F}_t \right] \quad (3.1)$$

of selling a stock at the ultimate maximum over all stopping times $\tau \in [t, T]$ when the asset price $(Y_t)_{t \in \mathbf{R}_+}$ is modeled by the regime switching model

$$dY_t = \mu(\beta_t)Y_t dt + \sigma(\beta_t)Y_t dB_t, \quad 0 \leq t \leq T, \quad (3.2)$$

driven by a finite-state continuous-time Markov chain $(\beta_t)_{t \in \mathbf{R}_+}$ with state space $\mathcal{M} := \{1, 2, \dots, m\}$ independent of the standard Brownian motion $(B_t)_{t \in \mathbf{R}_+}$ on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t \in \mathbf{R}_+}, \mathbb{P})$, where $(\mathcal{F}_t)_{t \in \mathbf{R}_+}$ is the filtration generated by $(B_t)_{t \in \mathbf{R}_+}$ and $(\beta_t)_{t \in \mathbf{R}_+}$ and $\mu : \mathcal{M} \rightarrow \mathbf{R}$, and $\sigma : \mathcal{M} \rightarrow (0, \infty)$ are deterministic functions. According to Chapter 2, under regime switching, the optimal boundary functions depend on the regime state of the system, and they may not be monotone if the drift coefficients have switching signs.

To solve the optimal stopping problem, we propose a recursive numerical algorithm of the problem (3.1) in the regime switching model (3.2). We

start by recalling the main notations and results needed. From Lemma 2.1 in Chapter 2, the optimal value function V_t in (3.1) can be written as

$$V_t = V(t, \beta_t, \hat{Y}_{0,t}/Y_t),$$

where the function $V : [0, T] \times \mathcal{M} \times [1, \infty) \rightarrow \mathbf{R}_+$ is given by

$$V(t, j, a) := \inf_{t \leq \tau \leq T} E \left[\frac{1}{Y_\tau} \max(aY_t, \hat{Y}_{t,T}) \mid \beta_t = j \right] \quad (3.3)$$

$$= \inf_{t \leq \tau \leq T} E \left[G \left(\tau, \beta_\tau, \frac{1}{Y_\tau} \max(aY_t, \hat{Y}_{t,\tau}) \right) \mid \beta_t = j \right] \quad (3.4)$$

for $t \in [0, T]$, $j \in \mathcal{M}$, $a \geq 1$, with $\hat{Y}_{s,t} := \max_{r \in [s,t]} Y_r$, $0 \leq s \leq t \leq T$, and

$$G(t, j, a) := E \left[\max(a, \hat{Y}_{t,T}/Y_t) \mid \beta_t = j \right], \quad t \in [0, T], \quad j \in \mathcal{M}. \quad (3.5)$$

From Proposition 3.1 in Chapter 2, given $\beta_t = j \in \mathcal{M}$ and $\hat{Y}_{0,t}/Y_t = a \in [1, \infty)$, $t \in [0, T]$, the optimal stopping time for (3.1), or equivalently for (3.4), is the first hitting time

$$\tau_D(t, j, a) := \inf \left\{ r \geq t : \left(r, \beta_r, \frac{\hat{Y}_{0,r}}{Y_r} \right) \in D \right\}$$

of the stopping set

$$D := \{(t, j, a) \in [0, T] \times \mathcal{M} \times [1, \infty) : V(t, j, a) = G(t, j, a)\} \quad (3.6)$$

by the process $(r, \beta_r, \hat{Y}_{0,r}/Y_r)_{r \in [t, T]}$.

The stopping set D defined in (3.6) is closed and its shape can be characterized in terms of the boundary functions $b_D(t, j)$ defined by

$$b_D(t, j) := \inf \{x \in [1, \infty) : (t, j, x) \in D\}, \quad t \in [0, T], \quad j \in \mathcal{M}.$$

Namely for any $(t, j, a) \in D$, by Proposition 3.2 of Chapter 2 we have $\{t\} \times \{j\} \times [a, \infty) \subset D$ and

$$D = \{(t, j, y) \in [0, T] \times \mathcal{M} \times [1, \infty) : y \geq b_D(t, j)\}.$$

In addition, by Proposition 5.1 of Chapter 2, in case $\mu(j) \geq 0$ for all $j \in \mathcal{M}$, the boundary function $b_D(t, j)$ is shown to be nonincreasing and continuous in $t \in [0, T]$, for all $j \in \mathcal{M}$, and to satisfy the Volterra integral equation (2.57).

If the condition $\mu(j) \geq 0$ is not satisfied for all $j \in \mathcal{M}$, then $t \mapsto b_D(t, j)$ may not be decreasing, cf. Figure 3.3 below, and in this case the Volterra equation (2.57) is not available because its proof relies on the Itô formula with local time of Peskir (2005), which requires $t \mapsto b_D(t, j)$ to be monotone.

Therefore we construct a recursive algorithm for the numerical solution of the optimal stopping problem (3.1), by determining the stopping set D from the values of $V(t, j, a)$ and $G(t, j, a)$, cf. Theorems 3.1.1 and 3.1.2 below. As this approach does not rely on the Volterra equation (2.57), it allows us in particular to determine the boundary function $b_D(t, j)$ without requiring the condition $\mu(j) \geq 0$ for all $j \in \mathcal{M}$ as in Chapter 2, cf. for example Figure 3.3 below.

In the sequel we let $\delta_n = T/n$, $\mathcal{T}_n = \{k\delta_n : k = 0, 1, \dots, n\}$, and

$$\lceil s \rceil_n := \min(t \in \mathcal{T}_n : t \geq s), \quad s \in [0, T], \quad n \geq 1. \quad (3.7)$$

In the following Theorem 3.1.1, which is proved in Section 3.2, the function $V_n(t, j, a)$ is computed by the backward induction (3.10) starting from the terminal time T .

Theorem 3.1.1. (i) For all $t \in [0, T]$, $j = 1, \dots, m$ and $a \geq 1$ we have

$$V(t, j, a) = \lim_{n \rightarrow \infty} V_n(t, j, a), \quad (3.8)$$

where $V_n(t, j, a)$ is the discrete infimum

$$V_n(t, j, a) := \inf_{[t]_n + \delta_n \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_{[t]_n} = j, \frac{\hat{Y}_{0,[t]_n}}{Y_{[t]_n}} = a \right], \quad (3.9)$$

taken over all $\mathcal{T}_n$ -valued stopping times τ_n , and $V_n(t, j, a) := V(T, j, a) = a$, $t \in (T - \delta_n, T]$.

(ii) The value of $V_n(t, j, a)$ in (3.9) can be computed by the backward induction

$$V_n(t - \delta_n, j, a) = E \left[G \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \wedge V_n \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \mid \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right], \quad (3.10)$$

$t \in [\delta_n, T]$, under the terminal condition $V_n(T, j, a) = G(T, j, a) = a$, $a \geq 1$, where $G(t, j, a)$ is defined in (3.5).

In Theorem 3.1.2 below, which is proved in Section 3.3, we provide a way to estimate the function $G(t, j, a)$ used in (3.10) as

$$G(t, j, a) = \Psi(t, t, j, a), \quad 0 \leq t \leq T, \quad a \geq 1, \quad j \in \mathcal{M}, \quad (3.11)$$

where for $0 \leq t \leq h \leq T$, $a \geq 1$, $j \in \mathcal{M}$,

$$\Psi(t, h, j, a) := E \left[\frac{\hat{Y}_{0,T}}{Y_h} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right], \quad (3.12)$$

can be approximated by a limiting sequence $\Psi_n(t, h, j, a)$ given by the backward induction (3.16) below. In the sequel we denote by

$$\varphi_r(x, y) := \sqrt{\frac{2}{\pi}} \frac{2y - x}{r^{3/2}} e^{-(2y-x)^2/2r}, \quad 0 \leq x \leq y, \quad r \in (0, T], \quad (3.13)$$

the joint probability density function of $\left(B_r, \sup_{0 \leq s \leq r} B_s \right)$, by $Q = [q_{i,j}]_{1 \leq i,j \leq m}$ the infinitesimal generator of $(\beta_t)_{t \in [0, T]}$, and we let

$$u(j) := \mu(j)/\sigma(j) - \sigma(j)/2, \quad j \in \mathcal{M}, \quad (3.14)$$

$0 \leq t \leq h \leq T$, $j \in \mathcal{M}$ and $a \geq 1$.

Theorem 3.1.2. For any $0 \leq t \leq h \leq T$, $j \in \mathcal{M}$ and $a \geq 1$, we have

$$\Psi(t, h, j, a) = \lim_{n \rightarrow \infty} \Psi_n(t, h, j, a), \quad (3.15)$$

where $\Psi_n(t, h, j, a)$ satisfies the backward induction

$$\begin{aligned} \Psi_n(t, h, j, a) := & e^{q_{j,j}\delta_n} \int_0^\infty dy \int_{-\infty}^y e^{(u(j)+\mathbf{1}_{\{t=h\}}\sigma(j))x-u^2(j)T/(2n)} \\ & \times \Psi_n(t + \delta_n, h + \mathbf{1}_{\{t=h\}}\delta_n, j, a \vee (\sigma(j)y) - \sigma(j)x) \varphi_{\delta_n}(x, y) dx \\ & + \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} dr \int_0^\infty dy \int_{-\infty}^y e^{(u(j)+\mathbf{1}_{\{t=h\}}\sigma(j))x-u^2(j)r/2} \\ & \times \Psi_n(t + \delta_n, h + \mathbf{1}_{\{t=h\}}\delta_n, l, a \vee (\sigma(j)y) - \sigma(j)x) e^{q_{j,j}r} \varphi_r(x, y) dx, \end{aligned} \quad (3.16)$$

$t \leq T - \delta_n$, $h \geq t$, $j \in \mathcal{M}$ and $a \geq 1$, with the terminal condition $\Psi_n(t, h, j, a) = a$, $n \geq 1$ for all t, h that $T - \delta_n < t \leq h \leq T$.

Theorem 3.1.2 can also be used for the minimization problem

$$U_t = \inf_{h \in [t, T]} E \left[\sup_{0 \leq s \leq T} \frac{Y_s}{Y_h} \middle| \mathcal{F}_t \right] = \inf_{h \in [t, T]} \Psi \left(t, h, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right), \quad (3.17)$$

in deterministic time $h \in [t, T]$ by providing a way to compute the function $\Psi(t, h, j, a)$ appearing in the right hand side of (3.17).

Theorems 3.1.1 and 3.1.2 also provide an alternative numerical solution in the geometric Brownian motion model of Du Toit and Peskir (2009), cf. also Chapter 2 in the particular case of constant drift μ and volatility σ . In this case, $V_n(t - \delta_n, a)$ is computed from (3.10) by the backward induction

$$\begin{aligned} V_n(t - \delta_n, a) = & \int_0^\infty dy \int_{-\infty}^y G \left(t, e^{\sigma(\frac{\log a}{\sigma} \vee y - x)} \right) \wedge V_n \left(t, e^{\sigma(\frac{\log a}{\sigma} \vee y - x)} \right) e^{\lambda x - \lambda^2 \delta_n x / 2} \\ & \varphi_{\delta_n}(x, y) dx, \end{aligned} \quad (3.18)$$

with

$$G(t, a) = E \left[a \vee e^{\sigma S_{T-t}^\lambda} \right] = \int_0^\infty dy \int_{-\infty}^y e^{\log a \vee (\sigma y) + \lambda x - \lambda^2 (T-t)/2} \varphi_{T-t}(x, y) dx, \quad (3.19)$$

for $t \in [0, T)$, $a \geq 1$, where

$$S_t^\lambda := \max_{0 \leq s \leq t} (B_s + \lambda s),$$

$\lambda := \mu/\sigma - \sigma/2$, and $\varphi_r(x, y)$ is given by (3.13). In comparison with Du Toit and Peskir (2009) our method does not require Monte Carlo simulations in the absence of regime switching, since in this case (3.19) can be computed in closed form, cf. (2.7) in Du Toit and Peskir (2009).

Numerical results

We close this section with numerical estimates for the boundary functions

$$b_D(t, j) := \inf\{x \in [1, \infty) : (t, j, x) \in D\}, \quad t \in [0, T], \quad j \in \mathcal{M},$$

of the stopping set D . In Figure 3.1 the recursive algorithm of Theorems 3.1.1 and 3.1.2 are compared with the solution of the Volterra integral equation (2.57) in Chapter 2 by dichotomy, for a two-state Markov chain $(\beta_t)_{t \in [0, T]}$ on $\{1, 2\}$ started at $\beta_0 = 1$, with $\mu(1) = 0.05$, $\sigma(1) = 0.25$, $\mu(2) = 0.06$, $\sigma(2) = 0.3$, $T = 10$, $Y_0 = 100$, $\delta_n = 0.05$, $n = 200$, and generator

$$\mathbf{Q} = \begin{bmatrix} -0.15 & 0.15 \\ 0.2 & -0.2 \end{bmatrix}.$$

As can be noted in Figure 3.1 below, the recursive method provides a more stable estimate of the boundary function, which is known to be decreasing in this range of parameters.

In the absence of regime switching this conclusion is confirmed in Figure 3.2 where we compare the convergence of the estimation of $b_D(t, 1)$ at a fixed time $t = 3.5$ by the recursive algorithm Theorems 3.1.1 and 3.1.2 and by the dichotomic solution of the Volterra integral equation of

Figure 3.1: Boundary function computed from Theorems 3.1.1-3.1.2 vs (2.57).

Du Toit and Peskir (2009), in terms of the decreasing values of the time discretization step of the interval of $[t, T]$, with $\mu = 0.2$, $\sigma = 0.5$, $T = 4$, and $Y_0 = 10$.

Figure 3.2: Convergence graph from Theorems 3.1.1-3.1.2 vs the approach by Volterra equation.

Figure 3.3 presents the graph of a non-monotone boundary function $t \mapsto b_D(t, 1)$ obtained from the recursive algorithm of Theorems 3.1.1 and 3.1.2 with $\mu(1) = 0.2$, $\sigma(1) = 0.5$, $\mu(2) = -1.0$, $\sigma(2) = 0.4$, $T = 4$, $Y_0 = 10$,

$\delta_n = 0.05$, $n = 80$, and

$$\mathbf{Q} = \begin{bmatrix} -0.8 & 0.8 \\ 0.1 & -0.1 \end{bmatrix}.$$

Note that in this example, the values of $\mu(i)$, $i = 1, 2$, do not have a constant sign and we have $b_D(t, 2) = 1$, $t \in [0, T]$, meaning that exercise is immediate as soon as $(\beta_t)_{t \in [0, T]}$ reaches state 2.

Figure 3.3: *Non-monotone boundary function under drifts of mixed signs.*

In the next section we prove Theorem 3.1.1, and in Section 3.3 we compute the expectation function G defined in (3.5).

3.2 Proof of Theorem 3.1.1

(i) First we note that for any stopping time $\tau \in [t, T]$ we have

$$\begin{aligned} E \left[\frac{(aY_t) \vee \hat{Y}_{t,T}}{Y_\tau} \mid \beta_t = j \right] &= E \left[\frac{a \vee (\hat{Y}_{t,T}/Y_t)}{Y_\tau/Y_t} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right] \\ &= E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right], \end{aligned} \quad (3.1)$$

$t \in [0, T]$, $j \in \mathcal{M}$, $a \in [1, \infty)$, since $\hat{Y}_{0,t}/Y_t$ is independent of

$$\left(\frac{Y_t}{Y_\tau}, \frac{\hat{Y}_{t,T}}{Y_\tau} \right) = \left(\exp \left(- \int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right), \exp \left(\hat{S}_{T-t}^t - \int_0^{\tau-t} \sigma(\beta_{u+t}) d\hat{B}_u^t \right) \right)$$

given β_t . By the definition (3.7) of $\lceil s \rceil_n$ that

$$\lceil s \rceil_n := \min \{ t \in \mathcal{T}_n : t \geq s \}, \quad s \in [0, T], \quad n \geq 1,$$

and by (3.3) we have

$$\begin{aligned} V(\lceil t \rceil_n, j, a) &= \inf_{\lceil t \rceil_n \leq \tau \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \beta_{\lceil t \rceil_n} = j, \frac{\hat{Y}_{0,\lceil t \rceil_n}}{Y_{\lceil t \rceil_n}} = a \right] \\ &\leq \inf_{\lceil t \rceil_n \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = j, \frac{\hat{Y}_{0,\lceil t \rceil_n}}{Y_{\lceil t \rceil_n}} = a \right] \\ &\leq \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = j, \frac{\hat{Y}_{0,\lceil t \rceil_n}}{Y_{\lceil t \rceil_n}} = a \right] \\ &= V_n(\lceil t \rceil_n, j, a), \end{aligned}$$

$t \in [0, T]$, $j \in \mathcal{M}$, $a \geq 1$, where we used (3.9) and the infimum is taken over all $\mathcal{T}_n$ -valued discrete stopping times τ_n . Therefore by the continuity of $V(t, j, a)$ with respect to t , cf. e.g. Peskir and Shiryaev (2006), Chap III, §7.1.1 page 130 and §7.4.1 pages 135-136, we obtain

$$\begin{aligned} V(t, j, a) &= \lim_{n \rightarrow \infty} V(\lceil t \rceil_n, j, a) \\ &\leq \liminf_{n \rightarrow \infty} V_n(\lceil t \rceil_n, j, a) \\ &= \liminf_{n \rightarrow \infty} V_n(t, j, a), \end{aligned} \tag{3.2}$$

since

$$\begin{aligned} V_n(\lceil t \rceil_n, j, a) &= \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = j, \frac{\hat{Y}_{0,\lceil t \rceil_n}}{Y_{\lceil t \rceil_n}} = a \right] \\ &= V_n(t, j, a), \quad t \in [0, T], j \in \mathcal{M}, a \geq 1 \end{aligned}$$

by (3.9).

On the other hand, by (3.1) we have

$$\begin{aligned}
\limsup_{n \rightarrow \infty} V_n(\lceil t \rceil_n, j, a) &= \limsup_{n \rightarrow \infty} \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = j, \frac{\hat{Y}_{0,\lceil t \rceil_n}}{Y_{\lceil t \rceil_n}} = a \right] \\
&= \limsup_{n \rightarrow \infty} \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = j \right] \\
&= \limsup_{n \rightarrow \infty} \sum_{l=1}^m [e^{(\lceil t \rceil_n - t)Q}]_{j,l} \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = l \right] \\
&\leq \limsup_{n \rightarrow \infty} \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} \sum_{l=1}^m [e^{(\lceil t \rceil_n - t)Q}]_{j,l} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = l \right] \\
&= \limsup_{n \rightarrow \infty} \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\tau_n}} \mid \beta_t = j \right], \tag{3.3}
\end{aligned}$$

$t \in [0, T]$, $j \in \mathcal{M}$, $a \in [1, \infty)$, where $Q = [q_{i,j}]_{1 \leq i, j \leq m}$ is the infinitesimal generator of $(\beta_t)_{t \in [0, T]}$. Next, we note that for every stopping time $\tau \in [t, T]$ we have

$$|\lceil \tau \rceil_n - \tau| < 1/n,$$

hence $(\lceil \tau \rceil_n)_{n \geq 1}$ converges to τ uniformly in $L^\infty(\Omega)$ and pointwise. Hence we have

$$\lim_{n \rightarrow \infty} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\lceil \tau \vee (t + \delta_n) \rceil_n}} \mid \beta_t = j \right] = E \left[\lim_{n \rightarrow \infty} \frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\lceil \tau \vee (t + \delta_n) \rceil_n}} \mid \beta_t = j \right], \tag{3.4}$$

for any stopping time $\tau \in [t, T]$, where we applied Lebesgue's dominated convergence theorem based on the bounds (3.7)-(3.8) stated at the end of this section. Hence from (3.3) and (3.4) we find, for any stopping time $\tau \in [t, T]$,

$$\begin{aligned}
\limsup_{n \rightarrow \infty} V_n(\lceil t \rceil_n, j, a) &\leq \limsup_{n \rightarrow \infty} \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\tau_n}} \mid \beta_t = j \right] \\
&\leq \lim_{n \rightarrow \infty} E \left[\frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\lceil \tau \vee (t + \delta_n) \rceil_n}} \mid \beta_t = j \right] \\
&= E \left[\lim_{n \rightarrow \infty} \frac{(aY_{\lceil t \rceil_n}) \vee \hat{Y}_{\lceil t \rceil_n, T}}{Y_{\lceil \tau \vee (t + \delta_n) \rceil_n}} \mid \beta_t = j \right]
\end{aligned}$$

$$\begin{aligned}
&= E \left[\frac{(aY_t) \vee \hat{Y}_{t,T}}{Y_\tau} \mid \beta_t = j \right] \\
&= E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right],
\end{aligned}$$

where we applied (3.1) and the pathwise continuity of $(Y_t)_{t \in [0, T]}$. Hence given that

$$V_n(t, j, a) = V_n(\lceil t \rceil_n, j, a)$$

by (3.9), we obtain

$$\begin{aligned}
\limsup_{n \rightarrow \infty} V_n(t, j, a) &= \limsup_{n \rightarrow \infty} V_n(\lceil t \rceil_n, j, a) \\
&\leq \inf_{t \leq \tau \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right] \\
&= V(t, j, a),
\end{aligned}$$

which completes the proof of (3.8) by (3.2).

