Adaptive robust control with statistical learning

Theerawat Bhudisaksang

St Edmund Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2022

Acknowledgements

The completion of this thesis would not have been possible without the invaluable assistance given by my supervisor, Álvaro Cartea. I am deeply indebted to him for his financial mathematics knowledge, writing skills, and advice on so many other topics, from presentations to the English language. He has been patient, understanding, and encouraging from the first to the last day of my study.

Thanks also to many of my friends who have supported me through the completion of my DPhil, which has helped me besides the academic but life in general. Special thanks to my office-mated who provide an invaluable discussions both in mathematics and life and make the DPhill life more proactive and fruitful.

I would like to express special thanks to my family who supports me unconditionally and guided me to reach my potential and my girlfriend who helps me with her invaluable advice and emotional support.

Finally, I would like to acknowledge financial support from the Development and Promotion of Science and Technology Talents Project (DPST) of the Government of Thailand.

Abstract

In stochastic control problems the agent chooses the optimal strategy to maximise or minimise the performance criterion. The performance criterion can be either the expectation of a reward function for the standard control problem or the non-linear expectation for the robust control problem. In parameterised stochastic control problems, the agent needs to know the value of the model parameters in the stochastic system to specify the optimal strategy correctly. However, it is hardly the case that the agent knows the values of the model parameters.

In this thesis, we aim to study a robust stochastic control problem where the agent does not know the values of the parameters of the underlying process. Therefore, we frame the stochastic control problem where we assume that the agent does not know the values of the model parameters. However, the agent uses the observable processes to estimate the values of the model parameters while simultaneously solving the stochastic control problem in a robust framework.

There are two key components in this new stochastic control problem. The first component is the parameter estimation part where the agent uses the realisation of the underlying process to estimate the unknown parameters in the stochastic system. We particularly focus on online parameter estimation. The online estimator is an important ingredient for our stochastic control problem because this type of estimator allows the agent to obtain the optimal strategy in feedback form. The second component is the stochastic control part which is the question of how to design a time-consistent stochastic control problem that allows the agent to also estimate the parameters and optimise her strategy simultaneously. In this thesis, we address each component of the problem above in the continuous-time setting and then the utility maximisation problem under this framework is studied carefully.

Contents

List of Figures

Chapter 1

Introduction

In this thesis, we study stochastic control problems where the agent does not have full knowledge of the value of parameters in the model and, as time evolves, uses new observations to estimate the parameters and, simultaneously, to update the optimal strategy. This problem is interesting from both theoretical and practical perspectives. The standard stochastic control problem usually assumes that the agent knows the value of the model parameters, which is a strong assumption and does not hold in practice. With this relaxation on the assumption of the knowledge of the parameters, we can apply our new stochastic control framework to many classical stochastic control problems such as utility maximisation where the agent does not have full knowledge of the value of model parameters in the stochastic system.

There are two key components in these stochastic control problems. First, estimate the value of the parameters as time evolves and more information is available. In this thesis, we focus on online parameter estimation. The online estimator is an important ingredient in the stochastic control problem that we study because an online estimator allows the agent to obtain a feedback form strategy (Markovian). Second, design a time-consistent stochastic control problem that allows the agent to estimate parameters online while deriving the optimal strategy. In this thesis, we address each component of the problem above in the continuous-time setting.

Next, we provide a brief introduction to each chapter of our thesis.

In Chapter 2, we introduce concepts that we use throughout the thesis. In particular, in Section 2.1, we list terminologies of descriptive set theory, which are essential to developing measurable selection theorems. In Section 2.2, we state definitions of key concepts such as analytic and Borel-measurable sets and standard measurable selection theorems under different assumptions which been used in the stochastic control problems. Moreover, we provide a sketch proof of the dynamic programming principle in a classical stochastic control setup to show how the measurable selection theorem is applied to prove the dynamic programming principle. Later, we discuss parameter estimation techniques in general and provide a standard result of the online parameter estimation in the discrete and continuous-time setup. This helps to illustrate the difference between the standard result and our result in Chapter 5 and provides insights into what assumptions are required for the proof of the convergence for our online estimator.

In Chapter 3, we study the online drift parameter estimation for the jump-diffusion process. The convergence of the estimator in Sirignano and Spiliopoulos (2017) is studied when the observed process is a diffusion process. Here, we provide a nontrivial extension for the convergence of the online estimator for the jump-diffusion case using an abstract version of the solution of the Poisson equation. Then, we show this convergence numerically for the Ornstein-Uhlenbeck process and the Bimodal process.

In Chapter 4, we propose "adaptive robust control in continuous-time", which we extend from Bielecki, Chen, and Cialenco (2017a) where they study the discrete-time framework. The proposed model is the stochastic control problem in which the agent does not have full knowledge of the values of the parameters in the model. The agent observes the underlying process to estimate the value of parameters and uses the updated parameter values to solve the optimal strategies of the stochastic control problem. The main difficulty of the adaptive robust control in continuous-time is the proof of the time-consistency and the dynamic programming of the model. We rely on a precise measurable selection theorem to prove these two results, and then we show the characterisation of the value function as a viscosity solution of a partial differential equation. As a financial example, we study the liquidation problem under the adaptive robust control framework.

In Chapter 5, we study the utility maximisation problem under the adaptive robust control framework discussed in Chapter 4. The utility maximisation problem is a classical stochastic problem. We are interested in the application of the adaptive robust control framework to this classical problem and compare the optimal strategy under the adaptive robust framework to the classical framework. We show the stochastic representation of the value function for the utility maximisation problem and we study the effect of the learning on the optimal strategy. Next, we analyze the optimal strategy to understand the role of each hyper-parameter and we present numerical results.

Chapter 2

Basic mathematics

In this chapter, we discuss the foundations of a dynamic programming principle in continuous-time and provide the notations and concepts that we use throughout this thesis. Section 1 presents the methodologies of the typical topological spaces that we use in classical probability theory and the definition of a probability kernel. In Section 2, we recall the results of measurable selection theorems in Bertsekas and Shreve (1996) and we state the dynamic programming principle for a standard stochastic control problem and provide a sketch proof of the theorem to illustrate the use of the measurable selection theorem. In the last Section, we introduce multiple online parameter estimation methods that are used in an adaptive robust control problem. We discuss these results in discrete and continuous-time setup.

2.1 Descriptive set theory

Here, we present some classical concepts of descriptive set theory, which are essential elements in an abstract dynamic programming and in a measurable selection theorem. First, we present the definition of the standard topological spaces and the definition of the probability kernel.

Probability theory has a setup in the space that is "nice" enough, so standard techniques can be used. In particular, separability of the space is required for constructing a conditional expectation. In general, most spaces we consider here are Polish spaces. For example, the space $\mathbb{R}^d$ equipped with the Euclidean norm is a Polish space.

Definition 2.1.1. (Polish Space) The space E is a Polish space if it is a completely metrisable topological space and a separable metric space.

The important property of a Polish space is when we define the set of probability measures on a Polish space E, denoted by $\mathcal{P}(E)$ equipped with the weak*-topology. Then, the space $\mathcal{P}(E)$ is also a Polish space. Now, we introduce the definition of Borel space, which we use in the measurable selection theorems.

Definition 2.1.2. (Borel Space) A topological space is said to be a Borel space if it is topologically homeomorphic to a Borel subset of a Polish space. (ii) Let E and F be two Borel spaces, E and F are said to be isomorphic, if there is a bijection ψ between $(E, \mathcal{B}(E))$ and $(F, \mathcal{B}(F))$ such that ψ and ψ^{-1} are both measurable.

If the topological spaces X and Y are homeomorphic, then they have the same topological structure. In stochastic control problems, generally, the problems involve a law of the future state process, and an agent takes an action to maximize his objective function which depends on the future value of the state. In the dynamic setting, the agent observes a realization of the state which alters the view of the future state and changes the law of the future state. We require a mathematical object to capture this.

Definition 2.1.3. (Probability kernel or stochastic kernel) Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be two measurable spaces, a probability kernel is a map $N : X \times \mathcal{Y} \to [0, 1]$ such that the map $x \to N(x, B)$ is a $\mathcal{X}$-measurable function for all $B \in \mathcal{Y}$ and for all $x \in X$, $N(x, \cdot)$ is a probability measure on $(Y, \mathcal{Y})$

The following definition and theorem are used in Chapter 4. The theorem shows that if a function is measurable in one variable and is continuous in the other variable, then, under additional topological assumptions, the function is jointly measurable. Proving the Section function of a function to be continuous or measurable is simpler than proving the joint measurability of the function directly.

Definition 2.1.4. Let (S, Σ) be a measurable space, and let X and Y be topological spaces. A function $f : S \times X \to Y$ is a **Carathéodory function** if:

1. For each $x \in X$, the function $f^x = f(\cdot, x) : S \to Y$ is $(\Sigma, \mathcal{B}_Y)$-measurable.

2. For each $s \in S$, the function $f_s = f(s, \cdot) : X \to Y$ is continuous.

Theorem 2.1.5. *Let (S, Σ) be a measurable space, X a separable metrizable space, and Y a metrizable space. Then every Carathéodory function $f : S \times X \to Y$ is jointly measurable.*

The full proof of the theorem above can be found in Kubińska (2005).

2.2 Measurable selection and dynamic programming principle

In this Section, we present the measurable selection theorems. First, we discuss the projection theorem and how it deduces the measurable selection theorem. Later, we give a definition for the semi-analytic function. We state multiple versions of measurable selection theorems. Later, we present a sketch proof of the dynamic programming principle in a standard stochastic control setup using a measurable selection theorem.

2.2.1 Measurable selection theorems

Consider the following question. If $A \subset [0,1] \times [0,1]$ is a Borel measurable set, can we show that the projection of A onto the interval $[0,1]$ is also a Borel measurable set?. The answer is provided by Souslin, who proved that it is not necessary for a projection of Borel set to be a Borel set. This stability under a projection requires a larger σ-field than the Borel measurable σ-field. Note that a projection of a set $A \in \Omega \times E$ onto Ω is

$$\{\omega \in \Omega \mid \exists x \in E \text{ such that } (\omega, x) \in \Omega \times E\}.$$

and there are multiple ways to enlarge the Borel set.

Indeed, we define an analytic set as a projection of a Borel set and with this weaker definition, the analytic set is stable under projection. Hence, from the definition of Borel space, the Polish space is a Borel space. We give a complete definition of analytic set later in this subsection. We use $\pi_\Omega(A)$ to represent a projection of the set A onto the set Ω.

Moreover, the stability of projection can hold, if one enlarges the σ-field to be a completed σ-field under a probability measure $\mathbb{P}$.

We state the results of the technical product paving and projection capacity. For each Borel measurable set, A, in $(\Omega \times E, \mathcal{F} \otimes \mathcal{B}(E))$ and for each probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, the projection of the set A onto Ω has a Borel measurable set that differs up to a $\mathbb{P}$-negligible set. In other words, we can find a Borel measurable set $\tilde{A} \subseteq \pi_\Omega(A)$ such that $\mathbb{P}^*(\pi_\Omega(A)) = \mathbb{P}(\tilde{A})$ where $\mathbb{P}^*$ is an outer measure of probability measure $\mathbb{P}$, for more details see El Karoui and Tan (2013a).

Remark 2.2.1. The choice of the measure $\mathbb{P}$ is arbitrary, it is possible to enlarge a Borel σ-field with the set N such that the outer measure $\mathbb{P}^*(N) = 0$. Moreover, since the projection set $\pi_\Omega(A)$ is not necessarily a member of $\mathcal{F} = \mathcal{B}(\Omega)$, to have a measurement on the set requires the use of an outer measure induced by the probability measure $\mathbb{P}$ instead.

With this observation, the first measurable selection theorem can be proved using the completed σ-field. The proof is just a corollary of the projection of Borel measurable sets.

Proposition 2.2.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and E be a locally bounded compact Polish space. We consider a measurable function $f(\omega, x)$ defined on $(\Omega \times E, \mathcal{F} \otimes \mathcal{B}(E))$ and $A \in \mathcal{F} \otimes \mathcal{B}(E)$. Then the function g defined by*

$$g(\omega) := \sup\{f(\omega, x) : (\omega, x) \in A\}, \quad \textit{with the convention} \quad \sup \emptyset = -\infty, \tag{2.2.1}$$

is $\mathcal{F}$-measurable when $\mathcal{F}$ is a $\mathbb{P}$-completed σ-field.

Proof. Let c be a real number. Note that $\{g(\omega) > c\} = \{\, \omega \mid \exists\, x \text{ such that } f(\omega, x) > c,\ (\omega, x) \in A\}$. Therefore, the set $\{g(\omega) > c\}$ is a projection of a Borel measurable set $\{f(\omega, x) > c,\ (\omega, x) \in A\}$. This completes the proof since $\mathcal{F}$ is a $\mathbb{P}$-completed σ-field.

□

For the σ-field $\mathcal{F}$, we denote the completed σ-field under a probability measure $\mathbb{P}$ as $\mathcal{F}^{\mathbb{P}}$, therefore, for $A \in \mathcal{F} \otimes \mathcal{B}(E)$, the projection of the set A onto Ω is measurable with respect to $\mathcal{F}^{\mathbb{P}}$. As we mentioned, the choice of the probability measure $\mathbb{P}$ is arbitrary, this leads to the definition of universal completion. First, we let $\mathfrak{P}(\Omega)$ be the set of probability measures on Ω

Definition 2.2.3. Let $(\Omega, \mathcal{F})$ be a measurable space, the universal completion of $\mathcal{F}$ is the σ-field defined as the intersection of $\mathcal{F}^{\mathbb{P}}$ for all probability measures $\mathbb{P} \in \mathcal{P}(\Omega)$ on $(\Omega, \mathcal{F})$ i.e.,

$$\mathcal{F}^U := \bigcap_{\mathbb{P} \in \mathfrak{P}(\Omega)} \mathcal{F}^{\mathbb{P}}.$$

A function f from $(\Omega, \mathcal{F})$ to E is called universally measurable, if $f^{-1}(A) \in \mathcal{F}^U$ for each $A \in \mathcal{B}(E)$. Now, we state the universally measurable selection theorem where the statement does not rely on a specific probability measure as in Proposition 2.2.2.

Theorem 2.2.4. *Let $(\Omega, \mathcal{F})$ be a measurable space, E be a Borel space with $\mathcal{E} := \mathcal{B}(E)$, and $A \in \mathcal{F} \otimes \mathcal{E}$ be a measurable subset in $\Omega \times \mathcal{F}$ and f is a $\mathcal{F} \otimes \mathcal{E}$- measurable function. Define*

$$g(\omega) := \sup\{f(\omega, x) : (\omega, x) \in A\}\,. \tag{2.2.2}$$

Then g is $\mathcal{F}^U$-universally measurable and for each $\epsilon > 0$, there exists a $\mathcal{F}^U$-universally measurable mapping Z_ϵ such that for all $\omega \in \pi_\Omega(A)$

$$f(\omega, Z_\epsilon(\omega)) \geq g(\omega) - \epsilon\,.$$

The full proof of the theorem above can be found in El Karoui and Tan (2013a). Now, we present a detail definition of analytic sets and the measurable selection theorem under this framework. The following result and definition are from El Karoui and Tan (2013a).

Definition 2.2.5. (Analytic set and analytically measurable function)

(i) Let E be a Borel space, then a subset B is an analytic set in E if there is another Borel space F and a Borel subset $A \subseteq E \times F$ such that $B = \pi_E(A)$. A subset $C \subseteq E$ is co-analytic if its complement C^c is analytic.

(ii) A function $g : E \to \mathbb{R} = \mathbb{R} \cup \infty$ is upper semianalytic (u.s.a.) if $\{x \in E : g(x) > c\}$ is analytic for every $c \in \mathbb{R}$.

(iii) Let E be a Borel set and $\mathcal{A}(E)$ denote a σ-field generated by all analytic subsets. A function $f : E \to F$, where F is a Borel set, is analytically measurable if $f^{-1}(C) \in \mathcal{A}(E)$ for every $C \in \mathcal{B}(F)$.

Note that all Borel measurable sets are analytic sets ($\mathcal{B}(E) \subseteq \mathcal{A}(E) \subseteq \mathcal{B}^U(E)$). Now, we state some results of a measurable selection theorem in the context of Borel spaces where the full statement of the proof can be found in El Karoui and Tan (2013a).

Theorem 2.2.6. *Let E and F be Borel spaces, let A be an analytic subset of $E \times F$, and $f : A \to R$ be an u.s.a function. Define $g(x) := \sup_{(x,y)\in A} f(x, y)$, then the following results hold:*

(i) The projection set $\pi_E(A)$ is an analytic subset in E.

(ii) There exists an analytically measurable function $\phi : \pi_E(A) \to F$ such that $(x, \phi(x)) \in A$, for every $x \in \pi_E(A)$.

(iii) The function $g : \pi_E(A) \to R = R \cup \infty$ is upper semianalytic.

(iv) For every $\epsilon > 0$, there is an analytically measurable function $\phi_\epsilon : \pi_E(A) \to F$ such that $f(x, \phi_\epsilon(x)) \geq g_\epsilon(x) := (g(x) - \epsilon)1_{g(x)<\infty} + \frac{1}{\epsilon}1_{g(x)=\infty}$ for every $x \in \pi_E(A)$.

Statement (iv) above is the essential tool for proving the time-consistency property for stochastic control problems. Note that if we use a stronger regularity assumption on the function f, we require a stronger assumption on the topological spaces.

2.2.2 Measurable selection to the dynamic programming principle

In this subsection, we connect the measurable selection theorems in the previous subsection to the dynamic programming principle using the proof of the dynamic programming principle in the classical setup. Here, we follow from Pham (2009) where they left some technical details to focus the intuition on the use of the measurable selection to prove dynamic programming principle in the classical setup. Let

$$dX_s = b(X_s, \alpha_s)\, ds + \sigma(X_s, \alpha_s)\, dW_s \,, \tag{2.2.3}$$

where X_t is valued in $\mathbb{R}^n$, W is a Brownian motion and α is a control process. We assume that the function b and σ are Lipschitz functions, so that the process X exists. We denote the set of control process at time t, x as $\mathcal{A}(t, x)$ where the set of control processes is the set of a progressively measurable function that satisfies

$$\mathbb{E}\left[\int_t^T \left|f(s, X_s^{t,x}, \alpha_s)\right|\right] < \infty \,. \tag{2.2.4}$$

The performance criterion is

$$J(t, x, \alpha) = \mathbb{E}\left[\int_t^T f(s, X_s^{t,s}, \alpha_s)\, ds + g(X_T^{t,x})\right] , \tag{2.2.5}$$

where f and g are under some growth assumptions. The value function v is given by

$$v(t, x) := \sup_{\alpha \in \mathcal{A}(t,x)} J(t, x, a) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^T f(s, X_s^{t,s}, \alpha_s)\, ds + g(X_T^{t,x})\right] . \tag{2.2.6}$$

We state the dynamic programming principle from Pham (2009).

Theorem 2.2.7. *(Dynamic programming principle) For each* $(t,x) \in [0,T] \times \mathbb{R}^n$*. Then, we have*

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^u f(s, X_s^{t,s}, \alpha_s)\, ds + v(u, X_u^{t,x})\right], \tag{2.2.7}$$

where $u \in [t,T]$*.*

Proof. Basically, we follow from the proof of Theorem 3.3.1 in Pham (2009). We only focus on the argument that showing

$$v(t,x) \geq \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^u f(s, X_s^{t,s}, \alpha_s)\, ds + v(u, X_u^{t,x})\right], \tag{2.2.8}$$

which can be shown by for each $\epsilon > 0$, we construct $\hat{\alpha}$ such that

$$v(t,x) \geq \mathbb{E}\left[\int_t^u f(s, X_s^{t,s}, \hat{\alpha}_s)\, ds + v(u, X_u^{t,x})\right] - \epsilon\,. \tag{2.2.9}$$

In his proof, Pham (2009) simplifies the proof by claiming that by the definition of the value functions, for any $\epsilon > 0$ and $\omega \in \Omega$, there exists $\alpha^{\epsilon,\omega} \in \mathcal{A}(s, X_s^{t,x}(\omega))$, which is an ϵ-control for $v(s, X_s^{t,x}(\omega)))$, i.e,

$$v(s, X_s^{t,x}(\omega))) - \epsilon \leq J(s, X_s^{t,x}(\omega), \alpha^{\epsilon,\omega})\,. \tag{2.2.10}$$

Although the above statement is true, it leaves multiple technical details. Here, we provide these missing steps. If we rewrite the problem as

$$v(s, X_s^{t,x}(\omega)) = \sup_{\alpha \in \mathcal{A}(s, X_s^{t,x}(\omega))} J(s, X_s^{t,x}, \alpha)\,, \tag{2.2.11}$$

can be written as Theorem 2.2.6 (with a substitution of the function). Therefore, with the proof of the regularity of the function J, the measurable selection in Theorem 2.2.6 can be applied and we have that the statement above from Pham (2009) is correct. $\square$

As we see from the above example, a measurable selection theorem is crucial for the prove of the dynamic programming principle of a stochastic control problem. We apply this technique again in Chapter 4.

2.3 Online parameter estimation

In the previous sections, we discuss the abstract set theory and measurable selection theorems which use to show the dynamic programming principle of a stochastic control problem. However, the adaptive robust control framework is not a standard stochastic control problem, the framework intends to use a statistical learning method to estimate an unknown parameter, and allows the agent to incorporate the information to make an optimal decision.

In particular, the agent implements an online estimation of the unknown parameter in the stochastic system. Therefore, we need to study the convergence of an online estimator. In this Section, we present a standard result of the convergence of these estimators for both results in the discrete-time setup and the continuous-time setup.

For the discrete-time setup, we follow the paper Bielecki, Chen, and Cialenco (2017b). They use a stochastic approximation technique to recursively estimate an unknown parameter in a discrete-time setting, and allow them to find the value of the parameter θ^*. In the continuous-time setup, we follow the paper Sirignano and Spiliopoulos (2017) as a standard result.

2.3.1 The discrete-time setup

The following material is from Bielecki, Chen, and Cialenco (2017b). Let $(\Omega, \mathcal{F})$ be a measurable space. Let $\boldsymbol{\Theta} \subset \mathbb{R}^d$ be a set of parameters and Z be a discrete-time stochastic process taking values in $\mathbb{R}^m$, we now define the raw filtration generated by process Z, $\mathcal{F}_t = \sigma(Z_s, s \leq t)$. Assume that we have uncertainty about the true probability measure $\mathbb{P}_{\theta^*}$, but we know that it is contained in the set of probability measures $\{\mathbb{P}_\theta,\ \theta \in \boldsymbol{\Theta}\}$. After we observe some realizations of the stochastic process Z, we can estimate the true parameter θ^*. The following assumptions allow us to use a stochastic approximation method to find a recursive form of our estimator.

Assumption 2.3.1. (Regularity assumption)

(i) For each $\theta \in \boldsymbol{\Theta}$, under probability measure $\mathbb{P}_\theta$, Z is a homogeneous Markov process.

(ii) Under the probability measure $\mathbb{P}_{\theta^*}$, Z is an ergodic Markov process.

(iii) Under each probability measure $\mathbb{P}_\theta$, there is a positive density function p_θ such that $\mathbb{P}_\theta(Z_1 \in A | Z_0 = x) = \int_A p_\theta(x, y)\, dy$.

We use definition (2.6) from Revuz (2008) to define an ergodic Markov process and the following result gives us a long-term behavior of the ergodic process.

Proposition 2.3.2. *Let Z be an ergodic process under the probability measure $\mathbb{P}$. Then, for any function g such that $\mathbb{E}_{\mathbb{P}}[g(Z_0, ..., Z_n)] < \infty$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} g(Z_i, ..., Z_{i+n}) = \mathbb{E}_{\mathbb{P}}[g(Z_0, ..., Z_n)] \quad \mathbb{P} - a.s\,. \tag{2.3.1}$$

For any θ and n, we define $\pi_n(\theta) := \log p_\theta(Z_{n-1}, Z_n)$ which accounts for the law of the random variable Z_n under the measure $\mathbb{P}_\theta$ given Z_{n-1}. First, we recall the definition of Kullback-Leibler divergence from $\mathbb{P}_{\theta^*}$ to $\mathbb{P}_\theta$, we have that

$$D_{KL}(\mathbb{P}_{\theta^*}||\mathbb{P}_\theta) := \int \log\left(\frac{d\mathbb{P}_{\theta^*}}{d\mathbb{P}_\theta}\right) \frac{d\mathbb{P}_{\theta^*}}{d\mathbb{P}_\theta}\, d\mathbb{P}_\theta.$$

Thus, we can rewrite the Kullback-Leibler divergence in terms of the density function as follows

$$0 \le D_{KL}(\mathbb{P}_{\theta^*}||\mathbb{P}_\theta) := \int \log\left(\frac{p_{\theta^*}(Z_0, x)}{p_\theta(Z_0, x)}\right) p_{\theta^*}(Z_0, x)\, dx,$$

where the left inequality follows from the Gibbs' inequality. Since the probability density $p_\theta(Z_0, x)$ is always positive and finite, we split the logarithm terms to obtain

$$\int \log p_{\theta^*}(Z_0, x)\, p_{\theta^*}(Z_0, x)\, dx \ge \int \log p_\theta(Z_0, x)\, p_{\theta^*}(Z_0, x)\, dx\,,$$

which can expressed in expectation form as

$$\mathbb{E}_{\mathbb{P}_{\theta^*}}[\pi_1(\theta^*)] \ge \mathbb{E}_{\mathbb{P}_{\theta^*}}[\pi_1(\theta)]\,.$$

The true parameter maximises the information ratio under the true probability measure. Therefore, under additional regularity assumption on the density p_θ, the differentiation and expectation can interchange. It suffices to consider the unknown equation

$$\mathbb{E}_{\theta^*}[\psi_1(\theta)] = 0\,, \tag{2.3.2}$$

where the solution is θ^*, ∇ is a gradient operator, and $\psi_n(\theta) = \nabla\pi_n(\theta)$. This is a stochastic equation since the parameter θ^* is unknown. However, under the measure $\mathbb{P}_{\theta^*}$, the process Z is ergodic and by Proposition 2.3.2 so the sequence $\frac{1}{n}\sum_{i=0}^{n-1}\psi_i(\theta)$ as an approximation of $\mathbb{E}_{\theta^*}[\psi_1(\theta)]$. Therefore, we need a stochastic approximation method to recursively find a point estimator. The method to solve the stochastic equation is extensive, for example see Rao (1999) and Bertsekas and Tsitsiklis (1996).

Remark 2.3.3. Assumption (i) gives us a well defined π_n function, and assumption (ii) enables us to obtain the Gibbs' inequality. The last assumption allows us to use stochastic approximation techniques for equation (2.3.2), and to have a recursive form for the estimator.

In Bielecki, Chen, and Cialenco (2017a), the estimator is a quasi-asymptotically linear recursive point estimator, see the definition of this estimator in Definition 4.2 in Bielecki, Chen, and Cialenco (2017a). Their estimator has the weak consistent property and is asymptotically normal. The advantage of the recursion form is that it is computationally efficient and hence allows us to define the augmented space in the discrete model. This motivates us in Chapter 3, which properties of the estimator should have for our extension in a continuous-time setting.

Example 2.3.4. Consider a stochastic process Z, taking values in $\mathbb{R}$, that is independent and identically distributed, and has a law $\mathcal{N}(\theta, \sigma^2)$ under the probability measure $\mathbb{P}_\theta$ with unknown θ, where $\theta \in \mathbb{R}$. We define our estimator to be the maximum likelihood estimator (MLE). If the realization is $Z_1, ..., Z_n$, the sample mean is given by $\hat{\theta}_n = (Z_1 + \cdots + Z_n)/n$, which is the MLE estimator for θ.

2.3.2 The continuous-time set up

The material in this subsection follows from Sirignano and Spiliopoulos (2017) where they prove the convergence of the estimator when the observed process is a diffusion process.

In particular, the ergodic process $X = (X_t)_{t\geq 0}$ under the probability measure $\mathbb{P}_{\theta^*}$ that takes values in $\mathbb{R}^n$, and which satisfies the stochastic differential equation (SDE)

$$dX_t = b(X_t, \theta^*)\, dt + \sigma(X_t)\, dB_t \tag{2.3.3}$$

where the functions b and σ are Lipschitz functions and θ^* is the unknown parameter and B is a Brownian motion process. We denote the corresponding invariant measure for the process X under the probability measure $\mathbb{P}_{\theta^*}$ as π^{θ^*}. The stochastic gradient descent in continuous-time (SGDCT) algorithm to estimate the unknown drift parameter θ^* is given by

$$\begin{aligned} d\theta_t = \beta_t \big[&\nabla_\theta b(X_t, \theta_t)(\sigma(X_t)\, \sigma(X_t)^\intercal)^{-1} dX_t \\ &- \nabla_\theta\, b(X_t, \theta_t)(\sigma(X_t)\, \sigma(X_t)^\intercal)^{-1} b(X_t, \theta_t)\, dt \big] \, . \end{aligned} \tag{2.3.4}$$

Here is a brief summary of their approach to estimate the parameter θ^*. We denote by $\langle \cdot, \cdot \rangle$ the dot product operator and define the function $\bar{g}(\theta) := \int g(x, \theta)\, \pi^{\theta^*}(dx)$, where π^{θ^*} is a unique invariant measure of the process X under the probability measure $\mathbb{P}_{\theta^*}$, and where

$$g(x,\theta) := \frac{1}{2} \langle b(x,\theta) - b(x,\theta^*) \,, (\sigma(x)\,\sigma(x)^\intercal)^{-1}(b(x,\theta) - b(x,\theta^*)) \rangle \tag{2.3.5}$$

is a function that specifies how close the function $b(x,\theta)$ is to the drift component $b(x,\theta^*)$ of the process X.

Theorem 2.3.5. ***Convergence of the gradient*** . *Let the process θ_t follow (2.3.4) and assume the process X in (2.3.3) is sufficent ergodic. Then*

$$\left|\nabla_\theta \bar{g}(\theta_t)\right| \to 0 \quad as \quad t \to \infty \,,$$

almost surely.

Remark 2.3.6. Note that the difference of the discrete-time and the continuous-time setup is in the parameterised model. In the discrete-time setup, the parameter specifies the density of the transition probability of the underlying process, whereas the parameter in the continuous-time setup specifies the drift function of the underlying process.

In both approaches, the online estimators follow the gradient of a criterion function. Also note that in both subsections, we require the assumption that the underlying process is the ergodic, which are used to show the convergence of the online estimators.

Chapter 3

Online drift estimation for jump-diffusion processes [1]

Typically, one judges the usefulness of models by how good they are in describing a system and by the accuracy of their predictions. In both cases, the degree of success of a model depends on a number of ingredients, which include: model assumptions, mathematical tractability, and the methodology to estimate the parameters of the model. In this chapter, we assume that the system is a jump-diffusion process and we show how to estimate its drift component with a learning algorithm that updates the parameter estimates in continuous-time. Specifically, one observes the process $X = (X_t)_{t\geq 0}$ that takes values in $\mathbb{R}^n$, and which satisfies the stochastic differential equation (SDE)

$$dX_t = b(X_t, \theta^*)\, dt + \sigma(X_t)\, dB_t + \xi(X_{t-}, dL_t)\,, \tag{3.0.1}$$

under the probability measure $\mathbb{P}_{\theta^*}$ where the jump component $\xi(X_{t-}, dL_t)$ is given by

$$\xi(X_{t-}, dL_t) = \int_0^t \int \xi(X_{t-}, z)\, \tilde{\mu}(dt, dz)\,. \tag{3.0.2}$$

Here, $\theta^* \in \mathbb{R}^d$ is an unknown parameter that one needs to estimate from observations of the process X. The process $B = (B_t)_{t\geq 0}$ is a standard m-dimensional Brownian motion under the probability measure $\mathbb{P}_{\theta^*}$ and the functions b, σ, ξ are such that $b : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$, $\xi : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$. In the jump component of X_t, $\tilde{\mu}$ is a compensated random measure of the Poisson measure μ associated with the pure-jump Lévy process $L = (L_t)_{t\geq 0}$ (independent of B) with Lévy-Khintchine triplet $(0, 0, \nu)$, such that $\int_{\mathbb{R}^r} |z|\ \nu(dz) < \infty$ and $\nu(\mathbb{R}^r) < \infty$.

[1] This chapter has been published by Bernoulli Journal, see Bhudisaksang and Cartea (2021b).