(ii) In order to prove (3.10), for each $n \geq 1$ and $t, h \in \mathcal{T}_n$, $0 \leq t \leq h \leq T$, we consider an optimal stopping time $\tau_n^{(h)}$ such that

$$E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n^{(h)}}} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right] = \inf_{h \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right], \quad (3.5)$$

where the infimum is taken over the discrete $\mathcal{T}_n$ -valued stopping times τ_n , and the existence of $\tau_n^{(h)}$ is guaranteed by Corollary 2.9 of Peskir and Shiryaev (2006) as in Proposition 3.1 of Chapter 2. We note the induction

$$\begin{aligned}
E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n^{(t)}}} \mid \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] &= E \left[\frac{\hat{Y}_{0,T}}{Y_t} \mid \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] \wedge E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n^{(t+\delta_n)}}} \mid \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] \\
&= G \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \wedge V_n \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right), \quad (3.6)
\end{aligned}$$

for $n \geq 1$, $t \in [0, T - \delta_n]$, $a \geq 1$, where

$$V_n(t, j, a) = \inf_{\lceil t \rceil_n + \delta_n \leq \tau_n \leq T} E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n}} \mid \beta_{\lceil t \rceil_n} = j, \frac{\hat{Y}_{0, \lceil t \rceil_n}}{Y_{\lceil t \rceil_n}} = a \right],$$

by (3.9) and

$$G(t, j, a) := E \left[\max \left(a, \hat{Y}_{t,T}/Y_t \right) \mid \beta_t = j \right], \quad t \in [0, T], \quad j \in \mathcal{M}.$$

By (3.5), this yields

$$\begin{aligned} V_n(t - \delta_n, j, a) &= E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n^{(t)}}} \mid \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right] \\ &= E \left[E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n^{(t)}}} \mid \beta_t, \frac{\hat{Y}_{0,t}}{Y_t}, \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right] \mid \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right] \\ &= E \left[E \left[\frac{\hat{Y}_{0,T}}{Y_{\tau_n^{(t)}}} \mid \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right] \mid \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right] \\ &= E \left[G \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \wedge V_n \left(t, \beta_t, \frac{\hat{Y}_{0,t}}{Y_t} \right) \mid \beta_{t-\delta_n} = j, \frac{\hat{Y}_{0,t-\delta_n}}{Y_{t-\delta_n}} = a \right], \end{aligned}$$

for any $t \in \mathcal{T}_n \setminus \{0\}$, where we applied (3.6) and the Markov property of $(\beta_t, \hat{Y}_{0,t}/Y_t)_{t \in [0, T]}$. To conclude, we note that the above extends to any $t \in [\delta_n, T]$ from the relation

$$V_n(t, j, a) = V_n(\lceil t \rceil_n, j, a),$$

for $t \in [0, T]$, $j \in \mathcal{M}$, $a \geq 1$. □

We close this section with the proof of the two bounds used in part (i) of the above proof of Theorem 3.1.1.

(a) Letting

$$\check{Y}_{t,T} := \min_{t \leq h \leq T} Y_h,$$

we check that, for any stopping time τ and $a \geq 1$,

$$\max \left(\frac{aY_{\lceil t \rceil_n}}{Y_{\lceil \tau \vee (t+\delta_n) \rceil_n}}, \frac{\hat{Y}_{\lceil t \rceil_n, T}}{Y_{\lceil \tau \vee (t+\delta_n) \rceil_n}} \right) \leq \frac{a\hat{Y}_{t,T}}{Y_{\lceil \tau \vee (t+\delta_n) \rceil_n}} \leq \frac{a\hat{Y}_{t,T}}{\check{Y}_{t,T}}, \quad n \geq 1. \quad (3.7)$$

(b) On the other hand we have

$$E \left[\hat{Y}_{t,T} / \check{Y}_{t,T} \mid \beta_t = j \right] < \infty$$

since, denoting by $\hat{B}_u^t$ the drifted Brownian motion

$$\hat{B}_u^t := B_{u+t} - B_t + \int_t^{t+u} \left(\frac{\mu(\beta_r)}{\sigma(\beta_r)} - \frac{\sigma(\beta_r)}{2} \right) dr, \quad u \in [0, T-t],$$

we have

$$E \left[\frac{\hat{Y}_{t,T}}{\check{Y}_{t,T}} \mid \beta_t = j \right] = E \left[e^{\sup_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t - \inf_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_t = j \right] \quad (3.8)$$

$$\begin{aligned} &\leq \sqrt{E \left[e^{2 \sup_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_t = j \right] E \left[e^{-2 \inf_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_t = j \right]} \\ &\leq e^{(T-t) \max_{i=1, \dots, m} |\sigma^2(i) - 2\mu(i)|} E \left[e^{2 \sup_{t \leq r \leq T} \int_0^{r-t} \sigma(\beta_{u+t}) d\hat{B}_u^t} \mid \beta_t = j \right] \\ &= e^{(T-t) \max_{i=1, \dots, m} |\sigma^2(i) - 2\mu(i)|} E \left[e^{2\hat{S}_{T-t}^t} \mid \beta_t = j \right] \\ &< \infty, \end{aligned} \quad (3.9)$$

where

$$\hat{S}_s^t := \sup_{0 \leq r \leq s} \int_0^r \sigma(\beta_{u+t}) d\hat{B}_u^t, \quad s \in [0, T-t].$$

3.3 Proof of Theorem 3.1.2

(i) We start by showing that $\Psi_n(t, h, j, a)$ defined in (3.16) can be written

as

$$\begin{aligned} &\Psi_n(t, h, j, a) \\ &= E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\ &\quad \left. \mid \beta_{[t]_n} = j \right], \end{aligned} \quad (3.1)$$

for any $0 \leq t \leq h \leq T$, $j \in \mathcal{M}$, and $a \geq 1$, or equivalently

$$\Psi_n(t, h, j, a) = \Phi_n(t, h, j, a) + \Upsilon_n(t, h, j, a), \quad (3.2)$$

where

$$\begin{aligned} & \Phi_n(t, h, j, a) \\ &:= E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\ & \quad \left. \mathbf{1}_{\{T_1([t]_n) > [t]_n + \delta_n\}} \middle| \beta_{[t]_n} = j \right], \end{aligned} \quad (3.3)$$

and

$$\begin{aligned} & \Upsilon_n(t, h, j, a) \\ &:= E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\ & \quad \left. \mathbf{1}_{\{T_1([t]_n) \leq [t]_n + \delta_n\}} \middle| \beta_{[t]_n} = j \right], \end{aligned} \quad (3.4)$$

where

$$T_1(t) := \inf\{s > t : \beta_s \neq \beta_t\}, \quad (3.5)$$

is the first jump time of $(\beta_s)_{s \in [t, T]}$ after t . When $[t]_n = [h]_n \leq T - \delta_n$, by (3.3) we have

$$\begin{aligned} & \Phi_n(t, h, j, a) \\ &= E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \middle| \beta_{[t]_n} = j \right] \\ &= e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y \\ & \quad E \left[\exp \left((\log a) \vee \left(\sigma(j)y \vee \left(\sigma(j)x + \sup_{[t]_n + \delta_n \leq s \leq T} \int_{[t]_n + \delta_n}^s \sigma(\beta_r) dB_r \right) \right) \right. \right. \\ & \quad \left. \left. + \int_{[t]_n + \delta_n}^T u(\beta_r) dB_r + u(j)x - \frac{1}{2} \int_{[t]_n + \delta_n}^T u^2(\beta_r) dr - \frac{1}{2} u^2(j) \delta_n \right) \right] \\ & \quad \varphi_{\delta_n}(x, y) dx \\ &= e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y \\ & \quad E \left[\exp \left(((\log a) \vee \sigma(j)y) \vee \left(\sigma(j)x + \sup_{[t]_n + \delta_n \leq s \leq T} \int_{[t]_n + \delta_n}^s \sigma(\beta_r) dB_r \right) \right) \right. \end{aligned}$$

$$\begin{aligned}
& + \int_{[t]_n + \delta_n}^T u(\beta_r) dB_r + u(j)x - \frac{1}{2} \int_{[t]_n + \delta_n}^T u^2(\beta_r) dr - \frac{1}{2} u^2(j) \delta_n \Big) \Big] \\
& \quad \varphi_{\delta_n}(x, y) dx \\
= & \quad e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y \\
& \quad E \left[\exp \left(((\log a) \vee (\sigma(j)y) - \sigma(j)x) \vee \sup_{[t]_n + \delta_n \leq s \leq T} \int_{[t]_n + \delta_n}^s \sigma(\beta_r) dB_r \right. \right. \\
& \quad \left. \left. + \sigma(j)x + \int_{[t]_n + \delta_n}^T u(\beta_r) dB_r + u(j)x - \frac{1}{2} \int_{[t]_n + \delta_n}^T u^2(\beta_r) dr - \frac{1}{2} u^2(j) \delta_n \right) \right] \\
& \quad \varphi_{\delta_n}(x, y) dx \\
= & \quad e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y e^{(u(j) + \sigma(j))x - u^2(j)T/(2n)} \\
& \quad \Psi_n(t + \delta_n, t + \delta_n, j, e^{(\log a) \vee (\sigma(j)y) - \sigma(j)x}) \varphi_{\delta_n}(x, y) dx,
\end{aligned}$$

where in the first equality we used the fact that the time to the first jump of $(\beta_s)_{s \in [t, \infty)}$ after t is exponentially distributed with parameter $-q_{j,j} > 0$ given $\beta_t = j$, cf. e.g. § 10.4 in Privault (2013). Similarly, when $[t]_n < [h]_n$ and $[t]_n \leq T - \delta_n$, we have

$$\begin{aligned}
& \Phi_n(t, h, j, a) \\
= & \quad E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\
& \quad \left. \Big| \beta_{[t]_n} = j \right] \\
= & \quad e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y \\
& \quad E \left[\exp \left((\log a) \vee \left(\sigma(j)y \vee \left(\sigma(j)x + \sup_{[t]_n + \delta_n \leq s \leq T} \int_{[t]_n + \delta_n}^s \sigma(\beta_r) dB_r \right) \right. \right. \right. \\
& \quad \left. \left. - \int_{[t]_n + \delta_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n + \delta_n}^T u(\beta_r) dB_r + u(j)x - \frac{1}{2} \int_{[t]_n + \delta_n}^T u^2(\beta_r) dr \right. \right. \\
& \quad \left. \left. - \frac{1}{2} u^2(j) \delta_n \right) \right] \varphi_{\delta_n}(x, y) dx \\
= & \quad e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y \\
& \quad E \left[\exp \left(((\log a) \vee (\sigma(j)y) - \sigma(j)x) \vee \sup_{[t]_n + \delta_n \leq s \leq T} \int_{[t]_n + \delta_n}^s \sigma(\beta_r) dB_r \right. \right.
\end{aligned}$$

$$\begin{aligned}
& - \int_{[t]_n + \delta_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n + \delta_n}^T u(\beta_r) dB_r + u(j)x - \frac{1}{2} \int_{[t]_n + \delta_n}^T u^2(\beta_r) dr \\
& - \frac{1}{2} u^2(j) \delta_n \Big) \Big] \varphi_{\delta_n}(x, y) dx \\
& = e^{q_{j,j} \delta_n} \int_0^\infty dy \int_{-\infty}^y e^{u(j)x - u^2(j)T/(2n)} \Psi_n(t + \delta_n, h, j, e^{(\log a) \vee (\sigma(j)y) - \sigma(j)x}) \\
& \varphi_{\delta_n}(x, y) dx. \tag{3.6}
\end{aligned}$$

Next, for $\Upsilon_n(t, h, j, a)$, considering the case $[t]_n = [h]_n \leq T - \delta_n$, by (3.4) we see that

$$\begin{aligned}
& \Upsilon_n(t, h, j, a) \\
& = E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \mathbf{1}_{\{T_1([t]_n) \leq [t]_n + \delta_n\}} \right. \\
& \quad \left. \Big| \beta_{[t]_n} = j \right] \\
& = \sum_{\substack{l=1 \\ l \neq j}}^m E \left[\mathbf{1}_{\{\beta_{T_1}=l\}} e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\
& \quad \left. \mathbf{1}_{\{T_1([t]_n) \leq [t]_n + \delta_n\}} \Big| \beta_{[t]_n} = j \right] \\
& = \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} dr \int_0^\infty dy \int_{-\infty}^y \\
& \quad E \left[\exp \left((\log a) \vee (\sigma(j)y) \vee \left(\sigma(j)x + \sup_{[t]_n + r \leq s \leq T} \int_{[t]_n + r}^s \sigma(\beta_z) dB_z \right) \right. \right. \\
& \quad \left. \left. + \int_{[t]_n + r}^T u(\beta_z) dB_z + u(j)x - \frac{1}{2} \int_{[t]_n + r}^T u^2(\beta_z) dz - u^2(j)r/2 \right) \right. \\
& \quad \left. \Big| \beta_r = l \right] e^{q_{j,j}r} \varphi_r(x, y) dx, \tag{3.7}
\end{aligned}$$

where we used the conditional probability distribution

$$\mathbb{P}(T_1 \in dt, \beta_{T_1} = j \mid \beta_0 = i) = \mathbf{1}_{[0, \infty)}(t) q_{i,j} e^{q_{i,i}t} dt, \quad i \neq j \in \mathcal{M},$$

computed from the exponential distribution with parameter $-q_{i,i}$ of the first jump time T_1 of the Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ started at $i \in \mathcal{M}$ and the transition matrix $(-q_{i,j} \mathbf{1}_{\{i \neq j\}} / q_{i,i})_{i,j \in \mathcal{M}}$ of the embedded Markov chain,

cf. e.g. § 10.7 of Privault (2013). Continue with (3.7) we have

$$\begin{aligned}
& \Upsilon_n(t, t, j, a) \\
&= \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} dr \int_0^\infty dy \int_{-\infty}^y \\
&\quad e^{(u(j)+\sigma(j))x - u^2(j)r/2} \Psi_n(t+r, t+r, l, e^{(\log a) \vee (\sigma(j)y) - \sigma(j)x}) e^{q_{j,j}r} \varphi_r(x, y) dx \\
&= \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} dr \int_0^\infty dy \int_{-\infty}^y \\
&\quad e^{(u(j)+\sigma(j))x - u^2(j)r/2} \Psi_n(t+\delta_n, t+\delta_n, l, e^{(\log a) \vee (\sigma(j)y) - \sigma(j)x}) e^{q_{j,j}r} \varphi_r(x, y) dx,
\end{aligned}$$

Similarly, when $\lceil t \rceil_n < \lceil h \rceil_n$ and $\lceil t \rceil_n \leq T - \delta_n$, we get

$$\begin{aligned}
& \Upsilon_n(t, h, j, a) \tag{3.8} \\
&= E \left[e^{(\log a) \vee \sup_{\lceil t \rceil_n \leq s \leq T} \int_{\lceil t \rceil_n}^s \sigma(\beta_r) dB_r - \int_{\lceil t \rceil_n}^{\lceil h \rceil_n} \sigma(\beta_r) dB_r + \int_{\lceil t \rceil_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{\lceil t \rceil_n}^T u^2(\beta_r) dr} \right. \\
&\quad \left. \mathbf{1}_{\{T_1(\lceil t \rceil_n) \leq \lceil t \rceil_n + \delta_n\}} \mid \beta_{\lceil t \rceil_n} = j \right] \\
&= \sum_{\substack{l=1 \\ l \neq j}}^m E \left[\mathbf{1}_{\{\beta_{T_1} = l\}} \exp \left((\log a) \vee \sup_{\lceil t \rceil_n \leq s \leq T} \int_{\lceil t \rceil_n}^s \sigma(\beta_r) dB_r - \int_{\lceil t \rceil_n}^{\lceil h \rceil_n} \sigma(\beta_r) dB_r \right. \right. \\
&\quad \left. \left. + \int_{\lceil t \rceil_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{\lceil t \rceil_n}^T u^2(\beta_r) dr \right) \mathbf{1}_{\{T_1(\lceil t \rceil_n) \leq \lceil t \rceil_n + \delta_n\}} \mid \beta_{\lceil t \rceil_n} = j \right] \\
&= \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} dr \int_0^\infty dy \int_{-\infty}^y \\
&\quad E \left[\exp \left((\log a) \vee (\sigma(j)y) \vee \left(\sigma(j)x + \sup_{\lceil t \rceil_n + r \leq s \leq T} \int_{\lceil t \rceil_n + r}^s \sigma(\beta_z) dB_z \right) \right. \right. \\
&\quad \left. \left. - \int_{\lceil t \rceil_n + r}^{\lceil h \rceil_n} \sigma(\beta_r) dB_r + \int_{\lceil t \rceil_n + r}^T u(\beta_z) dB_z + u(j)x - \frac{1}{2} \int_{\lceil t \rceil_n + r}^T u^2(\beta_z) dz \right. \right. \\
&\quad \left. \left. - \frac{1}{2} u^2(j)r \right) \mid \beta_r = l \right] e^{q_{j,j}r} \varphi_r(x, y) dx \\
&= \sum_{\substack{l=1 \\ l \neq j}}^m q_{j,l} \int_0^{\delta_n} e^{q_{j,j}r} dr \int_0^\infty dy \int_{-\infty}^y \\
&\quad e^{u(j)x - u^2(j)r/2} \Psi_n(t+\delta_n, h, l, e^{(\log a) \vee (\sigma(j)y) - \sigma(j)x}) \varphi_r(x, y) dx,
\end{aligned}$$

and we conclude to (3.2) by (3.6), (3.8) and (3.16).

(ii) In order to derive the limit in (3.15) we will use the measure $\tilde{\mathbb{P}}$ defined by

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} := \exp \left(- \int_t^T u(\beta_r) dB_r - \frac{1}{2} \int_t^T u^2(\beta_r) dr \right),$$

where $u(j) = \mu(j)/\sigma(j) - \sigma(j)/2$ is defined in (3.14) for $j \in \mathcal{M}$, and

$$\tilde{B}_r := B_r + \int_t^r u(\beta_r) dr \quad r \in [t, T],$$

is a standard Brownian motion under $\tilde{\mathbb{P}}$. From the definition of $\Psi(t, h, j, a)$ by (3.12) we have

$$\begin{aligned} \Psi(t, h, j, a) &= E \left[\frac{\hat{Y}_{0,t} \vee \hat{Y}_{t,T}}{Y_h} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right] \\ &= E \left[\frac{\left(\frac{\hat{Y}_{0,t}}{Y_t} Y_t \right) \vee \hat{Y}_{t,T}}{Y_h} \mid \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = a \right] \\ &= E \left[\frac{a \vee \exp \left(\sup_{t \leq s \leq T} \left(\int_t^s \sigma(\beta_r) dB_r + \int_t^s (\mu(\beta_r) - \frac{1}{2} \sigma^2(\beta_r)) dr \right) \right)}{\exp \left(\int_t^h \sigma(\beta_r) dB_r + \int_t^h (\mu(\beta_r) - \frac{1}{2} \sigma^2(\beta_r)) dr \right)} \mid \beta_t = j \right] \\ &= E \left[\frac{a \vee \exp \left(\sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) d\tilde{B}_r \right)}{\exp \left(\int_t^h \sigma(\beta_r) d\tilde{B}_r \right)} \mid \beta_t = j \right] \\ &= \tilde{E} \left[e^{(\log a) \vee \sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) d\tilde{B}_r - \int_t^h \sigma(\beta_r) d\tilde{B}_r + \int_t^T u(\beta_r) d\tilde{B}_r + \frac{1}{2} \int_t^T u^2(\beta_r) dr} \mid \beta_t = j \right] \\ &= \tilde{E} \left[e^{(\log a) \vee \sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) d\tilde{B}_r - \int_t^h \sigma(\beta_r) d\tilde{B}_r + \int_t^T u(\beta_r) d\tilde{B}_r - \frac{1}{2} \int_t^T u^2(\beta_r) dr} \mid \beta_t = j \right] \\ &= E \left[e^{(\log a) \vee \sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) dB_r - \int_t^h \sigma(\beta_r) dB_r + \int_t^T u(\beta_r) dB_r - \frac{1}{2} \int_t^T u^2(\beta_r) dr} \mid \beta_t = j \right], \end{aligned}$$

which allows us to remove the drift component in the supremum

$\sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) dB_r$. Now, for any $n \geq 1$ we have

$$\sup_{\lceil t \rceil_n \leq s \leq T} \int_{\lceil t \rceil_n}^s \sigma(\beta_r) dB_r$$

$$\begin{aligned}
&= \sup_{[t]_n \leq s \leq T} \int_t^s \sigma(\beta_r) dB_r - \int_t^{[t]_n} \sigma(\beta_r) dB_r \\
&\leq \sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) dB_r - \inf_{t \leq s \leq T} \int_t^s \sigma(\beta_r) dB_r,
\end{aligned}$$

and similarly by replacing $\sigma(\beta_r)$ with $u(\beta_r)$, therefore

$$\begin{aligned}
&\exp \left((\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r \right. \\
&\quad \left. - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr \right)
\end{aligned}$$

is upper bounded by

$$a \exp \left(\sup_{t \leq s \leq T} \int_t^s (2\sigma(\beta_r) + u(\beta_r)) dB_r - \inf_{t \leq s \leq T} \int_t^s (2\sigma(\beta_r) + u(\beta_r)) dB_r \right),$$

which is $\mathbb{P}$ -integrable as in (3.8). Hence by (3.1) and the dominated convergence we get

$$\begin{aligned}
&\lim_{n \rightarrow \infty} \Psi_n(t, h, j, a) \\
&= \lim_{n \rightarrow \infty} E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\
&\quad \left. \middle| \beta_{[t]_n} = j \right] \\
&= \lim_{n \rightarrow \infty} \sum_{l=1}^m \left([e^{(\lceil t \rceil_n - t)Q}]_{j,l} \right. \\
&\quad \times E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\
&\quad \left. \middle| \beta_{[t]_n} = l \right] \Big) \\
&= \lim_{n \rightarrow \infty} E \left[e^{(\log a) \vee \sup_{[t]_n \leq s \leq T} \int_{[t]_n}^s \sigma(\beta_r) dB_r - \int_{[t]_n}^{[h]_n} \sigma(\beta_r) dB_r + \int_{[t]_n}^T u(\beta_r) dB_r - \frac{1}{2} \int_{[t]_n}^T u^2(\beta_r) dr} \right. \\
&\quad \left. \middle| \beta_t = j \right] \\
&= E \left[e^{(\log a) \vee \sup_{t \leq s \leq T} \int_t^s \sigma(\beta_r) dB_r - \int_t^h \sigma(\beta_r) dB_r + \int_t^T u(\beta_r) dB_r - \frac{1}{2} \int_t^T u^2(\beta_r) dr} \middle| \beta_t = j \right] \\
&= \Psi(t, h, j, a).
\end{aligned}$$

□

Finally we note that Theorem 3.1.2 also applies to the characterization of the optimal stopping set D'

$$D' = \left\{ (t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : \Psi(t, t, j, x) = \inf_{h \in [t, T]} \Psi(t, h, j, x) \right\} \quad (3.9)$$

associated to (3.17), which contains the optimal stopping set

$$\begin{aligned} D &= \left\{ (t, j, x) \in [0, T] \times \mathcal{M} \times [1, \infty) : E \left[\frac{\hat{Y}_{0,T}}{Y_t} \middle| \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \right. \\ &\quad \left. = \inf_{\tau \in [t, T]} E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \middle| \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \right\} \end{aligned}$$

associated to (3.1), i.e. we have $D \subset D'$, since

$$\inf_{\tau \in [t, T]} E \left[\frac{\hat{Y}_{0,T}}{Y_\tau} \middle| \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right] \leq \inf_{h \in [t, T]} E \left[\frac{\hat{Y}_{0,T}}{Y_h} \middle| \beta_t = j, \frac{\hat{Y}_{0,t}}{Y_t} = x \right]$$

where the infimum in the left hand side is taken over all stopping times $\tau \in [t, T]$.

Chapter 4

Sensitivity analysis by Wiener-Malliavin calculus

In this chapter we compute the Greeks by Malliavin calculus in a generalized Black-Scholes model with regime switching that reflects the underlying changes in the state of the economy and extend the results for models with non-differentiable payoff function. Sensitivity analysis regarding Markov transition rate matrix is computed by a close form expression and a successive approximation method respectively.

Applying the Wiener-Malliavin calculus, we compute the sensitivities in the framework of the regime switching model (4.1) below. For any payoff function $\phi \in \mathcal{C}_b^1(\mathbf{R})$ with bounded derivative, value function V given by (4.2) below, and any $g : \{1, \dots, m\} \rightarrow (0, \infty)$, we show by Proposition 4.2.1 that

$$\Delta = \frac{\partial V(x_0, \alpha_0)}{\partial x_0} = \frac{e^{-rT}}{x_0} E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right], \quad (4.1)$$

where $\langle \sigma, g \rangle := \int_0^T \sigma(\alpha_s) g(\alpha_s) ds$, note that optimal choice of g for efficient Monte Carlo simulation is achievable, by Subsection 4.2.5, the choice of g such that $g(i) = \sigma(i)$ for all $i \in \{1, \dots, m\}$ minimizes the variance of

the weight of Δ . Other Greeks like Γ , Vega, Λ are similarly computed in Subsection 4.2.2-4.2.4. Moreover we note that,

- i) Given $\phi \in \mathcal{C}_b^1(\mathbb{R})$, differentiating the payoff function ϕ inside expectation tends to be easier than our approach applying Malliavin calculus.
- ii) In the right-hand of the expression (4.1), no derivative of the payoff function ϕ appears any more.

Based on these observations, it is tempted and necessary to relax the restriction of ϕ . Therefore, by Section 4.3, we extend the results to a class of non-differentiable payoff functions.

4.1 Formulation

In this chapter, we calculate the Greeks based on a generalized Black-Scholes model with regime switching as in the previous chapters. Consider a standard Brownian motion $(W_t)_{t \in \mathbb{R}_+}$ and a Markov chain $(\alpha_t)_{t \in \mathbb{R}_+}$, assume that they are independent and the Markov chain α_t has a state space $\mathcal{M} = \{1, \dots, m\}$. Consider the stochastic process $(X_t)_{t \in \mathbb{R}_+}$ given by the following SDE:

$$dX_t = X_t[\mu(\alpha_t)dt + \sigma(\alpha_t)dW_t], \quad 0 \leq t \leq T, \quad (4.1)$$

with $X_0 = x_0 > 0$, $\alpha_0 \in \mathcal{M}$, and $\mu : \mathcal{M} \rightarrow \mathbb{R}$, and $\sigma : \mathcal{M} \rightarrow (0, \infty)$ are deterministic functions. Denote the filtration generated by $(B_t)_{t \in \mathbb{R}_+}$ and $(\alpha_t)_{t \in \mathbb{R}_+}$ as $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$. Given a payoff function ϕ on $\mathbb{R}$, consider the value function $V(x_0, \alpha_0)$ defined as follows:

$$V(x_0, \alpha_0) = e^{-rT} E[\phi(X_T) \mid X_0 = x_0, \alpha_0], \quad (4.2)$$

where $(X_t)_{t \in [0, T]}$ follows SDE (4.1), $r > 0$ denotes the risk-free rate. Then we proceed the computation of the sensitivities of $V(x_0, \alpha_0)$ to the changes of coefficients in this model, such like $X_0 = x_0, \mu, \sigma$ and even Q -matrix. We apply the Malliavin calculus (cf. Nualart (1995), Protter (1990)) to sensitivity analysis. To begin with, we recall some Malliavin operators. For any random variable $F \in L^2(\Omega, P)$ of the form:

$$F = f \left(\int_0^T h_1(\alpha_s) dW_s, \dots, \int_0^T h_n(\alpha_s) dW_s, \int_0^T h_0(\alpha_s) ds \right), \quad (4.3)$$

for $n \in \mathbf{N}_+$, $f \in \mathcal{C}^1(\mathbf{R}^{n+1})$, and deterministic functions $h_i : \mathcal{M} \rightarrow \mathbf{R}, i \in \{0, 1, \dots, n\}$. Given $F \in L^2(\Omega, P)$ in form of (4.3), the gradient of F is defined by

$$D_t F := \sum_{k=1}^n \partial_k f \left(\int_0^T h_1(\alpha_s) dW_s, \dots, \int_0^T h_n(\alpha_s) dW_s, \int_0^T h_0(\alpha_s) ds \right) h_k(\alpha_t), \quad (4.4)$$

for $t \in [0, T]$. For any deterministic function $g : \mathcal{M} \rightarrow \mathbf{R}$ we define

$$W(g) := \int_0^T g(\alpha_t) dW_t, \quad (4.5)$$

and

$$D_g F := \int_0^T D_t F g(\alpha_t) dt \quad (4.6)$$

for any random variable $F \in L^2(\Omega, P)$ in form of (4.3).

4.2 Computation of the Greeks based on the partial Malliavin calculus

In this section, we proceed the computation of the sensitivities of $V(x_0, \alpha_0)$ to the changes of coefficients in this model, such like X_0, μ , and σ . First in this section, we assume that $\phi \in \mathcal{C}_b^1(\mathbf{R})$ and has bounded derivative, extension to non-differentiable payoff function will be shown in Section 4.3.

4.2.1 Variations in the initial price

The sensitivity of value function (4.2) to the initial price X_0 is given by the following proposition:

Proposition 4.2.1. For any payoff function $\phi \in \mathcal{C}_b^1(\mathbb{R})$ with bounded derivative, $(X_t)_{t \in [0, T]}$ defined in (4.1) and any $g : \mathcal{M} \rightarrow (0, \infty)$, we have

$$\frac{\partial V(x_0, \alpha_0)}{\partial x_0} = \frac{e^{-rT}}{x_0} E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right], \quad (4.7)$$

where $\langle \sigma, g \rangle := \int_0^T \sigma(\alpha_s) g(\alpha_s) ds$.

Proof. First we check that the right-hand side of (4.7) is well defined by the following points,

i) $\langle \sigma, g \rangle \geq \varepsilon > 0$ a.s. for some $\varepsilon > 0$, since $\sigma(i) > 0$ and $g(i) > 0$ for all $i \in \mathcal{M}$.

ii) We show that $\phi(X_T)W(g)/\langle \sigma, g \rangle$ is integrable. Since

$$\begin{aligned} E \left[\frac{W(g)^2}{\langle \sigma, g \rangle^2} \right] &= E \left[\frac{\int_0^T g^2(\alpha_s) ds}{\left(\int_0^T \sigma(\alpha_s) g(\alpha_s) ds \right)^2} \right] \\ &\leq \frac{\max\{g^2(i) : i \in \mathcal{M}\}}{T \min\{\sigma^2(i) g^2(i) : i \in \mathcal{M}\}} < \infty, \end{aligned} \quad (4.8)$$

and $E[\phi(X_T)^2] < \infty$ for ϕ is a bounded function, we claim the integrability of $\phi(X_T)W(g)/\langle \sigma, g \rangle$.

For the left-hand side, we need to show that $V(x_0, \alpha_0)$ is differentiable with respect to x_0 . Since ϕ has bounded derivative, there exists a $K_0 > 0$ that $|\partial \phi(x)/\partial x| < K_0$ for any $x \in \mathbb{R}$. We have the solution to the SDE (4.1):

$$X_t = x_0 \exp \left(\int_0^t \mu(\alpha_u) - \frac{1}{2} \sigma^2(\alpha_u) du + \int_0^t \sigma(\alpha_u) dW_u \right), \quad t \in [0, T], \quad (4.9)$$

it follows that

$$\frac{\partial \phi(X_T)}{\partial x_0} = \phi'(X_T) \frac{X_T}{x_0},$$

which is uniformly bounded by a integrable random variable $X_T K_0/x_0$.

Therefore by (4.2), $V(x_0, \alpha_0)$ is differentiable with respect to $X_0 = x_0$ and

$$\begin{aligned} \frac{\partial V(x_0, \alpha_0)}{\partial x_0} &= e^{-rT} \frac{\partial}{\partial x_0} E \left[\phi(X_T) \mid X_0 = x_0, \alpha_0 \right] \\ &= e^{-rT} E \left[\frac{\partial \phi(X_T)}{\partial x_0} \mid X_0 = x_0, \alpha_0 \right]. \end{aligned} \quad (4.10)$$

Then we continue to prove (4.7). It follows from (4.9) that $X_T \in L^2(\Omega, P)$ in form of (4.3), and we have

$$D_g X_T = X_T \langle \sigma, g \rangle > 0 \quad a.s. \quad (4.11)$$

Since the payoff function $\phi \in \mathcal{C}_b^1(\mathbb{R})$, $\phi(X_T) \in L^2(\Omega, P)$ and is in form of (4.3), we have the following chain rule:

$$D_g \phi(X_T) = \phi'(X_T) D_g X_T, \quad \text{for any } g : \mathcal{M} \rightarrow (0, \infty). \quad (4.12)$$

Hence by (4.9)-(4.12) we see that

$$\begin{aligned} \frac{\partial V(x_0, \alpha_0)}{\partial x_0} &= e^{-rT} E \left[\phi'(X_T) \frac{\partial X_T}{\partial x_0} \mid X_0 = x_0, \alpha_0 \right] \\ &= e^{-rT} E \left[\frac{D_g \phi(X_T)}{D_g X_T} \frac{X_T}{x_0} \mid X_0 = x_0, \alpha_0 \right] \\ &= \frac{e^{-rT}}{x_0} E \left[\frac{D_g \phi(X_T)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right]. \end{aligned} \quad (4.13)$$

Denote by $\mathcal{F}_\alpha$ a filtration generated by $\{\alpha_t; t \in [0, T]\}$, it follows from the integration by parts formula by Lemma 1.2.1 in Nualart (1995) and the independence between $(W_t)_{t \in [0, T]}$ and $(\alpha_t)_{t \in [0, T]}$ that,

$$\begin{aligned} &E \left[\frac{D_g \phi(X_T)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \\ &= E \left[E \left[\frac{D_g \phi(X_T)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \mathcal{F}_\alpha \right] \mid X_0 = x_0, \alpha_0 \right] \end{aligned}$$

$$\begin{aligned}
 &= E \left[E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \mathcal{F}_\alpha \right] \mid X_0 = x_0, \alpha_0 \right] \\
 &= E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right], \tag{4.14}
 \end{aligned}$$

for any $g : \mathcal{M} \rightarrow (0, \infty)$. By (4.13)-(4.14), we obtain,

$$\Delta = \frac{e^{-rT}}{x_0} E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right].$$

□

4.2.2 Variations in the initial price for second order

As in most context, the second derivative of the value function with respect to the initial price is denoted as Γ . Similarly, we have,

Proposition 4.2.2. For any payoff function $\phi \in \mathcal{C}_b^1(\mathbb{R})$ with bounded derivative, $g : \mathcal{M} \rightarrow (0, \infty)$, and $(X_t)_{t \in [0, T]}$ defined in (4.1), we have

$$\Gamma = \frac{\partial^2 V(x_0, \alpha_0)}{\partial x_0^2} = \frac{e^{-rT}}{x_0^2} E \left[\phi(X_T) \frac{W(g)^2 - \|g\| - W(g)\langle \sigma, g \rangle}{\langle \sigma, g \rangle^2} \right],$$

where $\langle \sigma, g \rangle := \int_0^T \sigma(\alpha_s)g(\alpha_s)ds$ and $\|g\| := \int_0^T g^2(\alpha_s)ds$.

Proof.

$$\begin{aligned}
 \frac{\partial^2 V(x_0, \alpha_0)}{\partial x_0^2} &= \frac{\partial}{\partial x_0} \left(\frac{e^{-rT}}{x_0} E \left[\frac{\phi(X_T)W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \right) \\
 &= - \frac{e^{-rT}}{x_0^2} E \left[\frac{\phi(X_T)W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \\
 &\quad + \frac{e^{-rT}}{x_0} \frac{\partial}{\partial x_0} E \left[\frac{\phi(X_T)W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \\
 &= - \frac{e^{-rT}}{x_0^2} E \left[\frac{\phi(X_T)W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \\
 &\quad + \frac{e^{-rT}}{x_0} E \left[\frac{W(g)}{\langle \sigma, g \rangle} \phi'(X_T) \frac{X_T}{x_0} \mid X_0 = x_0, \alpha_0 \right] \\
 &= - \frac{e^{-rT}}{x_0^2} E \left[\frac{\phi(X_T)W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right]
 \end{aligned}$$

$$\begin{aligned}
& + \frac{e^{-rT}}{x_0} E \left[\frac{W(g)}{\langle \sigma, g \rangle} \frac{D_g \phi(X_T)}{D_g X_T} \frac{X_T}{x_0} \mid X_0 = x_0, \alpha_0 \right] \\
& = - \frac{e^{-rT}}{x_0^2} E \left[\frac{\phi(X_T) W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \\
& + \frac{e^{-rT}}{x_0^2} E \left[\frac{D_g \phi(X_T) W(g)}{\langle \sigma, g \rangle^2} \mid X_0 = x_0, \alpha_0 \right], \quad (4.15)
\end{aligned}$$

where $\langle \sigma, g \rangle > 0$ a.s. By the chain rule of derivative and the independence between $(W_t)_{t \in \mathbb{R}_+}$ and $(\alpha_t)_{t \in \mathbb{R}_+}$, we have

$$\begin{aligned}
& E \left[D_g \phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle^2} \mid X_0 = x_0, \alpha_0 \right] \\
& = E \left[D_g \left(\frac{\phi(X_T) W(g)}{\langle \sigma, g \rangle^2} \right) \mid X_0 = x_0, \alpha_0 \right] \\
& \quad - E \left[\phi(X_T) D_g \left(\frac{W(g)}{\langle \sigma, g \rangle^2} \right) \mid X_0 = x_0, \alpha_0 \right] \\
& = E \left[\frac{\phi(X_T) W(g)^2}{\langle \sigma, g \rangle^2} \mid X_0 = x_0, \alpha_0 \right] - E \left[\frac{\phi(X_T) \int_0^T g^2(\alpha_s) ds}{\langle \sigma, g \rangle^2} \mid X_0 = x_0, \alpha_0 \right], \quad (4.16)
\end{aligned}$$

where in the last line we applied the integration by parts formula. Substituting (4.16) into (4.15), we see that

$$\frac{\partial^2 V(x_0, \alpha_0)}{\partial x_0^2} = \frac{e^{-rT}}{x_0^2} E \left[\phi(X_T) \frac{W(g)^2 - \|g\| - W(g) \langle \sigma, g \rangle}{\langle \sigma, g \rangle^2} \right],$$

where $\langle \sigma, g \rangle = \int_0^T \sigma(\alpha_s) g(\alpha_s) ds > 0$ a.s. and $\|g\| = \int_0^T g^2(\alpha_s) ds > 0$. □

4.2.3 Variations in the diffusion coefficient

The variations in the diffusion coefficient is denoted as Vega, precisely,

$$\text{Vega} = \frac{\partial V(x_0, \alpha_0)}{\partial \sigma}$$

where σ is a constant parameter. In our model, $\sigma(\cdot)$ is a function on $\mathcal{M}$, so there is no direct way of derivative. Instead, we introduce the

perturbed process:

$$dX_t^{\sigma,\varepsilon} = X_t^{\sigma,\varepsilon}[\mu(\alpha_t)dt + (\sigma(\alpha_t) + \varepsilon)dW_t], \quad (4.17)$$

where $\varepsilon \in \mathbb{R}$, $X_0^{\sigma,\varepsilon} := X_0 = x_0$. Accordingly, we define

$$V^{\sigma,\varepsilon}(x_0, \alpha_0) := E[e^{-rT}\phi(X_T^{\sigma,\varepsilon}) \mid X_0 = x_0, \alpha_0]. \quad (4.18)$$

Then the variation in the diffusion coefficient is given by

$$\text{Vega} = \frac{\partial V^{\sigma,\varepsilon}(x_0, \alpha_0)}{\partial \varepsilon} \Big|_{\varepsilon=0}. \quad (4.19)$$

Note that the solution to (4.17) is

$$X_t^{\sigma,\varepsilon} = x_0 \exp \left(\int_0^t \mu(\alpha_u) - \frac{1}{2}(\sigma(\alpha_u) + \varepsilon)^2 du + \int_0^t (\sigma(\alpha_u) + \varepsilon) dW_u \right), \quad (4.20)$$

for $t \in [0, T]$, we compute Vega by the following proposition.