Here is a brief summary of our approach to estimate the parameter θ^*. We denote by $\langle \cdot, \cdot \rangle$ the dot product operator and define the function $\bar{g}(\theta) := \int g(x,\theta)\, \pi^{\theta^*}(dx)$, where π^{θ^*} is a unique invariant measure of the process X under the probability measure $\mathbb{P}_{\theta^*}$, and where

$$g(x,\theta) := \frac{1}{2} \langle b(x,\theta) - b(x,\theta^*)\,, (\sigma(x)\,\sigma(x)^\intercal)^{-1}(b(x,\theta) - b(x,\theta^*)) \rangle \tag{3.0.3}$$

is a function that specifies how close the function $b(x,\theta)$ is to the drift component $b(x,\theta^*)$ of the process X.

Next, we employ the deterministic gradient descent method to construct a process θ_t that converges to a stationary point of the function $\bar{g}$. Recall that the deterministic descent direction $\nabla_\theta \bar{g}(\theta)$ moves the value of the parameter θ toward a stationary point (local or global) of the function $\bar{g}$. This stationary point is the estimate of the unknown parameter θ^*. In continuous-time, the updates of the gradient descent to a stationary point of the function $\bar{g}$ follow

$$\begin{aligned} d\theta_t &= -\,\beta_t\, \nabla_\theta \bar{g}(\theta_t)\, dt \\ &= -\,\beta_t\, \nabla_\theta g(X_t,\theta_t)\, dt + \beta_t\, (\nabla_\theta g(X_t,\theta_t) - \nabla_\theta \bar{g}(\theta_t))\, dt \\ &= \beta_t\, (\, \nabla_\theta b(X_t,\theta_t)\,)\, (\sigma(X_t)\, \sigma(X_t)^\intercal)^{-1}\, (b(X_t,\theta^*) - b(X_t,\theta_t))\, dt \\ &\quad + \beta_t\, (\nabla_\theta g(X_t,\theta_t) - \nabla_\theta \bar{g}(\theta_t))\, dt\,, \end{aligned} \tag{3.0.4}$$

where $\beta_t > 0$ is the learning rate.

One cannot implement the unbiased estimator in (3.0.4) because the drift term $b(x,\theta^*)$ depends on the unknown value of θ^*. Thus, we employ (3.0.1) to write

$$b(X_t,\theta^*)\, dt = dX_t - \int \xi(X_{t^-},z)\, \tilde{\mu}(dz,dt) - \sigma(X_t)\, dB_t\,, \tag{3.0.5}$$

which we substitute into (3.0.4), so the estimator

$$\begin{aligned} d\theta_t =& \beta_t \left[\nabla_\theta b(X_{t^-},\theta_{t^-})\, (\sigma(X_{t^-})\, \sigma(X_{t^-})^\intercal)^{-1}\, dX_t - \nabla_\theta\, b(X_t,\theta_t)\, (\sigma(X_t)\, \sigma(X_t)^\intercal)^{-1}\, b(X_t,\theta_t)\, dt \right] \\ &+ \beta_t\, (\nabla_\theta\, g(X_t,\theta_t) - \nabla_\theta\, \bar{g}(\theta_t))\, dt - \beta_t\, \nabla_\theta b(X_t,\theta_t)\, (\sigma(X_t)\, \sigma(X_t)^\intercal)^{-1}\, \sigma(X_t)\, dB_t \\ &- \beta_t\, \nabla_\theta b(X_{t^-},\theta_{t^-})\, (\sigma(X_{t^-})\, \sigma(X_{t^-})^\intercal)^{-1} \int \xi(X_{t-},z)\, \tilde{\mu}(dt,dz) \end{aligned} \tag{3.0.6}$$

does not depend on the unknown parameter θ^*. Note that, for an arbitrary function F, we have that $\int_0^t F(X_s,\theta_s)\, ds = \int_0^t F(X_{s^-},\theta_{s^-})\, ds$ and $\int_0^t F(X_s,\theta_s)\, dB_s = \int_0^t F(X_{s^-},\theta_{s^-})\, dB_s$ since the process X and θ have only countably many jumps and the Lebesgue measure of the set of times where $X_s \neq X_{s^-}$ or $\theta_s \neq \theta_{s^-}$ is 0.

In this chapter, we show that the integrals

$$\int_t^{\infty} \beta_s \left(\nabla_\theta g(X_s, \theta_s) - \nabla_\theta \bar{g}(\theta_s)\right) ds \tag{3.0.7}$$

and

$$\begin{aligned}\int_t^{\infty} & \beta_s \, \nabla_\theta b(X_s, \theta_s) \left(\sigma(X_s)\, \sigma(X_s)^\intercal\right)^{-1} \sigma(X_s) \, dB_s \\ & + \int_t^{\infty} \beta_s \, \nabla_\theta b(X_{s^-}, \theta_{s^-}) \left(\sigma(X_{s^-})\, \sigma(X_s)^\intercal\right)^{-1} \int \xi(X_{s-}, z) \, \tilde{\mu}(ds, dz)\end{aligned} \tag{3.0.8}$$

converge to zero as $t \to \infty$. Therefore, the expression in (3.0.6) is close to the deterministic descent direction $\nabla_\theta \bar{g}$ when t is large because the contribution of the last three terms in (3.0.6) becomes negligible. Thus, the stochastic gradient descent in continuous-time (SGDCT) algorithm to estimate the unknown drift parameter θ^* we propose in this chapter is given by

$$\begin{aligned} d\theta_t = \beta_t \big[& \nabla_\theta b(X_{t^-}, \theta_{t^-}) (\sigma(X_{t^-})\, \sigma(X_{t^-})^\intercal)^{-1} \, dX_t \\ & - \nabla_\theta \, b(X_t, \theta_t) (\sigma(X_t)\, \sigma(X_t)^\intercal)^{-1} \, b(X_t, \theta_t) \, dt \big] \, . \end{aligned} \tag{3.0.9}$$

One can see how the SGDCT estimator compares to the classical maximum likelihood estimator (MLE) when the process X is a diffusion. Above, if one excludes the jump component in (3.0.1), the process X becomes a diffusion, in which case the estimate of the parameter θ is the value that maximises the likelihood function

$$\mathcal{L}_t(\theta) = \int_0^t b(X_{s^-}, \theta) (\sigma(X_{s^-}) \sigma(X_{s^-})^\intercal)^{-1} dX_s - \frac{1}{2} \int_0^t b(X_s, \theta) \left(\sigma(X_s)\, \sigma(X_s)^\intercal\right)^{-1} b(X_s, \theta) ds. \tag{3.0.10}$$

Observe that the first order condition one obtains from maximising the likelihood (i.e., computing the gradient of $\mathcal{L}_t(\theta)$ with respect to θ and equating to zero) has the same form as the expression that appears inside the brackets in the first line of the estimator in (3.0.9).

There are two reasons that make a recursive estimator like that in (3.0.9) useful. First, online updates of the estimator are more computationally efficient than those based on offline updates. As new observations arrive, the online estimator is updated, while the offline algorithm solves an new optimisation problem to estimate the unknown parameter; e.g., find the maximiser of the MLE in (3.0.10). Therefore, the offline algorithm is in most cases redundant and inefficient, and in high dimensions it is computationally expensive and in some cases infeasible.

Second, the SGDCT estimator is well-suited for the recent advances in control theory that focus on 'adaptive robust control' to obtain strategies robust to model uncertainty, see Bielecki, Chen, and Cialenco (2017a). The adaptive control framework requires a recursive online estimator (i.e., an estimator that is continuously updated with the arrival of new information) and also requires that the coupled process $(X_t, \theta_t)_{t\geq 0}$ is Markov (to obtain the optimal strategy in feedback form). The SGDCT algorithm in (3.0.9) satisfies both requirements.

Finally, there is literature on offline estimators for diffusion and jump-diffusion processes that focuses on different aspects of the estimation algorithm and observation frequency. When an ergodic diffusion process is observed continuously, Kutoyants (2013) derives the statistical properties of the estimator of the drift parameter, such as consistency, asymptotic normality, and efficiency. These results are extended in Sorensen (1991) for an ergodic jump-diffusion process. The literature on parameter estimation with discrete observations is vast, see for example Yoshida (1992) for diffusions and Kessler (1997) for jump-diffusions.

On the other hand, the literature on parameter estimation for continuously observed processes is scant. In Levanony, Shwartz, and Zeitouni (1994), the authors assume that the process X follows a diffusion and develop an online MLE to obtain the drift parameter of the process X. In their approach, the online MLE requires one to compute the gradient of $\mathcal{L}_t(\theta)$ in (3.0.10) every time a new observation arrives, which is computationally expensive because the function $\nabla_\theta \mathcal{L}_t$ depends on the whole trajectory of the process X. Moreover, there are two papers that discuss the convergence of an online estimation of parameters in a diffusion-based model. Sirignano and Spiliopoulos (2017) introduce an online method to estimate parameters of a continuously observed diffusion, i.e., the process follows (3.0.1) without the jump component $\xi(X_{t-}, dL_t)$ – thus, our paper extends their work by including jumps in the stochastic process. The work of Surace and Pfister (2018) studies the problem in Sirignano and Spiliopoulos (2017) when the process X is a diffusion (i.e., no jumps) and it is partially observed.

The remainder of this chapter is organised as follows. In Section 2, we state conditions on the functions b, σ, ξ, so that the process X in (3.0.1) is exponentially ergodic. Then, we show that the term $\int_t^\infty \beta_s \left(\nabla_\theta g(X_s, \theta_s) - \nabla_\theta \bar{g}(\theta_s)\right) ds$ converges to zero as $t \to \infty$, and we prove the convergence of θ_t to a stationary point of the function $\bar{g}$. In Section 3, we illustrate the performance of the SGDCT estimator for the Ornstein–Uhlenbeck-type (OU-type) and bimodal processes. Last, Section 3.3 concludes and discusses extensions and applications of our work.

3.1 Online recursive estimator

In this Section, we state the assumptions that we use throughout this chapter and show our main results. We recall that the function

$$g(x,\theta) = \frac{1}{2}\left\langle b(x,\theta) - b(x,\theta^*)\,, (\sigma(x)\,\sigma(x)^\intercal)^{-1}(b(x,\theta) - b(x,\theta^*))\right\rangle, \tag{3.1.1}$$

specifies how close the function $b(x,\theta)$ is to the drift $b(x,\theta^*)$ of the process X. Note that the function $g(x,\theta) \geq 0$, and $g(x,\theta) = 0$ if $\theta = \theta^*$.[2]

A pivotal result in this chapter is to show that

$$\left|\Gamma_{k,\gamma}\right| \to 0 \quad \text{as} \quad k \to \infty\,, \text{ almost surely}\,, \tag{3.1.2}$$

where

$$\Gamma_{k,\gamma} := \int_{\underline{\tau}_k}^{\overline{\tau}_{k,\gamma}} \beta_s \left(\nabla_\theta g(X_s,\theta_s) - \nabla_\theta \bar{g}(\theta_s)\right) ds \tag{3.1.3}$$

is a deviation term. Here, the processes $\{\overline{\tau}_{k,\gamma}\}_{k\geq 1}$, $\{\underline{\tau}_k\}_{k\geq 1}$ are sequences of increasing stopping times and $|\theta|$ denotes the Euclidean norm of $\theta \in \mathbb{R}^d$, and recall that $\bar{g}(\theta) = \int g(x,\theta)\,\pi^{\theta^*}(dx)$. We refer to $\Gamma_{k,\gamma}$ in (3.1.3) as the deviation term because it represents the deviation of the stochastic gradient descent direction $\nabla_\theta g(X_s,\theta_s)$ from the deterministic descent direction $\nabla_\theta \bar{g}(\theta_s)$. Therefore, we need a careful analysis of the deviation term $\Gamma_{k,\gamma}$ to show that θ_t converges to the unknown parameter θ^*.

We show that the deviation term $\Gamma_{k,\gamma}$ can be written as a sum of four components: a weak solution of a non-local Poisson equation, a stochastic integral, a Riemann integral, and the covariation of two processes. Then, we prove that each component converges to zero as $k \to \infty$, so the estimator process θ_t converges to a stationary point of the function $\bar{g}$. That is, we show that $\lim_{t\to\infty}\left|\nabla_\theta \bar{g}(\theta_t)\right| = 0$ almost surely, and if the function $\bar{g}(\theta)$ is convex, the minimum we find is global, hence $\theta_t \to \theta^*$ as $t \to \infty$.

3.1.1 Main result and conditions

To streamline the discussion and results, we provide some preliminary assumptions. Throughout this chapter, the stochastic process X satisfies the SDE in (3.0.1).

[2]One could consider alternative functions to g to specify how close is $b(x,\theta)$ to $b(x,\theta^*)$. For example (in the one dimensional case), let $h(x,\theta) := \frac{1}{2}\left|(c(x,\theta) - c(x,\theta^*))/\sigma(x)\right|^2$, where c is an arbitrary function and $\bar{h}(\theta) = \int h(x,\theta)\,\pi^{\theta^*}(dx)$. Then, as in (3.0.4), the estimator of the unknown parameter would be given by $d\theta_t = -\beta_t\,\nabla_\theta \bar{h}(\theta_t)\,dt$. As discussed above, the drift term $c(x,\theta^*)$ depends on the unknown value of θ^*, so one needs to employ the dynamics of dX_t to write $b(X_t,\theta^*)$ and substitute it in the estimator. Thus, the function c should be such that one can factor out the term $b(X_t,\theta^*)$; for example, $c(x,\theta) = d(x)\,b(x,\theta)$ for a function $d(x)$.

Condition 3.1.1. The functions $\sigma(x)$ and $\xi(x, z)$ are globally Lipschitz in all variables and the function $b(x, \theta)$ is locally Lipschitz for all θ.

Condition 3.1.2. There exists a time $t > 0$, such that X_t admits a density $p_t^\theta(x, y)$ with respect to the Lebesgue measure on $\mathbb{R}^n$, bounded in $y \in \mathbb{R}^n$, and bounded in $x \in K$ for every compact set $K \subset \mathbb{R}^n$. Moreover, for every $x \in \mathbb{R}^n$, and every open ball $U \in \mathbb{R}^n$, there exists a point $z = z(x, U) \in \text{supp}(v)$, where $\text{supp}(\nu)$ is the support of the measure ν (see definition in Appendix A.1.5), such that $\xi(x, z) \in U$.

Condition 3.1.3. Exponential ergodicity of the process X.

(i) For all $p > 0$, $\int_{|z|>1} |z|^p \, \nu(dz) < \infty$ and $\nu(\mathbb{R}^r) < \infty$.

(ii) For all θ there exists a constant $C > 0$, such that $x\, b(x, \theta) \leq -C\, |x|^2$ as $|x| \to \infty$.

(iii) $\left|\xi(x, z)\right| / |x| \to 0$ as $|x| \to \infty$ for all z.

(iv) There exist positive constants C_1 and C_2, such that $C_1 \leq x^\intercal \sigma(x)\, \sigma(x)^\intercal x / |x|^2 \leq C_2$ for all $x \in \mathbb{R}^n / \{0\}$ and $\sigma(x)\, \sigma(x)^\intercal$ is bounded.

Under Conditions 3.1.1 and 3.1.3, the SDE in (3.0.1) admits a unique non-explosive càdlàg adapted solution that possesses the strong Markov property. Masuda (2007) employs Condition 3.1.2 to show that the process X is irreducible; a property required to show that X is ergodic. Also, the process X has a unique invariant measure and it is exponentially ergodic due to Condition 3.1.3. Finally, the process X is of finite jump activity because ν is a finite measure, so there is a finite number of jumps on every compact interval almost surely.

Remark 3.1.4. Condition (3.1.2) does hold when the process X is non-degenerate and the measure ν is a finite measure, see also Masuda (2008) and Kwon and Chaniio (1999).

Condition 3.1.5. The learning rate β_t is such that the following hold: $\int_0^\infty \beta_t \, dt = \infty$, $\int_0^\infty \beta_t^2 \, dt < \infty$, $\int_0^\infty |\beta_t'| \, dt < \infty$, where β_t' is the derivative of β_t with respect to t, and there is a $p > 0$, such that $\lim_{t\to\infty} \beta_t^2 \, t^{2\,p+1/2} = 0$.

Condition 3.1.6. The function $b(x, \theta)$ is twice-differentiable with respect to both x and θ. Also, there exist positive constants K and q, such that

$$\left| \frac{\partial \nabla_\theta g}{\partial x}(x, \theta) \right| + \left| \nabla_\theta g(x, \theta) \right| \leq K \left(1 + |\theta|\right) \left(1 + |x|^q\right) , \tag{3.1.4}$$

and

$$\sum_{i=1}^{2}\left|\frac{\partial^i \nabla_\theta g}{\partial \theta^i}(x,\theta)\right| + \left|\nabla_\theta b(x,\theta)\right| \le K\left(1+|x|^q\right). \tag{3.1.5}$$

Moreover, we assume that $\nabla_\theta \bar{g}(\theta)$ is globally Lipschitz.

Condition 3.1.7. There exists a positive function $\kappa(x)$, such that

$$\langle -\nabla_\theta g(x,\theta),\, \theta/|\theta|\,\rangle \le -\kappa(x)\,|\theta|\,. \tag{3.1.6}$$

Thus, there exists a function $\lambda(x)$ with at most polynomial growth in $|x|$, such that for all $\theta_1, \theta_2 \in \mathbb{R}^d$ and $x \in \mathbb{R}^n$, we have

$$\left|\tau(x,\theta_1) - \tau(x,\theta_2)\right| \le \left|\lambda(x)\right| |\theta_1 - \theta_2|\,, \tag{3.1.7}$$

where

$$\tau(x,\theta) := \left\langle \nabla_\theta b(x,\theta)\, \nabla_\theta b(x,\theta)^\intercal\, \theta/|\theta|\,,\, \theta/|\theta|\right\rangle^{1/2}. \tag{3.1.8}$$

Lastly, for all $x_2 \ge x_1$, we have $\xi(x_1,z) - \xi(x_2,z) \le x_2 - x_1$, where the greater or equal symbol $\ge$ indicates the inequality holds for all coordinates.

Condition 3.1.5 ensures that the size of the learning rate β is such that the SGDCT estimator for the jump-diffusion process in (3.0.9) converges to a stationary point of the function $\bar{g}(\theta)$. Condition 3.1.6 is needed for the proof of the convergence of the deviation term $\Gamma_{k,\gamma}$, i.e., $\Gamma_{k,\gamma} \to 0$ as $k \to \infty$, see (3.1.2). Condition 3.1.7 allows us to use a comparison theorem to show that the bound $\sup_{t>0} \mathbb{E}\left[|\theta_t|^p\right] < \infty$ for all $p > 0$.

3.1.2 Weak solution of a non-local Poisson equation.

In this subsection, we define a notion of a weak solution of a Poisson equation to construct a weak solution of the process X. The following terminology is from Kulik and Veretennikov (2012).

Definition 3.1.8. For a Markov process X, we say that a measurable function $f : \mathbb{R}^n \to \mathbb{R}^d$ belongs to the domain of the extended generator $\mathcal{A}$ of the process X if there exists a measurable function $g : \mathbb{R}^n \to \mathbb{R}^d$, such that the process $f(X_t) - \int_0^t g(X_s)\,ds$ is a local $\mathbb{F}^X$-martingale. Then, we say that the function f is the weak solution of the Poisson equation $\mathcal{A}f = g$.

We also need the following notation. The f-norm is denoted by

$$\|\omega\|_f := \sup_{|g|\le f}\left|\int g(y)\,\omega(dy)\right|,$$

where ω is a signed measure and $f \geq 1$ is a Borel measurable function. Also, we say that the function g is centered if $\int g(x,\theta)\,\pi^{\theta^*}(dx) = 0$.

Next, we state an auxiliary Lemma to show the ergodicity of the process X and to show the existence of moments of all orders for the processes X and θ.

Lemma 3.1.9. ***Ergodicity and moment bounds of the processes X and θ.*** *Suppose that Conditions 3.1.1 to 3.1.6 hold. Then, there exists a unique invariant measure π^{θ^*}, such that $\int |x|^p\,\pi^{\theta^*}(dx) < \infty$ and there exist positive constants C_1 and C_2, such that*

$$\left\| \mathbb{P}_{\theta^*,t}(x,\cdot) - \pi^{\theta^*}(\,\cdot\,) \right\|_{1+|x|^p} \leq C_1 \left(1 + |x|^p\right) e^{-C_2\, t}, \tag{3.1.9}$$

for all $p > 0$, $x \in \mathbb{R}^n$ and $t > 0$. Thus, $\sup_{t>0} \mathbb{E}_{x,\theta^*}\left[|X_t|^p\right] < \infty$ for all $p > 0$. In addition, if Condition 3.1.7 holds, we have that

$$\sup_{t>0} \mathbb{E}_{x,\theta^*}\left[|\theta_t|^p\right] < \infty \quad \textit{for all} \quad p > 0\,. \tag{3.1.10}$$

Moreover, $\sup_{t>0} \mathbb{E}_{x,\theta^*}\left[|X_{t-}|^p\right] < \infty$ and $\sup_{t>0} \mathbb{E}_{x,\theta^*}\left[|\theta_{t-}|^p\right] < \infty$ for all $p > 0$.

For a proof see Appendix A.1.1.

Now, we state our first theorem which is important to show that the deviation term $\Gamma_{k,\gamma} \to 0$ as $k \to \infty$, see (3.1.2).

Theorem 3.1.10. ***Weak Poisson equation.*** *Let the process X follow (3.0.1). Suppose that Conditions 3.1.1 to 3.1.6 hold. Let f be a centered function (i.e., $\int f(x,\theta)\,\pi^{\theta^*}(dx) = 0$ for all θ) and assume that for each θ, there exist positive constants C_θ and q, such that $\left|f(x,\theta)\right| \leq C_\theta\,\left(1 + |x|^q\right)$ for all x. Then, the function*

$$F(x,\theta) := \int_0^\infty \mathbb{E}_{x,\theta^*}\left[f(X_t,\theta)\right]\,dt \tag{3.1.11}$$

is well-defined and satisfies, in a weak sense, the Poisson equation

$$\mathcal{A}_{x,\theta^*} F(x,\theta) = -f(x,\theta)\,, \tag{3.1.12}$$

where $\mathcal{A}_{x,\theta^*}$ is the extended generator of the process X under the probability measure $\mathbb{P}_{\theta^*}$ (see Definition 3.1.8).

Proof. First, we show that the function $F(x,\theta)$ is well-defined. Recall that $\mathbb{E}_{x,\theta^*}[f(X_t,\theta)] = \int f(y,\theta)\,p_t(x,dy;\theta^*)$ and that $p_t(x,dy;\theta^*)$ is the distribution of X_t under the probability measure $\mathbb{P}_{\theta^*}$ and $X_0 = x$. Note that

$$\begin{aligned} \left|\mathbb{E}_{x,\theta^*}[f(X_t,\theta)]\right| &= \left| \int f(y,\theta)\,p_t(x,dy;\theta^*) - \int f(y,\theta)\,\pi^{\theta^*}(dy) \right| \\ &\leq C_1\,\left(1 + |x|^q\right)\,e^{-C_2\,t}, \end{aligned} \tag{3.1.13}$$

Next, we show that the function F is the solution to the Poisson equation in (3.1.12). It suffices to show that for all θ, the process $F(X_t,\theta)+\int_0^t f(X_s,\theta)\,ds$ is an $\mathbb{F}$-martingale under the probability measure $\mathbb{P}_{\theta^*}$, as in the definition of the weak solution of the Poisson equation. Thus, we first show that the first moment of $F(X_t,\theta)+\int_0^t f(X_s,\theta)\,ds$ is finite for all t.

From (3.1.11) and (3.1.13), we have that $F(X_t,\theta) \leq (C_1/C_2)\,(1+|X_t|^q)$ and

$$\mathbb{E}_{\theta^*}\left[\left|\int_0^t f(X_s,\theta)\,ds\right|\right] \leq \int_0^t \mathbb{E}_{\theta^*}\left[\left|f(X_s,\theta)\right|\right]\,ds \leq C_\theta \int_0^t \mathbb{E}_{\theta^*}\left[1+|X_s|^q\right]\,ds < \infty\,,$$

where the last inequality follows because X_t has a finite q-moment, see Lemma 3.1.9.

Hence, to show that the process $F(X_t,\theta)+\int_0^t f(X_s,\theta)\,ds$ is an $\mathbb{F}$-martingale, let $t \geq u \geq 0$ and write

$$\begin{aligned}
&\mathbb{E}_{\theta^*}\left[F(X_t,\theta)+\int_0^t f(X_s,\theta)\,ds \,\middle|\, \mathcal{F}_u\right]\\
&= \mathbb{E}_{\theta^*}\left[\left[\int_0^\infty \mathbb{E}_{x,\theta^*} f(X_{t+s},\theta)\,ds\right]+\int_0^t f(X_s,\theta)\,ds \,\middle|\, \mathcal{F}_u\right] && \text{(Definition in (3.1.11))}\\
&= \mathbb{E}_{\theta^*}\left[\int_0^\infty \mathbb{E}_{X_t,\theta^*}\left[f(X_{s+t},\theta)\,\middle|\mathcal{F}_t\right]ds+\int_0^t f(X_s,\theta)\,ds \,\middle|\, \mathcal{F}_u\right] && \text{(Markove property of X)}\\
&= \int_0^u f(X_s,\theta)\,ds + \mathbb{E}_{X_u,\theta^*}\left[\int_0^\infty f(X_{s+u},\theta)\,ds \,\middle|\, \mathcal{F}_u\right]\\
&= \int_0^u f(X_s,\theta)\,ds + F(X_u,\theta)\,,
\end{aligned}$$

where the second to last equation follows from conditional expectation. □

Remark 3.1.11. Note that in Theorem 3.1.10 above, we do not require any assumptions on the existence or the regularity of the density function of the process X. The regularity of the density function of the jump-diffusion process is difficult to analyse.

The following Lemma is useful in the next subsection.

Lemma 3.1.12. *For any $p>0$, there is a positive constant C_p, such that*

$$\int_{\mathbb{R}}\left(1+|x|^p\right)\,\pi^{\theta^*}(dx) \leq C_p\,.$$

In addition, we have that $\left|\nabla_\theta \bar{g}(\theta)\right| \leq C_q\,K\,\left(1+|\theta|\right)$ and $\left|\nabla^2_\theta \bar{g}(\theta)\right| \leq C_q\,K$, where K is the same positive constant as that in Condition 3.1.6.

Proof. Recall that $\bar{g}(\theta) = \int g(x,\theta)\,\pi^{\theta^*}(dx)$. The process X is exponentially ergodic and from Lemma 3.1.9 we have

$$\int_{\mathbb{R}} (1 + |x|^p)\,\pi^{\theta^*}(dx) \leq C_p\,.$$

Note that the invariant measure π^{θ^*} does not depend on θ, so we can interchange the gradient and the integral in the function $\nabla_\theta \bar{g}(\theta)$. From Condition 3.1.6, we also have that

$$\left|\nabla_\theta \bar{g}(\theta)\right| \leq \int_{\mathbb{R}} \left|\nabla_\theta g(x,\theta)\right|\,\pi^{\theta^*}(dx) \leq K\left(1+|\theta|\right)\int_{\mathbb{R}}\left(1+|x|^q\right)\,\pi^{\theta^*}(dx) \leq C_q\,K\left(1+|\theta|\right)\,,$$

and

$$\left|\nabla^2_\theta \bar{g}(\theta)\right| \leq \int_{\mathbb{R}} \left|\nabla^2_\theta g(x,\theta)\right|\,\pi^{\theta^*}(dx) \leq K\int_{\mathbb{R}}\left(1+|x|^q\right)\,\pi^{\theta^*}(dx) \leq C_q\,K\,.$$

□

3.1.3 Convergence of the deviation term

In this subsection, we propose a novel method to prove the convergence of the deviation term $\Gamma_{k,\gamma}$, see (3.1.2) and (3.1.3). When the process X is a diffusion without jumps, one of the steps to prove the convergence of the SGDCT estimator is to show that the function

$$G(x,\theta) := \int_0^\infty \mathbb{E}_{x,\theta^*}\left[\nabla_\theta g(X_t,\theta) - \nabla_\theta \bar{g}(\theta)\right]\,dt \tag{3.1.14}$$

is a classical solution of the Poisson equation (3.1.12) when $f(x,\theta) = \nabla_\theta g(x,\theta) - \nabla_\theta \bar{g}(\theta)$, see Pardoux and Veretennikov (2001). Thus, in the pure-diffusion case, one can apply Itô's Lemma to the function $\beta_t\,G$ to obtain the following decomposition

$$\beta_{\overline{\tau}_{k,\gamma}}\,G(X_{\overline{\tau}_{k,\gamma}},\theta_{\overline{\tau}_{k,\gamma}}) - \beta_{\mathcal{I}_k}\,G(X_{\mathcal{I}_k},\theta_{\mathcal{I}_k}) = \Gamma_{k,\gamma} + \text{martingale} + \text{Riemann integral}\,,$$

where the deviation term $\Gamma_{k,\gamma}$ is as in (3.1.3). With this decomposition, one shows that the deviation term $\Gamma_{k,\gamma} \to 0$ as $k \to \infty$, by proving that the martingale term, the Riemann integral, and the function $\beta_t\,G(X_t,\theta_t)$ converge to zero as $k \to \infty$.

However, when the process X follows a jump-diffusion, the Poisson equation (3.1.12) has a non-local term, and to the best of our knowledge there is no general result to show that the Poisson equation with a non-local operator admits a classical solution. It is difficult, or perhaps not possible, to show that the function G in (3.1.14) is second order differentiable with respect to x when X is a jump-diffusion process. Several papers in the literature also encounter a similar difficulty.

Uehara (2018) states that the properties of the function G are difficult to determine and may require Malliavin calculus to study the gradient estimation of the function G – the author assumes a stochastic process similar to ours. Moreover, Kulik and Veretennikov (2012) mention that there is no analogue to the results in Pardoux and Veretennikov (2003) when the process X has a jump component.

Therefore, in our setup we face the problem that the function G is not differentiable with respect to the variable x because the process X is a jump-diffusion. Hence, we cannot employ the classical Itô formula on the function G, so below we prove an 'extended Itô formula' that we employ in our analysis. We provide a brief summary of the steps we follow.

First, we show that G is a locally Lipschitz function with the following bound

$$\big|G(x,\theta) - G(y,\theta)\big| \le C \,\left(1 + |\theta|\right) |x-y| \,\left(1 + |x|^q + |y|^q\right) , \tag{3.1.15}$$

where C is a positive real number that does not depend on x, y, θ, $q > 0$. Then, because G is a weak solution of a Poisson equation, see (3.1.12), and because it is twice differentiable with respect to θ, we show that

$$\begin{aligned}\beta_{\overline{\tau}_{k,\gamma}}\, G(X_{\overline{\tau}_{k,\gamma}}, \theta_{\overline{\tau}_{k,\gamma}}) - & \beta_{\underline{\tau}_k}\, G(X_{\underline{\tau}_k}, \theta_{\underline{\tau}_k}) \\ & = \Gamma_{k,\gamma} + \text{martingale} + \text{Riemann integral} + \text{covariation}\,,\end{aligned} \tag{3.1.16}$$

where the covariation term in the equation above is a càdlàg process.

Second, we study the terms in the above equation in the limit $k \to \infty$. We show that the left-hand side of (3.1.16) converges to zero and each of the last three terms on the right-hand side of (3.1.16) also converges to zero. Therefore, the representation in (3.1.16) implies that the deviation term $\Gamma_{k,\gamma} \to 0$ as $k \to \infty$.

Now, we state a Lemma and then prove that the function G is locally Lipschitz. The Lemma uses the result that the function $\mathbb{E}_{x,\theta^*}[f(X_t,\theta)]$ is locally Lipschitz with respect to x, see Theorem 2.2 in Wang (2010).

Lemma 3.1.13. *Suppose Conditions 3.1.1 to 3.1.6 hold and let $C > 0$, $q > 0$. Assume a function $f : \mathbb{R}^{n+d} \to \mathbb{R}^d$ satisfies $\big|f(x,\theta) - f(y,\theta)\big| \le C\,\left(1 + |x|^q + |y|^q\right)|x-y|$ for all x, y and $t \ge 0$. Then there exist constants $\delta \in (0,1)$ and $C_\delta > 0$, such that for all $\theta \in \mathbb{R}^d$ the following inequality holds*

$$\begin{aligned}&\big|\mathbb{E}_{x,\theta^*}[f(X_t,\theta)] - \mathbb{E}_{y,\theta^*}[f(X_t,\theta)]\big| \\ &\le \begin{cases} (2+\delta^{-1})\, e^{-C_\delta\, t}\, C\,\left(1 + |x|^q + |y|^q\right)|x-y|\,, & |x-y| \le \delta\,, \\ 2\,\delta^{-1}\, C\,\left(1 + |x|^q + |y|^q\right)|x-y|\,, & |x-y| > \delta\,, \end{cases}\end{aligned}$$

where the constant C does not depend on f.