Proposition 4.2.3. For any payoff function $\phi \in \mathcal{C}_b^1(\mathbb{R})$ with bounded derivative, $g : \mathcal{M} \rightarrow (0, \infty)$, and $(X_t)_{t \in [0, T]}$ defined in (4.1), we have

$$\begin{aligned} \text{Vega} &= \frac{\partial V^{\sigma,\varepsilon}(x_0, \alpha_0)}{\partial \varepsilon} \Big|_{\varepsilon=0} \\ &= e^{-rT} E \left[\phi(X_T) \left(\frac{W_T W(g) - \int_0^T g(\alpha_s) ds - W(g) \int_0^T \sigma(\alpha_s) ds}{\langle \sigma, g \rangle} \right) \mid X_0 = x_0, \alpha_0 \right], \end{aligned} \quad (4.21)$$

where $\langle \sigma, g \rangle := \int_0^T \sigma(\alpha_s) g(\alpha_s) ds$.

Proof. By the integration by parts formula and chain rule we have,

$$\begin{aligned} \frac{\partial V^{\sigma,\varepsilon}(x_0, \alpha_0)}{\partial \varepsilon} &= \frac{\partial}{\partial \varepsilon} E \left[e^{-rT} \phi(X_T^{\sigma,\varepsilon}) \mid X_0 = x_0, \alpha_0 \right] \\ &= e^{-rT} E \left[\phi'(X_T^{\sigma,\varepsilon}) \frac{\partial X_T^{\sigma,\varepsilon}}{\partial \varepsilon} \mid X_0 = x_0, \alpha_0 \right] \\ &= e^{-rT} E \left[\frac{D_g \phi(X_T^{\sigma,\varepsilon})}{D_g X_T^{\sigma,\varepsilon}} X_T^{\sigma,\varepsilon} \left(W_T - \int_0^T \sigma(\alpha_u) du - \varepsilon T \right) \mid X_0 = x_0, \alpha_0 \right] \end{aligned}$$

$$\begin{aligned}
&= e^{-rT} E \left[\frac{D_g \phi(X_T^{\sigma, \varepsilon})}{\langle \sigma + \varepsilon, h \rangle} \left(W_T - \int_0^T \sigma(\alpha_u) du - \varepsilon T \right) \mid X_0 = x_0, \alpha_0 \right] \\
&= e^{-rT} E \left[\frac{1}{\langle \sigma + \varepsilon, h \rangle} D_g \left(\phi(X_T^{\sigma, \varepsilon}) \left(W_T - \int_0^T \sigma(\alpha_u) du - \varepsilon T \right) \right) \mid X_0 = x_0, \alpha_0 \right] \\
&\quad - e^{-rT} E \left[\frac{\phi(X_T^{\sigma, \varepsilon})}{\langle \sigma + \varepsilon, h \rangle} D_g \left(W_T - \int_0^T \sigma(\alpha_u) du - \varepsilon T \right) \mid X_0 = x_0, \alpha_0 \right] \\
&= e^{-rT} E \left[\frac{\phi(X_T^{\sigma, \varepsilon}) W(g)}{\langle \sigma + \varepsilon, h \rangle} \left(W_T - \int_0^T \sigma(\alpha_u) du - \varepsilon T \right) \mid X_0 = x_0, \alpha_0 \right] \\
&\quad - e^{-rT} E \left[\frac{\phi(X_T^{\sigma, \varepsilon}) \int_0^T g(\alpha_s) ds}{\langle \sigma + \varepsilon, h \rangle} \mid X_0 = x_0, \alpha_0 \right], \tag{4.22}
\end{aligned}$$

for any deterministic function $g : \mathcal{M} \rightarrow \mathbb{R}$, and $\varepsilon \in \mathbb{R}$. Note by (4.20) that

$$\lim_{\varepsilon \rightarrow 0} X_t^{\sigma, \varepsilon} = X_t, \quad t \in [0, T] \quad a.e.,$$

the expression Vega in (4.21) is obtained by passing ε to 0 in (4.22). $\square$

4.2.4 Variations in the drift coefficient

Similarly, we introduce the perturbed process for drift coefficient:

$$dX_t^{\mu, \varepsilon} = X_t^{\mu, \varepsilon} [(\mu(\alpha_t) + \varepsilon)dt + \sigma(\alpha_t)dW_t], \tag{4.23}$$

and the value function $V^{\mu, \varepsilon}(x_0, \alpha_0)$ driven by $X_t^{\mu, \varepsilon}$,

$$V^{\mu, \varepsilon}(x_0, \alpha_0) := e^{-rT} E[\phi(X_T^{\mu, \varepsilon}) \mid X_0 = x_0, \alpha_0]. \tag{4.24}$$

Accordingly, we have the solution to (4.23) as follows,

$$\begin{aligned}
X_t^{\mu, \varepsilon} &= x_0 \exp \left(\int_0^t \left(\mu(\alpha_u) + \varepsilon - \frac{1}{2} \sigma^2(\alpha_u) \right) du + \int_0^t \sigma(\alpha_u) dW_u \right) \\
&= X_t e^{\varepsilon t} \quad t \in [0, T]. \tag{4.25}
\end{aligned}$$

Through a similar computation as that of V_{age} , we have the derivative of $V^{\mu, \varepsilon}(x_0, \alpha_0)$ with respect to ε :

$$\frac{\partial V^{\mu, \varepsilon}(x_0, \alpha_0)}{\partial \varepsilon}$$

$$\begin{aligned}
 &= e^{-rT} E \left[\frac{D_g \phi(X_T^{\mu, \varepsilon})}{D_g X_T^{\mu, \varepsilon}} \frac{\partial X_T^{\mu, \varepsilon}}{\partial \mu} \middle| X_0 = x_0, \alpha_0 \right] \\
 &= T e^{-rT} E \left[\frac{D_g \phi(X_T^{\mu, \varepsilon})}{\langle \sigma, g \rangle} \middle| X_0 = x_0, \alpha_0 \right] \\
 &= T e^{-rT} E \left[\frac{\phi(X_T^{\mu, \varepsilon}) W(g)}{\langle \sigma, g \rangle} \middle| X_0 = x_0, \alpha_0 \right]
 \end{aligned}$$

for any deterministic function $g : \mathcal{M} \rightarrow \mathbb{R}$ and $\varepsilon \in \mathbb{R}$. Therefore, we have the following Proposition 4.2.4 for *Lambda*.

Proposition 4.2.4. For any payoff function $\phi \in \mathcal{C}_b^1(\mathbb{R})$ with bounded derivative, $g : \mathcal{M} \rightarrow (0, \infty)$, and $(X_t)_{t \in [0, T]}$ defined in (4.1), we have

$$\Lambda = T e^{-rT} E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \middle| X_0 = x_0, \alpha_0 \right].$$

where $\langle \sigma, g \rangle := \int_0^T \sigma(\alpha_s) g(\alpha_s) ds$.

4.2.5 Optimization of convergence

In this subsection, we aim at figuring out an optimal choice of g for efficient Monte Carlo simulation. Recall David and Privault (2009), based on the computation of *Theta* in a jump-diffusion model, Proposition 4.1 of David and Privault (2009) minimizes the variance of the weight of *Theta*. In our model, there exists such a problem too. We can minimize the weight by choosing a optimal $g \in \mathcal{G}$. We gain a similar conclusion to that in David and Privault (2009).

Proposition 4.2.5. The weight of *Delta* in (4.7) is $\Psi(g) := W(g)/\langle \sigma, g \rangle$, then the infimum on $\{Var[\Psi(g)] ; g : \mathcal{M} \rightarrow \mathbb{R}\}$ is attained when

$$g(i) = c\sigma(i) \quad i \in \mathcal{M},$$

with a constant $c > 0$, and is given by

$$\inf_{g \in \mathcal{G}} Var[\Psi(g)] = E \left[\left(\int_0^T \sigma^2(\alpha_t) dt \right)^{-1} \right].$$

Proof. Since $W(g)$ and $\langle \sigma, g \rangle$ are independent, we have $E[\Psi(g)] = 0$ and

$$Var[\Psi(g)] = E[\Psi(g)^2]. \quad (4.26)$$

For any $g : \mathcal{M} \rightarrow \mathbf{R}$, we see that $\int_0^T \sigma(\alpha_t)g(\alpha_t)dt > 0$, with the isometry property of $W(g)$, we have

$$E[\Psi(g)^2] = E \left[\frac{\int_0^T g^2(\alpha_t)dt}{\left(\int_0^T \sigma(\alpha_t)g(\alpha_t)dt \right)^2} \right]. \quad (4.27)$$

By Cauchy-Schwarz inequality,

$$\left(\int_0^T \sigma(\alpha_t)g(\alpha_t)dt \right)^2 \leq \int_0^T \sigma^2(\alpha_t)dt \int_0^T g^2(\alpha_t)dt. \quad (4.28)$$

Therefore, by (4.26)-(4.28), we obtain that

$$Var[\Psi(g)] \geq E \left[\left(\int_0^T \sigma^2(\alpha_t)dt \right)^{-1} \right],$$

and equality of both sides is achieved if and only if $g(\alpha_t) = c\sigma(\alpha_t)$ with a constant $c > 0$ for all $t \in [0, T]$. $\square$

In the case of David and Privault (2009), it is proved that the best function chosen in the weight is the constant number. However, in our case, $\{g(i) = c\sigma(i) ; i \in \mathcal{M}\}$ obtains a faster convergence than $\{g(i) = 1 ; i \in \mathcal{M}\}$ does, as shown in the following graph (4.1).

4.3 Extension to non-differentiable payoff function

In this section, by a similar arguments as Kawai and Takeuchi (2011) we show by Proposition 4.3.1 that the conclusion in Proposition 4.2.1 is able to be extended for non-differentiable payoff function in the class $\Lambda(\mathbf{R})$

Figure 4.1: *Computation of Delta with respect to the number of samples.*

defined as in Kawai and Takeuchi (2011),

$$\Lambda(\mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \mid f = \sum_{i=1}^n f_i \mathbf{1}_{\{A_i\}}, n \geq 1, f_i \in C_L(\mathbb{R}; \mathbb{R}), \right. \\ \left. A_i \text{ are intervals of } \mathbb{R} \right\}, \quad (4.29)$$

where

$$C_L(\mathbb{R}) := \{f \in \mathcal{C}(\mathbb{R}; \mathbb{R}) \mid |f(x) - f(y)| \leq k|x - y| \text{ for some } k \geq 0\}.$$

Proposition 4.3.1. For any payoff function $\phi \in \Lambda(\mathbb{R})$, process $(X_t)_{t \in [0, T]}$ defined in (4.1) and any $g : \mathcal{M} \rightarrow (0, \infty)$, we have

$$\Delta = \frac{e^{-rT}}{x_0} E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right], \quad (4.30)$$

where $\langle \sigma, g \rangle = \int_0^T \sigma(\alpha_s) g(\alpha_s) ds > 0$ a.s.

Proof. For any $\phi \in \Lambda(\mathbb{R})$, there exists $N \geq 1$ and a sequence $\{k_i \in$

$\mathbb{R}_+; i = 1, \dots, N\}$ and a list of disjoint sets $(A_1, \dots, A_N)$ such that

$$\phi(x) = \sum_{i=1}^N f_i(x) \mathbf{1}_{\{A_i\}}(x), \quad x \in \mathbb{R}, \quad (4.31)$$

where $f_i(x) \in C_L(\mathbb{R}; \mathbb{R})$ with

$$|f(x) - f(y)| \leq k_i |x - y|, \quad x, y \in A_i, \quad i = 1, \dots, N, \quad (4.32)$$

and

$$A_i = (a_{i-1}, a_i], \quad i = 1, \dots, N,$$

with $a_0 = -\infty$ and $a_N = \infty$.

(i) First, we prove that (4.30) holds for $\phi \in \Lambda(\mathbb{R}) \cap C^1(\mathbb{R})$. Recalling the definition (4.1) of $(X_t)_{t \in [0, T]}$, we let $X_0 = x_0 > 0$, and for any $\varepsilon \in \mathbb{R}$ we define $(X_t^\varepsilon)_{t \in [0, T]}$ by

$$X_t^\varepsilon := \frac{x_0 + \varepsilon}{x_0} X_t \quad t \in [0, T]. \quad (4.33)$$

Without letting ϕ bounded, we can also show that $\phi(X_T)$ is integrable. Since ϕ is continuous, by (4.32) we see that

$$\begin{aligned} & |\phi(X_T) - \phi(X_0)| \\ & \leq \max_{i \in 1, \dots, N} k_i \max_{i \in 1, \dots, N-1} |X_T - a_i| + \max_{i \in 1, \dots, N} k_i \max_{i \in 1, \dots, N-1} |x_0 - a_i| \\ & \quad + \sum_{i, j \in 1, \dots, N-1} |\phi(a_i) - \phi(a_j)| \\ & \leq \max_{i \in 1, \dots, N} k_i (X_T + \max_{i \in 1, \dots, N-1} |a_i|) + \max_{i \in 1, \dots, N} k_i \max_{i \in 1, \dots, N-1} |x_0 - a_i| \\ & \quad + \sum_{i, j \in 1, \dots, N-1} |\phi(a_i) - \phi(a_j)|, \end{aligned} \quad (4.34)$$

which is integrable, hence the integrability of $\phi(X_T)$ is proved. On the other hand, without the bounded derivative of ϕ , we also show that (4.10) holds, namely,

$$\lim_{\varepsilon \rightarrow 0} E \left[\frac{\phi(X_T^\varepsilon) - \phi(X_T)}{\varepsilon} \mid X_0 = x_0, \alpha_0 \right]$$

$$= E \left[\lim_{\varepsilon \rightarrow 0} \frac{\phi(X_T^\varepsilon) - \phi(X_T)}{\varepsilon} \mid X_0 = x_0, \alpha_0 \right]. \quad (4.35)$$

By (4.32) and definition (4.33) of X_T^ε we see that

$$\lim_{\varepsilon \rightarrow 0} \left| \frac{\phi(X_T^\varepsilon) - \phi(X_T)}{\varepsilon} \right| \leq \lim_{\varepsilon \rightarrow 0} \left(\max_{1 \leq i \leq n} k_i \right) \left| \frac{X_T^\varepsilon - X_T}{\varepsilon} \right| = \frac{1}{x_0} \max_{1 \leq i \leq n} k_i,$$

which is uniformly bounded. Therefore, (4.35) is proved by Lebesgue's dominated convergence theorem and we have

$$\begin{aligned} \frac{\partial}{\partial x_0} V(x_0, \beta_0) &= \lim_{\varepsilon \rightarrow 0} e^{-rT} E \left[\frac{\phi(X_T^\varepsilon) - \phi(X_T)}{\varepsilon} \mid X_0 = x_0, \alpha_0 \right] \\ &= e^{-rT} E \left[\lim_{\varepsilon \rightarrow 0} \frac{\phi(X_T^\varepsilon) - \phi(X_T)}{\varepsilon} \mid X_0 = x_0, \alpha_0 \right]. \end{aligned} \quad (4.36)$$

Relation (4.30) follows from (4.36) in the case $\phi \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R}; \mathbf{R})$ by repeating the arguments from (4.10) until the end of Subsection 4.2.1.

(ii) Finally, we extend from $\phi \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R})$ to the class $\Lambda(\mathbf{R})$. We will structure a sequence $(\phi_n)_{n \in \mathbf{N}} \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R})$ approaching to $\phi \in \Lambda(\mathbf{R})$. By e.g. Theorem 7.17 of Rudin (1976) or (3.6)- (3.7) in Kawai and Takeuchi (2011), it suffices to show that for all compact set $K \subset (0, \infty)$ we have

$$\lim_{n \rightarrow \infty} E[\phi_n(X_T) \mid X_0 = x_0, \alpha_0] = E[\phi(X_T) \mid X_0 = x_0, \alpha_0] \quad (4.37)$$

for any $x_0 \in K$ and

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{x_0 \in K} \left| \frac{\partial}{\partial x_0} E \left[\phi_n(X_T) \mid X_0 = x_0, \alpha_0 \right] - E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \right| \\ = 0. \end{aligned} \quad (4.38)$$

Since $\phi \in \Lambda(\mathbf{R})$ is continuous on every interval (a_{i-1}, a_i) , $i = 1, \dots, N$, there exists a pointwise increasing sequence $(\phi_n)_{n \in \mathbf{N}} \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R})$ such that

$$\lim_{n \rightarrow \infty} \phi_n(x) = \phi(x), \quad x \in \mathbf{R} \setminus \{a_0, a_1, \dots, a_N\}. \quad (4.39)$$

For any $x \in \mathbb{R} \setminus \{a_0, a_1, \dots, a_N\}$, there exist a $N_x > 0$ such that for $n \geq N_x$ we have $|\phi_n(x)| \leq |\phi(x)| + 1$. Since $(\phi_n)_{n \in \mathbb{N}} \in \Lambda(\mathbb{R}) \cap C^1(\mathbb{R})$, by (4.34) we see that, for any $n \geq N_x$,

$$\begin{aligned} |\phi_n(X_T)| &\leq |\phi_n(x)| + \max_{i \in 1, \dots, N} k_i (X_T + \max_{i \in 1, \dots, N-1} |a_i|) \\ &\quad + \max_{i \in 1, \dots, N} k_i \max_{i \in 1, \dots, N-1} |x - a_i| + \sum_{i, j \in 1, \dots, N-1} |\phi(a_i) - \phi(a_j)| \\ &\leq |\phi(x)| + 1 + \max_{i \in 1, \dots, N} k_i (X_T + \max_{i \in 1, \dots, N-1} |a_i|) \\ &\quad + \max_{i \in 1, \dots, N} k_i \max_{i \in 1, \dots, N-1} |x - a_i| + \sum_{i, j \in 1, \dots, N-1} |\phi(a_i) - \phi(a_j)|, \end{aligned}$$

which is integrable, hence we can apply the Lebesgue's dominated convergence theorem and by (4.39) we obtain (4.37). Regarding (4.38), it follows from (i) that (4.30) holds for any $\phi_n \in \Lambda(\mathbb{R}) \cap C^1(\mathbb{R})$, therefore

$$\begin{aligned} &\left| \frac{\partial}{\partial x_0} E \left[\phi_n(X_T) \mid X_0 = x_0, \alpha_0 \right] - E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \right|^2 \\ &= \left| E \left[\phi_n(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] - E \left[\phi(X_T) \frac{W(g)}{\langle \sigma, g \rangle} \mid X_0 = x_0, \alpha_0 \right] \right|^2 \\ &\leq \left(E \left[\left| (\phi(X_T) - \phi_n(X_T)) \frac{W(g)}{\langle \sigma, g \rangle} \right| \mid X_0 = x_0, \alpha_0 \right] \right)^2 \\ &\leq E \left[(\phi(X_T) - \phi_n(X_T))^2 \mid X_0 = x_0, \alpha_0 \right] E \left[\left| \frac{W(g)}{\langle \sigma, g \rangle} \right|^2 \mid X_0 = x_0, \alpha_0 \right], \end{aligned} \tag{4.40}$$

where by (4.8) we have

$$E \left[\left| \frac{W(g)}{\langle \sigma, g \rangle} \right|^2 \mid X_0 = x_0, \alpha_0 \right] < \infty.$$

Define a decreasing sequence $(f_n)_{n \in \mathbb{N}}$ of continuous functions:

$$f_n(x) := E \left[(\phi(X_T) - \phi_n(X_T))^2 \mid X_0 = x, \alpha_0 \right], \quad n \in \mathbb{N}, x \in K.$$

By Dini's theorem, $f_n(x)$ pointwise decreases to 0 on K . Therefore, by (4.40), we complete the proof for (4.38). $\square$

Chapter 5

Sensitivity analysis by integration by parts for Markov chains

In this chapter, we establish an integration by parts formula for the random functionals of a continuous-time Markov chain, based on partial differentiation with respect to jump times, while leaving the discrete-time embedded chain unchanged. As an application we compute sensitivities (Greeks) in asset price models with Markovian regime switching. In comparison with classical approaches based on the partial Malliavin calculus for Brownian motion, our method has a wider scope of application by taking both sources of Markovian and Gaussian noises into account. We also improve on some recent results in the literature on integration by parts for Markov chains.