For a proof see Appendix A.1.2.

Next, we show that the function G in (3.1.14) is locally Lipschitz in the variable x. We employ Conditions 3.1.1 to 3.1.7, which are easily checked, to show the dissipativity (stability with respect to initial condition) of the process X. The main idea of the proof of the proposition below is that for a small value of t, we use Lemma 3.1.13 to control the value $|\mathbb{E}_{x,\theta^*}[\tilde{g}(X_t,\theta)]-\mathbb{E}_{y,\theta^*}[\tilde{g}(X_t,\theta)]|$, where $\tilde{g}(x,\theta) := \nabla_\theta g(x,\theta)-\nabla_\theta \bar{g}(\theta)$. And for a large value of t, we use the exponential ergodic property of the process X. These two results imply the following proposition.

Proposition 3.1.14. *Suppose Conditions 3.1.1 to 3.1.6 hold and let $q>0$. Then, the function G in (3.1.14) is locally Lipschitz with respect to x and it is twice-differentiable with respect to θ. Thus, there exists a positive constant $C_{G,q}$, such that*

$$\big|G(x,\theta)-G(y,\theta)\big| \le C_{G,q}\,\big(1+|\theta|\big)\,|x-y|\,\big(1+|x|^q+|y|^q\big) \tag{3.1.17}$$

and

$$\sum_{i=1}^{2}\left|\frac{\partial^i G}{\partial\theta^i}(x,\theta)\right| \le C_{G,q}\left(1+|x|^q\right), \tag{3.1.18}$$

for all x, y, θ.

Proof. Let $\tilde{g}(x,\theta) := \nabla_\theta g(x,\theta)-\nabla_\theta\bar{g}(\theta)$. Then, due to Condition 3.1.6, write

$$\big|\tilde{g}(x,\theta)-\tilde{g}(y,\theta)\big| = \big|\nabla_\theta g(x,\theta)-\nabla_\theta g(y,\theta)\big| \le K\,\big(1+|\theta|\big)\,|x-y|\,\big(1+|x|^q+|y|^q\big)\,. \tag{3.1.19}$$

Let δ be the same constant as that in Lemma 3.1.13. Then, for $|x-y|\le\delta$, we have

$$\begin{aligned}\big|G(x,\theta)-G(y,\theta)\big| &\le \int_0^\infty \Big|\mathbb{E}_{x,\theta^*}\big[\tilde{g}(X_t,\theta)\big]-\mathbb{E}_{y,\theta^*}\big[\tilde{g}(X_t,\theta)\big]\Big|\,dt\\ &\le \int_0^\infty (2+\delta^{-1})\,e^{-C_\delta t}\,K\,\big(1+|\theta|\big)\,\big(1+|x|^q+|y|^q\big)|x-y|\,dt\\ &\le \frac{2+\delta^{-1}}{C_\delta}\,K\,\big(1+|\theta|\big)\,\big(1+|x|^q+|y|^q\big)\,|x-y|\,,\end{aligned} \tag{3.1.20}$$

where the second inequality follows from Lemma 3.1.13. Now, consider the case $|x-y|\ge\delta$ and recall that

$$\Big\|\mathbb{P}_t(x,\cdot)-\pi^{\theta^*}(\cdot)\Big\|_{1+|x|^q} \le C_1\,\big(1+|x|^q\big)\,e^{-C_2 t},$$

from Lemma 3.1.9. Use the triangle inequality to write

$$\big\|\mathbb{P}_t(x,\cdot)-\mathbb{P}_t(y,\cdot)\big\|_{1+|x|^q} \le 2\,C_1\,(1+|x|^q+|y|^q)\,e^{-C_2 t}\,.$$

Therefore, for $t_0 \geq (1/C_2)\,\log(C_1/(C_2\,\delta))$, we have

$$\begin{aligned}
&\int_{t_0}^{\infty} \left|\mathbb{E}_{x,\theta^*}[\tilde{g}(X_t,\theta)] - \mathbb{E}_{y,\theta^*}[\tilde{g}(X_t,\theta)]\right|\, dt \\
&\qquad \leq K\,\big(1+|\theta|\big) \int_{t_0}^{\infty} 2\,C_1\,\big(1+|x|^q+|y|^q\big)\, e^{-C_2\,t}\, dt \\
&\qquad = 2\,K\,\big(1+|\theta|\big)\,(C_1/C_2)\,\big(1+|x|^q+|y|^q\big)\, e^{-C_2\,t_0} \\
&\qquad \leq 2\,K\,\big(1+|\theta|\big)\,\big(1+|x|^q+|y|^q\big)\,|x-y|\;,
\end{aligned} \tag{3.1.21}$$

where the last inequality holds because $t_0 \geq (1/C_2)\,\log(C_1/(C_2\,\delta))$. From Lemma 3.1.13, write

$$\begin{aligned}
&\int_{0}^{t_0} \left|\mathbb{E}_{x,\theta^*}[\tilde{g}(X_t,\theta)] - \mathbb{E}_{y,\theta^*}[\tilde{g}(X_t,\theta)]\right|\, dt \\
&\qquad \leq K\,\big(1+|\theta|\big) \int_{0}^{t_0} \frac{2\,C_\delta}{\delta}\,\big(1+|x|^q+|y|^q\big)\,|x-y|\, dt \\
&\qquad \leq K\,\big(1+|\theta|\big)\,\tilde{C}_\delta\,\big(1+|x|^q+|y|^q\big)\,|x-y|\,,
\end{aligned} \tag{3.1.22}$$

where $\tilde{C}_\delta$ is a constant. Therefore, (3.1.17) holds by (3.1.20), (3.1.21), (3.1.22). Note that we can interchange the differentiation with the integration operator, therefore

$$\left|\frac{\partial^i G}{\partial \theta^i}(x,\theta)\right| \leq \int_0^{\infty} \left|\mathbb{E}_{x,\theta^*}\left[\frac{\partial^i \nabla_\theta g}{\partial \theta^i}(X_t,\theta) - \frac{\partial^i \nabla_\theta \bar{g}}{\partial \theta^i}(\theta)\right]\right| dt \leq C_q\,\big(1+|x|^q\big)\;, \tag{3.1.23}$$

where the last inequality follows from Lemma 3.1.12. □

In the next Lemma we construct predictable representation processes $\tilde{G}_1$ and $\tilde{G}_2$ for the process $G(X_t,\theta) + \int_0^t \big(\nabla_\theta g(X_s,\theta) - \nabla_\theta \bar{g}(\theta)\big)\, ds$, which are also measurable with respect to the variable θ.

Lemma 3.1.15. ***Martingale representation.*** *For each θ, let*

$$M_t^\theta := G(X_t,\theta) - G(X_0,\theta) + \int_0^t \big(\nabla_\theta g(X_s,\theta) - \nabla_\theta \bar{g}(\theta)\big)\, ds\,,$$

which is a martingale under $\mathbb{P}_{\theta^*}$ with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ generated by the process X, see Theorem 3.1.10 above. Then, there exist measurable functions $\tilde{G}_1(x,\theta)$ and $\tilde{G}_2(x,\theta,z)$ that are continuous with respect to θ, such that

$$M_t^\theta = \int_0^t \tilde{G}_1(X_s,\theta)\, dB_s + \int_0^t \int_{\mathbb{R}^r} \tilde{G}_2(X_{s^-},\theta,z)\, \tilde{\mu}(ds,dz)\,. \tag{3.1.24}$$

Recall that $\tilde{\mu}$ is the compensated random measure of the jump component in the process X, see (3.0.2). Thus, there exists a constant $C_{M,q} > 0$ independent of x and θ, such that for all $\theta \in \mathbb{R}^d$, the following two inequalities hold:

$$\begin{aligned} \left|\tilde{G}_1(x,\theta)\right| &\leq C_{M,q}\left(1+|\theta|\right)\left(1+|x|^q\right) \\ \left|\tilde{G}_2(x,\theta,z)\right| &\leq C_{M,q}\left(1+|\theta|\right)|x||z|\left(1+|x|^q|z|^q+|x|^q\right) \end{aligned} \tag{3.1.25}$$

for all x almost everywhere.

Proof. First, we show that there exists a Borel measurable function $\tilde{G}(x,\theta)$ that agrees with the weak derivative of G in x almost everywhere and that it is continuous with respect to θ. Let

$$\hat{G}_i(x,\theta) := \limsup_{h\to 0}\left\{\frac{\nabla_\theta G(x+h\,e_i,\theta)-\nabla_\theta G(x,\theta)}{h}\right\}, \tag{3.1.26}$$

where e_i is the unit vector (the vector whose i-th entry is 1 and 0 elsewhere) and the function $\hat{G}_i$ takes values in $\mathbb{R}^d\times\mathbb{R}^d$, which is a Borel measurable function and agrees with the weak derivative of $\nabla_\theta G$ with respect to x in the $i-$th coordinate almost everywhere. Therefore, we have that

$$\int \phi\,\hat{G}_i(x,\theta)\,dx = -\int \phi_i\nabla_\theta G(x,\theta)\,dx\,, \tag{3.1.27}$$

for all $1\leq i\leq n$ and $\phi\in C_c^\infty(\mathbb{R}^n)$ where ϕ_i is the derivative of the i-th coordinate of the function ϕ.

Next, define the function

$$\tilde{G}(x,\theta) := \tilde{G}(x,0)+\int_0^1\left[\hat{G}_1(x,\theta\,t)\,\theta,\cdots,\hat{G}_n(x,\theta\,t)\,\theta\right]dt\,,$$

where $\tilde{G}=[\tilde{G}_1,\cdots,\tilde{G}_n]$ takes values in $\mathbb{R}^d\times\mathbb{R}^n$ and

$$\tilde{G}_i(x,0) := \limsup_{h\to 0}\left\{\frac{G(x+h\,e_i,0)-G(x,0)}{h}\right\}.$$

Therefore, we have that

$$\int \phi\,\tilde{G}_i(x,0)\,dx = -\int \phi_i\,G(x,0)\,dx\,, \tag{3.1.28}$$

because the function $\tilde{G}_i(x,0)$ agrees with the weak derivative of the function $G(x,0)$ almost everywhere. For all θ and $\phi\in C_c^\infty(\mathbb{R}^n)$, by the definition of $\tilde{G}$ and Fubini's theorem, write

$$\begin{aligned}
\int \phi\, \tilde{G}_k(x,\theta)\, dx &= \int \phi \left(\tilde{G}_k(x,0) + \int_0^1 \hat{G}_k(x,\theta\, t)\, \theta\, dt \right) dx \\
&= \int \phi\, \tilde{G}_k(x,0)\, dx + \int \int_0^1 \phi\, \hat{G}_k(x,\theta\, t)\, \theta\, dt\, dx \\
&= - \int \phi_k\, G(x,0)\, dx + \int_0^1 \int \phi\, \hat{G}_k(x,\theta\, t)\, \theta\, dx\, dt \\
&= - \int \phi_k\, G(x,0)\, dx + \int_0^1 - \int \phi_k\, \nabla_\theta G(x,\theta\, t)\, \theta\, dx\, dt \\
&= - \int \phi_k\, G(x,0)\, dx - \int \phi_k \int_0^1 \nabla_\theta G(x,\theta\, t)\, \theta\, dt\, dx \\
&= - \int \phi_k\, G(x,0)\, dx - \int \phi_k \left(G(x,\theta) - G(x,0) \right) dt\, dx \\
&= - \int \phi_k\, G(x,\theta)\, dx\,,
\end{aligned} \tag{3.1.29}$$

where the third equation follows from (3.1.28), the fourth equation follows from (3.1.27). Therefore, from (3.1.29), for each θ, the function $\tilde{G}(x,\theta)$ agrees with the weak derivative in x of the function $G(x,\theta)$ almost everywhere.

The function $\tilde{G}$ is continuous with respect to the variable θ and Borel measurable with respect to x. Therefore, by the Carathéodory theorem, $\tilde{G}(x,\theta)$ is a Borel measurable function continuous in θ and it agrees with the weak derivative of G in x almost everywhere.

Next, we show that

$$\big[G(X,\theta), B\big]_t = \int_0^t \tilde{G}(X_s,\theta)\, d[X,B]_s\,, \tag{3.1.30}$$

where $[G(X,\theta), B]$ denotes the covariation of the processes $G(X,\theta)$ and B.

If the function G is differentiable with respect to x, then it is straightforward to show that (3.1.30) holds. However, we need to analyse the function G very carefully because it is only locally Lipschitz. Let $C_{\text{loc}}(\mathbb{R}^d)$ be the space of locally Lipschitz functions and let $\mathbb{D}$ be the space of càdlàg processes equipped with the metric of uniform convergence in probability (u.c.p.) topology. Then, the mapping

$$G(\cdot,\theta) \mapsto \big[G(X,\theta), B\big] \tag{3.1.31}$$

is a continuous mapping from $C_{\text{loc}}(\mathbb{R}^d)$ to $\mathbb{D}$ by a similar argument to that in the proof of Theorem 3.8 in Errami, Russo, and Vallois (2002). Consider $G^k = G * \phi_k$

where $*$ is the convolution operator and $\{\phi_k\}_{k\geq 1}$ is a sequence of classical mollifiers converging to the Dirac measure at 0. Then,

$$\Big[G^k(X,\theta), B\Big]_t = \int_0^t \nabla_x G^k(X_s,\theta)\, d[X,B]_s\,, \tag{3.1.32}$$

because the function G^k is differentiable with respect to x. Moreover, the continuity of the mapping in (3.1.31) implies that

$$\begin{aligned}\big[G(X,\theta), B\big]_t = \lim_{k\to\infty}\Big[G^k(X,\theta), B\Big]_t &= \lim_{k\to\infty}\int_0^t \nabla_x G^k(X_s,\theta)\, d[X,B]_s \\ &= \int_0^t \tilde{G}(X_s,\theta)\, d[X,B]_s\,,\end{aligned}$$

$\mathbb{P}_{\theta^*}$ almost surely, where the last equation follows from the dominated convergence argument – in Appendix A.1.3 we provide the proof.

This shows (3.1.30), so we write

$$\big[G(X,\theta), B\big]_t = \int_0^t \tilde{G}_1(X_s,\theta)\, ds\,, \tag{3.1.33}$$

where $\tilde{G}_1(x,\theta) := \tilde{G}(x,\theta)\,\sigma(x)$. Therefore, from the boundedness of the term $\sigma(x), \sigma^\intercal$ in the process X, the function $\tilde{G}_1$ satisfies (3.1.25) for all x almost everywhere by (3.1.17) and (3.1.26).

Next, consider the jump part of the process M_t^θ. From the definition of covariation, write

$$\begin{aligned}[G(X,\theta), Z]_t &= \lim_{\|P_n\|\to 0}\sum_{P_n}(G(X_{t_{i+1}},\theta) - G(X_{t_i},\theta))\,(Z_{t_{i+1}} - Z_{t_i}) \\ &= \int_0^t\int_{\mathbb{R}^r}\big[\,G(X_{s^-} + \xi(X_{s^-},z),\theta) - G(X_{s^-},\theta)\,\big]\;\tilde{\mu}(ds,dz) \\ &= \int_0^t\int_{\mathbb{R}^r}\tilde{G}_2(X_{s^-},\theta,z)\,\tilde{\mu}(ds,dz)\,,\end{aligned}$$

where P_n is a partition of the interval $[0,t]$ and $\tilde{G}_2(x,\theta,z) := G(x+\xi(x,z),\theta) - G(x,\theta)$. From (3.1.17), write

$$\begin{aligned}\Big|\tilde{G}_2(x,\theta,z)\Big| &= \big|G(x+\xi(x,z),\theta) - G(x,\theta)\big| \\ &\leq C_{G,q}\,\big(1+|\theta|\big)\,\big|\xi(x,z)\big|\,\Big(1+\big|x+\xi(x,z)\big|^q + |x|^q\Big) \\ &\leq C_{M,q}\,\big(1+|\theta|\big)\,|x||z|\,\big(1+|x|^q|z|^q+|x|^q\big)\;,\end{aligned}$$

where the last inequality follows from Condition 3.1.3. $\square$

We note that the existence of the martingale representation of the process M^θ is a result of Theorem 1.1 in Kunita (2004).

Although the processes X and θ are a jump processes, for an arbitrary function F, we have that

$$\begin{aligned}\int_0^t F(X_s,\theta_s)\,ds &= \int_0^t F(X_{s^-},\theta_{s^-})\,ds\\ \int_0^t F(X_s,\theta_s)\,dB_s &= \int_0^t F(X_{s^-},\theta_{s^-})\,dB_s\,,\end{aligned} \tag{3.1.34}$$

because the processes X and θ have only countably many jumps, and the Lebesgue measure of the set of times where $X_s \neq X_{s^-}$ or $\theta_s \neq \theta_{s^-}$ is 0. Therefore, we can choose either one to write the integrals. Throughout this thesis, we use $\int_0^t F(X_s,\theta_s)\,ds$ and $\int_0^t F(X_s,\theta_s)\,dB_s$.

Now, we show that the difference $G(X_t,\theta_t) - G(X_0,\theta_0)$ is the sum of a stochastic integral, a Riemann integral, and a covariation term. Our approach is based on Föllmer, Protter, and Shiryayev (1995).

Proposition 3.1.16. ***Extended Itô Lemma***. *Suppose Conditions 3.1.1 to 3.1.6 hold. Let the function G be as in (3.1.14) and let the functions $\tilde{G}_1$ and $\tilde{G}_2$ be as in Lemma 3.1.15. The following extended Itô Lemma holds:*

$$\begin{aligned}G(X_t,\theta_t) = G(X_0,\theta_0) &+ \int_0^t \mathcal{A}_{x,\theta^*} G(X_s,\theta_s)\,ds + \int_0^t \tilde{G}_1(X_s,\theta_s)\,dB_s\\ &+ \int_0^t \int_{\mathbb{R}} \tilde{G}_2(X_{s^-},\theta_{s^-},z)\,\tilde{\mu}(ds,dz) + \int_0^t \nabla_\theta G(X_{s^-},\theta_{s^-})\,d\theta_s\\ &+ \left[\int_0^t \tfrac{1}{2}\,\nabla_\theta^2 G_k(X_s,\theta_s)\,d[\theta,\theta]_s\right]_{k=1}^d + \left[\nabla_\theta G(X,\theta),\theta\right]_t^x\,,\end{aligned} \tag{3.1.35}$$

where G_k is a real-valued function for all $1 \le k \le d$, $G = [G_1,...,G_d]^\intercal$, and the covariation with respect to x, $\left[\nabla_\theta G(X,\theta),\theta\right]_t^x$, is defined as

$$\left[\nabla_\theta G(X,\theta),\theta\right]_t^x := \lim_{\|P_n\|\to 0} \sum_{P_n} (\nabla_\theta G(X_{t_{i+1}},\theta_{t_i}) - \nabla_\theta\, G(X_{t_i},\theta_{t_i}))\,(\theta_{t_{i+1}} - \theta_{t_i})\,, \tag{3.1.36}$$

and recall that $\mathcal{A}_{x,\theta^*}$ is the extended generator of the process X, see (3.1.12).

Proof. Let $\{P_n\}_{n\ge 1}$ be a sequence of partitions of the interval $[0,t]$, such that the maximal mesh size of P_n converges to zero as n goes to infinity. For a partition P_n, write

$$\begin{aligned}G(X_t,\theta_t) - G(X_0,\theta_0) &= \sum_{P_n} \left\{G(X_{t_{i+1}},\theta_{t_{i+1}}) - G(X_{t_i},\theta_{t_i})\right\}\\ &= \sum_{P_n} \left\{G(X_{t_{i+1}},\theta_{t_{i+1}}) - G(X_{t_{i+1}},\theta_{t_i}) + G(X_{t_{i+1}},\theta_{t_i}) - G(X_{t_i},\theta_{t_i})\right\}.\end{aligned} \tag{3.1.37}$$

Consider each component on the right-hand of (3.1.37) individually and because the function G is twice differentiable with respect to θ, use Taylor's expansion to write

$$\begin{aligned}
&G(X_{t_{i+1}},\theta_{t_{i+1}}) - G(X_{t_{i+1}},\theta_{t_i}) = \nabla_\theta G(X_{t_{i+1}},\theta_{t_i})(\theta_{t_{i+1}} - \theta_{t_i}) \\
&\quad + \left[\frac{1}{2}\,\mathrm{Tr}\left((\theta_{t_{i+1}} - \theta_{t_i})^\intercal\,\nabla_\theta^2 G_k(X_{t_{i+1}},\theta_{t_i})\,(\theta_{t_{i+1}} - \theta_{t_i})\right)\right]_{k=1}^d + R(X_{t_{i+1}},\theta_{t_i}) \\
&= \nabla_\theta G(X_{t_i},\theta_{t_i})(\theta_{t_{i+1}} - \theta_{t_i}) \\
&\quad + \left[\frac{1}{2}\,\mathrm{Tr}\left((\theta_{t_{i+1}} - \theta_{t_i})^\intercal\,\nabla_\theta^2 G_k(X_{t_i},\theta_{t_i})\,(\theta_{t_{i+1}} - \theta_{t_i})\right)\right]_{k=1}^d \\
&\quad + (\nabla_\theta G(X_{t_{i+1}},\theta_{t_i}) - \nabla_\theta G(X_{t_i},\theta_{t_i}))(\theta_{t_{i+1}} - \theta_{t_i}) \\
&\quad + \left[\frac{1}{2}\,\mathrm{Tr}\Big(\,(\theta_{t_{i+1}} - \theta_{t_i})^\intercal\,\left(\nabla_\theta^2 G(X_{t_{i+1}},\theta_{t_i}) - \nabla_\theta^2 G(X_{t_i},\theta_{t_i})\right)\,(\theta_{t_{i+1}} - \theta_{t_i})\Big)\right]_{k=1}^d \\
&\quad + R(X_{t_{i+1}},\theta_{t_i})\,,
\end{aligned}$$

where R is the remainder function and the operation $\mathrm{Tr}[\,\cdot\,]$ denotes the trace of a matrix. Then, by standard stochastic integration and Riemann integration, the following equations hold:

$$\begin{aligned}
\lim_{\|P_n\|\to 0}\sum_{P_n}\nabla_\theta G(X_{t_i},\theta_{t_i})(\theta_{t_{i+1}} - \theta_{t_i}) &= \int_0^t \nabla_\theta G(X_{s^-},\theta_{s^-})\,d\theta_s\,, \\
\lim_{\|P_n\|\to 0}\sum_{P_n}\tfrac{1}{2}\,\mathrm{Tr}\left((\theta_{t_{i+1}} - \theta_{t_i})^\intercal\,\nabla_\theta^2 G_k(X_{t_i},\theta_{t_i})\,(\theta_{t_{i+1}} - \theta_{t_i})\right) &= \int_0^t \tfrac{1}{2}\,\nabla_\theta^2 G_k(X_s,\theta_s)\,d[\theta,\theta]_s\,,
\end{aligned} \tag{3.1.38}$$

for all $1 \le k \le d$, because $\nabla_\theta G$ and $\nabla_\theta^2 G$ are continuous with respect to θ. Moreover, we have that

$$\lim_{\|P_n\|\to 0}\sum_{P_n} R(X_{t_{i+1}},\theta_{t_i}) = 0 \tag{3.1.39}$$

because the remainder function R depends on the third order of the increment of the process θ and the quadratic variation of the process θ is finite.

The locally Lipschitz property of the function G implies that for all $1 \le k \le d$

$$\begin{aligned}
&|\nabla_\theta^2 G_k(X_{t_{i+1}},\theta_{t_i}) - \nabla_\theta^2 G_k(X_{t_i},\theta_{t_i})| \\
&\qquad \le C_{G,q}\left(1 + \left|X_{t_{i+1}}\right|^q + \left|X_{t_i}\right|^q\right)\left|X_{t_{i+1}} - X_{t_i}\right| \\
&\qquad \le 2\,C_{G,q}\,(1 + \sup_{0\le s\le t}|X_s|^q)\left(\sum_{t_i,t_{i+1}}|\Delta X| + |X^c_{t_{i+1}} - X^c_{t_i}|\right),
\end{aligned} \tag{3.1.40}$$

where X^c is a continuous component of the process X and $\sum_{t_i,t_{i+1}}|\Delta X|$ is the sum of all jumps in the interval $[t_i, t_{i+1}]$. The following sum converges

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \left(\theta_{t_{i+1}} - \theta_{t_i}\right)^\intercal \left|X^c_{t_{i+1}} - X^c_{t_i}\right| \left(\theta_{t_{i+1}} - \theta_{t_i}\right) = 0 \tag{3.1.41}$$

because the process X^c is continuous and the quadratic variation of the process θ is finite. We denote by J the sum of the jumps of the process X in $[0, T]$. Then, the random variable J is finite with probability one because ν is a finite measure. Therefore,

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \sum_{t_i,t_{i+1}} (\theta_{t_{i+1}} - \theta_{t_i})^\intercal |\Delta X| \left(\theta_{t_{i+1}} - \theta_{t_i}\right) \leq J^2 \lim_{\|P_n\|\to 0} \sup_i \left|\theta_{t_{i+1}} - \theta_{t_i}\right|^2 = 0\,. \tag{3.1.42}$$

The last equation holds because the process θ is continuous and its quadratic variation is finite. From (3.1.40), (3.1.41), (3.1.42), we have that

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \frac{1}{2} \operatorname{Tr}\left((\theta_{t_{i+1}} - \theta_{t_i})^\intercal \left(\nabla^2_\theta G_k(X_{t_{i+1}}, \theta_{t_i}) - \nabla^2_\theta G_k(X_{t_i}, \theta_{t_i})\right) \left(\theta_{t_{i+1}} - \theta_{t_i}\right)\right) = 0\,. \tag{3.1.43}$$

Next, we show that the covariation term is well-defined. Recall that (see (3.1.38))

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \nabla_\theta G(X_{t_i}, \theta_{t_i})(\theta_{t_{i+1}} - \theta_{t_i}) = \int_0^t \nabla_\theta G(X_{s^-}, \theta_{s^-})\, d\theta_s\,,$$

with convergence in u.c.p because the processes X and θ are semimartingales and the function $\nabla_\theta G(x, \theta)$ is continuous, see Protter (1990). In Appendix A.1.2, we show that the coupled process (X, θ) satisfies the time reversal property, so we write

$$\begin{aligned}
&\lim_{\|P_n\|\to 0} \sum \nabla_\theta G(X_{t_{i+1}}, \theta_{t_i})\, (\theta_{t_{i+1}} - \theta_{t_i}) \\
&\qquad = \lim_{\|P_n\|\to 0} \left\{ \sum_{t_i \leq t} \nabla_\theta G(X_{t-t_{i+1}}, \theta_{t-t_i})\, (\theta_{t-t_{i+1}} - \theta_{t-t_i}) \right\} \circ \mathbf{R} \\
&\qquad = \lim_{\|P_n\|\to 0} -\left\{ \sum_{s_i = t - t_i \leq t} \nabla_\theta G(X_{s_i}, \theta_{s_{i+1}})\, (\theta_{s_{i+1}} - \theta_{s_i}) \right\} \circ \mathbf{R}\,,
\end{aligned} \tag{3.1.44}$$

where $\mathbf{R}$ is the time reversal operator (i.e., $X_s \circ \mathbf{R} = X_{t-s}$). Therefore, it suffices to show that the limit on the right-hand of (3.1.44) exists. Consider the sum

$$\begin{aligned}
\sum_{s_i \leq t} \nabla_\theta G(X_{s_i}, \theta_{s_{i+1}})\, (\theta_{s_{i+1}} - \theta_{s_i}) &= \sum_{s_i \leq t} \nabla_\theta G(X_{s_i}, \theta_{s_i})\, (\theta_{s_{i+1}} - \theta_{s_i}) \\
&\quad + \sum_{s_i \leq t} \left[\frac{1}{2} \operatorname{Tr}\left((\theta_{s_{i+1}} - \theta_{s_i})^\intercal\, \nabla^2_\theta G_k(X_{s_i}, \theta_{s_i'})\, (\theta_{s_{i+1}} - \theta_{s_i}) \right) \right]_{k=1}^d\,,
\end{aligned} \tag{3.1.45}$$

and note that

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \nabla_\theta G(X_{t_{i+1}}, \theta_{t_i})\, (\theta_{t_{i+1}} - \theta_{t_i})$$

is well-defined because the limits of both terms on the right-hand side of (3.1.45) exist; thus, the covariation with respect to x in (3.1.36) is well-defined. Therefore,

$$\begin{aligned}&\lim_{\|P_n\|\to 0} \sum_{P_n} \left(G(X_{t_{i+1}}, \theta_{t_{i+1}}) - G(X_{t_{i+1}}, \theta_{t_i})\right)\\ &\quad = \int_0^t \nabla_\theta G(X_{s^-}, \theta_{s^-})\, d\theta_s + \left[\int_0^t \tfrac{1}{2}\nabla_\theta^2 G_k(X_s, \theta_s)\, d[\theta, \theta]_s\right]_{k=1}^d + \left[\nabla_\theta G(X, \theta), \theta\right]_t^x ,\end{aligned} \tag{3.1.46}$$

from (3.1.36), (3.1.38), (3.1.43). Now, consider the second component of the sum on the right-hand side of (3.1.37). From Lemma 3.1.15, the following equation holds:

$$\begin{aligned}G(X_{t_{i+1}}, \theta_{t_i}) - G(X_{t_i}, \theta_{t_i}) = &\int_{t_i}^{t_{i+1}} \mathcal{A}_{x,\theta^*} G(X_s, \theta_{t_i})\, ds + \int_{t_i}^{t_{i+1}} \tilde{G}_1(X_s, \theta_{t_i})\, dB_s\\ &+ \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} \tilde{G}_2(X_{s^-}, \theta_{t_i}, z)\, \tilde{\mu}(ds, dz)\,.\end{aligned} \tag{3.1.47}$$

From Lemma 3.1.15, we have that

$$\left|\tilde{G}_1(X_s, \theta_u)\right| \le C_{M,q}\, \left(1 + |\theta_u|\right)\, \left(1 + |X_s|^q\right)\,, \tag{3.1.48}$$

on $d\mathbb{P}_{\theta^*} \times dt$ almost everywhere. The sequence $\tilde{G}_1(X_s, \theta_{t_i})$ converges to $\tilde{G}_1(X_s, \theta_s)$ pointwise because the function $\tilde{G}_1$ is continuous with respect to the variable θ and because the process θ_s is continuous almost everywhere. Use Lemma 3.1.9, the Lebesgue's dominated convergence theorem for stochastic integration, and (3.1.48), to write

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \int_{t_i}^{t_{i+1}} \tilde{G}_1(X_s, \theta_{t_i})\, dB_s = \int_0^t \tilde{G}_1(X_s, \theta_s)\, dB_s\,. \tag{3.1.49}$$

For the jump part, we use the same argument as that above to obtain (3.1.49), and write

$$\lim_{\|P_n\|\to 0} \sum_{P_n} \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} \tilde{G}_2(X_{s^-}, \theta_{t_i}, z)\, \tilde{\mu}(ds, dz) = \int_0^t \int_{\mathbb{R}} \tilde{G}_2(X_{s^-}, \theta_{s^-}, z)\, \tilde{\mu}(ds, dz)\,. \tag{3.1.50}$$

The last term on the right-hand side of (3.1.47) is

$$\int_{t_i}^{t_{i+1}} \mathcal{A}_{x,\theta^*} G(X_s, \theta_{t_i})\, ds = \int_{t_i}^{t_{i+1}} \left(\nabla_\theta g(X_s, \theta_{t_i}) - \nabla_\theta \bar{g}(\theta_{t_i})\right)\, ds\,.$$

By the assumptions on the function g, we have that

$$\left|\nabla_\theta g(x,\theta) - \nabla_\theta \bar{g}(\theta)\right| \le C\,\left(1 + |\theta|\right)\,\left(1 + |x|^q\right)\,,$$

for some constants $C > 0$ and $q > 0$. Therefore, apply the Lebesgue's dominated convergence theorem to the Riemann integral and write

$$\begin{aligned}\lim_{\|P_n\|\to 0} \sum_{P_n} \int_{t_i}^{t_{i+1}} \left(\nabla_\theta g(X_s,\theta_{t_i}) - \nabla_\theta \bar{g}(\theta_{t_i})\right) ds &= \int_0^t \left(\nabla_\theta g(X_s,\theta_s) - \nabla_\theta \bar{g}(\theta_s)\right)\, ds \\ &= \int_0^t \mathcal{A}_{x,\theta^*} G(X_s,\theta_s)\, ds\,.\end{aligned} \tag{3.1.51}$$

From (3.1.49), (3.1.50), (3.1.51), we have

$$\begin{aligned}\lim_{\|P_n\|\to 0} \sum_{P_n} \left\{G(X_{t_{i+1}},\theta_{t_i}) - G(X_{t_i},\theta_{t_i})\right\} &= \int_0^t \mathcal{A}_{x,\theta^*} G(X_s,\theta_s)\, ds + \int_0^t \tilde{G}_1(X_s,\theta_s)\, dB_s \\ &+ \int_0^t \int_{\mathbb{R}} \tilde{G}_2(X_{s^-},\theta_{s^-},z)\, \tilde{\mu}(ds,dz)\,.\end{aligned} \tag{3.1.52}$$

Therefore, (3.1.46) and (3.1.52) show that the extended Itô formula in (3.1.35) holds. □

Note that if the function G is twice differentiable with respect to x, then the co-variation term in (3.1.36) is a Riemann integral, and the formula in (3.1.35) coincides with the classical Itô formula.

Remark 3.1.17. Note that Lowther (2010) has a similar result as (3.1.35), but in more generic form. In our result, we require to know an explicit formula for zero continuous quadratic variation term as mentioned in Theorem 1.2 in Lowther (2010). Thus, our version of Itô's lemma allows us to analyse the term more precisely when we show the convergence of the estimator in the next section.

The function G in (3.1.14) satisfies the weak Poisson equation (3.1.12) because the function $\nabla_\theta g(x,\theta) - \nabla_\theta \bar{g}(\theta)$ is centered as required in Theorem 3.1.10. Then, in the next proposition we apply Proposition 3.1.14 and the extended Itô Lemma to prove the convergence of the deviation term $\Gamma_{k,\gamma}$.

Before proceeding, we define the following increasing stopping times. For an arbitrary $\kappa > 0$ and $\lambda = \lambda(\kappa) > 0$

$$\underline{\tau}_k := \inf\{t > \overline{\tau}_{k-1} : \left|\nabla_\theta \bar{g}(\theta_t)\right| \ge \kappa\}\,,$$

$$\overline{\tau}_k := \sup\left\{t > \underline{\tau}_k : \left|\nabla_\theta \bar{g}(\theta_{\underline{\tau}_k})\right|/2 \leq \left|\nabla_\theta \bar{g}(\theta_s)\right| \leq 2\left|\nabla_\theta \bar{g}(\theta_{\underline{\tau}_k})\right| \right.$$
$$\left. \text{for all } s \in [\underline{\tau}_k, t] \text{ and } \int_{\underline{\tau}_k}^t \beta_s\, ds \leq \lambda \right\},$$

where $k = 1, 2, ...$ and $\sigma_0 = 0$, and we denote $\overline{\tau}_{k,\gamma} = \overline{\tau}_k + \gamma$ for $\gamma > 0$. This sequence of stopping times quantifies the period of time for which $\left|\nabla_\theta \bar{g}(\theta)\right|$ is small, see Bertsekas and Tsitsiklis (2000).

Proposition 3.1.18. ***Convergence of deviation term***. *Recall that the deviation term is given by*

$$\Gamma_{k,\gamma} = \int_{\underline{\tau}_k}^{\overline{\tau}_{k,\gamma}} \beta_s \left(\nabla_\theta g(X_s, \theta_s) - \nabla_\theta \bar{g}(\theta_s)\right) ds\,.$$

Assume Conditions 3.1.1 to 3.1.7 hold and let $X_0 = x_0$. Then, with probability one, we have that

$$\left|\Gamma_{k,\gamma}\right| \to 0 \quad as \quad k \to \infty\,.$$

Proof. The idea of this proof is closely related to the proof of Lemma 3.1 in Sirignano and Spiliopoulos (2017). Recall that the function $\nabla_\theta g(x, \theta) - \nabla_\theta \bar{g}(\theta)$ is centered and that G is a solution to the weak Poisson equation (3.1.12) and that it is locally Lipschitz.

Let $\underline{\tau}$ and $\overline{\tau}$ be stopping times, such that $\underline{\tau} \leq \overline{\tau}$. First, we show that

$$\begin{aligned}
&\beta_{\overline{\tau}}\, G(X_{\overline{\tau}}, \theta_{\overline{\tau}}) - \beta_{\underline{\tau}}\, G(X_{\underline{\tau}}, \theta_{\underline{\tau}}) \\
&\quad = \int_{\underline{\tau}}^{\overline{\tau}} \beta_s'\, G(X_s, \theta_s)\, ds + \int_{\underline{\tau}}^{\overline{\tau}} \beta_s\, \mathcal{A}_{x,\theta^*} G(X_s, \theta_s)\, ds + \int_{\underline{\tau}}^{\overline{\tau}} \beta_s\, \tilde{G}_1(X_s, \theta_s)\, dB_s \\
&\qquad + \int_{\underline{\tau}}^{\overline{\tau}} \int_{\mathbb{R}} \beta_s\, \tilde{G}_2(X_s, \theta_s, z)\, \tilde{\mu}(ds, dz) + \int_{\underline{\tau}}^{\overline{\tau}} \beta_s\, \nabla_\theta G(X_{s^-}, \theta_{s^-})\, d\theta_s \\
&\qquad + \left[\int_{\underline{\tau}}^{\overline{\tau}} \tfrac{1}{2}\, \beta_s\, \nabla_\theta^2 G_k(X_s, \theta_s)\, d[\theta, \theta]_s\right]_{k=1}^d + \left[\beta\, \nabla_\theta G(X, \theta), \theta\right]_{\underline{\tau}}^{\overline{\tau}, x},
\end{aligned} \tag{3.1.53}$$

where β_t' is the derivative of β_t with respect to t and the quadratic covariation term

$$\left[\beta\, \nabla_\theta G(X, \theta), \theta\right]_{\underline{\tau}}^{\overline{\tau}, x} = \lim_{\|P_n\| \to 0} \sum_{P_n} (\beta_{t_{i+1}} \nabla_\theta G(X_{t_{i+1}}, \theta_{t_i}) - \beta_{t_i} \nabla_\theta G(X_{t_i}, \theta_{t_i}))(\theta_{t_{i+1}} - \theta_{t_i})\,. \tag{3.1.54}$$

Here, P_n is a partition of the interval $[\underline{\tau}, \overline{\tau}]$, and (3.1.54) is well-defined due to the time reversal argument in Proposition 3.1.16. Write

$$\begin{aligned}\beta_{\overline{\tau}}\, G(X_{\overline{\tau}}, \theta_{\overline{\tau}}) - &\beta_{\underline{\tau}}\, G(X_{\underline{\tau}}, \theta_{\underline{\tau}}) \\ &= \sum_{P_n} \beta_{t_{i+1}} \left(G(X_{t_{i+1}}, \theta_{t_{i+1}}) - G(X_{t_i}, \theta_{t_i})\right) + \left(\beta_{t_{i+1}} - \beta_{t_i}\right) G(X_{t_i}, \theta_{t_i}) \\ &= A_1^n + A_2^n\,,\end{aligned} \tag{3.1.55}$$

where $A_1^n = \sum_{P_n} \beta_{t_{i+1}} \left(G(X_{t_{i+1}}, \theta_{t_{i+1}}) - G(X_{t_i}, \theta_{t_i})\right)$ and $A_2^n = \left(\beta_{t_{i+1}} - \beta_{t_i}\right) G(X_{t_i}, \theta_{t_i})$. Next, consider

$$\begin{aligned}\lim_{n\to\infty} A_2^n &= \lim_{n\to\infty} \sum_{P_n} \beta'_{s_{i+1}} (t_{i+1} - t_i)\, G(X_{t_i}, \theta_{t_i}) \\ &= \int_{\underline{\tau}}^{\overline{\tau}} \beta'_s\, G(X_s, \theta_s)\, ds\,,\end{aligned} \tag{3.1.56}$$

where the second equation holds because $\sum_{P_n} \beta'_{s_{i+1}} \mathbf{1}_{\{s_{i+1} \in [t_i, t_{i+1})\}}\, G(X_{t_i}, \theta_{t_i})$ converges to $\beta'_s\, G(X_s, \theta_s)$ almost surely and by the dominated convergence theorem. Now, consider

$$\begin{aligned}\lim_{n\to\infty} A_1^n = &\lim_{n\to\infty} \sum_{P_n} \beta_{t_{i+1}} \left(G(X_{t_{i+1}}, \theta_{t_{i+1}}) - G(X_{t_i}, \theta_{t_i})\right) \\ = &\int_{\underline{\tau}}^{\overline{\tau}} \beta_s\, \mathcal{A}_{x,\theta^*} G(X_s, \theta_s)\, ds + \int_{\underline{\tau}}^{\overline{\tau}} \beta_s\, \tilde{G}_1(X_s, \theta_s)\, dB_s + \int_{\underline{\tau}}^{\overline{\tau}} \beta_s\, \nabla_\theta G(X_{s^-}, \theta_{s^-})\, d\theta_s \\ &+ \int_{\underline{\tau}}^{\overline{\tau}} \int_{\mathbb{R}} \beta_s\, \tilde{G}_2(X_{s^-}, \theta_{s^-}, z)\, \tilde{\mu}(ds, dz) + \left[\int_{\underline{\tau}}^{\overline{\tau}} \tfrac{1}{2}\, \beta_s\, \nabla_\theta^2 G_k(X_s, \theta_s)\, d[\theta, \theta]_s\right]_{k=1}^{d} \\ &+ \lim_{n\to\infty} \sum_{P_n} \beta_{t_{i+1}} \left(\left[\nabla_\theta G(X, \theta), \theta\right]^x_{t_{i+1}} - \left[\nabla_\theta G(X, \theta), \theta\right]^x_{t_i}\right)\,,\end{aligned} \tag{3.1.57}$$

where the second equation follows from Proposition 3.1.16 and from the dominated convergence theorem.

Next, we must show that

$$\lim_{n\to\infty} \sum_{P_n} \beta_{t_{i+1}} \left(\left[\nabla_\theta G(X, \theta), \theta\right]^x_{t_{i+1}} - \left[\nabla_\theta G(X, \theta), \theta\right]^x_{t_i}\right) \tag{3.1.58}$$

is well-defined and it is equal to $\left[\beta\, \nabla_\theta G(X, \theta), \theta\right]^{\overline{\tau}, x}_{\underline{\tau}}$. From the proof of Proposition 3.1.16, we have that

$$\begin{aligned}\beta_{t_{i+1}} &\left(\left[\nabla_\theta G(X, \theta), \theta\right]^x_{t_{i+1}} - \left[\nabla_\theta G(X, \theta), \theta\right]^x_{t_i}\right) \\ &= \beta_{t_{i+1}} \left[\int_{t_i}^{t_{i+1}} \nabla_\theta G(X_s, \theta_s)\, d^{*,x}\theta_s - \int_{t_i}^{t_{i+1}} \nabla_\theta G(X_{s^-}, \theta_{s^-})\, d\theta_s\right]\,,\end{aligned} \tag{3.1.59}$$

where $\int_{t_i}^{t_{i+1}} \nabla_\theta G(X_s,\theta_s)\, d^{*,x}\theta_s := \lim \sum_{t_i<s_i<t_{i+1}} \nabla_\theta G(X_{s_{i+1}},\theta_{s_i})\,(\theta_{s_{i+1}}-\theta_{s_i})$ is a backward integration. Take the limit $n\to\infty$ on both sides of the equation above and write

$$\begin{aligned}
\lim_{n\to\infty}\sum_{P_n} \beta_{t_{i+1}} &\Big(\big[\nabla_\theta G(X,\theta),\theta\big]^x_{t_{i+1}} - \big[\nabla_\theta G(X,\theta),\theta\big]^x_{t_i} \Big) \\
&= \left[\int_{\underline{\tau}}^{\overline{\tau}} \beta_s \nabla_\theta G(X_{s^-},\theta_{s^-})\, d^{*,x}\theta_s - \int_{\underline{\tau}}^{\overline{\tau}} \beta_s \nabla_\theta G(X_{s^-},\theta_{s^-})\, d\theta_s \right] \\
&= \big[\beta\, \nabla_\theta G(X,\theta),\theta\big]^{\overline{\tau},x}_{\underline{\tau}}\,.
\end{aligned} \tag{3.1.60}$$

The first equality follows from (3.1.59) and the dominated convergence theorem. The second equality follows from the definition of $\big[\beta\,\nabla_\theta G(X,\theta),\theta\big]^{\overline{\tau},x}_{\underline{\tau}}$. Therefore, (3.1.53) holds and we rewrite the deviation term in (3.1.3) as follows:

$$\begin{aligned}
\Gamma_{k,\gamma} &= \int_{\underline{\tau}_k}^{\overline{\tau}_{k,\gamma}} \beta_s\, \mathcal{A}_{x,\theta^*} G(X_s,\theta_s)\, ds \\
&= \beta_{\overline{\tau}_{k,\gamma}}\, G(X_{\overline{\tau}_{k,\gamma}},\theta_{\overline{\tau}_{k,\gamma}}) - \beta_{\underline{\tau}}\, G(X_{\underline{\tau}},\theta_{\underline{\tau}}) - \int_{\underline{\tau}}^{\overline{\tau}_{k,\gamma}} \beta_s'\, G(X_s,\theta_s)\, ds \\
&\quad - \int_{\underline{\tau}}^{\overline{\tau}_{k,\gamma}} \beta_s\, \tilde{G}_1(X_s,\theta_s)\, dB_s - \int_{\underline{\tau}}^{\overline{\tau}_{k,\gamma}} \int_{\mathbb{R}} \beta_s\, \tilde{G}_2(X_{s^-},\theta_{s^-},z)\, \tilde{\mu}(ds,dz) \\
&\quad - \int_{\underline{\tau}}^{\overline{\tau}_{k,\gamma}} \beta_s\, \nabla_\theta G(X_{s^-},\theta_{s^-})\, d\theta_s \\
&\quad - \left[\int_{\underline{\tau}}^{\overline{\tau}_{k,\gamma}} \tfrac{1}{2}\, \beta_s\, \nabla_\theta^2 G_k(X_s,\theta_s)\, d[\theta,\theta]_s \right]_{k=1}^{d} - \big[\beta\, \nabla_\theta G(X,\theta),\theta\big]^{\overline{\tau}_{k,\gamma},x}_{\underline{\tau}}\,.
\end{aligned} \tag{3.1.61}$$

Next, we prove the convergence of each term on the right-hand side of (3.1.61). The expectation operator under the probability measure $\mathbb{P}_{x_0,\theta^*}$ is denoted by $\mathbb{E}$. Let $J_t^{(1)} = \beta_t\,\big|G(X_t,\theta_t)\big|$. The inequality

$$\begin{aligned}
\mathbb{E}\left[\left|J_t^{(1)}\right|^2\right] &\le C_{M,q}\, \beta_t^2\, \mathbb{E}\left[\big(1+|\theta_t|\big)^2 \big(1+|X_t|^q\big)^2 \right] \\
&\le 4\, C_{M,q}\, \beta_t^2\, \mathbb{E}\left[\Big(1+|\theta_t|^2\Big) \Big(1+|X_t|^{2\,q}\Big) \right] \le \tilde{C}\, \beta_t^2\,,
\end{aligned}$$

follows from Proposition 3.1.14, and because the random variables X_t and θ_t have finite moments and because $\tilde{C}$ is a constant that depends on x_0. Now, apply the Borel–Cantelli argument as in Lemma 4.1 in Sirignano and Spiliopoulos (2017) to show that $J_t^{(1)}$ converges to zero with probability one. Next, consider the finite variation term on the right-hand side of (3.1.61)

$$J_{t,0}^{(2)} := \int_0^t \Big| \beta_s'\, G(X_s,\theta_s) + \beta_s\, \mathcal{L}_{\theta,\theta^*} G(X_s,\theta_s) \Big|\, ds\,,$$

where $\mathcal{L}_{\theta,\theta^*}$ is the infinitesimal generator of the process θ under the probability measure $\mathbb{P}_{\theta^*}$. The term $J^{(2)}_{t,0}$ obeys the bound

$$\sup_{t>0} \mathbb{E}\left|J^{(2)}_{t,0}\right| \leq \int_0^\infty C\,\left(|\beta_s'| + \beta_s^2\right)\left(1 + \mathbb{E}\,|X_s|^{2\,q} + \mathbb{E}\,|\theta_s|^2\right)\,ds < \infty\,,$$

where the last inequality follows from the properties of the learning rate β_t in Condition 3.1.5, and because the process X has a bounded $2\,q$-moment and the process θ has a bounded second moment. Thus, there exists an integrable random variable $J^{(2)}_{\infty,0}$, such that

$$J^{(2)}_{t,0} \to J^{(2)}_{\infty,0} \quad \text{as} \quad t \to \infty \tag{3.1.62}$$

with probability one.

Next, consider the stochastic integral

$$\begin{aligned} J^{(3)}_{t,0} := \int_0^t \beta_s\,\tilde{G}_1(X_s,\theta_s)\,dB_s &+ \int_0^t \int_{\mathbb{R}} \beta_s\,\tilde{G}_2(X_{s^-},\theta_{s^-},z)\,\tilde{\mu}(ds,dz) \\ &+ \int_0^t \beta_s\,\nabla_\theta G(X_s,\theta_s)\,\nabla_\theta b(X_s,\theta)\,\sigma(X_s)^{-1}\,dB_s\,. \end{aligned}$$

The Burkholder-Davis-Gundy (BDG) inequality implies that

$$\begin{aligned} \mathbb{E}\left[\sup_{0\leq s\leq t}\left|J^{(3)}_{s,0}\right|^2\right] \leq C_1\,\mathbb{E}\Bigg[\int_0^t \left|\beta_s\,\tilde{G}_1(X_s,\theta_s)\right|^2\,ds &+ \int_0^t\int_{\mathbb{R}} \left|\beta_s\,\tilde{G}_2(X_{s^-},\theta_{s^-},z)\right|^2\,\nu(dz)\,ds \\ + \int_0^t \left|\beta_s\,\nabla_\theta G(X_s,\theta_s)\nabla_\theta b(X_s,\theta)\,\sigma(X_s)^{-1}\right|^2\,ds\Bigg]\,, \end{aligned} \tag{3.1.63}$$

where C_1 is a positive constant. From Lemma 3.1.15, we have that

$$\begin{aligned} \left|\tilde{G}_1(X_s,\theta_s)\right| &\leq C_{M,q}\,\left(1+|\theta_s|\right)\,\left(1+|X_s|^q\right) \\ \left|\tilde{G}_2(X_{s^-},\theta_{s^-},z)\right| &\leq C_{M,q}\,\left(1+|\theta_{s^-}|\right)|X_{s^-}||z|\,\left(1+|X_{s^-}|^q|z|^q+|X_{s^-}|^q\right)\,. \end{aligned} \tag{3.1.64}$$

Therefore, use the BDG and the Cauchy–Schwarz inequalities to show that the expectation of the first and third terms on the right-hand of (3.1.63) are both bounded by $C\,K\int_0^t \beta_s^2\,ds$ because the processes X_s, θ_s, X_{s^-}, and θ_{s^-} have finite p-moments for all $p > 0$. For the remaining term, write

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^t \int_{\mathbb{R}} \left|\beta_s\,\tilde{G}_2(X_{s^-},\theta_{s^-},z)\right|^2 \nu(dz)\,ds\right] \\
&\qquad \le C_{M,q}^2\,\mathbb{E}\left[\int_0^t \int_{\mathbb{R}} \beta_s^2 \left(1+|\theta_{s^-}|^2\right)|X_{s^-}|^2\,|z|^2\left(1+|X_{s^-}|^{2\,q}\,|z|^{2\,q}+|X_{s^-}|^{2\,q}\right)\nu(dz)\,ds\right] \\
&\qquad \le \tilde{C}\,\mathbb{E}\left[\int_0^t \beta_s^2\left(1+|\theta_{s^-}|^2\right)\left(|X_{s^-}|^2+|X_{s^-}|^{2\,q+2}\right) ds\right] \\
&\qquad \le 8\,\tilde{C}\,\mathbb{E}\left[\int_0^t \beta_s^2\left(1+|\theta_{s^-}|^4\right)+\beta_s^2\left(1+|X_{s^-}|^{4\,q+4}\right) ds\right] \\
&\qquad < \hat{C}\int_0^t \beta_s^2\,ds\,,
\end{aligned}
\tag{3.1.65}
$$

where $\tilde{C}$ and $\hat{C}$ are positive constants that depend on x_0. The first inequality in (3.1.65) results from Lemma 3.1.15, the second inequality follows from Condition 3.1.3, and the last inequality follows because the processes X_s, θ_s, X_{s^-}, and θ_{s^-} have finite moments, see Lemma 3.1.9. Hence, the right-hand side of (3.1.63) is finite. Then, by Doob's martingale convergence theorem, there exists a square integrable random variable $J^{(3)}_{\infty,0}$, such that

$$
J^{(3)}_{t,0} \to J^{(3)}_{\infty,0} \quad \text{as} \quad t\to\infty \tag{3.1.66}
$$

with probability one.

Next, consider the following covariation in the interval $[t_i, t_{i+1}]$

$$
\begin{aligned}
&\left|\left[\beta\,\nabla_\theta G(X,\theta),\theta\right]_{t_i}^{t_{i+1},x}\right| \le \\
&\lim_{\|\tilde{P}_n\|\to 0}\left(\sum_{\tilde{P}_n}\beta_{s_{j+1}}\left|\nabla_\theta G(X_{s_{j+1}},\theta_{s_j})-\nabla_\theta G(X_{s_j},\theta_{s_j})\right|\left|\theta_{s_{j+1}}-\theta_{s_j}\right|\right. \\
&\left.+\left|\beta_{s_j}-\beta_{s_{j+1}}\right|\left|\nabla_\theta G(X_{s_j},\theta_{s_j})\right|\left|\theta_{s_{j+1}}-\theta_{s_j}\right|\right),
\end{aligned}
\tag{3.1.67}
$$

where $\tilde{P}_n$ is the partition of the interval $[t_i, t_{i+1}]$. Now, consider the second term on the right-hand side of (3.1.67)

$$
\begin{aligned}
&\lim_{\|\tilde{P}_n\|\to 0} \sum_{\tilde{P}_n} \left|\beta_{s_j} - \beta_{s_{j+1}}\right| \left|\nabla_\theta G(X_{s_j}, \theta_{s_j})\right| \left|\theta_{s_{j+1}} - \theta_{s_j}\right| \\
&\qquad\qquad \leq \lim_{\|\tilde{P}_n\|\to 0} \sum_{\tilde{P}_n} C_{G,q} \left|\beta_{s_j} - \beta_{s_{j+1}}\right| \left(1 + \sup_{t_i<s<t_{i+1}} |X_s|^q\right) \sup_j \left|\theta_{s_{j+1}} - \theta_{s_j}\right| \\
&\qquad\qquad \leq \lim_{\|\tilde{P}_n\|\to 0} C_{G,q} \left(1 + \sup_{t_i<s<t_{i+1}} |X_s|^q\right) \left|\beta_{t_i} - \beta_{t_{i+1}}\right| \sup_j \left|\theta_{s_{j+1}} - \theta_{s_j}\right| = 0\,,
\end{aligned} \tag{3.1.68}
$$

where the first inequality follows from (3.1.18) and the last inequality follows because the process θ_s is continuous. Now, consider the term

$$
\begin{aligned}
&\lim_{\|\tilde{P}_n\|\to 0} \sum_{\tilde{P}_n} \beta_{s_{j+1}} \left|\nabla_\theta G(X_{s_{j+1}}, \theta_{s_j}) - \nabla_\theta G(X_{s_j}, \theta_{s_j})\right| \left|\theta_{s_{j+1}} - \theta_{s_j}\right| \\
&\qquad \leq \lim_{\|\tilde{P}_n\|\to 0} \sum_{t_i<s_j<t_{i+1}} C_{G,q}\, \beta_{s_{j+1}} \left(1 + \sup_{t_i<s<t_{i+1}} |X_s|^q\right) \left|X_{s_{j+1}} - X_{s_j}\right| \left|\theta_{s_{j+1}} - \theta_{s_j}\right| \\
&\qquad \leq C_{G,q}\, \beta_{t_i} \left(1 + \sup_{t_i<s<t_{i+1}} |X_s|^q\right) \int_{t_i}^{t_{i+1}} \left|d[X,\theta]_s\right| \\
&\qquad = C_{G,q}\, \beta_{t_i}^2 \left(1 + \sup_{t_i<s<t_{i+1}} |X_s|^q\right) \int_{t_i}^{t_{i+1}} \left|\nabla_\theta b(X_s, \theta_s)\right|\, ds\,,
\end{aligned} \tag{3.1.69}
$$

where the first inequality follows from (3.1.18) and the second inequality follows from the definition of covariation. Therefore, for $t > r$, from (3.1.69), we take the sum and the expectation on the left-hand side of (3.1.67) to obtain

$$
\begin{aligned}
&\mathbb{E}\left[\left|\left[\beta\, \nabla_\theta G(X,\theta), \theta\right]_r^{t,x}\right|\right] \\
&\leq C_{G,q} \sum_{i=\lfloor r\rfloor}^{\infty} \beta_i^2\, \mathbb{E}\left[\left(1 + \sup_{i<s<i+1} |X_s|^q\right) \int_i^{i+1} |\nabla_\theta b(X_s, \theta_s)|\, ds\right] \\
&\leq 2\, C_{G,q} \sum_{i=\lfloor r\rfloor}^{\infty} \beta_i^2\, \mathbb{E}\left[\left(1 + \sup_{i<s<i+1} |X_s|^{2\,q}\right)\right] \mathbb{E}\left[\int_i^{i+1} |\nabla_\theta b(X_s, \theta_s)|^2\, ds\right] \\
&\leq 2\, C_{G,q}\, K \sum_{i=\lfloor r\rfloor}^{\infty} \beta_i^2 \left(1 + \sqrt{(1+i)}\right) \left[\int_i^{i+1} \mathbb{E}|\nabla_\theta b(X_s, \theta_s)|^2\, ds\right] \\
&\leq \tilde{C} \sum_{i=\lfloor r\rfloor}^{\infty} \beta_i^2 \left(1 + \sqrt{(1+i)}\right) < \frac{2\,\tilde{C}}{\sqrt{1+r}}\,,
\end{aligned}
$$

where the second inequality follows from the Cauchy–Schwartz inequality, the third inequality follows from Lemma A.1.1 and the constant K is from Lemma A.1.1, and the last inequality follows from Condition 3.1.6 on $\nabla_\theta b(x,\theta)$ and because $\mathbb{E}|X_s|^{2q}$ is finite. The inequality above implies that there exists a finite random variable $J_{\infty,0}$, such that

$$\left|\left[\beta\, \nabla_\theta G(X,\theta),\theta\right]_0^{t,x}\right| \to J_{\infty,0} \quad \text{as} \quad t\to\infty \tag{3.1.70}$$

with probability one. From (3.1.61), we write that

$$\left|\Gamma_{k,\gamma}\right| \le J^{(1)}_{\overline{\tau}_{k,\gamma}} + J^{(1)}_{\underline{\tau}_k} + J^{(2)}_{\overline{\tau}_{k,\gamma},\underline{\tau}_k} + \left|J^{(3)}_{\overline{\tau}_{k,\gamma},\underline{\tau}_k}\right| + \left|\left[\beta\, \nabla_\theta G(X,\theta),\theta\right]_{\underline{\tau}_k}^{\overline{\tau}_{k,\gamma},x}\right|. \tag{3.1.71}$$

From (3.1.62), (3.1.66), (3.1.70), each term on the right-hand side converges to 0 as $k\to\infty$, which implies that $\left|\Gamma_{k,\gamma}\right| \to 0$ as $k\to\infty$. □

3.1.4 Convergence of jump SGDCT

In this subsection, we prove a sequence of lemmas and show the convergence of the estimator in (3.0.9) to a stationary point of the function $\bar{g}$; we do so in two steps. First, we analyse the deviation term $\Gamma_{k,\gamma}$ in (3.1.3) when $k\to\infty$. In our case, the jump-diffusion process X is Markov, therefore we need to use the ergodicity of the process X to quantify the deviation term $\Gamma_{k,\gamma}$ via the Poisson equation in (3.1.12), see Pardoux and Veretennikov (2003). The technique dates back to Delyon (1996) and Metivier and Priouret (1984).

In Proposition 3.1.18, we showed that the absolute value of the deviation term $\Gamma_{k,\gamma}$ is asymptotically close to zero. Below, we use a similar argument as that in Proposition 3 in Bertsekas and Tsitsiklis (2000) to prove that $\lim_{t\to\infty}\left|\nabla_\theta \bar{g}(\theta_t)\right| = 0$ almost surely.

Lemma 3.1.19. *Suppose Conditions 3.1.1 to 3.1.7 hold. Choose a constant $\tilde{\lambda}$, such that for a given $\kappa>0$ we have $3\,\tilde{\lambda} + \tilde{\lambda}/4\,\kappa = 1/2\, L_{\nabla\bar{g}}$, where $L_{\nabla\bar{g}}$ is the Lipschitz constant of $\nabla\bar{g}$. For k large enough and for $\eta>0$ small enough, we have*

$$\int_{\underline{\tau}_k}^{\overline{\tau}_{k,\eta}} \beta_s\, ds > \tilde{\lambda} \qquad \textit{and} \qquad \tilde{\lambda}/2 \le \int_{\underline{\tau}_k}^{\overline{\tau}_k} \beta_s\, ds \le \tilde{\lambda}\,.$$

Proof. Most of the proof follows Lemma 3.2 in Sirignano and Spiliopoulos (2017). It suffices to show that for given $\epsilon>0$, there is k large enough such that

$$\left|\int_{\underline{\tau}_k}^{\overline{\tau}_k}\int \beta_s\, \frac{\kappa}{\left|\nabla_\theta \bar{g}(\theta_{\underline{\tau}_k})\right|}\, \nabla_\theta b(X_{s^-},\theta_{s^-})\, \left(\sigma(X_{s^-})\,\sigma(X_{s^-})^{\intercal}\right)^{-1} \xi(X_{s^-},z)\, \tilde{\mu}(ds,dz)\right| < \epsilon\,, \tag{3.1.72}$$

almost surely. By the BDG inequality, we obtain

$$\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^t\int \beta_s \frac{\kappa}{\left|\nabla_\theta \bar{g}(\theta_{\underline{\tau}_k})\right|}\,\nabla_\theta b(X_{s^-},\theta_{s^-})\,(\sigma(X_{s^-})\,\sigma(X_{s^-})^\intercal)^{-1}\,\xi(X_{s^-},z)\,\tilde{\mu}(ds,dz)\right|\right]$$
$$\leq C\,\mathbb{E}\left[\int_0^t\int \beta_s^2\left|\nabla_\theta b(X_{s^-},\theta_{s^-})\,(\sigma(X_{s^-})\,\sigma(X_{s^-})^\intercal)^{-1}\,\xi(X_{s^-},z)\right|^2\nu(dz)\,ds\right]^{1/2}$$
$$<\infty\,,$$

where the last inequality follows because the functions b, σ, ξ have polynomial growth and $\mathbb{E}|X_{s^-}|^{2\,q}$ is finite. Therefore, the martingale process

$$\int_0^t\int \beta_s \frac{\kappa}{\left|\nabla_\theta \bar{g}(\theta_{\underline{\tau}_k})\right|}\,\nabla_\theta b(X_{s^-},\theta_{s^-})\,(\sigma(X_{s^-})\,\sigma(X_{s^-})^\intercal)^{-1}\,\xi(X_{s^-},z)\,\tilde{\mu}(ds,dz)$$

is an L^2- martingale process and the martingale convergence theorem implies (3.1.72). □

The following Lemma shows that the value of the function $\bar{g}$ decreases when the value of k is large enough.