Defining a gradient for the functionals based on Markov chains, we apply a similar technique as Wiener-Malliavin calculus to develop our version of integration by parts formula for Markov chain setting, c.f. Proposition 5.3.1 in Section 5.3. Besides, another proof of integration by parts formula for birth processes is provided by Section 5.6, in this proof we mainly

apply the basic techniques of fundamental calculus and achieve the same formula of integration by parts, in particular, we differentiate the joint density function of interjump times to compute the gradient. Also, the integration by parts formula is applied for sensitivity analysis regarding the initial price in a regime switching model, and finally we extend the results for non-differentiable payoff functions.

5.1 Poisson process setting

In this section we start by reviewing integration by parts for a standard Poisson process $(N_t)_{t \in \mathbb{R}_+}$ with jump times $(T_n)_{n \geq 1}$, and whose interjump times $(\tau_k)_{k \geq 1}$ are independent exponentially distributed random variables, cf. e.g. § 7.3 of Privault (2009).

Consider the space $\mathcal{S}_T$ of Poisson functionals of the form

$$F = g_0 \mathbf{1}_{\{N_T=0\}} + \sum_{k=1}^n \mathbf{1}_{\{N_T=k\}} g_k(T_1, \dots, T_k), \quad (5.1)$$

under the continuity condition

$$g_k(T_1, \dots, T_k) = g_{k+1}(T_1, \dots, T_k, t), \quad k = 1, \dots, n-1, \quad (5.2)$$

for all $t \geq T$, where $g_0 \in \mathbb{R}$ and $g_k \in \mathcal{C}_b^1(\mathbb{R}_+^k)$, $k = 1, \dots, n$.

We define the operator D on $F \in \mathcal{S}_T$ of the form (5.1) by

$$D_t F = - \sum_{k=1}^n \mathbf{1}_{\{N_T=k\}} \sum_{l=1}^k \mathbf{1}_{[0, T_l]}(t) \partial_l g_k(T_1, \dots, T_k), \quad (5.3)$$

where $\partial_l g_k$ denotes the partial derivative of g_k with respect to its l -th variable. We also let

$$\mathcal{U} = \left\{ \sum_{i=1}^n F_i f_i : f_1, \dots, f_n \in \mathcal{C}([0, T]), F_1, \dots, F_n \in \mathcal{S}_T, n \geq 1 \right\},$$

and we define δ on $\mathcal{U}$ by

$$\delta(Ff) = F \int_0^T f(t) d(N_t - t) - \int_0^T f(t) D_t F dt, \quad (5.4)$$

$F \in \mathcal{S}_T$, $f \in \mathcal{C}([0, T])$. The following duality relation (5.5) between D and δ relies on standard integration by parts and on the expression

$$E[F] = e^{-T} g_0 + e^{-T} \sum_{k=1}^n \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} g_k(t_1, \dots, t_k) dt_1 \cdots dt_k,$$

for $F \in \mathcal{S}_T$ of the form (5.1), cf. Proposition 7.3.3 of Privault (2009).

Proposition 5.1.1. The operators D defined by (5.3) and δ defined by (5.4) satisfy the duality relation

$$E \left[\int_0^T D_t F u_t dt \right] = E[F \delta(u)], \quad F \in \mathcal{S}_T, \quad u \in \mathcal{U}. \quad (5.5)$$

As a consequence of Proposition 5.1.1 the operators D and δ are closable and (5.5) extends to their closed domains $\text{Dom}(D)$ and $\text{Dom}(\delta)$.

Letting $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ denote the filtration generated by $(N_t)_{t \in \mathbb{R}_+}$ on the probability space $(\Omega, \mathbb{P})$, we note that for any $\mathcal{F}_T$ -measurable random variable $F \in L^2(\Omega)$ we have

$$\tilde{D}_t F = 0, \quad t \geq T,$$

and that δ coincides with the stochastic integral with respect to the compensated Poisson process, i.e.

$$\tilde{\delta}(u) = \int_0^\infty u_t d(N_t - t),$$

for all $\mathcal{F}_t$ -adapted square-integrable process $u \in L^2(\Omega \times \mathbb{R}_+)$.

5.2 Birth process setting

In this section we implement the construction of Privault (1995a), Privault (1995b), in which the gradient of functionals of a point process is

defined by a time change based on the intensity of the point process.

We consider a pure birth process $(N_t^\alpha)_{t \in \mathbb{R}_+}$ whose interjump times $(\tau_k^\alpha)_{k \in \mathbb{N}}$ are independent exponentially distributed random variables with respective parameters $(\alpha_k)_{k \in \mathbb{N}}$ satisfying the condition (5.18).

In other words, the jump times $T_n^\alpha = \tau_0^\alpha + \dots + \tau_{n-1}^\alpha$ of $(N_t^\alpha)_{t \in \mathbb{R}_+}$ can be written as

$$T_n^\alpha = \sum_{k=0}^{n-1} \frac{\tau_k}{\alpha_k}, \quad n \geq 1, \quad (5.6)$$

and

$$N_t^\alpha = \min \{n \in \mathbb{N} : t < T_{n+1}^\alpha\}, \quad t \in \mathbb{R}_+. \quad (5.7)$$

Defining

$$\lambda(t) := \int_0^t \alpha_{N_s^\alpha} ds, \quad t \in \mathbb{R}_+, \quad (5.8)$$

it follows from (5.6)-(5.8) that N_t^α can be written as the time-changed Poisson process

$$N_t^\alpha = N_{\lambda(t)}, \quad t \in \mathbb{R}_+, \quad (5.9)$$

and we have

$$\lambda(T_n^\alpha) = T_n, \quad n \in \mathbb{N}. \quad (5.10)$$

We will use the time-changed gradient D_t^α defined as

$$D_t^\alpha F := \alpha_{N_t^\alpha} D_{\lambda(t)} F, \quad t \in \mathbb{R}_+, \quad (5.11)$$

for F in the space $\mathcal{S}_T^\alpha$ of functionals of the form

$$F = f_0 \mathbf{1}_{\{N_T^\alpha=0\}} + \sum_{k=1}^n \mathbf{1}_{\{N_T^\alpha=k\}} g_k(T_1^\alpha, \dots, T_k^\alpha), \quad (5.12)$$

under the continuity condition

$$g_k(T_1^\alpha, \dots, T_k^\alpha) = g_{k+1}(T_1^\alpha, \dots, T_k^\alpha, t), \quad k = 0, 1, \dots, n-1.$$

for all $t \geq T$, where $g_0 \in \mathbb{R}$ and $g_k \in \mathcal{C}_b^1(\mathbb{R}_+^k)$, $k = 1, \dots, n$.

We note that given $F \in \mathcal{S}_T^\alpha$ of the form (5.12), the function f_n defined by

$$f_n(t_1, \dots, t_n) := g_0 \mathbf{1}_{\{T < t_1\}} + \sum_{k=1}^n \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_k \leq T < t_{k+1} \leq \dots \leq t_n\}} g_k(t_1, \dots, t_k),$$

is weakly differentiable and satisfies

$$F = f_n(T_1^\alpha, \dots, T_n^\alpha), \quad (5.13)$$

i.e. for all $k = 0, 1, \dots, n-1$ we have

$$g_k(T_1^\alpha, \dots, T_k^\alpha) = f_n(T_1^\alpha, \dots, T_k^\alpha, t_{k+1}, \dots, t_n), \quad T \leq t_{k+1} \leq \dots \leq t_n.$$

Proposition 5.2.1. For $F \in \mathcal{S}_T^\alpha$ of the form (5.13) we have

$$D_t^\alpha F = - \sum_{k=1}^n \mathbf{1}_{[0, T_k^\alpha]}(t) \partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) \quad t \in \mathbb{R}_+.$$

Proof. By (5.6) we have

$$\begin{aligned} f_n(T_1^\alpha, \dots, T_n^\alpha) &= h_n(\alpha_0 T_1^\alpha, \alpha_0 T_1^\alpha + \alpha_1(T_2^\alpha - T_1^\alpha), \dots, \\ &\quad \sum_{m=1}^n \alpha_{m-1}(T_m^\alpha - T_{m-1}^\alpha)) \\ &= h_n(T_1, \dots, T_n), \end{aligned}$$

where h_n is the function

$$h_n(x_1, \dots, x_n) := f_n\left(\frac{1}{\alpha_0}x_1, \frac{1}{\alpha_0}x_1 + \frac{1}{\alpha_1}(x_2 - x_1), \dots, \sum_{m=1}^n \frac{1}{\alpha_{m-1}}(x_m - x_{m-1})\right),$$

i.e.

$$f_n(y_1, \dots, y_n) = h_n\left(\alpha_0 y_1, \alpha_0 y_1 + \alpha_1(y_2 - y_1), \dots, \sum_{m=1}^n \alpha_{m-1}(y_m - y_{m-1})\right). \quad (5.14)$$

By (5.6) and (5.14), we see that

$$\partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) \quad (5.15)$$

$$\begin{aligned}
&= \alpha_{k-1} \sum_{i=k}^n \partial_i h_n \left(\alpha_0 T_1^\alpha, \alpha_0 T_1^\alpha + \alpha_1 (T_2^\alpha - T_1^\alpha), \dots, \sum_{m=1}^n \alpha_{m-1} (T_m^\alpha - T_{m-1}^\alpha) \right) \\
&\quad - \mathbf{1}_{\{k < n\}} \alpha_k \sum_{i=k+1}^n \partial_i h_n (\alpha_0 T_1^\alpha, \alpha_0 T_1^\alpha + \alpha_1 (T_2^\alpha - T_1^\alpha), \dots, \\
&\quad \sum_{m=1}^n \alpha_{m-1} (T_m^\alpha - T_{m-1}^\alpha)) \tag{5.16} \\
&= \alpha_{k-1} \sum_{i=k}^n \partial_i h_n (T_1, T_2, \dots, T_n) - \mathbf{1}_{\{k < n\}} \alpha_k \sum_{i=k+1}^n \partial_i h_n (T_1, T_2, \dots, T_n)
\end{aligned}$$

for $k = 1, \dots, n$. Hence by (5.15) we have

$$\begin{aligned}
&\sum_{k=1}^n \mathbf{1}_{[0, T_k^\alpha]}(t) \partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) = \sum_{k=1}^n \mathbf{1}_{(T_{k-1}^\alpha, T_k^\alpha]}(t) \sum_{j=1}^n \partial_j f_n(T_1^\alpha, \dots, T_n^\alpha) \\
&= \sum_{k=1}^n \mathbf{1}_{(T_{k-1}^\alpha, T_k^\alpha]}(t) \left(\sum_{j=1}^{n-1} \left(\alpha_{j-1} \sum_{i=j}^n \partial_i h_n(T_1, \dots, T_n) \right. \right. \\
&\quad \left. \left. - \alpha_j \sum_{i=j+1}^n \partial_i h_n(T_1, \dots, T_n) \right) + \alpha_{n-1} \partial_n h_n(T_1, \dots, T_n) \right) \\
&= \sum_{k=1}^n \mathbf{1}_{(T_{k-1}^\alpha, T_k^\alpha]}(t) \alpha_{k-1} \sum_{i=k}^n \partial_i h_n(T_1, \dots, T_n).
\end{aligned} \tag{5.17}$$

Finally by (5.10), (5.11), (5.12), (5.14) and (5.17) we find

$$\begin{aligned}
D_t^\alpha F &= \alpha_{N_t^\alpha} D_{\lambda(t)} F \\
&= -\alpha_{N_t^\alpha} \sum_{k=1}^n \mathbf{1}_{[0, T_k^\alpha]}(\lambda(t)) \partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) \\
&= -\alpha_{N_t^\alpha} \sum_{k=1}^n \mathbf{1}_{[0, T_k^\alpha]}(t) \partial_k h_n(T_1, \dots, T_n) \\
&= -\alpha_{N_t^\alpha} \sum_{k=1}^n \mathbf{1}_{(T_{k-1}^\alpha, T_k^\alpha]}(t) \sum_{i=k}^n \partial_i h_n(T_1, \dots, T_n) \\
&= -\sum_{k=1}^n \mathbf{1}_{(T_{k-1}^\alpha, T_k^\alpha]}(t) \alpha_{k-1} \sum_{i=k}^n \partial_i h_n(T_1, \dots, T_n) \\
&= -\sum_{k=1}^n \mathbf{1}_{[0, T_k^\alpha]}(t) \partial_k f_n(T_1^\alpha, \dots, T_n^\alpha).
\end{aligned}$$

□

In the following proposition, cf. Proposition 6 of Privault (1995b) or Proposition 1 of Privault (1995a), we denote by $(\mathcal{F}_t^\alpha)_{t \in \mathbf{R}_+}$ the filtration generated by $(N_t^\alpha)_{t \in \mathbf{R}_+}$, and we let

$$\text{Dom}(D^\alpha) = \{F \in \text{Dom}(D) : D^\alpha \in L^2(\Omega \times [0, T])\}.$$

The stochastic integral

$$\int_0^T u_t(dN_t^\alpha - \alpha_{N_t^\alpha} dt)$$

is defined for $(\mathcal{F}_t^\alpha)_{t \in \mathbf{R}_+}$ -adapted processes $(u_t)_{t \in \mathbf{R}_+}$ via the isometry formula

$$E \left[\left(\int_0^T u_t(dN_t^\alpha - \alpha_{N_t^\alpha} dt) \right)^2 \right] = E \left[\int_0^T |u_t|^2 \alpha_{N_t^\alpha} dt \right].$$

Proposition 5.2.2. Given $F \in \text{Dom}(D^\alpha)$ and u a square-integrable $(\mathcal{F}_t^\alpha)_{t \in \mathbf{R}_+}$ -adapted process we have the integration by parts formula

$$E \left[\int_0^T D_t^\alpha F u_t dt \right] = E \left[F \int_0^T u_t(dN_t^\alpha - \alpha_{N_t^\alpha} dt) \right].$$

Proof. We consider $F \in \mathcal{S}_T^\alpha$ of the form

$$F = f_n(T_1^\alpha, \dots, T_n^\alpha) = g_n(T_1, \dots, T_n).$$

To apply Proposition 5.1.1 to F we need to check that g_n satisfies the continuity condition (5.2) on the domain $[0, CT]$. Indeed, for any $r \geq C$ we have

$$\begin{aligned} g_n(x_1, \dots, x_n) &= f_n \left(\frac{1}{\alpha_0} x_1, \dots, \sum_{m=1}^n \frac{1}{\alpha_{m-1}} (x_m - x_{m-1}) \right) \\ &= f_{n+1} \left(\frac{1}{\alpha_0} x_1, \dots, \sum_{m=1}^n \frac{1}{\alpha_{m-1}} (x_m - x_{m-1}), \sum_{m=1}^n \frac{1}{\alpha_{m-1}} (x_m - x_{m-1}) \right. \\ &\quad \left. + \frac{1}{\alpha_m} (rT - x_m) \right) \end{aligned}$$

$$= g_{n+1}(x_1, \dots, x_n, rT),$$

for all $x_1, \dots, x_n \in [0, T]$ and $F \in \mathcal{S}_T$, since from (5.18) we have

$$\sum_{m=1}^n \frac{1}{\alpha_{m-1}}(x_m - x_{m-1}) + \frac{1}{\alpha_m}(rT - x_m) \geq \frac{rT}{C} \geq T, \quad n \geq 1.$$

Since $\lambda(T) \leq CT$ by (5.8), (5.11), (5.18) and Proposition 5.1.1 applied on $[0, CT]$, for all $(\mathcal{F}_t^\alpha)_{t \in \mathbb{R}_+}$ -adapted u in

$$\mathcal{U}^\alpha := \left\{ \sum_{i=1}^n F_i u_i : u_1, \dots, u_n \in \mathcal{C}([0, T]), F_1, \dots, F_n \in \mathcal{S}_T^\alpha, n \geq 1 \right\},$$

we have

$$\begin{aligned} E \left[\int_0^T D_t^\alpha F u_t dt \right] &= E \left[\int_0^\infty \mathbf{1}_{[0, T]}(\lambda_t^{-1}) D_t F u_{\lambda_t^{-1}} dt \right] \\ &= E \left[\int_0^{CT} \mathbf{1}_{[0, \lambda(T)]}(t) D_t F u_{\lambda_t^{-1}} dt \right]. \\ &= E \left[F \int_0^{CT} \mathbf{1}_{[0, \lambda(T)]}(t) u_{\lambda_t^{-1}} (dN_t - dt) \right] \\ &= E \left[F \int_0^\infty \mathbf{1}_{[0, T]}(\lambda_t^{-1}) u_{\lambda_t^{-1}} (dN_t - dt) \right] \\ &= E \left[F \int_0^T u_t (dN_{\lambda(t)} - \alpha_{N_t^\alpha} dt) \right]. \end{aligned}$$

□

5.3 Markov chain setting

Consider a continuous-time Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ with state space $\{1, \dots, m\}$ and transition rate matrix $Q = (q_{i,j})_{1 \leq i, j \leq m}$. Let T_n^β denote the n -th jump time of $(\beta_t)_{t \in [0, T]}$ with $T_0^\beta = 0$, and let

$$N_t^\beta := \min \left\{ n \in \mathbb{N} : t < T_{n+1}^\beta \right\}$$

denote the number of jumps of $(\beta_t)_{t \in [0, T]}$ up to time $t \in \mathbb{R}_+$. The Markov chain $(\beta_t)_{t \in \mathbb{R}_+}$ can be represented as

$$\beta_t = Z_{N_t^\beta}, \quad t \in [0, T], \quad \text{where} \quad Z_n := \beta_{T_n^\beta}, \quad n \in \mathbb{N},$$

denotes the embedded chain of $(\beta_t)_{t \in \mathbb{R}_+}$, with $Z_0 = \beta_0$. In addition it is known that, given $(Z_n)_{n \in \mathbb{N}}$, the interjump times $(T_{n+1}^\beta - T_n^\beta)_{n \in \mathbb{N}}$ form a sequence of independent exponentially distributed random variables with respective parameters q_{Z_n} , where

$$q_n := \sum_{\substack{l=1 \\ l \neq n}}^r q_{n,l} = -q_{n,n}, \quad n \in \{1, \dots, m\}.$$