Lemma 3.1.20. *Suppose Conditions 3.1.1 to 3.1.7 hold and assume that there are an infinite number of intervals $I_k = [\underline{\tau}_k, \overline{\tau}_k)$. Then, there is a fixed constant $\gamma > 0$, such that for k large enough, we have*

$$\bar{g}(\theta_{\overline{\tau}_k}) - \bar{g}(\theta_{\underline{\tau}_k}) \leq -\gamma\,. \tag{3.1.73}$$

Proof. The function $\bar{g}$ is C^2, so we use Itô's Lemma to write

$$\begin{aligned}\bar{g}(\theta_{\overline{\tau}_k}) - \bar{g}(\theta_{\tau_k}) = &-\int_{\underline{\tau}_k}^{\overline{\tau}_k}\beta_s\left|\nabla_\theta\bar{g}(\theta_s)\right|^2 ds + \int_{\underline{\tau}_k}^{\overline{\tau}_k}\beta_s\langle\nabla_\theta\bar{g}(\theta_s),\nabla_\theta b(X_s,\theta_s)\sigma(X_s)^{-1}\rangle dB_s\\ &+\int_{\underline{\tau}_k}^{\overline{\tau}_k}\int\langle\nabla_\theta\bar{g}(\theta_{s^-})\,,\nabla_\theta b(X_{s^-},\theta_{s^-})\,(\sigma(X_{s^-})\,\sigma(X_{s^-})^\intercal)^{-1}\,\xi(X_{s^-},z)\rangle\tilde{\mu}(ds,dz)\\ &+\int_{\underline{\tau}_k}^{\overline{\tau}_k}\frac{\beta_s^2}{2}\,\mathrm{Tr}\left[(b(X_s,\theta_s)\,\sigma(X_s)^{-1})(b(X_s,\theta_s)\,\sigma(X_s)^{-1})^T\nabla_\theta^2\,\bar{g}(\theta_s)\right]\,ds\\ &+\int_{\underline{\tau}_k}^{\overline{\tau}_k}\beta_s\langle\nabla_\theta\bar{g}(\theta_s),\nabla_\theta\bar{g}(\theta_s)-\nabla_\theta g(X_s,\theta_s)\,\rangle\,ds\,.\end{aligned}$$

Recall that the operation $\mathrm{Tr}[\,\cdot\,]$ denotes the trace of a matrix.

Now, use the same argument as in Lemma 3.4 in Sirignano and Spiliopoulos (2017) to show that

$$\int_{\underline{\tau}_k}^{\overline{\tau}_k}\beta_s\left|\nabla_\theta\,\bar{g}(\theta_s)\right|^2\,ds \leq -\frac{\left|\nabla_\theta\bar{g}(\theta_{\tau_k})\right|}{8}\,\lambda\,,$$

$$\int_{\underline{\tau}_k}^{\overline{\tau}_k} \beta_s \, \langle \nabla_\theta \bar{g}(\theta_s), \nabla_\theta b(X_s, \theta_s) \, \sigma(X_s)^{-1} \, \rangle \, dB_s \leq \left| \nabla_\theta \bar{g}(\theta_{\tau_k}) \right| \epsilon \, ,$$

and

$$\int_{\underline{\tau}_k}^{\overline{\tau}_k} \frac{\beta_s^2}{2} \operatorname{Tr}[\, (b(X_s, \theta_s) \, \sigma(X_s, \theta_s)^{-1} \,) \, (b(X_s, \theta_s) \, \sigma(X_s, \theta_s)^{-1})^{\intercal} \, \nabla_\theta^2 \, \bar{g}(\theta_s) \,] \, ds \to 0 \tag{3.1.74}$$

as $k \to \infty$ with probability one. Use a similar argument to that in the proof of Lemma 3.1.19 and Doob's martingale convergence theorem to write

$$\int_{\underline{\tau}_k}^{\overline{\tau}_k} \int \langle \nabla_\theta \bar{g}(\theta_{s^-}) \, , \nabla_\theta b(X_{s^-}, \theta_{s^-}) \, (\sigma(X_{s^-}) \, \sigma(X_{s^-})^{\intercal})^{-1} \, \xi(X_{s^-}, z) \rangle \, \tilde{\mu}(ds, dz) \to 0 \tag{3.1.75}$$

as $k \to \infty$ with probability one. Next, we employ the same argument as that in Lemma 3.1.18 above with a higher order in θ to show that

$$\int_{\underline{\tau}_k}^{\overline{\tau}_k} \beta_s \, \langle \nabla_\theta \bar{g}(\theta_s), \nabla_\theta \bar{g}(\theta_s) - \nabla_\theta g(X_s, \theta_s) \, \rangle \, ds \to 0 \qquad \text{as} \qquad k \to \infty \, .$$

The remainder of the proof follows from Lemma 3.4 in Sirignano and Spiliopoulos (2017). □

Next, we show that the value of $\bar{g}(\theta_{\underline{\tau}_k}) - \bar{g}(\theta_{\overline{\tau}_{k-1}})$ eventually starts to decrease as $k \to \infty$.

Lemma 3.1.21. *Suppose Conditions 3.1.1 to 3.1.7 hold and assume that there is an infinite number of intervals $I_k = [\underline{\tau}_k, \overline{\tau}_k)$. There is a fixed constant $\gamma_1 < \gamma$, such that for k large enough, one has*

$$\bar{g}(\theta_{\underline{\tau}_k}) - \bar{g}(\theta_{\overline{\tau}_{k-1}}) \leq \gamma_1 \, . \tag{3.1.76}$$

Proof. See Lemma 3.5 in Sirignano and Spiliopoulos (2017). □

The next theorem shows the convergence of the SGDCT estimator.

Theorem 3.1.22. ***Convergence of the gradient**. Let the process θ_t follow (3.0.9) and assume Conditions 3.1.1 to 3.1.7 hold. Then*

$$\left| \nabla_\theta \bar{g}(\theta_t) \right| \to 0 \quad \text{as} \quad t \to \infty \, ,$$

almost surely.

Proof. Apply Lemmas 3.1.20 and 3.1.21 above, and Theorem 2.4 in Sirignano and Spiliopoulos (2017). □

We remark that the crucial part of the proof of the convergence of the SGDCT estimator is the convergence of the deviation term $\Gamma_{k,\gamma}$ in (3.1.3), which we proved in Proposition 3.1.18.

3.1.5 Alternative specification of SGDCT estimator

Here, we propose an alternative SGDCT estimator that employs the continuous part of the process X, which we denote by X^c, to estimate the unknown parameter θ^*. Specifically, the dynamics of X^c satisfy

$$dX_t^c := b(X_t, \theta^*)\,dt + \sigma(X_t)\,dB_t - \int \xi(X_{t-}, z)\,\nu(dz)\,dt\,. \tag{3.1.77}$$

Recall that the last term on the right-hand side of (3.1.77) is the compensator of the jump term in (3.0.1). We proceed as above to write

$$b(X_t, \theta^*)\,dt = dX_t^c + \int \xi(X_{t-}, z)\,\nu(dz)\,dt - \sigma(X_t)\,dB_t\,, \tag{3.1.78}$$

which we substitute into (3.0.4), and the SGDCT becomes

$$\begin{aligned} d\theta_t = \beta_t\, \big[& \nabla_\theta b(X_t, \theta_t)\,(\sigma(X_t)\,\sigma(X_t)^\intercal)^{-1}\,dX_t^c \\ & + \nabla_\theta b(X_t, \theta_t)\,(\sigma(X_t)\,\sigma(X_t)^\intercal)^{-1} \int \xi(X_{t-}, z)\,\nu(dz)\,dt \\ & - \nabla_\theta\, b(X_t, \theta_t)\,(\sigma(X_t)\,\sigma(X_t)^\intercal)^{-1}\,b(X_t, \theta_t)\,dt \big]\,. \end{aligned} \tag{3.1.79}$$

The proofs of the convergence of the estimator in (3.1.79) are very similar to those for the convergence of the estimator in (3.0.9). However, to implement the estimator in (3.1.79), one must filter the jumps in the process X to obtain the process X^c; we return to this point below when we implement the estimator in (3.1.79) and the process X follows (3.0.1).

We do not have a theoretical result to compare the performance of the two estimators: one where we employ X and the other where we employ X^c. However, when we study the performance of the two estimators in Section 3.2, we find that the SGDTC that employs X^c performs better: converges quicker and the standard deviation of the estimator of θ^* is smaller.

3.2 Performance of SGDCT estimator

In this Section, we use simulations to study the performance of the SGDCT estimators for two jump-diffusion processes. We employ an Euler scheme to simulate paths of the process X over a time horizon $[0, T]$, with $T > 0$ and initial value $X_0 = 2$, for two specifications of the functions b and σ in (3.0.1). An agent observes the path of the process X and employs the estimators in (3.0.9) and (3.1.79) to estimate the unknown parameter θ^*.

Here, we denote by $|\Delta_t X| = |X_{t+\Delta t} - X_t|$ the increment of the process X over the time-step Δt. Therefore, we write the discrete-time dynamics of the estimator θ_t in (3.0.9), which uses the raw data, as follows:

$$\theta_{t+\Delta t} - \theta_t = \beta_t \, \nabla_\theta b(X_t, \theta_t) \, (\sigma^2(X_t))^{-1} \, \left[(X_{t+\Delta t} - X_t) - \, b(X_t, \theta_t) \, \Delta t \right] ,$$

and the estimator θ_t in (3.1.79), which uses the filtered data, as follows:

$$\begin{aligned}\theta_{t+\Delta t} - \theta_t = \beta_t \, \nabla_\theta b(X_t, \theta_t) \, (\sigma^2(X_t))^{-1} \\ \times \left[(X^c_{t+\Delta t} - X^c_t) + \int \xi(X_t, z) \, \nu(dz) \, \Delta t - \, b(X_t, \theta_t) \, \Delta t \right] .\end{aligned} \tag{3.2.1}$$

To implement (3.2.1), the agent employs the methodology in Gloter, Loukianova, and Mai (2018) to disentangle the diffusion and jump components of the innovations of the process X as follows.[3] When the value of the increment $|\Delta_t X|$ is larger than a threshold $v_t = \rho \, (\Delta t)^{1/3}$, with $\rho > 0$, we assume that the innovation contains a jump, so we set the increment of the process to zero (i.e., $|\Delta_t X| = 0$) and write the discrete-time dynamics of the estimator θ_t in (3.2.1) as follows:

$$\begin{aligned}\theta_{t+\Delta t} - \theta_t &= \beta_t \, \nabla_\theta b(X_t, \theta_t) \, (\sigma^2(X_t))^{-1} \\ &\quad \times \left[(X^c_{t+\Delta t} - X^c_t) + \int \xi(X_t, z) \, \nu(dz) \, \Delta t - \, b(X_t, \theta_t) \, \Delta t \right] \\ &\approx \beta_t \, \nabla_\theta b(X_t, \theta_t) \, (\sigma^2(X_t))^{-1} \\ &\quad \times \left[\Delta_t X \, \mathbf{1}_{\{|\Delta_t X| \le v_t\}} + \int \xi(X_t, z) \, \nu(dz) \, \Delta t - \, b(X_t, \theta_t) \, \Delta t \right] ,\end{aligned}$$

where the indicator function $\mathbf{1}_{\{|\Delta_t X| \le v_t\}}$ returns 1 if $|\Delta_t X| \le v_t$ or 0 otherwise. In the simulations below we assume that the value of the threshold parameter ρ is 1.

3.2.1 Ornstein-Uhlenbeck-type process

The process X_t follows:

$$dX_t = -\theta^* \, X_t \, dt + \sigma \, dB_t + dL_t \tag{3.2.2}$$

where the unknown parameter $\theta^* = 2.7$, $\sigma > 0$, $B = (B_t)_{t \ge 0}$ is a standard Brownian motion independent of the compensated compound Poisson process $L = (L_t)_{t \ge 0}$ with

[3] In the finance literature there are several algorithms designed to filter jumps, see e.g., Lee and Mykland (2008), Cartea and Karyampas (2012), and the references therein.

arrival rate $\lambda > 0$ and the jump sizes are i.i.d. normally distributed $N(0.1, 1)$.[4] Substitute the function $b(x, \theta) = -\theta\, x$ into (3.0.9), and write the first SGDCT estimator, which uses the raw data, as

$$d\theta_t = \beta_t \left[-\frac{X_{t-}}{\sigma^2}\, dX_t - \frac{\theta_t\, X_t^2}{\sigma^2}\, dt \right] , \tag{3.2.3}$$

and write the second SGDCT estimator, which uses the filtered data, as

$$d\theta_t = \beta_t \left[-\frac{X_t}{\sigma^2}\, dX_t^c - \frac{\lambda\, 0.1\, X_{t-}}{\sigma^2}\, dt - \frac{\theta_t\, X_t^2}{\sigma^2}\, dt \right] , \tag{3.2.4}$$

both with initial condition $\theta_0 = 4$, $\sigma = 1$, and learning rate $\beta_t = 1/(1 + t)$. We simulate 1,000 paths, and at every time-step of a path we employ the methodology described above to filter the jumps in the process X when we employ (3.2.4).

We study the performance of each online estimator for various choices of: time horizon T, arrival rate of jumps λ, and number of time-steps in the Euler scheme. The number of time-steps determines the frequency of the observations; thus, as the number of time-steps increases (for a fixed T), the agent observes the process more frequently. For each path, we update the estimate θ_t every time a new observation arrives and denote by θ_T the final estimate of the parameter θ^*.

Table 3.1 reports the mean and the standard deviation of the final parameter estimates θ_T where the process θ follows (3.2.3). For example, when the arrival rate of the jumps is $\lambda = 10$, $T = 80$, and the number of time-steps (i.e., observations) is 120,000, the mean and the standard deviation of the 1,000 final estimates θ_T are 2.80 and 0.29, respectively. We observe that as the number of time-steps increases, the standard deviation of the estimates θ_T decreases because the agent employs more observations to estimate the unknown parameter θ^*.

Similarly, Table 3.2 reports the mean and the standard deviation of the final parameter estimates θ_T when the agent employs the estimator in (3.2.4). For example, when the arrival rate of the jumps is $\lambda = 10$, $T = 80$, and the number of time-steps (i.e., observations) is 120,000, the mean and the standard deviation of the 1,000 final estimates θ_T are 2.72 and 0.09, respectively. Note that in all the simulations we run, this estimator converges quicker, and with a lower standard deviation, than the estimator that employs the raw data; i.e., without filtering the jumps – we return to this point at the end of the Section.

[4]If the value of the volatility parameter is unknown, one can employ various methods to obtain an estimate of σ, see Cartea and Karyampas (2012).

		$\lambda = 10$		$\lambda = 30$		$\lambda = 50$	
		θ_T		θ_T		θ_T	
T	Steps (observations)	mean	std dev	mean	std dev	mean	std dev
15	75000	3.04	0.69	3.34	0.92	3.62	1.07
	150000	3.14	0.65	3.52	0.99	3.78	1.10
	225000	3.07	0.67	3.39	0.89	3.78	1.12
25	125000	3.02	0.54	3.20	0.64	3.47	0.78
	250000	2.98	0.52	3.30	0.76	3.48	0.84
	375000	2.97	0.53	3.14	0.78	3.41	0.94
35	175000	2.82	0.42	3.09	0.67	3.47	0.89
	350000	2.93	0.44	3.08	0.64	3.19	0.78
	525000	2.76	0.44	3.04	0.58	3.22	0.72
50	250000	2.80	0.35	2.88	0.52	3.17	0.63
	500000	2.79	0.33	2.97	0.57	3.16	0.76
	750000	2.77	0.39	2.89	0.55	3.08	0.57
80	400000	2.80	0.29	2.85	0.41	3.03	0.58
	800000	2.82	0.32	2.84	0.40	2.95	0.51
	1200000	2.77	0.27	2.83	0.43	2.99	0.51

Table 3.1: Mean and standard deviation of θ_T (1,000 simulations) with estimator (3.2.3) when X follows the OU-type process in (3.2.2), where $\sigma = 1$, $\lambda = 10, 30, 50$, $\theta^* = 2.7$, and $\theta_0 = 4$.

As time evolves, we track the accuracy of the estimate θ_t for $t \in [0, T]$ with $T = 150$. We perform 500 simulations and the number of time-steps in each simulation is 1,500,000. The time t percent error of the estimate is the average of $100\,|\theta_t - \theta^*|/\theta^*$ for the simulated paths, see Figures 3.1 and 3.2 for the estimator using raw data and filtered data respectively. Clearly, as the agent collects more information about the process X, the estimate of the unknown parameter θ^* improves.

3.2.2 Bimodal process

The process X_t follows:

$$dX_t = -\theta^* \, X_{t-}^3 \, dt + \sigma \, dB_t + dL_t \,, \tag{3.2.5}$$

where the unknown drift coefficient $\theta^* = 2.7$, $\sigma = 2$, $B = (B_t)_{t\geq 0}$ is a Brownian motion and $L = (L_t)_{t\geq 0}$ is a compensated Poisson process with arrival rate $\lambda > 0$. Again, the second equation of (3.2.5) uses to align the equation with (3.0.1). Substitute the

		$\lambda = 10$		$\lambda = 30$		$\lambda = 50$	
		θ_T		θ_T		θ_T	
T	Steps (observations)	mean	std dev	mean	std dev	mean	std dev
15	75000	2.73	0.21	2.67	0.17	2.65	0.18
	150000	2.73	0.22	2.71	0.17	2.71	0.17
	225000	2.76	0.21	2.74	0.20	2.71	0.16
25	125000	2.71	0.16	2.72	0.15	2.70	0.14
	250000	2.72	0.16	2.69	0.15	2.72	0.14
	375000	2.75	0.17	2.72	0.13	2.72	0.16
35	175000	2.73	0.15	2.70	0.12	2.68	0.12
	350000	2.70	0.14	2.69	0.14	2.70	0.12
	525000	2.73	0.15	2.71	0.12	2.70	0.12
50	250000	2.68	0.10	2.68	0.11	2.68	0.12
	500000	2.72	0.11	2.70	0.10	2.69	0.10
	750000	2.74	0.11	2.71	0.10	2.67	0.10
80	400000	2.71	0.08	2.70	0.07	2.68	0.08
	800000	2.71	0.09	2.69	0.08	2.68	0.09
	1200000	2.72	0.09	2.70	0.09	2.68	0.09

Table 3.2: Mean and standard deviation of θ_T (1,000 simulations) with the second estimator (3.2.4) when X follows the OU-type process in (3.2.2), where $\sigma = 1$, $\lambda = 10, 30, 50$, $\theta^* = 2.7$, and $\theta_0 = 4$.

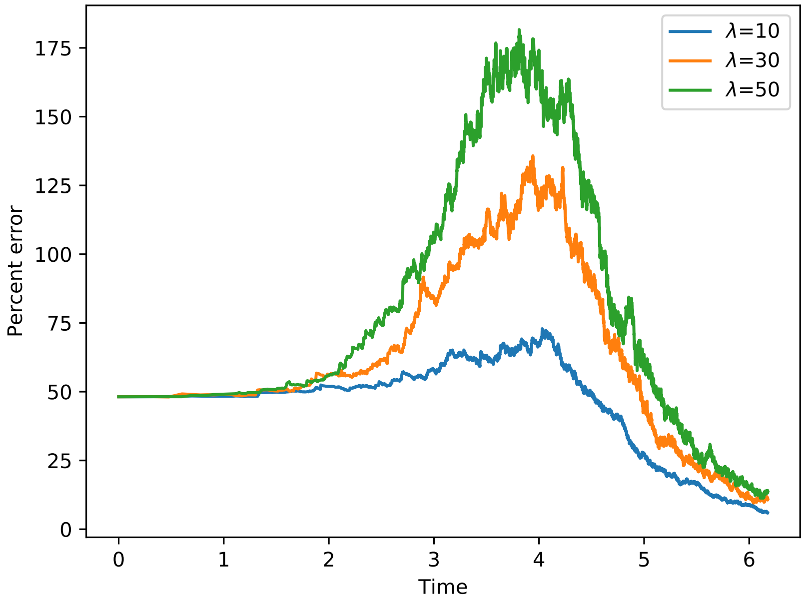

Figure 3.1: The y-axis shows the mean of the percent error $100\,|\hat{\theta} - \theta^*|/\theta^*$ for 500 simulations of the path of X in (3.2.2). The x-axis is log of time, $T = 150$ with 1,500,000 observations. We use the estimator in (3.2.3), which uses raw data.

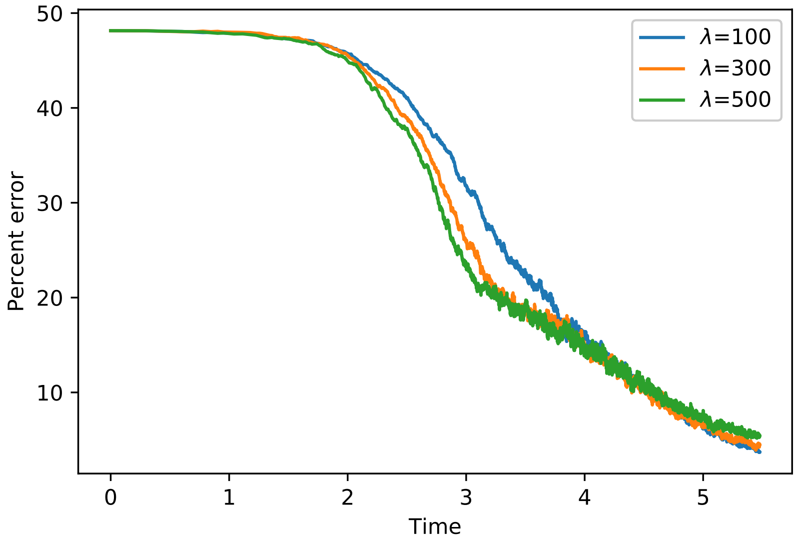

Figure 3.2: The y-axis shows the mean of the percent error $100\,|\hat{\theta} - \theta^*|/\theta^*$ for 500 simulations of the path of X in (3.2.2). The x-axis is log of time, $T = 150$ with 1,500,000 observations. We use the estimator in (3.2.4), which uses filtered data

		$\lambda = 10$		$\lambda = 30$		$\lambda = 50$	
		θ_T		θ_T		θ_T	
T	Steps (observations)	mean	std dev	mean	std dev	mean	std dev
15	75000	3.18	0.50	3.57	0.82	3.93	1.01
	150000	3.21	0.50	3.47	0.78	4.02	0.89
	225000	3.10	0.47	3.52	0.87	3.96	0.96
25	125000	2.99	0.42	3.31	0.61	3.68	0.74
	250000	2.97	0.43	3.31	0.66	3.62	0.76
	375000	3.02	0.43	3.28	0.65	3.53	0.81
35	175000	2.91	0.37	3.13	0.54	3.38	0.75
	350000	2.90	0.33	3.26	0.53	3.40	0.67
	525000	2.94	0.35	3.18	0.51	3.39	0.76
50	250000	2.88	0.30	3.01	0.48	3.23	0.61
	500000	2.86	0.32	3.00	0.43	3.23	0.55
	750000	2.88	0.30	3.06	0.44	3.19	0.57
80	400000	2.77	0.22	2.89	0.39	3.06	0.50
	800000	2.83	0.27	2.96	0.40	3.01	0.50
	1200000	2.79	0.25	2.92	0.40	3.04	0.49

Table 3.3: Mean and standard deviation of θ_T (1,000 simulations) with estimator that uses raw data for the Bimodal process in (3.2.5) with $\sigma = 2$, $\lambda = 10, 30, 50$, $\theta^* = 2.7$, and $\theta_0 = 4$.

function $b(x, \theta) = -\theta\, x^3$ into (3.0.9), the first SGDCT estimator has the form

$$d\theta_t = \beta_t \left[-\frac{X_{t-}^3}{\sigma^2}\, dX_t - \frac{\theta_t\, X_t^6}{\sigma^2}\, dt \right], \tag{3.2.6}$$

and the second SGDCT estimator is

$$d\theta_t = \beta_t \left[-\frac{X_t^3}{\sigma^2}\, dX_t^c - \frac{\lambda\, X_{t-}^3}{\sigma^2}\, dt - \frac{\theta_t\, X_t^6}{\sigma^2}\, dt \right], \tag{3.2.7}$$

both with initial condition $\theta_0 = 4$, and learning rate $\beta_t = 1/(1+t)$. As above, Tables 3.3 and 3.4 report the results of 1,000 simulations, and Figures 3.3 and 3.4 show the percent error of the estimation as time evolves for the estimator using raw data and filtered data, respectively. The interpretation of the results is similar to that provided in the examples above.

Finally, we remark that in the two examples we discuss, the estimator that employs the filtered data converges considerably faster, and with lower standard deviation, than the estimator that uses the raw data. Clearly, the jumps affect the performance

		$\lambda = 10$		$\lambda = 30$		$\lambda = 50$	
		θ_T		θ_T		θ_T	
T	Steps (observations)	mean	std dev	mean	std dev	mean	std dev
15	75000	2.73	0.18	2.65	0.20	2.65	0.21
	150000	2.75	0.18	2.75	0.19	2.71	0.18
	225000	2.74	0.20	2.71	0.18	2.71	0.17
25	125000	2.71	0.14	2.70	0.13	2.55	0.22
	250000	2.73	0.13	2.71	0.14	2.68	0.13
	375000	2.70	0.15	2.69	0.15	2.71	0.14
35	175000	2.73	0.13	2.65	0.17	2.55	0.20
	350000	2.71	0.11	2.71	0.11	2.66	0.13
	525000	2.73	0.13	2.70	0.12	2.71	0.12
50	250000	2.71	0.12	2.64	0.12	2.52	0.25
	500000	2.72	0.10	2.69	0.09	2.66	0.11
	750000	2.72	0.10	2.70	0.10	2.71	0.11
80	400000	2.69	0.08	2.62	0.12	2.53	0.19
	800000	2.70	0.09	2.68	0.10	2.68	0.09
	1200000	2.70	0.08	2.68	0.08	2.67	0.08

Table 3.4: Mean and standard deviation of θ_T (1,000 simulations) with estimator that uses filtered data for the Bimodal process in (3.2.5) with $\sigma = 2$, $\lambda = 10, 30, 50$, $\theta^* = 2.7$, and $\theta_0 = 4$.

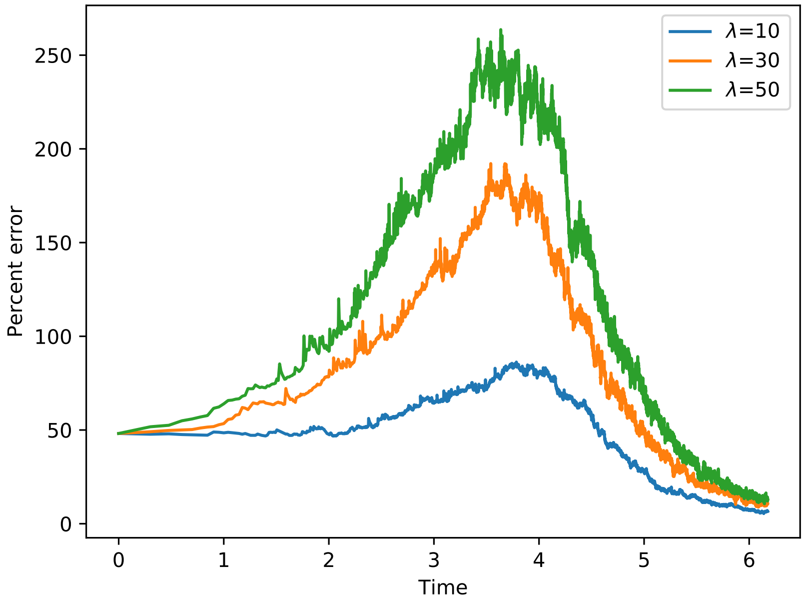

Figure 3.3: The y-axis shows the mean of the percent error $100\,|\hat{\theta} - \theta^*|/\theta^*$ for 500 simulations of the path of X in (3.2.5). The x-axis is log of time, $T = 150$ with 1,500,000 observations. We use the estimator in (3.2.6), which uses raw data.

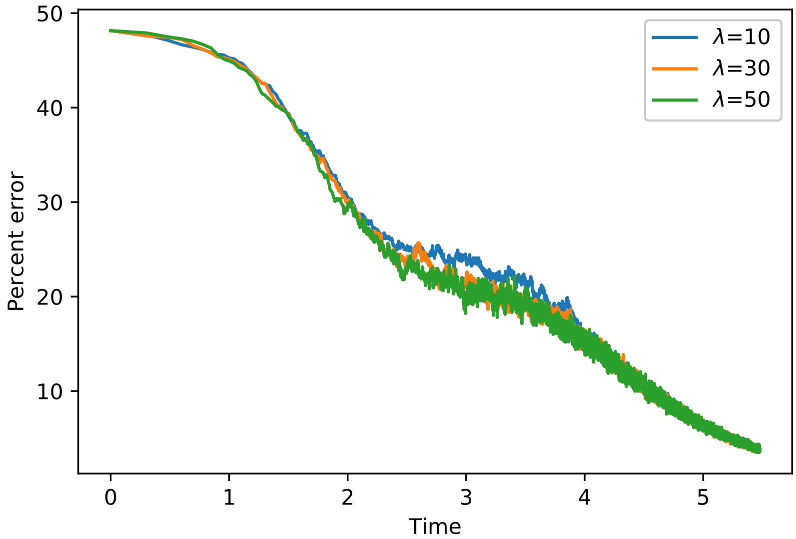

Figure 3.4: The y-axis shows the mean of the percent error $100\,|\hat{\theta} - \theta^*|/\theta^*$ for 500 simulations of the path of X in (3.2.5). The x-axis is log of time, $T = 150$ with 1,500,000 observations. We use the estimator in (3.2.7), which uses filtered data.

of the estimator. Everything else being equal, when the value of the arrival rate λ increases, the unbiased noise that stems from the compensated jump component in the process X increases the variance of the estimator θ_t. Thus, the estimator that uses the raw data is noisier and takes longer to converge.

3.3 Conclusions and future research

We showed the convergence of the SGDCT algorithm to estimate the drift parameter of a jump-diffusion process. Our result is an extension to the work of Sirignano and Spiliopoulos (2017). We proposed two estimators: an estimator that uses the raw data, and another that filters the jumps of the jump-diffusion process before estimating the drift parameter.

Our estimators are desirable because online updates are more computationally efficient than offline updates. As new observations arrive, the online estimator is updated, while offline algorithms solve a new optimisation problem to estimate the unknown parameter.

We employed simulations to illustrate the performance of the estimators. We see that when the agent filters the jumps in the process, the estimator converges quicker than when the agent employs the raw data (i.e., without filtering jumps).

The SGDCT estimator is well-suited for the recent advances in control theory that focus on adaptive robust control to obtain strategies robust to model uncertainty, see Bielecki, Chen, and Cialenco (2017a). The adaptive control framework requires an estimator that is continuously updated with the arrival of new information and also requires that the coupled process $(X_t, \theta_t)_{t\geq 0}$ is Markov. The SGDCT algorithm we developed here meets both requirements.

The online estimators may be employed in many financial problems where prices, or other observed variables, follow jump-diffusion processes. For example, in high-frequency trading, many models assume that the midprice follows, over a short-period of time, an OU-type process. A similar application is in pairs trading, where a linear combination of two or more assets follows an OU-type process, see Cartea, Jaimungal, and Penalva (2015). In both cases, the success of these high-frequency trading strategies relies on being able to estimate the drift of a process. An important advantage of our estimator is that the estimates of the drift are quickly updated with the arrival of every innovation in the prices of the assets, so the strategies can be implemented in real time.

Finally, there are a number of future research directions. For example, one may analyse the statistical properties of the jump-diffusion SGDCT we presented here. If the function b is convex, one could follow the same argument as that in Sirignano and Spiliopoulos (2020) to prove the central limit theorem of the estimator in (3.0.9). Although most of the proof of this extension would follow from standard results on the property of jump-diffusion processes, one would still require the proof of the convergence of the term in (3.1.2). Also, it would be interesting to study the jump-diffusion filtering problem where the agent partially observes the process.

Chapter 4

Adaptive robust control [1]

In classical stochastic control problems, agents look for the best policy to optimise a value function that depends on a stochastic process $X = (X_t)_{t\geq 0}$. An extensively studied problem is one in which the process X is a diffusion and it is the unique solution to the stochastic differential equation (SDE)

$$dX_t = b(X_t, \theta)\, dt + \sigma(X_t)\, dB_t^{\mathbb{P}}\,,$$

where the functions b and σ are Lipschitz, $B^{\mathbb{P}} = (B_t^{\mathbb{P}})_{t\geq 0}$ is a standard Brownian motion under a reference measure $\mathbb{P}$, and $\theta \in \mathbb{R}^d$. A standard method to solve the agent's problem involves two key steps. First, show that the value function admits the dynamic programming principle (DPP). Second, characterise the value function as the solution to a non-linear partial differential equation (PDE) and derive the agent's optimal decisions; see e.g., Pham (2009).

In the classical approach, the agent assumes that the reference measure $\mathbb{P}$ of the process X is known, but if the reference measure is incorrectly specified, the agent will find a sub-optimal policy. A common approach to making the agent's control problem robust to model misspecification is to consider a set of alternative probability measures $\mathcal{P}$ and then specify a criterion to choose the 'optimal' measure from this alternative set. The criterion is generally one where the agent adopts conservative strategies.