In the sequel we assume that the sequence $(\alpha_k)_{k \geq 0}$ satisfies the bound

$$0 < \alpha_k \leq C, \quad k \in \mathbb{N}. \quad (5.18)$$

for some $C > 0$. In this section we will use the partial gradient

$$D_t^\beta F := q_{\beta_t} D_{\lambda(t)} F, \quad t \in \mathbb{R}_+, \quad (5.19)$$

defined on the space $\mathcal{S}_T^\beta$ of functionals of the form

$$F = f_0 \mathbf{1}_{\{N_T^\beta=0\}} + \sum_{k=1}^n \mathbf{1}_{\{N_T^\beta=k\}} f_k(T_1^\beta, \dots, T_k^\beta, Z_1, \dots, Z_k), \quad (5.20)$$

where $f_0 \in \mathbb{R}$ and $f_k(\cdot, z_1, \dots, z_k) \in \mathcal{C}_b^1(\mathbb{R}_+^k)$, $z_1, \dots, z_k \in \mathcal{M}$, $k = 1, \dots, n$, $n \geq 1$. From Proposition 5.2.2 we obtain the following integration by parts formula, in which $\text{Dom}(D^\beta)$ and the filtration $(\mathcal{F}_t^\beta)_{t \in \mathbb{R}_+}$ are defined analogously to $\text{Dom}(D^\alpha)$ and $(\mathcal{F}_t^\alpha)_{t \in \mathbb{R}_+}$. As above, the stochastic integral

$$\int_0^T u_t (dN_t^\beta - q_{\beta_t} dt)$$

is defined for $(\mathcal{F}_t^\beta)_{t \in \mathbb{R}_+}$ -adapted processes $(u_t)_{t \in \mathbb{R}_+}$ via the isometry

$$E \left[\left(\int_0^T u_t (dN_t^\beta - q_{\beta_t} dt) \right)^2 \right] = E \left[\int_0^T |u_t|^2 q_{\beta_t} dt \right].$$

Proposition 5.3.1. Given $F \in \text{Dom}(D^\beta)$ and u a square-integrable $(\mathcal{F}_t^\beta)_{t \in \mathbb{R}_+}$ -adapted process we have the integration by parts formula

$$E \left[\int_0^T D_t^\beta F u_t dt \right] = E \left[F \int_0^T u_t (dN_t^\beta - q_{\beta_t} dt) \right]. \quad (5.21)$$

5.4 Sensitivity analysis in a regime switching model

In this section we apply the integration by parts formula (5.21) to the computation of Greeks. Consider a geometric Brownian motion $(Y_t)_{t \in [0, T]}$ with regime switching, given by the SDE

$$dY_t = Y_t \mu(\beta_t) dt + Y_t \sigma(\beta_t) dB_t, \quad 0 \leq t \leq T, \quad Y_0 = y > 0, \quad (5.22)$$

where $(B_t)_{t \in \mathbb{R}_+}$ is a standard Brownian motion and $(\beta_t)_{t \in [0, T]}$ is a finite-state continuous-time Markov chain independent of $(B_t)_{t \in \mathbb{R}_+}$, and $\mu : \mathcal{M} \rightarrow \mathbb{R}$, $\sigma : \mathcal{M} \rightarrow (0, \infty)$ are deterministic functions on the state space $\mathcal{M} := \{1, 2, \dots, m\}$ of $(\beta_t)_{t \in [0, T]}$.

Consider the value function

$$V(y, b) := e^{-rT} E[\phi(Y_T) \mid Y_0 = y, \beta_0 = b], \quad y > 0,$$

where $r > 0$ denotes the risk-free rate. In Proposition 5.4.1 we compute the sensitivity

$$\Delta := \frac{\partial V}{\partial y}(y, b) \quad Y_0 > 0, \quad \beta_0 \in \mathcal{M}, \quad (5.23)$$

with respect to the initial price y when the payoff function $\phi \in \mathcal{C}_b^1(\mathbb{R})$ has bounded derivative (an extension to non-differentiable payoff functions is given in Section 5.5). For convenience we will use the following notation:

$$W_\sigma := \int_0^T \sigma(\beta_u) dB_u, \quad \|\sigma\|^2 := \int_0^T \sigma^2(\beta_u) du, \quad T_\sigma := \int_0^T \sigma(\beta_u) du. \quad (5.24)$$

In the next proposition the adapted process $(u_t)_{t \in [0, T]}$ can be taken of the form

$$u_t := \sum_{k=1}^{N_T^\beta} \eta(k, \beta_k) (t - T_{k-1}^\beta)^\gamma \mathbf{1}_{(T_{k-1}^\beta, T_k^\beta]}(t), \quad (5.25)$$

for some $\gamma \in (-1, -2/3)$. Examples of functions $\eta(k, z)$ include $\eta(k, z) = \mathbf{1}_{\{z=z_0\}}$ where $z_0 \in \mathcal{M}$ is a given state, and $\eta(k, z) = \mathbf{1}_{\{k=1\}}$. In particular, $\eta(k, z) = \eta > 0$ can be taken constant in the two-state case.

Proposition 5.4.1. Assume that $\phi \in \mathcal{C}_b^1(\mathbb{R})$ and consider an adapted process $(u_t)_{t \in [0, T]}$ such that

$$T_\mu := \int_0^T u_t(\mu(\beta_T) - \mu(\beta_t))dt$$

satisfies $E[\mathbf{1}_{\{|T_\mu|>0\}}|T_\mu|^{-3}] < \infty$. Then we have

$$\Delta = \frac{e^{-rT}}{y} E \left[\phi(Y_T) \left(\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \left(\int_0^T u_t dN_t^\beta - \int_0^T v_t dt \right) + \mathbf{1}_{\{T_\mu=0\}} \frac{B_T}{T_\sigma} \right) \middle| Y_0 = y, \beta_0 = b \right],$$

where

$$v_t := \left(q_{\beta_t} - \mathbf{1}_{\{|T_\mu|>0\}} \frac{u_t}{T_\mu} (\mu(\beta_T) - \mu(\beta_t)) + \frac{(1 + W_\sigma) \|\sigma\|^2 - W_\sigma^2 (\sigma^2(\beta_t) - \sigma^2(\beta_T))}{2\|\sigma\|^4} \right) u_t,$$

$t \in [0, T]$.

Proof. We write Δ as $\Delta = \Delta_1 + \Delta_2$, where

$$\Delta_1 := e^{-rT} \frac{\partial}{\partial y} E [\phi(Y_T) \mathbf{1}_{\{|T_\mu|>0\}} \mid Y_0 = y, \beta_0 = b] \quad (5.26)$$

and

$$\Delta_2 := e^{-rT} \frac{\partial}{\partial y} E [\phi(Y_T) \mathbf{1}_{\{T_\mu=0\}} \mid Y_0 = y, \beta_0 = b]. \quad (5.27)$$

Since $\phi \in \mathcal{C}_b^1$ and has bounded derivative, taking $Y_0 = y$ and $\beta_0 = b$ we have

$$\begin{aligned} \Delta_1 &= e^{-rT} \frac{\partial}{\partial y} E [\mathbf{1}_{\{|T_\mu|>0\}} \phi(Y_T)] \\ &= e^{-rT} E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{\partial}{\partial y} \phi(Y_T) \right] \end{aligned} \quad (5.28)$$

$$= e^{-rT} E \left[\mathbf{1}_{\{|T_\mu| > 0\}} \int_{-\infty}^{\infty} \frac{\partial}{\partial y} \phi(y e^{X_T+x}) \varphi_{\|\sigma\|}(x) dx \right], \quad (5.29)$$

where

$$X_T := \int_0^T \mu(\beta_u) du = \sum_{i=1}^{N_T^\beta} \mu(Z_{i-1})(T_i - T_{i-1}) + \mu(Z_{N_T^\beta})(T - T_{N_T^\beta}), \quad (5.30)$$

and

$$\varphi_{\|\sigma\|}(x) := \frac{1}{\sqrt{2\pi\|\sigma\|^2}} \exp \left(-\frac{1}{2\|\sigma\|^2} (x + \|\sigma\|^2/2)^2 \right), \quad x \in \mathbf{R}, \quad (5.31)$$

is the probability density function of $W_\sigma - \|\sigma\|^2/2$ given $\|\sigma\|^2$. For any $F \in \mathcal{S}_T^\beta$ in form of (5.20), let

$$D_u^\beta F := \int_0^T u_t D_t^\beta F dt. \quad (5.32)$$

Note that X_T in (5.30) is expressed in form of (5.20), by Proposition 5.2.1 we have

$$D_t^\beta \phi(y e^{X_T+x}) = y e^{X_T+x} \phi'(y e^{X_T+x}) D_t^\beta X_T, \quad (5.33)$$

and

$$\begin{aligned} D_t^\beta X_T &= - \sum_{k=1}^{N_T^\beta} \mathbf{1}_{[0, T_k]}(t) (\mu(Z_{k-1}) - \mu(Z_k)) \\ &= - \sum_{k=1}^{N_T^\beta} \mathbf{1}_{(T_{k-1}, T_k]}(t) (\mu(Z_{k-1}) - \mu(Z_{N_T^\beta})) \\ &= \mu(\beta_T) - \mu(\beta_t). \end{aligned} \quad (5.34)$$

By similar arguments, using the notation D_u^β we have

$$\begin{aligned} D_u^\beta e^{X_T+x} &= e^{X_T+x} T_\mu, \quad D_u^\beta T_\mu = - \int_0^T u_t^2 (\mu(\beta_T) - \mu(\beta_t)) dt, \\ D_u^\beta \mathbf{1}_{\{|T_\mu| > 0\}} &= 0, \quad a.e. \end{aligned} \quad (5.35)$$

where we note $\mathbf{1}_{\{|T_\mu|=0\}} = \mathbf{1}_{\{N_T^\beta=0\}} + \sum_{k=1}^n \mathbf{1}_{\{N_T^\beta=k\}} \mathbf{1}_{\{Z_1=\dots=Z_k\}}$ a.e., which is in form of (5.20), besides we have

$$D_t^\beta \varphi_{\|\sigma\|}(x) = \frac{\|\sigma\|^2 - x^2 + \|\sigma\|^4/4}{2\|\sigma\|^4} (\sigma^2(\beta_t) - \sigma^2(\beta_T)) \varphi_{\|\sigma\|}(x), \quad t \in \mathbf{R}_+. \quad (5.36)$$

By (5.33)-(5.35), taking $Y_0 = y$ and $\beta_0 = b$ we have

$$\begin{aligned} E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{\partial}{\partial y} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \right] &= E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{D_u^\beta \phi(ye^{X_T+x})}{D_u^\beta(ye^{X_T+x})} \varphi_{\|\sigma\|}(x) \right] \\ &= \frac{1}{y} E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \varphi_{\|\sigma\|}(x) D_u^\beta \phi(ye^{X_T+x}) \right] \\ &= \frac{1}{y} E^{y,b} \left[D_u^\beta \left(\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \right) \right. \\ &\quad \left. - \phi(ye^{X_T+x}) D_u^\beta \left(\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \varphi_{\|\sigma\|}(x) \right) \right]. \end{aligned} \quad (5.37)$$

Denoting by $(\mathcal{F}_t^B)_{t \in [0, T]}$ the filtration generated by $(B_t)_{t \in [0, T]}$, by Proposition 5.3.1 we obtain

$$\begin{aligned} &E \left[D_u^\beta \left(\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \right) \right] \\ &= E \left[E \left[D_u^\beta \left(\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \right) \middle| \mathcal{F}_T^B \right] \right] \\ &= E \left[E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \int_0^T u_t (dN_t^\beta - q_{\beta_t} dt) \middle| \mathcal{F}_T^B \right] \right] \\ &= E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \int_0^T u_t (dN_t^\beta - q_{\beta_t} dt) \right]. \end{aligned} \quad (5.38)$$

By (5.33)-(5.36), we see that

$$\begin{aligned} &E \left[\phi(ye^{X_T+x}) D_u^\beta \left(\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \varphi_{\|\sigma\|}(x) \right) \right] \\ &= E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} D_u^\beta \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) + \mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu^2} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) D_u^\beta T_\mu \right] \\ &= E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{\phi(ye^{X_T+x})}{2T_\mu \|\sigma\|^4} (\|\sigma\|^2 - x^2 + \|\sigma\|^4/4) \varphi_{\|\sigma\|}(x) \int_0^T (\sigma^2(\beta_t) - \sigma^2(\beta_T)) u_t dt \right. \\ &\quad \left. - \mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu^2} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \int_0^T u_t (\mu(\beta_T) - \mu(\beta_t)) dt \right]. \end{aligned} \quad (5.39)$$

Combining (5.29), (5.37)-(5.39) and applying Fubini's theorem, we have

$$\begin{aligned}
\Delta_1 &= \frac{e^{-rT}}{y} \int_{-\infty}^{\infty} E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(ye^{X_T+x}) \varphi_{\|\sigma\|}(x) \left(\int_0^T u_t dN_t^\beta - \int_0^T u_t q_{\beta_t} dt \right. \right. \\
&\quad \left. \left. + \frac{1}{T_\mu} \int_0^T u_t (\mu(\beta_T) - \mu(\beta_t)) dt \right. \right. \\
&\quad \left. \left. - \frac{\|\sigma\|^2 - x^2 + \|\sigma\|^4/4}{2\|\sigma\|^4} \int_0^T (\sigma^2(\beta_t) - \sigma^2(\beta_T)) u_t dt \right) \right] dx \\
&= \frac{e^{-rT}}{y} E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(Y_T) \left(\int_0^T u_t dN_t^\beta - \int_0^T u_t q_{\beta_t} dt \right. \right. \\
&\quad \left. \left. + \frac{1}{T_\mu} \int_0^T u_t (\mu(\beta_T) - \mu(\beta_t)) dt \right. \right. \\
&\quad \left. \left. - \frac{(1 + W_\sigma)\|\sigma\|^2 - W_\sigma^2}{2\|\sigma\|^4} \int_0^T (\sigma^2(\beta_t) - \sigma^2(\beta_T)) u_t dt \right) \right] \\
&= \frac{e^{-rT}}{y} E \left[\mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \phi(Y_T) \left(\int_0^T u_t dN_t^\beta - \int_0^T v_t dt \right) \right], \tag{5.40}
\end{aligned}$$

where

$$\begin{aligned}
v_t &:= \left(q_{\beta_t} - \mathbf{1}_{\{|T_\mu|>0\}} \frac{u_t (\mu(\beta_T) - \mu(\beta_t))}{T_\mu} \right. \\
&\quad \left. + \frac{(1 + W_\sigma)\|\sigma\|^2 - W_\sigma^2}{2\|\sigma\|^4} (\sigma^2(\beta_t) - \sigma^2(\beta_T)) \right) u_t,
\end{aligned}$$

$t \in [0, T]$. Regarding the computation of Δ_2 we note that by the classical Malliavin calculus for Brownian motion we have

$$\Delta_2 = e^{-rT} \frac{\partial}{\partial y} E [\mathbf{1}_{\{T_\mu=0\}} \phi(Y_T)] = \frac{e^{-rT}}{y} E \left[\mathbf{1}_{\{T_\mu=0\}} \phi(Y_T) \frac{B_T}{T_\sigma} \right]. \tag{5.41}$$

The proof is complete by (5.40)-(5.41). $\square$

To end this section, we carry on a numerical simulation to compare the computation by Proposition 4.2.1 in Chapter 4 and that by Proposition 5.4.1. As shown by the Figure 5.1 below, the former (in blue) outperforms the latter (in red) since the blue line reaches a stable value by less than 30000 samples while the red one approaches it by consuming 50000 samples.

Figure 5.1: *Computation of Delta with respect to the number of samples.*

5.5 Extension to non-differentiable payoff functions

In this section, following the approach of in Kawai and Takeuchi (2011) we show that Proposition 5.4.1 can be extended to non-differentiable payoff functions in the class

$$\Lambda(\mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} : f = \sum_{i=1}^n f_i \mathbf{1}_{A_i}, n \geq 1, f_i \in C_L(\mathbb{R}), A_i \text{ intervals of } \mathbb{R} \right\}, \quad (5.42)$$

where

$$C_L(\mathbb{R}) := \{f \in C(\mathbb{R}) : |f(x) - f(y)| \leq k|x - y| \text{ for some } k \geq 0\}. \quad (5.43)$$

In Proposition 5.5.1 the adapted process $(u_t)_{t \in [0, T]}$ can also be taken of the form (5.25) with $\gamma \in (-1, -2/3)$.

Proposition 5.5.1. Assume that $\phi \in \Lambda(\mathbf{R})$ and consider an adapted process $(u_t)_{t \in [0, T]}$ such that $E[\mathbf{1}_{\{|T_\mu| > 0\}} |T_\mu|^{-3}] < \infty$. Then we have

$$\Delta = \frac{e^{-rT}}{y} E \left[\phi(Y_T) \left(\mathbf{1}_{\{|T_\mu| > 0\}} \frac{1}{T_\mu} \left(\int_0^T u_t dN_t^\beta - \int_0^T v_t dt \right) + \mathbf{1}_{\{T_\mu = 0\}} \frac{B_T}{T_\sigma} \right) \middle| Y_0 = y, \beta_0 = b \right]. \quad (5.44)$$

Proof. Since $\phi \in \Lambda(\mathbf{R})$, there exists $N \geq 1$ and a sequence $(k_1, \dots, k_N) \subset \mathbf{R}_+^N$ and a family $(A_1, \dots, A_N)$ of disjoint sets such that

$$\phi(x) = \sum_{i=1}^N f_i(x) \mathbf{1}_{A_i}(x), \quad x \in \mathbf{R},$$

where $f_i(x) \in C_L(\mathbf{R})$ with

$$|f(x) - f(y)| \leq k_i |x - y|, \quad x, y \in A_i, \quad i = 1, \dots, N.$$

We denote $A_i = (a_{i-1}, a_i]$, $i = 1, \dots, N$ with $a_0 = -\infty$ and $a_N = \infty$. Let

$$Y_t^\varepsilon := \frac{y + \varepsilon}{y} Y_t \quad t \in [0, T], \quad \varepsilon \in \mathbf{R},$$

where $(Y_t)_{t \in [0, T]}$ is defined in (5.22) with $Y_0 = y > 0$.

(i) Assuming $\phi \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R})$ we show that $\phi(Y_T)$ is integrable, with

$$\lim_{\varepsilon \rightarrow 0} E \left[\frac{\phi(Y_T^\varepsilon) - \phi(Y_T)}{\varepsilon} \middle| Y_0 = y, \beta_0 = b \right] \quad (5.45)$$

$$= E \left[\lim_{\varepsilon \rightarrow 0} \frac{\phi(Y_T^\varepsilon) - \phi(Y_T)}{\varepsilon} \middle| Y_0 = y, \beta_0 = b \right]. \quad (5.46)$$

Since ϕ is continuous, we see that

$$\begin{aligned} & |\phi(Y_T) - \phi(Y_0)| \\ & \leq \max_{i=1, \dots, N} k_i \max_{i=1, \dots, N-1} |Y_T - a_i| + \max_{i=1, \dots, N} k_i \max_{i=1, \dots, N-1} |y - a_i| \\ & + \sum_{i,j=1, \dots, N-1} |\phi(a_i) - \phi(a_j)| \\ & \leq \max_{i=1, \dots, N} k_i (Y_T + \max_{i=1, \dots, N-1} |a_i|) + \max_{i=1, \dots, N} k_i \max_{i=1, \dots, N-1} |y - a_i| \end{aligned}$$

$$+ \sum_{i,j=1,\dots,N-1} |\phi(a_i) - \phi(a_j)|,$$

which is integrable, hence the integrability of $\phi(Y_T)$ is proved. On the other hand,

$$\lim_{\varepsilon \rightarrow 0} \left| \frac{\phi(Y_T^\varepsilon) - \phi(Y_T)}{\varepsilon} \right| \leq \left(\max_{1 \leq i \leq N} k_i \right) \lim_{\varepsilon \rightarrow 0} \left| \frac{Y_T^\varepsilon - Y_T}{\varepsilon} \right| = \frac{1}{y} \max_{1 \leq i \leq N} k_i,$$

which is uniformly bounded. Therefore, (5.45) follows by dominated convergence.