We highlight two shortcomings of the classical robust stochastic control approach. One, it may lead to optimal policies that are too conservative. For example, if the agent's level of confidence on each measure in the set $\mathcal{P}$ is the same, the outcome is to choose the optimal policy under the probability measure that produces the worst result with respect to the performance criterion of the agent. Two, in the classical

[1] This chapter has been published by SIAM Journal on Control and Optimization, see Bhudisaksang and Cartea (2021a).

robust approach, the agent does not update her views on the set of measures $\mathcal{P}$ as more information is revealed when the process X evolves in time; i.e., the set of alternative measures is fixed throughout the horizon of the control problem.

In recent work, Bielecki, Chen, Cialenco, et al. (2017) propose an 'adaptive robust' framework in which the agent incorporates the evolution of the underlying stochastic process as an ingredient in a control problem robust to model misspecification. They assume that the underlying dynamics follow a discrete-time homogeneous Markov process under a reference measure, construct the set of alternative measures $\mathcal{P}$ via a composition of probability kernels to incorporate the arrival of new information, and use a measurable selection theorem to prove that the value function of the agent satisfies the DPP.

In this chapter, we assume that X is a jump-diffusion process taking values in $\mathbb{R}^n$ and it is the unique strong solution to the SDE

$$dX_t = b(X_t, \theta^*)\, dt + \sigma(X_t)\, dB_t^{\mathbb{P}_{\theta^*}} + \xi(X_{t-})\, dL_t\,. \tag{4.0.1}$$

Here, $\theta^* \in \mathbb{R}^d$ is an unknown parameter, the process $B^{\mathbb{P}_{\theta^*}} = (B_t^{\mathbb{P}_{\theta^*}})_{t\geq 0}$ is a standard Brownian motion under the probability measure $\mathbb{P}_{\theta^*}$, and $b : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$, $\xi : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$. The jump component of X_t is the Lévy process $L = (L_t)_{t\geq 0}$ (independent of $B^{\mathbb{P}_{\theta^*}}$), which takes values in $\mathbb{R}^r$ and with Lévy-Khintchine triplet $(0, 0, \nu)$ such that $\int_{\mathbb{R}^r} |z|\ \nu(dz) < \infty$ and $\nu(\mathbb{R}^r) < \infty$, and μ is an associated Poisson measure of the process L, where $\tilde{\mu}$ denotes the compensated Poisson measure. Note that the diffusion component σ, the jump component ξ, and the jump process L do not depend on the parameter θ^*.

At time t, the agent's estimate of the parameter θ^* is denoted by $\hat{\theta}_t$, and as time evolves and new information arrives, the agent updates the estimate $\hat{\theta}_t$. We model the processes X and $\hat{\theta}$ jointly, so we let $\tilde{X} = (\hat{\theta}, X)^\intercal$, which satisfies the SDE

$$d\tilde{X}_t = \boldsymbol{b}(t, \tilde{X}_t, \theta^*)\, dt + \boldsymbol{\sigma}(t, \tilde{X}_t)\, dB_t^{\mathbb{P}_{\theta^*}} + \boldsymbol{\xi}(t, \tilde{X}_{t-})\, dL_t\,. \tag{4.0.2}$$

In addition, the agent considers the controlled diffusion process $Y = (Y_t)_{t\geq 0}$ valued in $\mathbb{R}^{\tilde{d}}$, which satisfies

$$dY_t = \bar{b}(Y_t, \tilde{X}_t, \alpha_t)\, dt + \bar{\sigma}(Y_t, \tilde{X}_t, \alpha_t)\, d\tilde{X}_t\,, \tag{4.0.3}$$

where $\bar{b} : \mathbb{R}^{\tilde{d}} \times \mathbb{R}^{n+d} \times \mathbb{R}^k \to \mathbb{R}^{\tilde{d}}$, $\bar{\sigma} : \mathbb{R}^{\tilde{d}} \times \mathbb{R}^{n+d} \times \mathbb{R}^k \to \mathbb{R}^{\tilde{d}} \times \mathbb{R}^{n+d}$, and the agent controls the process $\alpha = (\alpha_t)_{t\geq 0}$, which takes values in $\mathbb{R}^k$.

Note that the second term on the right-hand side of (4.0.3) is $d\tilde{X}$ whose drift depends on the parameter θ^* (see (4.0.2)); thus, the drift of the process Y depends on

θ^*. In this chapter, we focus on the case where the agent does not have full knowledge of the drift term of the process X, while the agent has complete information about the volatility and jump terms of the process X.

In the adaptive robust framework, the agent considers the set of alternative measures $\mathcal{P}(t, x, G)$, where x represents the state of the underlying stochastic process X, and G is a function of the value of x and time t. The function G specifies the model uncertainty that stems from the estimation process $\hat{\theta}$. We remark that in general it is difficult (perhaps not possible) to construct the function G to be a confidence interval in a statistical sense for the estimator process $\hat{\theta}$, see the definition of a confidence interval in Bielecki, Chen, and Cialenco (2017a). However, when the process X is a geometric Brownian motion and the estimator process is the maximum likelihood estimator (MLE), we can construct the function G that specifies the confidence interval of the estimator in a statistical framework. In the remainder of this chapter, it suffices to denote $\mathcal{P}(t, x, G)$ as $\mathcal{P}(t, x)$ because the function G is defined and fixed at the initial time of the agent's control problem. The agent's performance criterion is

$$J(t, x, y, \mathbb{P}, \alpha) := \mathbb{E}^{\mathbb{P}}\left[\int_t^T f(s, Y_s, \alpha_s)\, ds + g(Y_T)\right], \tag{4.0.4}$$

where $\mathbb{P}$ is a probability measure, f and g are continuous functions, $\tilde{X}_t = x$, $Y_t = y$ and recall that α_t is the control process. The value function of the adaptive robust control problem is

$$\begin{aligned} w(t, x, y, \alpha) &:= \sup_{\mathbb{P} \in \mathcal{P}(t,x)} J(t, x, y, \mathbb{P}, \alpha)\,, \\ v(t, x, y) &:= \inf_{\alpha \in \mathcal{A}_0} w(t, x, y, \alpha)\,, \end{aligned} \tag{4.0.5}$$

where $\mathcal{A}_0$ is the set of admissible control processes.

We extend the discrete-time adaptive robust control framework in Bielecki, Chen, Cialenco, et al. (2017). In their work, the authors solve a similar problem to that in (4.0.5), where the underlying discrete-time stochastic process is ergodic and Markov, while here we assume that the underlying stochastic process is driven by a continuous-time jump-diffusion process. In the discrete-time setup of Bielecki et al., it is difficult to obtain a numerical solution of the value function v because in each back-propagation of v, one needs to compute the expectation of the criterion function J in (4.0.4); see Chen and Ludkovski (2019). In contrast, here we show that the value function v in (4.0.5) satisfies the dynamic programming principle (DPP), characterise the value function as the solution of a non-linear PDE, and we prove the uniqueness of the viscosity solution. The PDE we derive is not standard because it contains a non-linear operator that depends on the state variable x (see e.g., Pham (1998)); this is discussed in detail in Section 2. Thus, the continuous-time framework for the adaptive robust control is easier to solve (numerically) because there exists extensive literature on solving non-linear PDEs. As example of the adaptive robust strategy, we derive the solution for the optimal execution problem in electronic markets; see e.g., Cartea, Jaimungal, and Penalva (2015).

Most of the literature that considers 'parameter uncertainty' in diffusion-based models assumes that the drift parameter and the volatility parameter of the diffusion process lie in a known fixed interval, which is in contrast with our adaptive robust model where both the size of the uncertainty interval and the estimates of the parameters are updated as time evolves. When the unknown parameters lie in a fixed interval, performance criteria in the form of (4.0.4) are time-consistent because the agent does not update the estimates of the unknown parameters. For example, Epstein and Ji (2014) consider a utility maximisation problem for a controlled diffusion process in which the drift and the volatility terms of the diffusion lie in a fixed interval, and Bannör et al. (2016) investigate parameter uncertainty in energy markets; see also Denis and Kervarec (2013), Biagini and Pınar (2017), and Bergen et al. (2018), all of which assume that one or more parameters of the model lie in a fixed set and the estimates are not updated. We also mention other literature working on robust optimal stopping problem where they use a similar measurable selection technique, see Bayraktar and Yao (2014) and Ekren, Touzi, and Zhang (2014).

Moreover, in the work of Ismail and Pham (2019) and Pham, Wei, and Zhou (2018), an agent maximises a mean-variance criterion (which is not time consistent) in a one-step optimal asset allocation problem. In both works, the authors assume that the unknown parameters lie in a fixed interval and estimates of the parameters are not updated as time evolves.

There are also a number of papers that consider Knightian uncertainty (also referred to as model uncertainty), but do not include learning as proposed in this chapter. In the Knightian uncertainty approach, the agent considers a set of alternative measures, equivalent to the agent's reference measure, to make decisions that are robust to model uncertainty. There is a penalty for choosing an alternative model – the penalty is a function of the entropy between the reference measure and the measure of the alternative model. Entropic penalties are convenient because they preserve time-consistency in optimisation problems, so it is straightforward to show that the DPP holds and one can employ standard control techniques to obtain the optimal strategy robust to model misspecification. See for example the work of

Hansen and Sargent (2011), Jaimungal and Sigloch (2012), Skiadas (2013), Cartea and Jaimungal (2017), Cartea and Sánchez-Betancourt (2021), all of which assume model uncertainty with respect to the drift of a diffusion process, while the work of Cartea, Donnelly, and Jaimungal (2017) assumes model uncertainty with respect to the drift of a diffusion and to the intensity of the arrival of jumps.

To illustrate the performance of the adaptive robust approach we analyse a classical problem in finance. We derive the optimal acquisition strategy for an agent who purchases a large block of shares over a trading window, where the drift parameter of the stock price dynamics is not known by the agent. The performance of the adaptive robust strategy is compared with that of strategies in which the agent employs a wrong value of the drift parameter or employs a robust strategy. The robust strategy assumes that the drift parameter lies in a fixed interval and there is no learning as time evolves. Our results show that when the agent has enough time to learn the value of the unknown parameter, the adaptive robust strategy we develop in this chapter performs better (lower average and lower variance of acquisition costs of block of shares) than when the agent employs a robust strategy or uses the incorrect parameter estimate. The superior performance of the robust adaptive strategy stems from learning the value of the drift parameter during the trading window.

The remainder of this chapter is organised as follows. Section 4.1 develops the set of alternative measures used by the agent and introduces the continuous-time adapted robust control problem. We show that the agent's dynamic optimisation problem in (4.0.5) admits the DPP and we characterise the value function as a viscosity solution of a non-linear PDE. In Section 3, we present an application of the adaptive robust control problem to financial problems. Lastly, Section 4 concludes and discusses future research directions.

4.1 Model

To streamline the presentation of the model, we provide a few definitions and other ingredients of the adaptive robust framework.

We define the terms "drift characteristic", "volatility characteristic", and "jump characteristic"of a stochastic process under a probability measure. Let $V = (V_t)_{t\geq 0}$ be a stochastic process in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We call $\mathbb{P}$ a semimartingale measure of the process V, if there exists a triple $(B^{\mathbb{P}}, C, \nu^{\mathbb{P}})$, such that $B^{\mathbb{P}}$ and C are the finite-variation part and the covariation process of the continuous local martingale of $V - \sum_{0\leq s\leq \cdot}(\Delta V_s - h(\Delta V_s))$ where h is a bounded function that satisfies

$h(x) = x$ in a neighbourhood of the origin, and $\nu^{\mathbb{P}}$ is the predictable compensator of μ^V, where

$$\mu^V(\omega, dt, dx) = \sum_{s\geq 0} \mathbf{1}_{\{\Delta Y_s(\omega)\neq 0\}} \mathbf{1}_{(s,\Delta V_s(\omega))}(dt, dx)\,. \tag{4.1.1}$$

Now, assume that for the triple $(B^{\mathbb{P}}, C, \nu^{\mathbb{P}})$, there exist the processes $(\gamma^{\mathbb{P}}, a, F^{\mathbb{P}}_{\omega,s})$, such that

$$B^{\mathbb{P}}_t(\omega) = \int_0^t \gamma^{\mathbb{P}}_s(\omega)\, ds\,, \quad C_t(\omega) = \int_0^t a_s(\omega)\, ds\,, \quad \text{and} \quad \nu^{\mathbb{P}}(\omega, ds, dx) = F^{\mathbb{P}}_{\omega,s}(dx)\, ds\,,$$

for $d\mathbb{P} \times dt$ almost-all $(\omega, s) \in \Omega \times (0, T]$. Then, $\mathbb{P}$ is a semimartingale measure of the process V with absolutely continuous characteristics, in which case we say that the processes $\gamma^{\mathbb{P}}$, a, and $F^{\mathbb{P}}_{\omega,s}$ are the drift characteristic, the volatility characteristic, and the jump characteristic, respectively, of the process V under the probability measure $\mathbb{P}$.

The remainder of this Section proceeds as follows. Subsection 2.1 uses an explicit formula of the jump-diffusion process to motivate our definition of the alternative measures. Subsection 2.2 defines the set of alternative measures $\mathcal{P}(t, x)$ and formulates the adaptive robust framework. Subsection 2.3 proves the measurability of the value function. And subsections 2.4 and 2.5 show that the adaptive robust control problem in continuous-time is time-consistent, prove the DPP of the value function and derive the non-linear PDE it satisfies.

4.1.1 Set of alternative measures

Recall that the process X follows (4.0.1) and that the agent does not know the value of the parameter $\theta^* \in \mathbb{R}^d$. As time evolves, the agent observes the realisation of the process X and updates $\hat{\theta}$ (i.e., the estimate of the parameter θ^*), which satisfies the SDE

$$d\hat{\theta}_t = \beta_t \left[\tilde{b}(X_t, \hat{\theta}_t, \theta^*)\, dt + \tilde{\sigma}(X_t, \hat{\theta}_t)\, dB^{\mathbb{P}_{\theta^*}}_t + \tilde{\xi}(X_{t-}, \hat{\theta}_t)\, dL_t \right]. \tag{4.1.2}$$

Here, the functions $\tilde{b}$ and $\tilde{\sigma}$ specify the drift and the volatility of the estimator process $\hat{\theta}$, and $\beta > 0$ is the learning rate that depends on t. The function $\tilde{b}$ satisfies

$$\tilde{b}(x, \hat{\theta}, \theta_1) - \tilde{b}(x, \hat{\theta}, \theta_2) = \frac{(b(x, \theta_1) - b(x, \theta_2))\, \tilde{\sigma}(x, \hat{\theta})}{\sigma(x)}\,, \tag{4.1.3}$$

for all $x, \hat{\theta}$, and all $\theta_1,\, \theta_2 \in \mathbb{R}^d$.

Therefore, under each probability measure $\mathbb{P}_\theta$ for $\theta \in \mathbb{R}^d$, the agent considers the augmented process $\tilde{X} = (X, \hat{\theta})^\top \in \mathbb{R}^{n+d}$, where X follows (4.0.1) and $\hat{\theta}$ follows (4.1.2), so $\tilde{X}$ satisfies

$$d\tilde{X}_t = \boldsymbol{b}(t, \tilde{X}_t, \theta)\, dt + \boldsymbol{\sigma}(t, \tilde{X}_t)\, dB_t^{\mathbb{P}_\theta} \; + \boldsymbol{\xi}(t, \tilde{X}_{t-})\, dL_t\,, \tag{4.1.4}$$

where $\boldsymbol{b}(t, \tilde{x}, \theta) := [\, b(x, \theta),\, \beta_t\, \tilde{b}(x, \hat{\theta}, \theta)\,]^\intercal$, $\boldsymbol{\sigma}(t, \tilde{x}) := [\, \sigma(x),\, \beta_t\, \tilde{\sigma}(x, \hat{\theta})\,]^\intercal$, $\boldsymbol{\xi}(t, \tilde{x}) := [\, \xi(x), \beta_t\, \tilde{\xi}(x, \hat{\theta})\,]^\intercal$, and $\tilde{x} = (x, \hat{\theta})$. The coefficients in (4.1.4) depend on time because the learning rate β_t is time-dependent.

When X follows an arithmetic or a geometric Brownian motion, the maximum likelihood estimators for the drift of X have the form in (4.1.2). Also, if the process X is ergodic, the stochastic gradient descent in continuous-time is also as in (4.1.2); see Sirignano and Spiliopoulos (2017) for diffusions and Bhudisaksang and Cartea (2021b) for jump-diffusions. However, we note that in general, the estimator of the parameter in the drift term does not have the form in (4.1.2).

Let $G : [0, T] \times \mathbb{R}^{n+d} \to 2^{\mathbb{R}^d}$ be an exogenous function that specifies the model uncertainty of the estimation process. Assume that at time t the agent's estimate of the parameter θ^* is the progressively measurable process $\tilde{\theta}_t \in G(t, \tilde{X}_t)$. Then, by Girsanov's theorem, there is a probability measure $\mathbb{P}_{\tilde{\theta}}$, such that $dX_t = b(X_t, \tilde{\theta}_t)\, dt + \sigma(X_t)\, dB_t^{\mathbb{P}_{\tilde{\theta}}} + \xi(X_{t-})\, dL_t$, and we write the dynamics of the joint process $\tilde{X}$, under the probability measure $\mathbb{P}_{\tilde{\theta}}$, as

$$d\tilde{X}_t = \boldsymbol{b}(t, \tilde{X}_t, \tilde{\theta}_t)\, dt + \boldsymbol{\sigma}(t, \tilde{X}_t)\, dB_t^{\mathbb{P}_{\tilde{\theta}}} + \boldsymbol{\xi}(t, X_{t-})\, dL_t\,.$$

This motivates our definition of the set of alternative probability measures $\mathcal{P}(t, x, G)$ that contains the measure $\mathbb{P}$ whose drift characteristic, volatility characteristic, and jump characteristic of the process X are:

$$\boldsymbol{b}(s, \tilde{X}_s, \tilde{\theta}_s) + \int_{\mathbb{R}^r} (\hat{h}(\boldsymbol{\xi}(t, \tilde{X}_{s-})\, z) - \boldsymbol{\xi}(s, \tilde{X}_{s-})\, h(z))\, \nu(dz)\,,$$

$\boldsymbol{\sigma}(s, \tilde{X}_s)\, \boldsymbol{\sigma}(s, \tilde{X}_s)^\intercal$, and $F^{\mathbb{P}}_{\omega,t}(dx) := \int_{\mathbb{R}^r} 1_{\{\boldsymbol{\xi}(s, \tilde{X}_{s-}(\omega))\, z \,\in\, dx\}}\, \nu(dz)$, respectively. The construction of these probability measures is similar to that of a weak formulation of the control problem in El Karoui and Tan (2013b). We provide a formal definition of the set of alternative measures $\mathcal{P}(t, x, G)$ after the following subsection.

4.1.2 Model setup

In this subsection, we present the adaptive robust control problem in continuous-time and define the performance criterion of the agent – we employ a weak formulation of the control problem to simplify some proofs. Denote by $\mathcal{B}(Y)$ a Borel σ-field of the Polish space Y and let $\Omega = D([0,T],\mathbb{R}^{n+d})$ be the space of all càdlàg paths $\omega = (\omega_t)_{t\geq 0}$, $\mathcal{F} = \mathcal{B}(\Omega)$, and let $\tilde{X}$ be a canonical process, i.e., $\tilde{X}_t(\omega) = \omega_t$. Note that for any probability measure $\mathbb{P}$ and $A \in \mathcal{F}$, we have that $\mathbb{P}(\tilde{X} \in A) = \mathbb{P}(A)$. Denote by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ the raw filtration generated by the canonical process X. Let $\mathfrak{P}(\Omega)$ denote the set of probability measures on Ω and denote the set of semimartingale probability measures with absolutely continuous characteristics of the process $\tilde{X}$ by $\mathfrak{P}^{ac}_{sem}(\Omega)$. Next, we define the measure kernel

$$\bar{\mathbb{P}}(\omega, t, dx) := \int_{\mathbb{R}^r} 1_{\{\boldsymbol{\xi}(t,\tilde{X}_{t-}(\omega))\, z \,\in\, dx\}}\, \nu(dz)\,.$$

Note that the kernel $\bar{\mathbb{P}}$ is a Borel measurable function from $\Omega \times [0,T] \times \mathcal{B}(\mathbb{R}^{n+d})$ to $\mathbb{R}$.

Now, we define the learnable property.

Definition 4.1.1. Let $t \in (0,T]$. A semimartingale probability measure $\mathbb{P} \in \mathfrak{P}^{ac}_{sem}(\Omega)$ is called *learnable by* G on $(t,T]$ if: the drift, volatility, and jump characteristics of $\mathbb{P}$ are

$$\gamma^{\mathbb{P}}_s \in \boldsymbol{b}^*(s,\tilde{X}_s)\,, \qquad a_s = \boldsymbol{\sigma}(s,\tilde{X}_s)\,\boldsymbol{\sigma}(s,\tilde{X}_s)^\intercal\,, \qquad \text{and} \qquad F^{\mathbb{P}}_{\omega,s}(dx) = \bar{\mathbb{P}}(\omega, s, dx)\,,$$

for $d\mathbb{P} \times dt$ almost-all $(\omega, s) \in \Omega \times (t,T]$, where

$$\boldsymbol{b}^*(s,\tilde{X}_s) = \left\{ \boldsymbol{b}(s,\tilde{X}_s,\theta) + \int_{\mathbb{R}^r} (\hat{h}(\boldsymbol{\xi}(s,\tilde{X}_{s-})\, z) - \boldsymbol{\xi}(s,\tilde{X}_{s-})\, h(z))\, \nu(dz) \mid \theta \in G(s,\tilde{X}_s) \right\}.$$

Here, the functions $h : \mathbb{R}^r \to \mathbb{R}^r$ and $\hat{h} : \mathbb{R}^{n+d} \to \mathbb{R}^{n+d}$ are bounded and satisfy $h(x) = x$, $\hat{h}(x) = x$ when $|x| < 1$ and 0 otherwise.

The mapping $\boldsymbol{b}^* : [0,T] \times \mathbb{R}^{n+d} \to 2^{\mathbb{R}^d}$ is continuous with respect to the Hausdorff metric by a standard calculation; see Definition B.2.2 in the Appendix.

Definition 4.1.2. Set of alternative measures. The set of alternative measures $\mathcal{P}(t,x)$ consists of all probability measures $\mathbb{P}$ that satisfy the following two properties:

(i) $\mathbb{P} \in \mathfrak{P}^{ac}_{sem}(\Omega)$ and $\mathbb{P}(\tilde{X}_t = x) = 1$.

(ii) $\mathbb{P}$ is *learnable by* G on the interval $(t,T]$.

The following assumption is crucial to prove measurability of the set of alternative probability measures. This assumption is standard in the literature, see e.g., Assumption 4.1 in Nutz and Handel (2013).

Assumption 4.1.3. For every $t \in \mathbb{R}_+$, we have that

$$\{(s, \omega, \gamma, \rho, F) \in [t, T] \times \Omega \times \mathbb{R}^{n+d} \times \mathbb{R}^{(n+d)\times(n+d)} \times \mathfrak{P}(\mathbb{R}^{n+d}) : \gamma \in \boldsymbol{b}^*(s, \omega_s),$$
$$\rho = a_s(\omega_s), F = \bar{\mathbb{P}}(\omega, s, \cdot)\} \in \mathcal{B}([t, T]) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^{n+d}) \otimes \mathcal{B}(\mathbb{R}^{(n+d)\times(n+d)}) \otimes \mathcal{B}(\mathfrak{P}(\mathbb{R}^{n+d})).$$

We denote the concatenation of a path by $\omega \otimes_t \tilde{\omega} := w_s \mathbf{1}_{s\le t} + (w_t + \tilde{w}_s - \tilde{w}_t)\mathbf{1}_{s>t}$. Let $\mathbb{E}^{\mathbb{P}^{\tau,\bar{\omega}}}[\xi^{\tau,\bar{\omega}}] := \mathbb{E}^{\mathbb{P}^{\bar{\omega}}_{\tau}}[\xi] = \mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_\tau](\bar{\omega})$ where $\xi^{\tau,\bar{\omega}}(\omega) := \xi(\bar{\omega} \otimes_{\tau(\bar{\omega})} \omega)$ and $\{\mathbb{P}^{\bar{\omega}}_{\tau}\}_{\bar{\omega}\in\Omega}$ is a regular conditional probability distribution given $\mathcal{F}_\tau$.

When the process X in (4.0.2) is a diffusion with no jump component, the set of alternative measures $\mathcal{P}(t, x)$ contains a semi-martingale probability measure $\mathbb{P}$, such that the drift characteristic $\gamma^{\mathbb{P}}_s$ of $\mathbb{P}$ is in the set $\boldsymbol{b}^*(s, \tilde{X}_s) = \left\{ \boldsymbol{b}(s, \tilde{X}_s, \theta) \mid \theta \in G(s, \tilde{X}_s)\right\}$.

Next, we define the agent's adaptive robust control problem.

The set of admissible controls, denoted by $\mathcal{A}_0$, consists of all progressively measurable processes that take values in a compact set $A \in \mathbb{R}^k$. Recall from (4.0.4) that the performance criterion of the agent is $J : [0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\tilde{d}} \times \mathfrak{P}(\Omega) \times \mathcal{A}_0 \to \mathbb{R}$, and is given by

$$J(t, x, y, \mathbb{P}, \alpha) := \mathbb{E}^{\mathbb{P}}\left[\int_t^T f(s, Y_s, \alpha_s)\, ds + g(Y_T)\right], \tag{4.1.5}$$

where α is a control in the admissible set $\mathcal{A}_0$, $\tilde{X}_t = x$, $Y_t = y$, and f and g are continuous functions in all variables.

Assume the controls α and $\tilde{\alpha}$ are in the set $\mathcal{A}_0$. We define the distance metric in the space $\mathcal{A}_0$ as follows:

$$\Delta(\alpha, \tilde{\alpha}) := \mathbb{E}^{\mathbb{P}_{\hat{\theta}}}\left[\int_0^T |\alpha_t - \tilde{\alpha}_t|^2\, dt\right],$$

where $\mathbb{P}_{\hat{\theta}}$ is the semimartingale measure for which the drift characteristic of $\tilde{X}$ is

$$\boldsymbol{b}(s, \tilde{X}_s, \hat{\theta}_s) + \int_{\mathbb{R}} \left(\hat{h}(\boldsymbol{\xi}(s, \tilde{X}_{s-})\, z) - \boldsymbol{\xi}(s, \tilde{X}_{s-})\, h(z)\right) \nu(dz).$$

Thus, we denote by $\mathcal{B}_{\mathcal{A}_0}$ the set of a Borel measurable set of the set $\mathcal{A}_0$ generated by the distance metric Δ. By elementary calculations, the set $\mathcal{A}_0$ is a Polish space. We denote the space of all probability measures on $(\Omega, \mathcal{F})$ by $\mathfrak{P}(\Omega)$ and equip it with the weak*-topology.

The agent's value function in the adaptive robust framework is

$$\begin{aligned} w(t, x, y, \alpha) &:= \sup_{\mathbb{P}\in\mathcal{P}(t,x)} J(t, x, y, \mathbb{P}, \alpha), \\ v(t, x, y) &:= \inf_{\alpha\in\mathcal{A}_0} w(t, x, y, \alpha). \end{aligned} \tag{4.1.6}$$

Assumption 4.1.4. The functions J, w, v in (4.1.6) are finite.

Assumption 4.1.4 is satisfied if f and g are bounded functions and if all other assumptions that guarantee that Y is integrable on $[0,T] \times \Omega$ are also satisfied. In what follows, we assume that Assumptions 4.1.3 and 4.1.4 hold.

In summary, the agent chooses a control process α to minimise the performance criterion J, but there is uncertainty about the estimate of the parameter $\hat{\theta}$. Therefore, the agent considers the worst case scenario by choosing a measure from the alternative set $\mathcal{P}(t,x)$ that depends on the state of the underlying process and the estimator process. On the other hand, if the agent chooses to maximise the performance criterion, the control problem is as in (4.1.6), but of inf-sup type (instead of sup-inf).

4.1.3 Measurability of functions and set of probability measures

Here, we show that the set $\{(\omega, t, \mathbb{P}) \in \Omega \times [0,T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\}$ is Borel measurable. This allows us to use a measurable selection theorem on the supremum problem in (4.1.6). Then, we check the measurability of: the performance criterion J, the function w, and the value function v. The measurability of these functions allows the use of measurable selection theorems for both the infimum and supremum problems in (4.1.6).

Lemma 4.1.5. *The set* $\{(\omega, t, \mathbb{P}) \in \Omega \times [0,T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\}$ *is a Borel measurable set.*

For a proof, see Appendix B.1.1.

The class of shifted control processes is constructed from the concatenation of a path as follows: For $\alpha \in \mathcal{A}_0$ and $(t, \bar{\omega}) \in [0,T] \times \Omega$, we set

$$\alpha_s^{t,\bar{\omega}}(\omega) = \alpha_s(\bar{\omega} \otimes_t \omega)\,, \quad (s, \bar{\omega}) \in [0,T] \times \Omega\,.$$

For a stopping time τ and for any $\alpha \in \mathcal{A}_0$, we denote the mapping

$$\alpha^\tau : (\Omega, \mathcal{F}) \to (\mathcal{A}_0, \mathcal{B}_{\mathcal{A}_0})\,, \qquad \bar{\omega} \mapsto \alpha^{\tau(\bar{\omega}), \bar{\omega}}\,.$$

We employ the shifted control processes to show the pseudo-Markov property and DPP of the problem in (4.1.6). Next, we show the pseudo-Markov property of the process $\tilde{X}$.

Lemma 4.1.6. ***Pseudo-Markov property.*** *Let τ be a stopping time in $[0,T]$. We have that*

$$J(\tau,\tilde{X}_\tau,Y_\tau,\mathbb{P}^\tau,\alpha^\tau)=\mathbb{E}^{\mathbb{P}}\left[\int_\tau^T f(s,Y_s,\alpha_s)\,ds+g(Y_T)\mid\mathcal{F}_\tau\right],\quad \mathbb{P}-a.s.. \tag{4.1.7}$$

For a proof, see Appendix B.1.2.

Next, we check the measurability of the functions J, w, v in (4.1.6), we prove a Lemma that allows us to use measurable selections for both the supremum and the infimum problems in (4.1.6). First, we provide the following definition. For a Borel space H, we denote by $\mathcal{L}_H$ the smallest σ-field containing Borel subsets of H and closed under the Souslin operation.[2] See also the definitions of analytic function and universally measurable function in Appendices B.2.3 and B.2.4.

Below, Lemma 4.1.7 shows that there exists a universally measurable selector for the infimum problem in (4.1.6). In general, this is not true. If the function w is only upper semianalytic, there is no guarantee that there exists a universally measurable ϵ-optimal selector for the infimum problem; see Nowak (2010). The function w is $\mathcal{L}_X\otimes B_Y$-measurable because we assume that the performance criterion J is continuous in the variable y. Therefore, there exists a universally measurable ϵ-optimal selector for the infimum problem.

Now, we show the regularity of the performance criterion J with respect to all its variables and the regularity of the value function v with respect to the state (t,x,y).

Lemma 4.1.7. ***Regularity of the value function.*** *Let J, w, v be as in (4.1.6). The function J is Borel measurable from $[0,T]\times\mathbb{R}^{n+d}\times\mathbb{R}^{\tilde{d}}\times\mathfrak{P}(\Omega)\times\mathcal{A}_0$ to $\mathbb{R}$, and continuous with respect to t,x,y,α. The function w is upper semianalytic, and the value function v is universally measurable from $[0,T]\times\mathbb{R}^{n+d}\times\mathbb{R}^{\tilde{d}}$ to $\mathbb{R}$.*

For a proof, see Appendix B.1.3.

4.1.4 Time-consistency and the DPP of the adaptive robust problem.

Here, we show that the set of probability measures $\mathcal{P}(t,x)$ has the stability property under conditioning and the stability property under concatenation, both of which we employ to prove the time-consistency property of problem (4.1.6).

First, we state two Lemmas that we need for the proof of the time-consistency property for the problem in (4.1.6). The proofs of the following two Lemmas follow directly from Lemma 4.28 in Hollender (2016).

[2]See Definition 7.15 in Bertsekas and Tsitsiklis (1996).

Lemma 4.1.8. ***Stability under conditioning.*** *Let $\mathbb{P} \in \mathcal{P}(t,x)$ and let τ denote a stopping time taking values in $[t,T]$. For each $\bar{\omega} \in \Omega$, the probability measure $\mathbb{P}^{\tau,\bar{\omega}}$ is in the set $\mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))$ $\mathbb{P}$-a.s..*

Lemma 4.1.9. ***Stability under concatenation.*** *Let $\mathbb{P} \in \mathcal{P}(t,x)$ and let τ denote a stopping time taking values in $[t,T]$. Let ν be a Borel measurable kernel, such that $\nu : \Omega \to \mathfrak{P}(\Omega)$ and $\nu(\bar{\omega}) \in \mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))$. Then,*

$$\bar{\mathbb{P}}(A) = \int \int 1_{\{\bar{\omega} \otimes_{\tau(\bar{\omega})} \omega \in A\}} \nu(d\omega; \bar{\omega}) \, \mathbb{P}(d\bar{\omega})$$

is a probability measure and it is in the set $\mathcal{P}(t,x)$.

Now, we prove the time-consistency property of problem (4.1.6). This is our first main result.

Theorem 4.1.10. ***Time-consistency property.*** *Let $(t,x,y) \in [0,T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\tilde{d}}$, and let α be a control process in the admissible set $\mathcal{A}_0$. Let τ be a stopping time that takes values in the interval $[t,T]$. The following equation holds:*

$$\begin{aligned}
&\sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) \, ds + g(Y_T) \right] \\
&= \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) \, ds + \sup_{\mathbb{Q} \in \mathcal{P}(\tau, \tilde{X}_\tau)} \mathbb{E}^{\mathbb{Q}} \left[\int_\tau^T f(s, Y_s, \alpha_s^\tau) \, ds + g(Y_T) \right] \right].
\end{aligned}$$

Proof. Use Lemmas 4.1.5, 4.1.8, 4.1.9, and the arguments in Nutz and Handel (2013) to show time-consistency of the stochastic control problem (4.1.6). For each $\mathbb{P} \in \mathcal{P}(t,x)$, by conditional expectations, one obtains

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) \, ds + g(Y_T) \right] \\
&\quad = \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) \, ds + \mathbb{E}^{\mathbb{P}} \left[\int_\tau^T f(s, Y_s, \alpha_s) \, ds + g(Y_T) \mid \mathcal{F}_\tau \right] \right] \\
&\quad = \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) \, ds + \mathbb{E}^{\mathbb{P}^{\tau(\bar{\omega}),\bar{\omega}}} \left[\int_\tau^T f(s, Y_s, \alpha_s^\tau) \, ds + g(Y_T) \right] \right] \\
&\quad \le \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) \, ds + \sup_{\mathbb{Q} \in \mathcal{P}(\tau, \tilde{X}_\tau)} \mathbb{E}^{\mathbb{Q}} \left[\int_\tau^T f(s, Y_s, \alpha_s^\tau) \, ds + g(Y_T) \right] \right],
\end{aligned} \tag{4.1.8}$$

where the last inequality follows because $\mathbb{P}^{\tau,\bar{\omega}}$ is in the set $\mathcal{P}(\tau, \tilde{X}_\tau)$ for all $\bar{\omega}$ $\mathbb{P}$-a.s.; see Lemma 4.1.8. Now, take the supremum on both sides of (4.1.8) and write

$$\sup_{\mathbb{P}\in\mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}}\left[\int_t^T f(s, Y_s, \alpha_s)\, ds + g(Y_T)\right]$$

$$\leq \sup_{\mathbb{P}\in\mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s, Y_s, \alpha_s)\, ds + \sup_{\mathbb{Q}\in\mathcal{P}(\tau,\tilde{X}_\tau)} \mathbb{E}^{\mathbb{Q}}\left[\int_\tau^T f(s, Y_s, \alpha_s^\tau)\, ds + g(Y_T)\right]\right]. \tag{4.1.9}$$

Next, we prove the reverse inequality of (4.1.9). For any stopping time τ and $\alpha \in \mathcal{A}_0$, write the function w as

$$\begin{aligned}
&w\left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha^{\tau(\bar{\omega}),\bar{\omega}}\right)\\
&= \sup_{\mathbb{Q}\in\mathcal{P}(\tau(\bar{\omega}),\tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))} J\left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \mathbb{Q}, \alpha^{\tau(\bar{\omega}),\bar{\omega}}\right)\\
&= \sup_{\mathbb{Q}\in\mathcal{P}(\tau(\bar{\omega}),\tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))} \hat{J}(\bar{\omega}, \mathbb{Q})\,.
\end{aligned}$$

From the proof of Lemma 3.1 in Neufeld and Nutz (2014), the function $\hat{J}$ is Borel measurable. From Lemma 4.1.5, the set $\{(\bar{\omega}, \mathbb{Q}) \mid \mathbb{Q} \in \mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))\}$ is analytic. Therefore, by a measurable selection theorem, there exists a universally measurable ν^ϵ from Ω to $\mathfrak{P}(\Omega)$, such that

$$w\left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha^{\tau(\bar{\omega}),\bar{\omega}}\right) < \epsilon + J\left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \nu^\epsilon(\bar{\omega}), \alpha^{\tau(\bar{\omega}),\bar{\omega}}\right)$$

for all $\bar{\omega} \in \Omega$. From Lemma 7.27 in Bertsekas and Shreve (1996), we have that for any probability measure $\mathbb{P} \in \mathcal{P}(t, x)$, there is a Borel measurable function $\nu^{\epsilon,\mathbb{P}}$, such that $\nu^{\epsilon,\mathbb{P}}(\bar{\omega}) = \nu^\epsilon(\bar{\omega})$ for $\mathbb{P}$- a.s.. Then,

$$w\left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha^{\tau(\bar{\omega}),\bar{\omega}}\right) < \epsilon + J\left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \nu^{\epsilon,\mathbb{P}}(\bar{\omega}), \alpha^{\tau(\bar{\omega}),\bar{\omega}}\right),$$

$\mathbb{P}$ almost surely. Therefore, we have the following inequality:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s, Y_s, \alpha_s)\, ds + \sup_{\mathbb{Q}\in\mathcal{P}(\tau,\tilde{X}_\tau)} \mathbb{E}^{\mathbb{Q}}\left[\int_\tau^T f(s, Y_s, \alpha_s^\tau)\, ds + g(Y_T)\right]\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s, Y_s, \alpha_s)\, ds + w\left(\tau, \tilde{X}_\tau, Y_\tau, \alpha^\tau\right)\right]\\
&< \mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s, Y_s, \alpha_s)\, ds + J\left(\tau, \tilde{X}_\tau, Y_\tau, \nu^{\epsilon,\mathbb{P}}, \alpha^\tau\right)\right] + \epsilon\\
&= \mathbb{E}^{\mathbb{P}^{\nu^\epsilon}}\left[\int_t^T f(s, Y_s, \alpha_s)\, ds\right] + \epsilon\,,
\end{aligned}$$

where $\mathbb{P}^{\nu^\epsilon}(A) := \int\int 1_{\{\bar{\omega}\otimes_{\tau(\bar{\omega})}\omega\in A\}}\nu^{\epsilon,\mathbb{P}}(d\omega;\bar{\omega})\,\mathbb{P}(d\bar{\omega})$ is in the set $\mathcal{P}(t,x)$; see Lemma 4.1.9. $\square$

Our proof of time-consistency is standard, but the result of Theorem 4.1.10 is not immediately available in the literature. For example, Nutz and Handel (2013) and Bayraktar and Yao (2014) look at the time-consistency property of problems in robust optimal control and robust stopping; however, their results cannot be applied directly in our framework to show time consistency of the problem we study.

Next, we prove the DPP and write the value function recursively. The proof relies on the time-consistency property of the adaptive robust problem in (4.1.6), which we showed in Theorem 4.1.10. We use the measurable selection theorem in Soner and Touzi (2002) because the set of admissible controls $\mathcal{A}_0$ is a separable metric space, we recall that $\mathcal{A}_0$ consists of all progressively measurable processes that take values in a compact set $A \in R^k$.

Theorem 4.1.11. ***Dynamic programming principle.*** *Let $(t,x,y) \in [0,T]\times\mathbb{R}^{n+d}\times\mathbb{R}^{\tilde{d}}$, let τ be a stopping time taking values in $[t,T]$, $\tilde{X}_t = x$ and $Y_t = y$. The value function of the adaptive robust problem in (4.1.6) satisfies*

$$v(t,x,y) = \inf_{\alpha\in\mathcal{A}_0}\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s,Y_s,\alpha_s)\,ds + v(\tau,\tilde{X}_\tau,Y_\tau)\right]. \tag{4.1.10}$$

Proof. Recall that $w(t,x,y,\alpha) := \sup_{\mathbb{P}\in\mathcal{P}(t,x)} J(t,x,y,\mathbb{P},\alpha)$. From the definition of the performance criterion J, we have that

$$\begin{aligned}\sup_{\mathbb{P}\in\mathcal{P}(t,x)} J(t,x,y,\mathbb{P},\alpha) &= \sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_t^T f(s,Y_s,\alpha_s)\,ds + g(Y_T)\right]\\ &= \sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s,Y_s,\alpha_s)\,ds + w(\tau,\tilde{X}_\tau,Y_\tau,\alpha^\tau)\right]\\ &\geq \sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s,Y_s,\alpha_s)\,ds + v(\tau,\tilde{X}_\tau,Y_\tau)\right],\end{aligned} \tag{4.1.11}$$

where the second line in (4.1.11) follows from Theorem 4.1.10 and the last inequality follows from the definition of the value function v. Take infimum over $\alpha\in\mathcal{A}_0$ in (4.1.11) and write

$$v(t,x,y) \geq \inf_{\alpha\in\mathcal{A}_0}\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_t^\tau f(s,Y_s,\alpha_s)\,ds + v(\tau,\tilde{X}_\tau,Y_\tau)\right]. \tag{4.1.12}$$

Now we show the reverse inequality in (4.1.12). Write the value function v as follows:

$$v(\tau(\bar{\omega}),\tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}),Y_{\tau(\bar{\omega})}(\bar{\omega})) = \inf_{\alpha\in\mathcal{A}_0} w(\tau(\bar{\omega}),\tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}),Y_{\tau(\bar{\omega})}(\bar{\omega}),\alpha) = \inf_{\alpha\in\mathcal{A}_0}\hat{w}(\bar{\omega},\alpha),$$

where $\hat{w}(\bar{\omega}, \alpha) := w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha)$. From Lemma 4.1.7, the function w is $\mathcal{B}([0,T]) \otimes \mathcal{L}(\mathbb{R}^{n+d}) \otimes \mathcal{B}(\mathbb{R}^{\tilde{d}}) \otimes \mathcal{B}_{\mathcal{A}_0}$-measurable. Then, the function $\hat{w}$ is $\mathcal{L}_\Omega \otimes \mathcal{B}_{\mathcal{A}_0}$-measurable because the set $[0,T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\tilde{d}} \times \mathcal{A}_0$ is Borel and the mappings $\bar{\omega} \mapsto \tau(\bar{\omega})$, $\bar{\omega} \mapsto \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega})$, and $\bar{\omega} \mapsto \tilde{Y}_{\tau(\bar{\omega})}(\bar{\omega})$ are Borel measurable. Therefore, from Lemma 4.1.7, there exists a universally measurable selection, such that $\varphi^\epsilon : (\Omega, \mathcal{F}) \to (\mathcal{A}_0, \mathcal{B}_{\mathcal{A}_0})$ and

$$v(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega})) + \epsilon > w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \varphi^\epsilon(\bar{\omega}))\,.$$

There exists a Borel measurable function $\varphi^{\epsilon,\mathbb{P}^*}$ such that $\varphi^{\epsilon,\mathbb{P}^*}(\bar{\omega}) = \varphi^\epsilon(\bar{\omega})$ $\mathbb{P}^*$-a.s. for each probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F})$ in the set $\mathcal{P}(t,x)$; see Lemma 7.27 in Bertsekas and Shreve (1996). For a control process $\alpha \in \mathcal{A}_0$, we define the new control process

$$\alpha^{\epsilon,\mathbb{P}^*}(t, \bar{\omega}) = \alpha(t, \bar{\omega})\, 1_{\tau(\bar{\omega}) > t} + \varphi^{\epsilon,\mathbb{P}^*}(\bar{\omega})\, 1_{\tau(\bar{\omega}) \leq t}\,. \tag{4.1.13}$$

Now, we check the measurability of the process $\alpha^{\epsilon,\mathbb{P}^*}$. The process $\alpha^{\epsilon,\mathbb{P}^*}$ is progressively measurable because $\mathcal{A}_0$ is a separable metric space and from Lemma 2.1 in Soner and Touzi (2002). By elementary calculations, the process $\alpha^{\epsilon,\mathbb{P}^*}$ is in the set $\mathcal{A}_0$. Let τ be a stopping time in the interval $[t,T]$. By Theorem 4.1.10, we have that

$$\begin{aligned} w\left(t, x, y, \alpha^{\epsilon,\mathbb{P}^*}\right) &= \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f\left(s, Y_s, \alpha_s^{\epsilon,\mathbb{P}^*}\right) ds + w\left(\tau, \tilde{X}_\tau, Y_\tau, (\alpha^{\epsilon,\mathbb{P}^*})^\tau\right) \right] \\ &= \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f\left(s, Y_s, \alpha_s\right) ds + w\left(\tau, \tilde{X}_\tau, Y_\tau, (\alpha^{\epsilon,\mathbb{P}^*})^\tau\right) \right] \\ &\leq \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s)\, ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right] + \epsilon\,, \end{aligned} \tag{4.1.14}$$

where the last inequality follows from

$$v(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega})) + \epsilon > w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \varphi^{\epsilon,\mathbb{P}^*}(\bar{\omega}))\,,$$

for $\mathbb{P}^*$ -a.s., and because all probability measures in the set $\mathcal{P}(t,x)$ are absolutely continuous with respect to the probability measure $\mathbb{P}^*$. Therefore, for an arbitrary control process $\alpha \in \mathcal{A}_0$, the value function v obeys the bound

$$v(t,x,y) \leq \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s)\, ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right] + \epsilon\,.$$

This shows the reverse inequality of (4.1.12), which completes the proof. □

4.1.5 Viscosity characterisation of the value function

In what follows, assume that Assumptions 2.3 and 2.4 hold. We show that the value function in (4.1.6) is the unique solution of a non-linear PDE. We recall that the processes $\tilde{X}$ and Y follow (4.0.2) and (4.0.3), respectively. Henceforth, the process $\tilde{Y} := [\tilde{X}, Y]^\intercal$ follows

$$d\tilde{Y}_t = \bar{\boldsymbol{b}}(t, \tilde{Y}_t, \tilde{\theta}_t, \alpha_t)\, dt + \bar{\boldsymbol{\sigma}}(t, \tilde{Y}_t, \alpha_t)\, dB_t^{\mathbb{P}_{\tilde{\theta}}} + \bar{\boldsymbol{\xi}}(t, \tilde{Y}_t, \alpha_t)\, dL_t\,, \tag{4.1.15}$$

where $\bar{\boldsymbol{b}}(t, \tilde{y}, \theta, a) := [\, \boldsymbol{b}(t, \tilde{x}, \theta),\, \boldsymbol{b}(t, \tilde{x}, \theta)\, \bar{\sigma}(\tilde{y}, a) + \bar{b}(\tilde{y}, a)\,]^\intercal$
, $\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) := [\, \boldsymbol{\sigma}(t, \tilde{x}),\, \boldsymbol{\sigma}(t, \tilde{x})\, \bar{\sigma}(\tilde{y}, a)\,]^\top$, $\bar{\boldsymbol{\xi}}(t, \tilde{y}, a) := [\, \boldsymbol{\xi}(t, \tilde{x}),\, \boldsymbol{\xi}(t, \tilde{x})\, \bar{\sigma}(\tilde{y}, a)\,]^\intercal$ and $\tilde{y} = (y, \tilde{x})$. Denote by $|x|$ the Euclidean norm of $x \in \mathbb{R}^m$, where m is a positive integer.

Assumption 4.1.12. The functions $\bar{\boldsymbol{b}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\xi}}$ are continuous with respect to a and the function f is Lipschitz with respect to a.

Assumption 4.1.13. For any compact set $\boldsymbol{\Theta} \subset \mathbb{R}^d$ and $\tilde{\boldsymbol{Y}} \subset \mathbb{R}^{n+d+\tilde{d}}$, there exists a positive constant K_L such that for all $a \in A$

$$\begin{aligned} \left|\bar{\boldsymbol{b}}(t, \tilde{y}_1, \tilde{\theta}_1, a) - \bar{\boldsymbol{b}}(t, \tilde{y}_2, \tilde{\theta}_1, a)\right| + \left|\bar{\boldsymbol{\sigma}}(t, \tilde{y}_1, a) - \bar{\boldsymbol{b}}(t, \tilde{y}_2, a)\right| + \left|\bar{\boldsymbol{\xi}}(t, \tilde{y}_1, a) - \bar{\boldsymbol{\xi}}(t, \tilde{y}_2, a)\right| \\ \leq K_L \left|\tilde{y}_1 - \tilde{y}_2\right|\,, \end{aligned} \tag{4.1.16}$$

for all $t \in [0, T]$, $\tilde{y}_1, \tilde{y}_2 \in \mathbb{R}^{1+d+\tilde{d}}$, and $\tilde{\theta} \in \boldsymbol{\Theta}$ and

$$\left|\bar{\boldsymbol{b}}(t, \tilde{y}_1, \tilde{\theta}_1, a) - \bar{\boldsymbol{b}}(t, \tilde{y}_1, \tilde{\theta}_2, a)\right| \leq K_L \left|\tilde{\theta}_1 - \tilde{\theta}_2\right|\,, \tag{4.1.17}$$

for all $t \in [0, T]$, $\tilde{y}_1 \in \tilde{\boldsymbol{Y}}$ and $\tilde{\theta}_1$, $\tilde{\theta}_2 \in \mathbb{R}^d$.

The two assumptions above are essential in the proof of the characterisation of the value function v. The assumptions below are to ensure the set G is not too large which allows us to prove the stability of moment lemma and to guarantee that the solutions to the SDEs (4.0.2) and (4.0.3) are unique.

Assumption 4.1.14. There exists a positive real number K_G such that for $\theta_1, \theta_2 \in G(t, x)$, we have that $|\theta_1 - \theta_2| \leq K_G\, |x|$.

The assumption on the size of the set G is needed in the Lemma below. Note that if the set G is too large (recall that $G(t, x) \in 2^{\mathbb{R}^d}$), then

$$\sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}}\Big[\sup_{0 \leq s \leq \eta} \left|\tilde{Y}_{t+s} - \tilde{y}\right|^q \Big] \tag{4.1.18}$$

does not converge to zero quickly enough as the value of η goes to zero, in which case the value function v in (4.1.6) cannot be characterised as a solution of a non-linear PDE.

Now, we use Lemma 3.4 in Fadina, Neufeld, and Schmidt (2018) to prove that the q-moment of the process $\tilde{Y}$ is locally stable under the non-linear expectation in (4.1.18).

Lemma 4.1.15. ***Stability of moment.*** *If Assumptions 4.1.12, 4.1.13, 4.1.14 hold, then for $\alpha_t \in \mathcal{A}_0$, for $q \geq 1$ and $(t, \tilde{y}) \in [0,T] \times \mathbb{R}^{n+d+\tilde{d}}$, there exists $\epsilon(q) \in (0,1)$ such that for all $\eta < \epsilon(q)$ we have*

$$\sup_{\mathbb{P}\in\mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\sup_{0\leq s\leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q \right] \leq C \left(\eta^q + \eta^{q/2}\right), \tag{4.1.19}$$

where the constant C is independent of t and h, and the stopping time $\tau_\eta := \inf\{u : |\tilde{Y}_{t+u} - \tilde{y}| \geq 1\} \wedge \eta$.

Proof. Recall Hölder's inequality

$$(a+b+c+d)^q \leq c_q \left(a^q + b^q + c^q + d^q\right), \tag{4.1.20}$$

where $q \geq 1$ and c_q is the smallest real number such that the inequality holds for all $a, b, c, d \geq 0$. Note that the value of the constant c_q depends only on the parameter q. For $\mathbb{P} = \mathbb{P}_{\tilde{\theta}} \in \mathcal{P}(t,x)$, the process $\tilde{Y}$ has the representation

$$\begin{aligned} \tilde{Y}_{t+s} &= \tilde{y} + \int_t^{t+s} \bar{\boldsymbol{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u)\, du + \int_t^{t+s} \bar{\boldsymbol{\sigma}}(u, \tilde{Y}_u, \alpha_u)\, dB_u^{\mathbb{P}_{\tilde{\theta}}} + \int_t^{t+s} \bar{\boldsymbol{\xi}}(u, \tilde{Y}_u, \alpha_u)\, dL_u\,, \\ &= \tilde{y} + \int_t^{t+s} \bar{\boldsymbol{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u)\, du + \int_t^{t+s} \int_{\mathbb{R}} \bar{\boldsymbol{\xi}}(u, \tilde{Y}_u, \alpha_u)\, z\, \nu(dz)\, du + \tilde{Y}^c_{t+s} + \tilde{Y}^d_{t+s}\,, \end{aligned} \tag{4.1.21}$$

where $\tilde{Y}^{c,\mathbb{P}}$ and $\tilde{Y}^{d,\mathbb{P}}$ are the continuous and discontinuous local martingale parts of

$\tilde{Y}$, respectively, under the probability measure $\mathbb{P}$. Use (4.1.20) to write

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}_{t+s}-\tilde{y}\right|^q\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\int_t^{t+s}\bar{\boldsymbol{b}}(u,\tilde{Y}_u,\tilde{\theta}_u,\alpha_u)\,du+\int_t^{t+s}\int_{\mathbb{R}}\bar{\boldsymbol{\xi}}(u,\tilde{Y}_u,\alpha_u)\,z\,\nu(dz)\,du+\tilde{Y}^c_{t+s}+\tilde{Y}^d_{t+s}\right|^q\right] \\
&\le c_q\left(\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\int_t^{t+s}\bar{\boldsymbol{b}}(u,\tilde{Y}_u,\tilde{\theta}_u,\alpha_u)\,du\right|^q\right]+\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}^{c,\mathbb{P}}_{t+s}\right|^q\right]+\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}^{d,\mathbb{P}}_{t+s}\right|^q\right]\right. \\
&\quad\left.+\,\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\int_t^{t+s}\int_{\mathbb{R}}\bar{\boldsymbol{\xi}}(u,\tilde{Y}_u,\alpha_u)\,z\,\nu(dz)\,du\right|^q\right]\right).
\end{aligned} \tag{4.1.22}$$

Let $\tilde{\boldsymbol{Y}} := \{y\in\mathbb{R}^{n+d+\tilde{d}}\,|\,|y-\tilde{y}|\le 1\}$ and $\tilde{\boldsymbol{\Theta}} := \bigcup_{\boldsymbol{y}\in\tilde{Y}}G(t,x)$. The sets $\tilde{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{\Theta}}$ are compact because the mapping $y\mapsto G(t,x)$ is continuous. From Assumption 4.1.13, there exists a positive constant $\tilde{K}_L$ that depends on $\tilde{y}$ such that $\bar{\boldsymbol{b}}(t,\tilde{y}_1,\tilde{\theta},a)\le\tilde{K}_L\,(1+|\tilde{y}_1|)$, $\bar{\boldsymbol{\sigma}}(t,\tilde{y}_1,a)\le\tilde{K}_L\,(1+|\tilde{y}_1|)$ and $\bar{\boldsymbol{\xi}}(t,\tilde{y}_1,a)\le\tilde{K}_L\,(1+|\tilde{y}_1|)$ for all $\tilde{y}_1\in\tilde{\boldsymbol{Y}}$, $\tilde{\theta}\in\tilde{\boldsymbol{\Theta}}$, and $a\in A$. Define $\hat{K}:=\max\{1,K_G,\tilde{K}_L\}$, where K_G is the constant from Assumption 4.1.14. Consider the first term on the right-hand side in (4.1.22).

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\left(\int_t^{t+\tau_\eta}\left|\bar{\boldsymbol{b}}(u,\tilde{Y}_u,\tilde{\theta}_u,\alpha_u)\right|\,du\right)^q\right]\le \\
&\quad\le\mathbb{E}^{\mathbb{P}}\left[\left(\int_t^{t+\tau_\eta}\left|\bar{\boldsymbol{b}}(u,\tilde{Y}_u,\hat{\theta}_u,\alpha_u)\right|+\hat{K}\left|\hat{\theta}_u-\tilde{\theta}_u\right|\,du\right)^q\right] && \text{from (4.1.17)} \\
&\quad\le\hat{K}^q\,\eta^q\,\mathbb{E}^{\mathbb{P}}\left[\left(1+\hat{K}\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}_{t+s}\right|\right)^q\right] && \text{from (4.1.14)} \\
&\quad\le\tilde{c}_q\,\hat{K}^{2q}\,\eta^q\,\mathbb{E}^{\mathbb{P}}\left[1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}-\tilde{y}|^q+|\tilde{y}|^q\right], && \text{from (4.1.20)}
\end{aligned} \tag{4.1.23}$$

where the second inequality follows from the Lipschitz assumption on $\bar{\boldsymbol{b}}$, the penultimate inequality follows from the linear growth property on $\bar{\boldsymbol{b}}$, and the last inequality follows from Hölder's inequality in (4.1.20). Next, consider the last term in (4.1.22)

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_h}\left|\int_t^{t+s}\int_{\mathbb{R}}\bar{\boldsymbol{\xi}}(u,\tilde{Y}_u,\alpha_u)\,z\,\nu(dz)\,du\right|^q\right]\\
&\le \mathbb{E}^{\mathbb{P}}\left[\left|\int_t^{t+\tau_\eta}\int_{\mathbb{R}}\bar{\boldsymbol{\xi}}(u,\tilde{Y}_u,\alpha_u)\,z\,\nu(dz)\,du\right|^q\right]\\
&\le \mathbb{E}^{\mathbb{P}}\left[\left\{\int_t^{t+\tau_\eta}\int_{\mathbb{R}}\left(1+\hat{K}\left|\tilde{Y}_{t+u}\right|\right)|z|\;\nu(dz)\,du\right\}^q\right] \qquad (4.1.24)\\
&\le C_J\,\eta^q\,\mathbb{E}^{\mathbb{P}}\left[\left(1+\hat{K}\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}_{t+s}\right|\right)^q\right]\\
&\le \tilde{c}_q\,C_J\,\hat{K}^q\,\eta^q\,\mathbb{E}^{\mathbb{P}}\left[1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}-\tilde{y}|^q+|\tilde{y}|^q\right],
\end{aligned}$$

where the second inequality follows from the linear growth property on $\bar{\boldsymbol{\xi}}$, and the last inequality follows from Hölder's inequality in (4.1.20). By the Burkholder-Davis-Gundy (BDG) inequality in Cohen and Elliott (2015), we have that

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}^{c,\mathbb{P}}_{t+s}\right|^q\right] &\le \tilde{c}_q\,\mathbb{E}^{\mathbb{P}}\left[\left(\int_t^{t+\tau_\eta}\left|\bar{\boldsymbol{\sigma}}(t+u,\tilde{Y}_{t+u},\alpha_{t+u})\,\bar{\boldsymbol{\sigma}}(t+u,\tilde{Y}_{t+u},\alpha_{t+u})^{\mathsf{T}}\right|\,du\right)^{q/2}\right]\\
&\le \tilde{c}_q\,\mathbb{E}^{\mathbb{P}}\left[\hat{K}^q\,\eta^{q/2}\,(1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}|)^q\right]\\
&\le \tilde{c}_q\,c_q\,\hat{K}^q\,\eta^{q/2}\,\mathbb{E}^{\mathbb{P}}\left[1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}-\tilde{y}|^q+|\tilde{y}|^q\right],
\end{aligned} \qquad (4.1.25)$$

where $\tilde{c}_q$ is a constant that depends on q, the last inequality follows from (4.1.20), and $\hat{K}$ does not depend on h. Next, consider the purely discontinuous local martingale, the BDG inequality and the linear growth of the function $\bar{\boldsymbol{\xi}}$ imply

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}^{d,\mathbb{P}}_{t+s}\right|^q\right] &\le \tilde{c}_q\,\mathbb{E}^{\mathbb{P}}\left[\left\langle \tilde{Y}^{d,\mathbb{P}}_{t+\tau_\eta}\right\rangle^{q/2}\right], \\
&\le \tilde{c}_q\,\mathbb{E}^{\mathbb{P}}\left[\left|\int_t^{t+\tau_\eta}\int_{\mathbb{R}}\left|\bar{\boldsymbol{\xi}}(u,\tilde{Y}_u,\alpha_u)\right|^2|z|^2\,\nu(dz)\,du\right|^{q/2}\right] \\
&\le \tilde{c}_q\,C_J\,\mathbb{E}^{\mathbb{P}}\left[\hat{K}^q\,\eta^{q/2}\left(1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}|\right)^q\right] \\
&\le \tilde{c}_q\,c_q\,C_J\,\hat{K}^q\,\eta^{q/2}\,\mathbb{E}^{\mathbb{P}}\left[1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}-\tilde{y}|^q+|\tilde{y}|^q\right].
\end{aligned} \tag{4.1.26}$$

Choose the value of the parameter ϵ, so that $1-C_1\,\eta^q-C_2\,\eta^{q/2}:=1-(\tilde{c}_q\,\hat{K}^{2q}\,h^q+\tilde{c}_q\,C_J\,\hat{K}^q\,\eta^q+\tilde{c}_q\,c_q\,\hat{K}^q\,\eta^{q/2}+\tilde{c}_q\,c_q\,C_J\,\hat{K}^q\,\eta^{q/2})>0$. Therefore, if $\eta<\epsilon$, from inequality (4.1.23) and (4.1.25), we have that

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}\left|\tilde{Y}_{t+s}-\tilde{y}\right|^q\right] &\le \left(C_1\,\eta^q+C_2\,\eta^{q/2}\right)\,\mathbb{E}^{\mathbb{P}}\left[1+\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}-\tilde{y}|^q+|\tilde{y}|^q\right] \\
&\le \left(C_1\,\eta^q+C_2\,\eta^{q/2}\right)\,\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_\eta}|\tilde{Y}_{t+s}-\tilde{y}|^q\right] \\
&\quad+\left(C_1\,\epsilon^q+C_2\,\epsilon^{q/2}\right)\,[1+|\tilde{y}|^q]\,.
\end{aligned}$$

Let

$$C=\frac{\left(C_1\,\eta^q+C_2\,\eta^{q/2}\right)\,(1+|\tilde{y}|^q)}{1-\left(C_1\,\epsilon^q+C_2\,\epsilon^{q/2}\right)}\,.$$

Now, for $\mathbb{P}\in\mathcal{P}(t,x,G)$ the following inequality holds:

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s\le\tau_n}\left|\tilde{Y}_{t+s}-\tilde{y}\right|^q\right]\le C\,\left(\eta^{q/2}+\eta^q\right)\,. \tag{4.1.27}$$

□

We require that the set G satisfies the following structural assumption for the proof of the characterisation of the value function as a solution of a non-linear PDE.

Assumption 4.1.16. We assume that $G(t,x)$ is a closed set, in the standard topology, for all $(t,x)\in[0,T]\times\mathbb{R}^{n+d}$. And we assume that the Hausdorff distance between the sets $G(t,x)$ and $G(s,y)$ has a Lipschitz-like property. That is, there exists a constant K_H such that $d_{\text{Haus}}(G(t,x),G(s,y))\le K_H\,(|t-s|+|x-y|\,)$ for all $(t,x),(s,y)\in[0,T]\times\mathbb{R}^{n+d}$.

The assumption above is essential to characterise the value function as a solution of a non-linear PDE, but not to prove the DPP of the value function. We assume that $G(t,x)$ is a closed set, so the maximal is attained in the set $G(t,x)$. The second statement of the assumption requires the function G to be well-behaved. This condition allows us to quantify the non-linear operator that depends on the set $G(t,x)$.

Denote by $\varphi \in C_b^{2,3}([0,T]\times\mathbb{R}^{n+d+\tilde{d}})$ the set of functions on $[0,T]\times\mathbb{R}^{n+d+\tilde{d}}$ that have bounded continuous derivatives up to the second and third order in t and x, respectively. We use the next Lemma to characterise the value function as a solution of a non-linear PDE.