(ii) Next, we note that (5.44) holds for $\phi \in \Lambda(\mathbb{R}) \cap C^1(\mathbb{R})$, as by (5.45) we have

$$\begin{aligned} \frac{\partial V}{\partial y}(y, b) &= e^{-rT} \lim_{\varepsilon \rightarrow 0} E \left[\frac{\phi(Y_T^\varepsilon) - \phi(Y_T)}{\varepsilon} \middle| Y_0 = y, \beta_0 = b \right] \\ &= e^{-rT} E \left[\lim_{\varepsilon \rightarrow 0} \frac{\phi(Y_T^\varepsilon) - \phi(Y_T)}{\varepsilon} \middle| Y_0 = y, \beta_0 = b \right] \\ &= e^{-rT} E \left[\mathbf{1}_{\{T_\mu > 0\}} \frac{\partial}{\partial y} \phi(Y_T) \middle| Y_0 = y, \beta_0 = b \right] \\ &\quad + e^{-rT} E \left[\mathbf{1}_{\{T_\mu = 0\}} \frac{\partial}{\partial y} \phi(Y_T) \middle| Y_0 = y, \beta_0 = b \right], \end{aligned}$$

which shows that (5.44) holds for $\phi \in \Lambda(\mathbb{R}) \cap C^1(\mathbb{R})$ by repeating the arguments from (5.28) to the end of proof of Proposition 5.4.1.

(iii) Finally, we extend the result from a sequence $(\phi_n)_{n \in \mathbb{N}} \subset \Lambda(\mathbb{R}) \cap C^1(\mathbb{R})$ to $\phi \in \Lambda(\mathbb{R})$. By e.g. Theorem 9.34 of Heinonen et al. (2000) or (3.6)-(3.7) in Kawai and Takeuchi (2011), it suffices to show that for all compact sets $K \subset (0, \infty)$ we have

$$\lim_{n \rightarrow \infty} \sup_{y \in K} |E[\phi_n(Y_T) \mid Y_0 = y, \beta_0 = b] - E[\phi(Y_T) \mid Y_0 = y, \beta_0 = b]| = 0, \quad (5.47)$$

and

$$\lim_{n \rightarrow \infty} \sup_{y \in K} \left| \frac{\partial}{\partial y} E[\phi_n(Y_T) \mid Y_0 = y, \beta_0 = b] - \frac{\partial}{\partial y} E[\phi(Y_T) \mid Y_0 = y, \beta_0 = b] \right|^2 = 0. \quad (5.48)$$

where the weight J given by

$$J := \mathbf{1}_{\{|T_\mu|>0\}} \frac{1}{T_\mu} \left(\int_0^T u_t dN_t^\beta - \int_0^T v_t dt \right) + \mathbf{1}_{\{T_\mu=0\}} \frac{B_T}{T_\sigma}$$

is square-integrable under the condition $E[\mathbf{1}_{\{|T_\mu|>0\}} |T_\mu|^{-3}] < \infty$. Since $\phi \in \Lambda(\mathbf{R})$ is continuous on every interval (a_{i-1}, a_i) , $i = 1, \dots, N$, there exists a pointwise increasing sequence $(\phi_n)_{n \in \mathbb{N}} \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R})$ such that

$$\lim_{n \rightarrow \infty} \phi_n(x) = \phi(x), \quad x \in \mathbf{R} \setminus \{a_1, \dots, a_{N-1}\}. \quad (5.49)$$

Consider the increasing sequence $(f_n)_{n \in \mathbb{N}}$ of continuous functions defined by

$$f_n(y) := E \left[\phi_n(Y_T) \middle| Y_0 = y, \beta_0 = b \right], \quad n \in \mathbb{N}, \quad y \in K,$$

which satisfies

$$\lim_{n \rightarrow \infty} f_n(y) = E \left[\phi(Y_T) \middle| Y_0 = y, \beta_0 = b \right], \quad y \in K,$$

uniformly in $y \in K$ by dominated convergence and Dini's theorem, which proves (5.47). Regarding (5.48), since $(\phi_n)_{n \in \mathbb{N}} \in \Lambda(\mathbf{R}) \cap C^1(\mathbf{R})$ and by (ii) we have

$$\begin{aligned} & \left| \frac{\partial}{\partial y} E[\phi_n(Y_T) \mid Y_0 = y, \beta_0 = b] - E[\phi(Y_T)J \mid Y_0 = y, \beta_0 = b] \right| \quad (5.50) \\ &= |E[\phi_n(Y_T)J \mid Y_0 = y, \beta_0 = b] - E[\phi(Y_T)J \mid Y_0 = y, \beta_0 = b]| \\ &\leq E[(\phi(Y_T) - \phi_n(Y_T))|J| \mid Y_0 = y, \beta_0 = b] \\ &\leq \sqrt{E[(\phi(Y_T) - \phi_n(Y_T))^2 \mid Y_0 = y, \beta_0 = b]} \sqrt{E[|J|^2 \mid Y_0 = y, \beta_0 = b]}. \end{aligned}$$

Similarly to the above we conclude by noting that sequence $(g_n)_{n \in \mathbb{N}}$ of continuous functions defined by

$$g_n(y) := E \left[(\phi(Y_T) - \phi_n(Y_T))^2 \middle| Y_0 = y, \beta_0 = b \right], \quad n \in \mathbb{N}, \quad y \in K.$$

is pointwise decreasing to 0, uniformly on K . □

5.6 Another proof of integration by parts formula for birth processes

For a pure birth process $(N_t^\alpha)_{t \in \mathbb{R}_+}$ whose interjump times $(\tau_k^\alpha)_{k \in \mathbb{N}}$ are independent exponentially distributed random variables with respective parameters $(\alpha_k)_{k \in \mathbb{N}}$ and the jump times $T_n^\alpha := \tau_0^\alpha + \cdots + \tau_{n-1}^\alpha$. For any F in the space $\mathcal{S}_T^\alpha$ of functionals of the form (5.12) that

$$F = f_0 \mathbf{1}_{\{N_T^\alpha=0\}} + \sum_{n=1}^m \mathbf{1}_{\{N_T^\alpha=n\}} f_n(T_1^\alpha, \dots, T_n^\alpha), \quad (5.51)$$

where $f_0 \in \mathbb{R}$, $m \geq 1$, and $f_n \in \mathcal{C}_b^1([0, T]^n)$, $1 \leq n \leq m$, with the continuity condition

$$f_n(T_1^\alpha, \dots, T_n^\alpha) = f_{n+1}(T_1^\alpha, \dots, T_n^\alpha, T).$$

Then perturb $(T_i^\alpha)_{i=1,2,\dots,k}$ by introducing $(T_i^{\alpha,\varepsilon})_{i=1,2,\dots,n}$ and $(\tau_i^{\alpha,\varepsilon})_{i=1,2,\dots,n}$ as follows,

$$T_i^{\alpha,\varepsilon} := T_i^\alpha + \varepsilon \int_0^{T_i^\alpha} u(s) ds, \quad \tau_i^{\alpha,\varepsilon} := T_{i+1}^{\alpha,\varepsilon} - T_i^{\alpha,\varepsilon}, \quad i \in \mathbb{N}, \quad (5.52)$$

with $T_0^{\alpha,\varepsilon} = T_0^\alpha = 0$, where $\varepsilon \in \mathbb{R}_+$ and u_t is a bounded function on $t \in [0, T]$. Define $F^\varepsilon \in \mathcal{S}_T^\alpha$ by

$$F^\varepsilon = f_0 \mathbf{1}_{\{N_T^\alpha=0\}} + \sum_{n=1}^m \mathbf{1}_{\{N_T^\alpha=n\}} f_n(T_1^{\alpha,\varepsilon}, \dots, T_n^{\alpha,\varepsilon}), \quad (5.53)$$

for any $\varepsilon \in \mathbb{R}_+$. Then we have the following Proposition 5.6.1 and 5.6.2.

Proposition 5.6.1. With the gradient $(D_t^\alpha; t \in \mathbb{R}_+)$ defined by (5.11), $F \in \mathcal{S}_T^\alpha$ given by (5.51), F^ε defined by (5.53), and bounded function $(u_t; t \in [0, T])$, we have,

$$-\frac{\partial}{\partial \varepsilon} E[F^\varepsilon] \Big|_{\varepsilon=0} = E \left[\int_0^T D_t^\alpha F u_t dt \right]. \quad (5.54)$$

Proof. By (5.53), we have

$$\begin{aligned} -\frac{\partial}{\partial \varepsilon} E[F^\varepsilon] \Big|_{\varepsilon=0} &= -\frac{\partial}{\partial \varepsilon} E \left[f_0 \mathbf{1}_{\{N_T^\alpha=0\}} + \sum_{n=1}^m \mathbf{1}_{\{N_T^\alpha=n\}} f_n(T_1^{\alpha,\varepsilon}, \dots, T_n^{\alpha,\varepsilon}) \right] \Big|_{\varepsilon=0} \\ &= -\sum_{n=1}^m E \left[\mathbf{1}_{\{N_T^\alpha=n\}} \frac{\partial}{\partial \varepsilon} f_n(T_1^{\alpha,\varepsilon}, \dots, T_n^{\alpha,\varepsilon}) \Big|_{\varepsilon=0} \right], \end{aligned} \quad (5.55)$$

where we can exchange the order of $\partial/\partial \varepsilon$ and E because $f_n \in \mathcal{C}^1([0, T]^n)$ which implies that $\partial f_n(T_1^{\alpha,\varepsilon}, \dots, T_n^{\alpha,\varepsilon})/\partial \varepsilon$ is uniformly bounded. By (5.52) and Proposition (5.2.1), continuing with (5.55) we have

$$\begin{aligned} -\frac{\partial}{\partial \varepsilon} E[F^\varepsilon] \Big|_{\varepsilon=0} &= -\sum_{n=1}^m E \left[\mathbf{1}_{\{N_T^\alpha=n\}} \sum_{k=1}^n \left(\partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) \int_0^{T_k^\alpha} u_t dt \right) \right] \\ &= -\sum_{n=1}^m E \left[\mathbf{1}_{\{N_T^\alpha=n\}} \sum_{k=1}^n \left(\partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) \int_0^T \mathbf{1}_{(0, T_k^\alpha)}(t) u_t dt \right) \right] \\ &= -E \left[\int_0^T \sum_{n=1}^m \mathbf{1}_{\{N_T^\alpha=n\}} \sum_{k=1}^n \mathbf{1}_{(0, T_k^\alpha)}(t) \partial_k f_n(T_1^\alpha, \dots, T_n^\alpha) u_t dt \right] \\ &= E \left[\int_0^T D_t^\alpha F u_t dt \right]. \end{aligned}$$

□

Proposition 5.6.2. For any F in form of (5.51), F^ε in form of (5.53), and a bounded function $(u_t; t \in [0, T])$, we have

$$E \left[\int_0^T D_t^\alpha F u_t dt \right] = E \left[F \left(\int_0^T u(t) dN_t - \int_0^T u(t) \alpha_{N_t^\alpha} dt \right) \right]. \quad (5.56)$$

Proof. By (5.51) we have

$$E[F] = f_0 P(N_T^\alpha = 0) + \sum_{k=1}^m E \left[\mathbf{1}_{\{N_T^\alpha=k\}} f_k(T_1, \dots, T_k) \right] = f_0 e^{-\alpha_0 T} + \sum_{k=1}^m \Upsilon^0(k), \quad (5.57)$$

where for any $k \in \{1, 2, \dots\}$, and $\varepsilon \in \mathbf{R}$, Υ^ε is defined for any $\varepsilon \in \mathbf{R}$ as follows,

$$\Upsilon^\varepsilon(k) := \int \cdots \int_{\mathbf{R}_+^{k+1}} \mathbf{1}_{\left\{ \sum_{i=0}^{k-1} y_i \leq T, \sum_{i=0}^k y_i \geq T \right\}} f_k \left(y_0, y_0 + y_1, \dots, \sum_{i=0}^{k-1} y_i \right)$$

$$\psi^\varepsilon(y_0, y_1, \dots, y_k) dy_0 \cdots dy_k, \quad (5.58)$$

where for any $k \in \{1, 2, \dots\}$, $\psi^\varepsilon(x_0, x_1, \dots, x_k)$ on $\mathbb{R}_+^{k+1}$ denotes the joint distribution of $(\tau_0^{\alpha, \varepsilon}, \tau_1^{\alpha, \varepsilon}, \dots, \tau_k^{\alpha, \varepsilon})$. Since $u \in \mathcal{U}$, u is a bounded function in $\mathbb{R}_+$ s.t. there exists a constant $K > 0$ that $|u(s)| < K$ for any $s \in \mathbb{R}_+$. Define

$$\kappa(x, \varepsilon, \theta) = x + \varepsilon \int_\theta^{x+\theta} u(s) ds$$

for $\varepsilon \in (-\frac{1}{K}, \frac{1}{K})$, $x \geq 0$, $\theta \geq 0$, note that $\kappa(x, \varepsilon, \theta)$ is increasing in x , by (5.57)-(5.58) we find that

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} E[F^\varepsilon] \Big|_{\varepsilon=0} &= \lim_{\varepsilon \rightarrow 0} \frac{E[F^\varepsilon] - E[F]}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{f_0 e^{-\alpha_0 \kappa^{-1}(T, \varepsilon, 0)} - f_0 e^{-\alpha_0 T}}{\varepsilon} + \lim_{\varepsilon \rightarrow 0} \frac{\sum_{k=1}^m [\Upsilon^\varepsilon(k) - \Upsilon^0(k)]}{\varepsilon} \\ &= \alpha_0 f_0 e^{-\alpha_0 T} \int_0^T u(s) ds + \sum_{k=1}^m \lim_{\varepsilon \rightarrow 0} \frac{\Upsilon^\varepsilon(k) - F(k)}{\varepsilon}. \end{aligned} \quad (5.59)$$

On the other hand, we apply Lemma 5.6.1 to get

$$\begin{aligned} &\frac{\partial}{\partial \varepsilon} \psi^\varepsilon(y_0, y_1, \dots, y_k) \Big|_{\varepsilon=0} \\ &= \frac{\partial}{\partial \varepsilon} \left(\prod_{j=0}^k \frac{\alpha_j e^{-\kappa^{-1}(y_j, \varepsilon, T_j) \alpha_j}}{1 + \varepsilon u(\kappa^{-1}(y_j, \varepsilon, T_j))} \right) \Big|_{\varepsilon=0} \\ &= \sum_{j=0}^k \left(\frac{\partial}{\partial \varepsilon} \left(\frac{\alpha_j e^{-\kappa^{-1}(y_j, \varepsilon, T_j) \alpha_j}}{1 + \varepsilon u(\kappa^{-1}(y_j, \varepsilon, T_j))} \right) \Big|_{\varepsilon=0} \left(\frac{\alpha_j e^{-\kappa^{-1}(y_j, \varepsilon, T_j) \alpha_j}}{1 + \varepsilon u(\kappa^{-1}(y_j, \varepsilon, T_j))} \right)^{-1} \Big|_{\varepsilon=0} \right) \\ &\quad \cdot \left(\prod_{j=0}^k \frac{\alpha_j e^{-\kappa^{-1}(y_j, \varepsilon, T_j) \alpha_j}}{1 + \varepsilon u(\kappa^{-1}(y_j, \varepsilon, T_j))} \right) \Big|_{\varepsilon=0} \\ &= \sum_{j=0}^k \left(-u \left(\sum_{i=0}^j y_i \right) + \alpha_j \int_{\sum_{i=0}^{j-1} y_i}^{\sum_{i=0}^j y_i} u(s) ds \right) \prod_{j=0}^k \alpha_j e^{-y_j \alpha_j}. \end{aligned} \quad (5.60)$$

Substituting (5.58) and (5.60), the second term of the above expression (5.59) becomes

$$\sum_{k=1}^m \left[\int \cdots \int_{\mathbb{R}_+^{k+1}} \mathbf{1}_{\{\sum_{i=0}^{k-1} y_i \leq T, \sum_{i=0}^k y_i \geq T\}} f_k \left(y_0, y_0 + y_1, \dots, \sum_{i=0}^{k-1} y_i \right) \right]$$

$$\begin{aligned}
 & \left. \frac{\partial}{\partial \varepsilon} \psi^\varepsilon(y_0, y_1, \dots, y_k) \right|_{\varepsilon=0} dy_0 \cdots dy_k \Bigg] \\
 &= \sum_{k=1}^m \left[\int \cdots \int_{\mathbb{R}_+^{k+1}} \mathbf{1}_{\left\{ \sum_{i=0}^{k-1} y_i \leq T, \sum_{i=0}^k y_i \geq T \right\}} f_k \left(y_0, y_0 + y_1, \dots, \sum_{i=0}^{k-1} y_i \right) \right. \\
 & \quad \left. \sum_{j=0}^k \left(-u \left(\sum_{i=0}^j y_i \right) + \alpha_j \int_{\sum_{i=0}^{j-1} y_i}^{\sum_{i=0}^j y_i} u(s) ds \right) \prod_{j=0}^k \alpha_j e^{-y_j \alpha_j} dy_0 \cdots dy_k \right] \\
 &= \sum_{k=1}^m \left[\int \cdots \int_{\mathbb{R}_+^{k+1}} \mathbf{1}_{\left\{ \sum_{i=0}^{k-1} y_i \leq T, \sum_{i=0}^k y_i \geq T \right\}} f_k \left(y_0, y_0 + y_1, \dots, \sum_{i=0}^{k-1} y_i \right) \right. \\
 & \quad \left[\sum_{j=0}^{k-1} \left(-u \left(\sum_{i=0}^j y_i \right) + \alpha_j \int_{\sum_{i=0}^{j-1} y_i}^{\sum_{i=0}^j y_i} u(s) ds \right) + \alpha_k \int_{\sum_{i=0}^{k-1} y_i}^T u(s) ds \right] \\
 & \quad \left. \prod_{j=0}^k \alpha_j e^{-y_j \alpha_j} dy_0 \cdots dy_k \right], \tag{5.61}
 \end{aligned}$$

where we applied the following result for any $\theta' > 0$,

$$\begin{aligned}
 & \int_0^\infty \mathbf{1}_{\{y_k + \theta' \geq T\}} \left(-u(y_k + \theta') + \alpha_k \int_T^{y_k + \theta'} u(s) ds \right) \alpha_k e^{-y_k \alpha_k} dy_k \\
 &= - \int_{T-\theta'}^\infty u(y_k + \theta') \alpha_k e^{-y_k \alpha_k} dy_k + \alpha_k^2 \int_{T-\theta'}^\infty e^{-y_k \alpha_k} dy_k \int_T^{y_k + \theta'} u(s) ds \\
 &= - \int_{T-\theta'}^\infty u(y_k + \theta') \alpha_k e^{-y_k \alpha_k} dy_k + \alpha_k^2 \int_{T-\theta'}^\infty e^{-y_k \alpha_k} dy_k \int_{T-\theta'}^{y_k} u(s + \theta') ds \\
 &= - \int_{T-\theta'}^\infty u(y_k + \theta') \alpha_k e^{-y_k \alpha_k} dy_k + \int_{T-\theta'}^\infty u(s + \theta') ds \int_s^\infty \alpha_k^2 e^{-y_k \alpha_k} dy_k \\
 &= - \int_{T-\theta'}^\infty u(y_k + \theta') \alpha_k e^{-y_k \alpha_k} dy_k + \int_{T-\theta'}^\infty u(s + \theta') \alpha_k e^{-s \alpha_k} ds \\
 &= 0.
 \end{aligned}$$

Hence with the density function of $(\tau_1, \dots, \tau_k)$ and by (5.51), (5.61) becomes

$$\begin{aligned}
 & \sum_{k=1}^m E \left[\left(\sum_{j=0}^{k-1} \left(-u(T_{j+1}) + \alpha_j \int_{T_j}^{T_{j+1}} u(s) ds \right) + \alpha_k \int_{T_k}^T u(s) ds \right) \right. \\
 & \quad \left. f_k(T_1, \dots, T_{N_T^\alpha}) \mathbf{1}_{\{N_T^\alpha = k\}} \right]
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k=1}^m E \left[\left(- \sum_{j=0}^{N_T^\alpha - 1} u(T_{j+1}) + \sum_{j=0}^{N_T^\alpha - 1} \alpha_j \int_{T_j}^{T_{j+1}} u(s) ds + \alpha_{N_T^\alpha} \int_{T_{N_T^\alpha}}^T u(s) ds \right) \right. \\
 &\quad \left. f_k(T_1, \dots, T_{N_T^\alpha}) \mathbf{1}_{\{N_T^\alpha = k\}} \right] \\
 &= \sum_{k=1}^m E \left[\left(- \int_0^T u(t) dN_t + \int_0^T u(s) \alpha_{N_t^\alpha} ds \right) f_k(T_1, \dots, T_{N_T^\alpha}) \mathbf{1}_{\{N_T^\alpha = k\}} \right] \\
 &= E \left[\left(\int_0^T u(s) \alpha_{N_t^\alpha} ds - \int_0^T u(t) dN_t \right) \sum_{k=1}^m f_k(T_1, \dots, T_{N_T^\alpha}) \mathbf{1}_{\{N_T^\alpha = k\}} \right] \\
 &= E \left[\left(\int_0^T u(s) \alpha_{N_t^\alpha} ds - \int_0^T u(t) dN_t \right) \sum_{k=1}^m F \mathbf{1}_{\{N_T^\alpha > 0\}} \right]. \tag{5.62}
 \end{aligned}$$

Since the first term in (5.59) is

$$\alpha_0 f_0 e^{-\alpha_0 T} \int_0^T u(s) ds = E \left[\left(\int_0^T u(s) \alpha_{N_t^\alpha} ds - \int_0^T u(t) dN_t \right) F \mathbf{1}_{\{N_T^\alpha = 0\}} \right],$$

combining (5.59), (5.61)-(5.62), we have

$$-\frac{\partial}{\partial \varepsilon} E[F^\varepsilon] \Big|_{\varepsilon=0} = E \left[F \left(\int_0^T u(t) dN_t - \int_0^T u(t) \alpha_{N_t^\alpha} dt \right) \right]. \tag{5.63}$$

Therefore (5.56) is proved by (5.63) and Proposition 5.6.1. $\square$

We finish this section by proving the Lemma (5.6.1) stating the density function of $\tau_i^{\alpha, \varepsilon}$ in (5.52). Recall that there exists a constant $K > 0$ that $|u(s)| < K$ for any $s \in \mathbb{R}_+$ since $u \in \mathcal{U}$.

Lemma 5.6.1. For any $\varepsilon \in (-K^{-1}, K^{-1})$, $i \in \{1, 2, \dots\}$, given $T_i = \theta > 0$, the density function of $\tau_i^{\alpha, \varepsilon}$ in (5.52) denoted as $\xi_{\tau_i^{\alpha, \varepsilon}}(x)$ is expressed as

$$\xi_{\tau_i^{\alpha, \varepsilon}}(x) = \frac{\alpha_i e^{-\kappa^{-1}(x, \varepsilon, \theta) \alpha_i}}{1 + \varepsilon u(\kappa^{-1}(x, \varepsilon, \theta))}, \tag{5.64}$$

where $\kappa(x, \varepsilon, \theta) = x + \varepsilon \int_\theta^{x+\theta} u(s) ds$ for $\varepsilon \in (-K^{-1}, K^{-1})$, $x \in \mathbb{R}_+$, and $\xi_{\tau_i^{\alpha, \varepsilon}}(x)$ has the derivative at $\varepsilon = 0$ that

$$\frac{\partial \xi_{\tau_i^{\alpha, \varepsilon}}(x)}{\partial \varepsilon} \Big|_{\varepsilon=0} = \left(-u(x + \theta) + \alpha_i \int_\theta^{x+\theta} u(s) ds \right) \alpha_i e^{-x \alpha_i}. \tag{5.65}$$

Proof. Since $|u(s)| < K$ for any $s \in \mathbb{R}_+$ and $\kappa'(x, \varepsilon, \theta) = 1 + \varepsilon u(x + \theta) > 0$, so $\kappa^{-1}(x, \varepsilon, \theta)$ exists in $x \in [0, \infty)$. From the density function of τ_i^α that $\xi_{\tau_i}(x) = \alpha_i e^{-x\alpha_i}$, we have the density function of $\tau_i^{\alpha, \varepsilon}$ in (5.52):

$$\xi_{\tau_i^{\alpha, \varepsilon}}(x) = \alpha_i e^{-\kappa^{-1}(x, \varepsilon, \theta)\alpha_i} \left| \frac{\partial \kappa^{-1}(x, \varepsilon, \theta)}{\partial x} \right| = \frac{\alpha_i e^{-\kappa^{-1}(x, \varepsilon, \theta)\alpha_i}}{1 + \varepsilon u(\kappa^{-1}(x, \varepsilon, \theta) + \theta)}, \quad (5.66)$$

note also that $1 + \varepsilon u(\kappa^{-1}(x, \varepsilon, \theta) + \theta) > 0$ for $|\varepsilon| < K^{-1}$. Then we compute its derivative, note that

$$\begin{aligned} \left. \frac{\partial \kappa^{-1}(x, \varepsilon, \theta)}{\partial \varepsilon} \right|_{\varepsilon=0} &= \lim_{\varepsilon \rightarrow 0} \frac{\kappa^{-1}(x, \varepsilon, \theta) - x}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\kappa^{-1}(x, \varepsilon, \theta) - \kappa(\kappa^{-1}(x, \varepsilon, \theta), \varepsilon, \theta)}{\partial \varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\kappa^{-1}(x, \varepsilon, \theta) - (\kappa^{-1}(x, \varepsilon, \theta) + \varepsilon \int_{\theta}^{\kappa^{-1}(x, \varepsilon, \theta) + \theta} u(s) ds)}{\partial \varepsilon} \\ &= - \int_{\theta}^{\kappa^{-1}(x, 0, \theta) + \theta} u(s) ds. \end{aligned} \quad (5.67)$$

With (5.67) we have

$$\begin{aligned} \left. \frac{\partial f_{\tau_i^{\alpha, \varepsilon}}(x)}{\partial \varepsilon} \right|_{\varepsilon=0} &= \frac{\partial}{\partial \varepsilon} \left(\frac{\alpha_i e^{-\kappa^{-1}(x, \varepsilon, \theta)\alpha_i}}{1 + \varepsilon u(\kappa^{-1}(x, \varepsilon, \theta) + \theta)} \right) \Big|_{\varepsilon=0} \\ &= - \frac{u(\kappa^{-1}(x, \varepsilon, \theta) + \theta)}{(1 + \varepsilon u(\kappa^{-1}(x, \varepsilon, \theta) + \theta))^2} \alpha_i e^{-\kappa^{-1}(x, \varepsilon, \theta)\alpha_i} \Big|_{\varepsilon=0} \\ &\quad - \frac{\alpha_i^2 e^{-\kappa^{-1}(x, \varepsilon, \theta)\alpha_i}}{1 + \varepsilon u(\kappa^{-1}(x, \varepsilon, \theta) + \theta)} \frac{\partial \kappa^{-1}(x, \varepsilon, \theta)}{\partial \varepsilon} \Big|_{\varepsilon=0} \\ &= -u(x + \theta) \alpha_i e^{-x\alpha_i} - \alpha_i^2 e^{-x\alpha_i} \frac{\partial \kappa^{-1}(x, \varepsilon, \theta)}{\partial \varepsilon} \Big|_{\varepsilon=0} \\ &= -u(x + \theta) \alpha_i e^{-x\alpha_i} + \alpha_i^2 e^{-x\alpha_i} \int_{\theta}^{\kappa^{-1}(x, \varepsilon, \theta) + \theta} u(s) ds \Big|_{\varepsilon=0} \\ &= \left(-u(x + \theta) + \alpha_i \int_{\theta}^{x+\theta} u(s) ds \right) \alpha_i e^{-x\alpha_i}. \end{aligned}$$

□

Chapter 6

Conclusion

Based on the regime switching models, this thesis presents the solutions to optimal stopping problems and computations of sensitivities by establishing the integration by parts formula for Markov chains.

We show that, the optimal stopping problem (2.1) in Chapter 2 can be characterized by a free-boundary problem and the optimal stopping time is the hitting time of the optimal stopping set, whose boundary curves are determined by a Volterra integral. Uniqueness of the solution to this free-boundary problem is justified under the condition $\mu(i) \geq 0$ for all $i \in \mathcal{M}$. We also achieve some special cases for immediate decision, when all drift parameters $\mu(i)$ are negative, $i \in \mathcal{M}$, immediate exercise is optimal; On the other hand, exercise at maturity becomes optimal when $\mu(i) \geq \sigma^2(i)$ for all $i \in \mathcal{M}$, where $\sigma(i)$ are the volatility parameters.

To solve this problem for all settings of parameters $(\mu(i), \sigma(i))$, $i \in \mathcal{M}$, we design a recursive algorithm for the same problem as above. This method computes the optimal value function through backward inductions and allows us to determine the boundary functions of the optimal stopping

set without having to solve their associated Volterra integral equations.

For the sensitivity analysis based on this regime switching model, Malliavin calculus can be applied to compute all kinds of Greeks and we also extend our results for non-differentiable payoff functions hence covering the majority of commonly used payoff functions.

Integration by parts formula for Markov chains is established similarly to that of Malliavin calculus for Brownian motion. We can also apply it for the sensitivity analysis of regime switching models.

Many extensions can be done along these directions for further studies. Optimal stopping based on regime switching models is applicable with other value functions. The backward induction designed for computation of the reward function actually computes the expectation of a large class of functionals of Markov chains, therefore some general arguments should be developed in this direction. Since the integration by parts formulas are well established for Wiener process, poisson process, as well as Markov chains, we expect to develop more versions of integration by parts for other processes.

References

- Bally, V., Bavouzet-Morel, M. P., and Messaoud, M. “Integration by parts formula for locally smooth laws and applications to sensitivity computations.” *Ann. Appl. Probab.*, 17:33–66 (2007).
- Bather, J. “Optimal Stopping Problems for Brownian Motion.” *Advances in Applied Probability*, 2:259–286 (1970).
- Bavouzet-Morel, M. and Messaoud, M. “Computation of Greeks using Malliavins calculus in jump type market models.” *Electronic Journal of Probability*, 11:276–300 (2006).
- Bayazit, D. and Nolder, C. “Sensitivities of options via Malliavin calculus: applications to markets of exponential variance Gamma and normal inverse Gaussian processes.” *Quantitative Finance*, 13:1257–1287 (2013).
- Biane, P. “Chaotic representations for finite Markov chains.” *Stochastics and Stochastics Reports*, 30:61–68 (1989).
- Bollen, N. P. B. “Valuing options in regime-switching models.” *Journal of Derivatives*, 6:38–49 (1998).
- Boyce, W. M. “Stopping rules for selling bonds.” *The Bell Journal of Economics and Management Science*, 1:27–53 (1970).

- Buffington, J. and Elliott, R. J. “American options with regime switching.” *International Journal of Theoretical and Applied Finance*, 5:497–514 (2002).
- Chow, Y. S., Robbins, H., and Siegmund, D. “Great expectations: the theory of optimal stopping.” *Houghton Mifflin, Boston, Mass* (1971).
- David, D. and Privault, N. “Numerical computation of Theta in a jump-diffusion model by integration by parts.” *Quantitative Finance*, 9 (2009).
- Davis, M. H. A. and Johansson, M. P. “Malliavin Monte Carlo Greeks for jump diffusions.” *Stochastic Processes and their Applications*, 116:101–129 (2006).
- Debelley, V. and Privault, N. “Sensitivity analysis of European options in jump diffusion models.” *Manuscript, Université de La Rochelle* (2004a).
- . “Sensitivity analysis of European options in jump-diffusion models via the Malliavin calculus on the Wiener space.” *Preprint* (2004b).
- Di Masi, G. B., Kabanov, Y. M., and Runggaldier, W. J. “Mean variance hedging of options on stocks with Markov volatility.” *Theory of Probability and Applications*, 39:173–181 (1994).
- Du Toit, J. and Peskir, G. “The trap of complacency in predicting the maximum.” *The Annals of Applied Probability*, 35:340–365 (2007).
- . “Predicting the time of the ultimate maximum for Brownian motion with drift.” *Mathematical Control Theory and Finance*, 95–112 (2008).
- . “Selling a stock at the ultimate maximum.” *Ann. Appl. Probab.*, 19:983–1014 (2009).

- Du Toit, J., Peskir, G., and Shiryaev, A. N. “Predicting the last zero of Brownian motion with drift.” *Stochastics*, 229–245 (2008).
- Duan, J. C., Popova, I., and Ritchken, P. “Option pricing under regime switching.” *Quantitative Finance*, 2:116–132 (2002).
- El-Khatib, Y. and Privault, N. “Computations of Greeks in markets with jumps via the Malliavin calculus.” *Finance and Stochastics*, 4:161–179 (2004).
- Eloe, P., Liu, R., Yatsuki, M., Yin, G., and Zhang, Q. “Optimal selling rules in a regime-switching exponential Gaussian diffusion model.” *SIAM J. Appl. Math.*, 69:810–829 (2008).
- Engelbert, H. “On the theory of optimal stopping rules for Markov processes.” *Theory of Probability and Its Applications*, 18:304–311 (1973).
- Fournié, E., Lasry, J. M., Lebuchoux, J., Lions, P. L., and Touzi, N. “Applications of Malliavin calculus to Monte Carlo methods in finance.” *Finance and Stochastics*, 3:391–412 (1999).
- Graversen, S. E. and Shiryaev, A. N. “An extension of P. Lévy’s distributional properties to the case of a Brownian motion with drift.” *Bernoulli*, 6:615–620 (2000).
- Guo, X. “Inside Information and Stock Fluctuations.” *Ph.D. thesis, Rutgers University* (1999).
- . “An explicit solution to an optimal stopping problem with regime switching.” *J. Appl. Probab.*, 38:464–481 (2001).
- Guo, X. and Zhang, Q. “Optimal selling rules in a regime switching model.” *IEEE Trans. Automat. Control*, 50:1450–1455 (2005).

- Hamilton, J. D. “A new approach to the economic analysis of non-stationary time series.” *Econometrica*, 57:357–384 (1989).
- Hardy, M. R. “A regime-switching model of long-term stock returns.” *North American Actuarial Journal*, 5:41–53 (2001).
- Heinonen, B. S., Bruckner, J. B., and Bruckner, A. M. “Elementary real analysis.” *Prentice-Hall Inc.* (2000).
- Heston, S. L. “A closed-form solution for options with stochastic with applications to bond and currency options.” *Review of Financial Studies*, 6:237–343 (1993).
- Kawai, R. and Kohatsu-Higa, A. “Computation of Greeks and multidimensional density estimation for asset price models with time-changed Brownian motion.” *Applied Mathematical Finance*, 17:301–321 (2010).
- Kawai, R. and Takeuchi, A. “Greeks formulas for an asset price model with Gamma processes.” *Mathematical Finance*, 21:723–742 (2011).
- . “Computation of Greeks for asset price dynamics driven by stable and tempered stable processes.” *Quantitative Finance*, 13:1303–1316 (2013).
- Kroeker, J. “Wiener analysis of functionals of a Markov chain: Application to neural transformations of random signals.” *Biol. Cybernetics*, 36:243–248 (1980).
- Le, H. and Wang, C. “A finite time horizon optimal stopping problem with regime switching.” *SIAM J. Control Optim.*, 48:5193–5213 (2010).
- Liu, R., Zhang, Q., and Yin, G. “Option pricing in a regime-switching model using the fast Fourier transform.” *J. Appl. Math. Stoch. Anal.*, 2006:1–22 (2006).

- Liu, Y. and Privault, N. “A recursive algorithm for optimal selling in regime switching models.” *Preprint* (2015a).
- . “Selling at the ultimate maximum in a regime switching model.” *Submitted* (2015b).
- Ly Vath, V. and Pham, H. “Explicit solution to an optimal switching problem in the two-regime case.” *SIAM J. Control Optim.*, 46:395–426 (2007).
- Malliavin, P. and Thalmaier, A. “Stochastic calculus of variations in mathematical finance.” *Springer Finance. Springer-Verlag, Berlin* (2006).
- Nualart, D. “The Malliavin Calculus and Related Topics.” *Probability and its Application. Springer-Verlag* (1995).
- Papanicolaou, A. and Sircar, R. “A regime-switching Heston model for VIX and S&P 500 implied volatilities.” *Quantitative Finance*, 14:1811–1827 (2014).
- Peskir, G. “A change-of-variable formula with local time on curves.” *J. Theoret. Probab.*, 18:499–535 (2005).
- Peskir, G. and Shiryaev, A. “Optimal stopping and free-boundary problems.” *Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel* (2006).
- Privault, N. “Calcul des variations stochastique pour les martingales.” *C. R. Acad. Sci. Paris Sér. I Math.*, 321:923–928 (1995a).
- . “Stochastic calculus of variations for martingales.” In *I. M. Davies, K. D. Elworthy, and A. Truman, editors, Proceedings of the Fifth Gregynog Symposium, Gregynog*, 385–400 (1995b).

-
- . “Stochastic Analysis in Discrete and Continuous Settings.” *Lecture Notes in Math. Springer-Verlag, Berlin*, 1982:309 (2009).
- . “Understanding Markov Chains - Examples and applications.” *Springer* (2013).
- Privault, N. and Schoutens, W. “Discrete chaotic calculus and covariance identities.” *Stochastics and Stochastics Reports*, 72:289–315 (2002).
- Privault, N. and Wei, X. “A Malliavin calculus approach to sensitivity analysis in insurance.” *Insurance Math. Econom.*, 35:679–690 (2004).
- Protter, P. “Stochastic integration and differential equations. A new approach.” *Springer-Verlag, Berlin* (1990).
- Rudin, W. “Principle of mathematical analysis.” *3rd edition. McGraw-Hill Inc.* (1976).
- Shepp, L. “A model for stock price fluctuations based on information.” *IEEE Transactions on Information Theory*, 48:1372–1378 (2002).
- Shiryaev, A., Xu, Z. Q., and Zhou, X. Y. “Thou shalt buy and hold.” *Quantitative Finance*, 8:765–776 (2008).
- Shiryaev, A. N. “Optimal stopping rules.” *Springer-Verlag, New York* (1978).
- Siu, T. K. “Integration by parts and martingale representation for a Markov chain.” *Abstr. Appl. Anal. page Art. ID 438258* (2014).
- Yam, S. C. P., Yung, S. P., and Zhou, W. “Two rationales behind the Buy-and-hold or Sell-at-once strategy.” *J. Appl. Probab.*, 46:651–668 (2009).

-
- . “Optimal selling time in stock market over a finite time horizon.” *Acta Math. Appl. Sin. Engl. Ser.*, 28:557–570 (2012).
- . “A unified Bang-Bang principle with respect to R-invariant performance benchmarks.” *Theory of Probability and Applications*, 57:357–366 (2013).
- Yao, D. D., Zhang, Q., and Zhou, X. Y. “A regime-switching model for European options.” *International Series in Operation Research and Management Science*, 94:281–300 (2006).
- Yin, G., Zhang, Q., and Yin, K. “Constrained stochastic estimation algorithms for a class of hybrid stock market models.” *Journal of Optimization theory and Applications*, 118:157–182 (2003).
- Zhang, Q. “Stock trading: An optimal selling rule.” *SIAM Journal on Control and Optimization*, 40:64–87 (2001).
- Zhang, Q. and Yin, G. “Nearly optimal asset allocation in hybrid stock investment models.” *Journal of Optimization theory and Applications*, 121:419–444 (2004).