Lemma 4.1.17. *Let $\varphi \in C_b^{2,3}([0,T]\times\mathbb{R}^{n+d+\tilde{d}})$ and let u be a small enough positive real number. If Assumptions 4.1.12, 4.1.13, 4.1.14, 4.1.16 hold, then there exists a probability measure $\mathbb{P}\in\mathcal{P}(t,x)$ such that for all $\alpha\in\mathcal{A}_0$,*

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\tilde{\theta}_{t+\tau_s},\alpha_{t+\tau_s})\\
&\quad+\tfrac{1}{2}\,\mathrm{Tr}\Big(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})^\intercal\,\partial^2_{\tilde{y}\tilde{y}}\varphi(t,\tilde{y})\Big)\\
&\quad+\int_{\mathbb{R}}(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,z)-\varphi(t,\tilde{y})-\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,h(z))\,\nu(dz)\Big]\\
&\geq -C_A\,s^{1/2}+\inf_{a\in A}\sup_{\tilde{\theta}\in G(t,x)}\Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a)\,+\tfrac{1}{2}\,\mathrm{Tr}\Big(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^\intercal\,\partial^2_{\tilde{y}\tilde{y}}\varphi(t,\tilde{y})\Big)\\
&\quad+\int_{\mathbb{R}}(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z)-\varphi(t,\tilde{y})-\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z))\,\nu(dz)\Big],
\end{aligned}
\tag{4.1.28}
$$

where the stopping times $\tau_s := \inf\{m : |\tilde{Y}_{t+m}-\tilde{y}|\geq 1\}\wedge s$ for all $s\in[0,u]$, and C_A is a constant that depends on t and x.

Proof. From Assumption 4.1.16, there exists an optimal a^* that minimises the supremum term and there exists an optimal $\tilde{\theta}^{a_*}$ that maximises the right-hand side of (4.1.28). The process $\gamma_{t+s}=\mathrm{proj}_{\tilde{\theta}^{a_*}}(G(t+s,\tilde{X}_{t+s}))$ is $\mathcal{F}_{t+s}$-adapted, where $\mathrm{proj}_x(A)$ is an element in the set A such that the Euclidean distance between the element and x is smallest. The process γ is right-continuous because the process $\tilde{X}$ is right-continuous and due to the Lipschitz property of the mapping G; see Assumption 4.1.16. Therefore, the process γ is progressively measurable. Let $\mathbb{P}^\gamma$ be a measure such that

$$
\beta^{\mathbb{P}^\gamma}_{t+s}=\boldsymbol{b}(t+s,\tilde{X}_{t+s},\gamma_{t+s})+\int_{\mathbb{R}}(\hat{h}(\boldsymbol{\xi}(t+s,\tilde{X}_{t+s})\,z)-\boldsymbol{\xi}(t+s,\tilde{X}_{t+s})\,h(z))\,\nu(dz)\,.
$$

It is easy to see that $\mathbb{P}^\gamma\in\mathcal{P}(t,x)$ because of the way we construct the process γ.

Next, we show that $\mathbb{P}^\gamma$ satisfies (4.1.28). First, note that

$$|\gamma_{t+s} - \tilde{\theta}^{a_*}| \leq d_{\text{Haus}}(G(t+s, \tilde{X}_{t+s}), G(t,x)) \leq K_H\,(\,s + |\tilde{X}_{t+s} - x|\,)\,. \tag{4.1.29}$$

Let $C_1 = \max\{\left|\,\partial_x\varphi(t,x)\,\right|, \left|\,\partial^2_{xx}\varphi(t,x)\,\right|\}$. Consider

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^\gamma}\left[\left|\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \gamma_{t+\tau_s}, \alpha_{t+\tau_s}) - \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta}^{a_*},\alpha_{t+\tau_s})\right|\right]\\
&\quad\leq C_1\,\mathbb{E}^{\mathbb{P}^\gamma}\left[\left|\bar{\boldsymbol{b}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \gamma_{t+\tau_s}, \alpha_{t+\tau_s}) - \bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta}^{a_*},\alpha_{t+\tau_s})\right|\right]\\
&\quad\leq C_1\,(K_L + K_H)\,\mathbb{E}^{\mathbb{P}^\gamma}\left[\left|\tilde{Y}_{t+\tau_s} - \tilde{y}\right| + \tau_s\right]\\
&\quad\leq C_2\,s^{1/2}\,,
\end{aligned}$$

where the second line follows from the bounded derivative of φ, the third line is a result of (4.1.29) and Assumptions 4.1.13, 4.1.14, 4.1.16, and the last inequality follows from Lemma 4.1.15 with $q = 1$. Moreover, C_1, C_2 are constants that depend on t and $\tilde{y}$. Next, consider

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^\gamma}\Bigg[\frac{1}{2}\,\bigg|\operatorname{Tr}\Big(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})^\intercal\,\partial^2_{\tilde{y}\tilde{y}}\varphi\Big)\\
&\qquad-\;\operatorname{Tr}\Big(\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})^\intercal\,\partial^2_{\tilde{y}\tilde{y}}\varphi\Big)\bigg|\Bigg]\\
&\leq \mathbb{E}^{\mathbb{P}^\gamma}\Bigg[\frac{1}{2}\left|\partial^2_{\tilde{y}\tilde{y}}\varphi(t,\tilde{y})\right|\,\bigg|\operatorname{Tr}\Big(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})^\intercal\Big)\\
&\qquad-\;\operatorname{Tr}\big(\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})^\intercal\big)\bigg|\Bigg]\\
&\leq C_1\,\mathbb{E}^{\mathbb{P}^\gamma}\Bigg[\bigg|\operatorname{Tr}\Big(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})^\intercal\Big)\\
&\qquad-\;\operatorname{Tr}\big(\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})^\intercal\big)\bigg|\Bigg]\,,
\end{aligned}\tag{4.1.30}$$

where the second line follows from the bounded derivative of the function φ. From Assumption 4.1.12, we have that $|\bar{\boldsymbol{\sigma}}(t+s,\tilde{Y}_{t+s},a) - \bar{\boldsymbol{\sigma}}(t,\tilde{y},a)| \leq K_L\left(s + |\tilde{Y}_{t+s} - \tilde{y}|\right)$ for all $a \in A$, therefore, the right-hand side of (4.1.30) obeys the bound

$$\begin{aligned}
&C_1\mathbb{E}^{\mathbb{P}^\gamma}\left[\left|\operatorname{Tr}\Big(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})^\intercal\Big) - \operatorname{Tr}\big(\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})^\intercal\big)\right|\right]\\
&\quad\leq C_1\,\max\{2\,K_L, K_L^2\}\,\mathbb{E}^{\mathbb{P}^\gamma}\left[\left|\bar{\boldsymbol{\sigma}}(t,\tilde{y},\alpha_{t+\tau_s})\right|\,\left(\tau_s + |\tilde{Y}_{t+\tau_s} - \tilde{y}|\right) + \left(\tau_s + |\tilde{Y}_{t+\tau_s} - \tilde{y}|\right)^2\right]
\end{aligned}$$

$$\leq C_A\, s^{1/2},$$

where the last inequality follows from Lemma 4.1.15 with $q = 1$ and C_A is a constant that depends on t and $\tilde{y}$. For the remaining term on the left-hand side of (4.1.28), we have that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\Bigg[\int_{\mathbb{R}} \Big(\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})\, z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}}\varphi(t, \tilde{y})\, \bar{\boldsymbol{\xi}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})\, h(z)\Big)\ \nu(dz) \\
&- \int_{\mathbb{R}} \big(\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, \alpha_{t+\tau_s})\, z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}}\varphi(t, \tilde{y})\, \bar{\boldsymbol{\xi}}(t, \tilde{y}, \alpha_{t+\tau_s})\, h(z)\big)\ \nu(dz)\Bigg] \\
&\leq C\, \mathbb{E}^{\mathbb{P}}\Bigg[\int_{\mathbb{R}} \Big(\Big|z\, (\bar{\boldsymbol{\xi}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) - \bar{\boldsymbol{\xi}}(t, \tilde{y}, \alpha_{t+\tau_s}))\Big| \\
&\quad + \Big|\Big(\bar{\boldsymbol{\xi}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) - \bar{\boldsymbol{\xi}}(t, \tilde{y}, \alpha_{t+\tau_s})\Big)\Big|\, \big|h(z)\big|\Big)\ \nu(dz)\Bigg] \\
&\leq C\, K_L\, \mathbb{E}^{\mathbb{P}}\Big[\tau_s + \Big|\tilde{Y}_{t+\tau_s} - \tilde{y}\Big|\Big] \leq C_J\, s^{1/2}\,,
\end{aligned}$$

where the first inequality follows from the property of the function φ, the second inequality follows from (4.1.16) and C_J is a constant that depends on t and $\tilde{y}$.

□

Next, we show that the value function v in (4.1.6) is a viscosity solution of a Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation.

Theorem 4.1.18. ***Viscosity solution.*** *If Assumptions 4.1.12, 4.1.13, 4.1.14, 4.1.16 hold and the value function v in (4.1.6) is continuous, then v is a viscosity solution of the PDE:*

$$\begin{aligned}
\partial_t v(t, \tilde{y}) + \inf_{a \in A} \sup_{\tilde{\theta} \in G(t,x)} &\Bigg[f(t, \tilde{y}, a) + \partial_{\tilde{y}} v \cdot \bar{\boldsymbol{b}}(t, \tilde{y}, \tilde{\theta}, a) + \tfrac{1}{2}\ \mathrm{Tr}\left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)\, \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\intercal}\, \partial^2_{\tilde{y}\tilde{y}} v\right) \\
&+ \int_{\mathbb{R}} \big(v(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a)\, z) - v(t, \tilde{y}) - \partial_{\tilde{y}} v(t, \tilde{y})\, \bar{\boldsymbol{\xi}}(t, \tilde{y}, a)\, h(z)\big)\ \nu(dz)\Bigg] = 0\,,
\end{aligned} \tag{4.1.31}$$

subject to the terminal condition $v(T, \tilde{y}) = g(\tilde{y})$, and recall that A is a compact set taking values in $\mathbb{R}^k$.

Proof. Let $(t, \tilde{y}) \in [0, T] \times \mathbb{R}^{n+d+\tilde{d}}$ and let $\varphi \in C_b^{2,3}([0, T] \times \mathbb{R}^{n+d+\tilde{d}})$ be such that $v(t, \tilde{y}) = \varphi(t, \tilde{y})$, and $v \geq \varphi$ on $[0, T] \times \mathbb{R}^{n+d+\tilde{d}}$. Recall the DPP holds, therefore

$$\begin{aligned}
0 &= \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, \alpha_{t+s})\, ds + v(t+s, \tilde{Y}_{t+s}) - v(t, \tilde{y})\right] \\
&\geq \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, \alpha_{t+s})\, ds + \varphi(t+s, \tilde{Y}_{t+s}) - \varphi(t, \tilde{y})\right].
\end{aligned} \tag{4.1.32}$$

We recall that the stopping times are $\tau_u = \inf\{s : |\tilde{Y}_{t+s} - \tilde{y}| \geq 1\} \wedge u$. Next, we show that v is a viscosity supersolution of (4.1.31). Let $\mathbb{P} \in \mathcal{P}(t,x)$ and $\alpha \in \mathcal{A}_0$, by Itô's Lemma we have that

$$
\begin{aligned}
&\varphi(t+\tau_u, \tilde{Y}_{t+\tau_u}) - \varphi(t,\tilde{y}) \\
&= \int_0^{\tau_u} \partial_t \varphi(t+s, \tilde{Y}_{t+s})\, ds + \int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s})\, dM^{\mathbb{P}}_{t+s} \\
&+ \int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s})\, \bar{\boldsymbol{b}}(t+s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s})\, ds \\
&+ \int_0^{\tau_u} \tfrac{1}{2} \operatorname{Tr}\Big(\bar{\boldsymbol{\sigma}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})\, \bar{\boldsymbol{\sigma}}(t+\tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^{\intercal}\, \partial^2_{\tilde{y}\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s}) \Big)\, ds \\
&+ \int_0^{\tau_u} \int_{\mathbb{R}} \Big(\varphi(t+s, \tilde{Y}_{t+s} + \bar{\boldsymbol{\xi}}(t+s, \tilde{Y}_{t+s}, \alpha_{t+s})\, z) - \varphi(t+s, \tilde{Y}_{t+s}) \\
&\quad - \partial_{\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s})\, h(z) \Big)\, \nu(dz)\, ds\,.
\end{aligned}
\tag{4.1.33}
$$

The expectation of the stochastic integral term is zero because the function $\varphi \in C_b^{2,3}([0,T] \times \mathbb{R}^{n+d+\tilde{d}})$ and the process $M^{\mathbb{P}}$ is a martingale under the probability measure $\mathbb{P}$. Now, consider the expectation of each term on the right-hand side of (4.1.33). Begin with the third term:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s})\, \bar{\boldsymbol{b}}(t+s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s})\, ds \right] \\
&\quad \geq \mathbb{E}^{\mathbb{P}}\Bigg[\int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t,\tilde{y})\, \bar{\boldsymbol{b}}(t+s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \\
&\qquad\qquad - \left| \partial_{\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s}) - \partial_{\tilde{y}} \varphi(t,\tilde{y}) \right| \left| \bar{\boldsymbol{b}}(t+s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \right| \, ds \Bigg].
\end{aligned}
$$

From Lemma 4.1.15 and Theorem 4.1 in Fadina, Neufeld, and Schmidt (2018), write

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} \left| \partial_{\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s}) - \partial_{\tilde{y}} \varphi(t,\tilde{y}) \right| \left| \bar{\boldsymbol{b}}(t+s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \right| \, ds \right] \\
&\leq C_1 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right),
\end{aligned}
$$

where C_1 is a constant that depends only on $\tilde{y}$. Perform similar calculations for the first and the fourth terms on the right-hand side of (4.1.33) (i.e., $\partial_t \varphi(t+s, \tilde{Y}_{t+s})$ and $\partial^2_{\tilde{y}\tilde{y}} \varphi(t+s, \tilde{Y}_{t+s})$) and write

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\varphi(t+\tau_u, \tilde{Y}_{t+\tau_u}) - \varphi(t,\tilde{y}) \right] \geq -C_2 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) \\
&+ \mathbb{E}^{\mathbb{P}}\Bigg[\int_0^{\tau_u} \Bigg(\partial_t \varphi(t,\tilde{y}) + \Big[\partial_{\tilde{y}} \varphi(t,\tilde{y})\, \bar{\boldsymbol{b}}(t+s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s})
\end{aligned}
$$

$$
\begin{aligned}
&+\ \tfrac{1}{2}\ \mathrm{Tr}\left(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+s},\alpha_{t+s})\,\bar{\boldsymbol{\sigma}}(t+s,\tilde{Y}_{t+s},\alpha_{t+s})^{\intercal}\partial^2_{\tilde{y}\tilde{y}}\varphi(t,\tilde{y})\right)\\
&+\ \int_0^{\tau_u}\int_{\mathbb{R}}(\varphi(t+s,\tilde{Y}_{t+s}+\bar{\boldsymbol{\xi}}(t+s,\tilde{Y}_{t+s},\alpha_{t+s})\,z)-\varphi(t+s,\tilde{Y}_{t+s})\\
&-\ \partial_{\tilde{y}}\varphi(t+s,\tilde{Y}_{t+s})\,h(z))\,\nu(dz)\Big]\Big)\,ds\Big]\\
&\geq -C_2\left(u^{3/2}+u^2+u^{5/2}+u^3\right)+\mathbb{E}^{\mathbb{P}}\left[\tau_u\,\partial_t\varphi(t,\tilde{y})\right]\\
&\quad+\int_0^u\Big(\mathbb{E}^{\mathbb{P}}\Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\tilde{\theta}_{t+\tau_s},\alpha_{t+\tau_s})\\
&\quad+\tfrac{1}{2}\ \mathrm{Tr}\left(\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,\bar{\boldsymbol{\sigma}}(t+\tau_s,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})^{\intercal}\partial^2_{\tilde{y}\tilde{y}}\varphi(t,\tilde{y})\right)\\
&\quad+\int_{\mathbb{R}}(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{Y}_{t+\tau_s},\alpha_{t+\tau_s})\,z)-\varphi(t,\tilde{y})-\partial_{\tilde{y}}\varphi(t,\tilde{y})\,h(z))\,\nu(dz)\Big]\Big)\ ds\,,
\end{aligned}
$$

where C_2 is a constant that depends only on $\tilde{y}$. From Lemma 4.1.17, if u is small enough, there exists a probability measure $\tilde{\mathbb{P}}\in\mathcal{P}(t,x)$, such that for all $\alpha\in\mathcal{A}_0$, the inequality in (4.1.28) holds. Hence, for all $\alpha\in\mathcal{A}_0$ the following equation holds:

$$
\begin{aligned}
&\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\varphi(t+\tau_u,\tilde{Y}_{t+\tau_u})-\varphi(t,\tilde{y})\right]\geq -C_2\left(u^{3/2}+u^2+u^{5/2}+u^3\right)-C_A\,u^{3/2}\\
&+\mathbb{E}^{\tilde{\mathbb{P}}}\left[\tau_u\,\partial_t\varphi(t,\tilde{y})\right]+\int_0^u\Bigg(\inf_{a\in A}\sup_{\tilde{\theta}\in G(t,x)}\Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a)+\tfrac{1}{2}\,\mathrm{Tr}\left(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^{\intercal}\partial^2_x\varphi\right)\\
&+\int_{\mathbb{R}}(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z)-\varphi(t,\tilde{y})-\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z))\,\nu(dz)\Big]\Bigg)\ ds\,.
\end{aligned}
\tag{4.1.34}
$$

Now, consider the running reward of the performance criterion. Use Lemma 4.1.15 to write

$$
\begin{aligned}
&\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u}\left|f(t+s,\tilde{Y}_{t+s},\alpha_{t+s})-f(t,\tilde{y},\alpha_{t+s})\right|\,ds\right]\\
&\qquad\leq\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u}K\left(s+|\tilde{Y}_{t+s}-\tilde{y}|\right)\,ds\right]\leq C_3\left(u^{3/2}+u^2\right)\,.
\end{aligned}
$$

Use the triangle inequality to write

$$
\begin{aligned}
\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u}f(t+s,\tilde{Y}_{t+s},\alpha_{t+s})\,ds\right]&\geq\sup_{\mathbb{P}\in\mathcal{P}(t,x)}\mathbb{E}^{\tilde{\mathbb{P}}}\left[\int_0^{\tau_u}f(t,\tilde{y},\alpha_{t+s})\,ds\right]-C_3\left(u^{3/2}+u^2\right)\\
&\geq\inf_{a\in A}\mathbb{E}^{\mathbb{P}}\left[\tau_u\,f(t,\tilde{y},a)\right]-C_3\left(u^{3/2}+u^2\right)\,,
\end{aligned}
\tag{4.1.35}
$$

where C_3 is a constant that depends only on x. Therefore, from equations (4.1.32), (4.1.34), (4.1.35), we have that

$$
\begin{aligned}
0 \geq \inf_{a\in A} u\,\mathbb{E}^{\tilde{\mathbb{P}}}\left[(\tau_u/u)\,f(t,\tilde{y},a)\right] + u\,\Big(\mathbb{E}^{\tilde{\mathbb{P}}}\left[(\tau_u/u)\,\partial_t\varphi(t,\tilde{y})\right] +& \\
\inf_{a\in A}\sup_{\tilde{\theta}\in G(t,x)}\Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a) + \tfrac{1}{2}\operatorname{Tr}(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^{\intercal}\,\partial^2_{\tilde{y}\tilde{y}}\varphi)& \\
+\int_{\mathbb{R}}\left(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z) - \varphi(t,\tilde{y}) - \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z)\right)\,\nu(dz)\Big]\Big)& \\
- C_4\left(u^{3/2}+u^2+u^{5/2}+u^3\right),&
\end{aligned}
\tag{4.1.36}
$$

where $C_4 = C_2 + C_3 + C_A$. As u is small enough and $\mathbb{E}^{\tilde{\mathbb{P}}}\left[\tau_u/u\right] \to 1$, write

$$
\begin{aligned}
0 \geq \partial_t\varphi(t,\tilde{y}) + \inf_{a\in A}\sup_{\tilde{\theta}\in G(t,x)}\Big[f(t,\tilde{y},a) + \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a) + \tfrac{1}{2}\operatorname{Tr}\left(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^{\intercal}\,\partial^2_{\tilde{y}\tilde{y}}\varphi\right)& \\
+\int_{\mathbb{R}}\left(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z) - \varphi(t,\tilde{y}) - \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z)\right)\,\nu(dz)\Big]&.
\end{aligned}
$$

Thus, v is a viscosity supersolution of the PDE.

Next, we show that v is a viscosity subsolution of the PDE (4.1.31). Let $(t,\tilde{y}) \in [0,T]\times\mathbb{R}^{n+d+\tilde{d}}$ and let $\varphi \in C_b^{2,3}([0,T]\times\mathbb{R}^{n+d+\tilde{d}})$ be such that $v(t,\tilde{y}) = \varphi(t,\tilde{y})$, and $v \leq \varphi$ on $[0,T]\times\mathbb{R}^{n+d+\tilde{d}}$. As v satisfies the DPP, we have

$$
\begin{aligned}
0 &= \inf_{\alpha\in\mathcal{A}_0}\sup_{\mathbb{P}\in\mathcal{P}(t,x,G)}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} f(t+s,\tilde{Y}_{t+s},\alpha_{t+s})\,ds + v(t+s,\tilde{Y}_{t+s}) - v(t,\tilde{y})\right] \\
&\leq \inf_{\alpha\in\mathcal{A}_0}\sup_{\mathbb{P}\in\mathcal{P}(t,x,G)}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} f(t+s,\tilde{Y}_{t+s},\alpha_{t+s})\,ds + \varphi(t+s,\tilde{Y}_{t+s}) - \varphi(t,\tilde{y})\right].
\end{aligned}
$$

Then, for a constant control process $\alpha = a \in A$, we have that

$$
0 \leq \sup_{\mathbb{P}\in\mathcal{P}(t,x,G)}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} f(t+s,\tilde{Y}_{t+s},a)\,ds + \varphi(t+s,\tilde{Y}_{t+s}) - \varphi(t,\tilde{y})\right].
$$

As above in (4.1.34), we obtain the following inequality for all $\mathbb{P}\in\mathcal{P}(t,x)$:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\varphi(t+\tau_u,\tilde{Y}_{t+\tau_u}) - \varphi(t,\tilde{y})\right] \leq C_1\left(u^{3/2}+u^2+u^{5/2}+u^3\right) \\
&+\mathbb{E}^{\mathbb{P}}\left[\tau_u\,\partial_t\varphi(t,\tilde{y})\right] + \int_0^u\Bigg(\sup_{\tilde{\theta}\in G(t,x)}\Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a) + \tfrac{1}{2}\operatorname{Tr}\left(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^{\intercal}\,\partial^2_{xx}\varphi(t,\tilde{y})\right) \\
&+\int_{\mathbb{R}}(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z) - \varphi(t,\tilde{y}) - \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z))\,\nu(dz)\Big]\Bigg)\,ds\,.
\end{aligned}
$$

Similarly,

$$\sup_{\mathbb{P}\in\mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}}\left[\int_0^{\tau_u} \left| f(t+s,\tilde{Y}_{t+s},a) - f(t,\tilde{y},a)\right| \, ds\right] \leq C_2 \left(u^{3/2}+u^2\right) \, ,$$

where C_1 and C_2 are constants that depend only on $\tilde{y}$. We combine these two inequalities to obtain

$$\begin{aligned}
0 \leq\, & u\, \mathbb{E}^{\mathbb{P}}\left[(\tau_u/u)\right] f(t,\tilde{y},a) + u \Big(\mathbb{E}^{\mathbb{P}}\left[(\tau_u/u)\right] \partial_t \varphi(t,\tilde{y}) \\
& + \sup_{\tilde{\theta}\in G(t,x)} \Big[\partial_{\tilde{y}}\varphi(t,\tilde{y})\, \bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a) + \tfrac{1}{2}\operatorname{Tr}\left(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^{\intercal}\,\partial^2_{\tilde{y}\tilde{y}}\varphi\right) \\
& + \int_{\mathbb{R}} \left(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z) - \varphi(t,\tilde{y}) - \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z)\right)\nu(dz)\Big]\Big) \\
& + (C_1+C_2)\left(u^{3/2}+u^2+u^{5/2}+u^3\right).
\end{aligned} \tag{4.1.37}$$

Thus, we have that

$$\begin{aligned}
0 \leq \partial_t\varphi(t,\tilde{y}) + \inf_{a\in A}\sup_{\tilde{\theta}\in G(t,x)} \Big[& f(t,\tilde{y},a) + \partial_{\tilde{y}}\varphi(t,\tilde{y})\, \bar{\boldsymbol{b}}(t,\tilde{y},\tilde{\theta},a) + \tfrac{1}{2}\operatorname{Tr}\left(\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)\,\bar{\boldsymbol{\sigma}}(t,\tilde{y},a)^{\intercal}\,\partial^2_{\tilde{y}\tilde{y}}\varphi\right) \\
& + \int_{\mathbb{R}} \left(\varphi(t,\tilde{y}+\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,z) - \varphi(t,\tilde{y}) - \partial_{\tilde{y}}\varphi(t,\tilde{y})\,\bar{\boldsymbol{\xi}}(t,\tilde{y},a)\,h(z)\right)\nu(dz)\Big] \, ,
\end{aligned}$$

because a is an arbitrary constant control in A. Therefore, v is a viscosity subsolution of the PDE in (4.1.31). □

The parameter $\tilde{\theta}$ does not appear in the second order term and in the non-local term in (4.1.31) because the parameter $\tilde{\theta}$ only appears in the drift term of the process $\tilde{Y}$. Next, we show that there exists an unique viscosity solution for the PDE in (4.1.31). This ensures that we can use a numerical method to approximate the value function v when we cannot find an analytic solution.

Proposition 4.1.19. ***Uniqueness of viscosity solution***. *If Assumptions 4.1.12, 4.1.13, 4.1.14, 4.1.16 hold, then the non-linear PDE in (4.1.31) has an unique uniformly continuous viscosity solution with at most linear growth in x.*

Proof. Let v_1 and v_2 be a uniformly continuous viscosity subsolution (resp. supersolution) of (4.1.31) with $v_1(T,\tilde{y}) \leq v_2(T,\tilde{y})$ for $\tilde{y}\in\mathbb{R}^{n+d+\tilde{d}}$. We show that $v_1(t,\tilde{y}) \leq v_2(t,\tilde{y})$ for all $(t,\tilde{y})\in[0,T]\times\mathbb{R}^{n+d+\tilde{d}}$. We follow an argument from Pham (1998) and for $\beta,\epsilon,\delta,\lambda>0$ let us define the function Φ in $[0,T]\times\mathbb{R}^{n+d+\tilde{d}}\times\mathbb{R}^{n+d+\tilde{d}}$

$$\Phi(t,\tilde{y}_1,\tilde{y}_2) = v_1(t,\tilde{y}_1) - v_2(t,\tilde{y}_2) - \frac{\beta}{t} - \frac{1}{2\,\epsilon}\left|\tilde{y}_1-\tilde{y}_2\right|^2 - \delta\, e^{\lambda(T-t)}\left(\left|\tilde{y}_1\right|^2 + \left|\tilde{y}_2\right|^2\right). \tag{4.1.38}$$

Let the function Φ admit a maximum at $(\bar{t}, \bar{y}_1, \bar{y}_2) \in [0,T] \times \mathbb{R}^{n+d+\tilde{d}} \times \mathbb{R}^{n+d+\tilde{d}}$. As in Theorem 4.1 in Pham (1998), we only consider the case when $\bar{t} < T$. Then, it suffices to show that the inequality

$$\sup_{a\in A}\Bigg[\sup_{\tilde{\theta}\in G(\bar{t},\bar{x}_1)} \bar{\boldsymbol{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a)\left(\frac{1}{\epsilon}(\bar{y}_1 - \bar{y}_2) + 2\,\delta\, e^{\lambda\,(T-\bar{t})}\,\bar{y}_1\right) \\ - \sup_{\tilde{\theta}\in G(\bar{t},\bar{x}_2)} \bar{\boldsymbol{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}, a)\left(\frac{1}{\epsilon}(\bar{y}_1 - \bar{y}_2) - 2\,\delta\, e^{\lambda\,(T-\bar{t})}\,\bar{y}_2\right)\Bigg] \\ \leq C\left(\frac{|\bar{y}_1 - \bar{y}_2|^2}{\epsilon} + 2\,\delta\, e^{\lambda\,(T-\bar{t})}\left(1 + |\bar{y}_1|^2 + |\bar{y}_2|^2\right)\right) \tag{4.1.39}$$

holds, where C is a positive constant that does not depend on $\epsilon, \delta, \lambda$, and β because all the conditions and all the remaining terms in equation (4.6) in Pham (1998) are satisfied. Next, proceed as in Theorem 4.1 in Pham (1998). For each $\tilde{\theta} \in G(t, \bar{x}_1)$, there exists $\tilde{\theta}^p \in G(\bar{t}, \bar{x}_2)$, such that $\left|\tilde{\theta} - \tilde{\theta}^p\right| \leq |\tilde{y}_1 - \tilde{y}_2|$,

$$\bar{\boldsymbol{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a)\left(\frac{1}{\epsilon}(\bar{y}_1 - \bar{y}_2) + 2\,\delta\, e^{\lambda\,(T-\bar{t})}\,\bar{y}_1\right) - \bar{\boldsymbol{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}^p, a)\left(\frac{1}{\epsilon}(\bar{y}_1 - \bar{y}_2) - 2\,\delta\, e^{\lambda\,(T-\bar{t})}\,\bar{y}_2\right) \\ \leq \left|\bar{\boldsymbol{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a) - \bar{\boldsymbol{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}^p, a)\right|\left|\frac{(\bar{y}_1 - \bar{y}_2)}{\epsilon}\right| + 2\,\delta\, e^{\lambda\,(T-\bar{t})}\left(\left|\bar{\boldsymbol{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a)\right||\bar{y}_1| + \left|\bar{\boldsymbol{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}, a)\right||\bar{y}_2|\right) \\ \leq K_L\frac{|\bar{y}_1 - \bar{y}_2|^2}{\epsilon} + 2\,\delta\, e^{\lambda\,(T-\bar{t})}\left(|\bar{y}_1| + |\bar{y}_2| + |\bar{y}_1|^2 + |\bar{y}_2|^2\right) \\ \leq C\left(\frac{|\bar{y}_1 - \bar{y}_2|^2}{\epsilon} + 2\,\delta\, e^{\lambda\,(T-\bar{t})}\left(1 + |\bar{y}_1|^2 + |\bar{y}_2|^2\right)\right), \tag{4.1.40}$$

where the third inequality follows from Assumption 4.1.13 and the linear growth of the function $\bar{\boldsymbol{b}}$, and C is a positive constant that does not depend on $\epsilon, \delta, \lambda, \beta$. Therefore, (4.1.40) implies that (4.1.39) holds because a is arbitrary. Thus, $v_1(t, \tilde{y}) \leq v_2(t, \tilde{y})$ for all $(t, \tilde{y}) \in [0,T] \times \mathbb{R}^{n+d+\tilde{d}}$. □

4.2 Example and numerical results

In this Section, we analyse a classical problem in finance to illustrate the adaptive robust control framework. We derive the optimal acquisition strategy for an agent who employs market orders to purchase a large block of shares in an order driven market; see Cartea, Jaimungal, and Penalva (2015).

4.2.1 Optimal acquisition

At time $t = 0$, an investor must purchase $Q_0 > 0$ shares by the terminal date $T > 0$. The investor sends buy market orders to the limit order book (LOB) of the equity exchange at the speed ν_t. The controlled inventory target Q_t^ν denotes the remaining shares to be purchased over the remaining trading horizon $[t, T]$ for $t \geq 0$, and the target satisfies the SDE $dQ_t^\nu = -\nu_t\, dt$.

The investor's orders have temporary price impact; i.e., the orders receive worse prices than the midprice, denoted by S_t, because market orders walk the LOB. The price impact is temporary (i.e., the LOB replenishes after each trade) and we assume that the impact is linear in the speed of trading. For example, over a small time step Δt the investor purchases $\nu_t\, \Delta t$ shares and instead of paying the midprice S_t per share, the investor pays $(S_t + k\, \nu_t)\, \nu_t\, \Delta t$, where $k \geq 0$ is the temporary price impact parameter; see Cartea, Jaimungal, and Penalva (2015). As the investor buys shares, the process X^ν keeps track of the cumulative cost, which satisfies the SDE $dX_t^\nu = (S_t + k\, \nu_t)\, \nu_t\, dt$. We assume that the midprice of the stock satisfies the SDE

$$dS_t = \theta^*\, dt + \sigma\, dB_t^{\mathbb{P}_{\theta^*}} + J_{1+M_t}\, dM_t\,,$$

where M is a Poisson process with intensity λ, $M_0 = 0$, and $\{J_1, J_2, \dots\}$ are identically independent random variables that take the value ϵ with probability $1/2$ and the value $-\epsilon$ with probability $1/2$, where $\epsilon > 0$. The estimate $\hat{\theta}$ satisfies the SDE

$$d\hat{\theta}_t = \beta_t \left[\frac{\tilde{\theta}_t - \hat{\theta}_t}{\sigma^2}\, dt + \frac{1}{\sigma}\, dB_t^{\mathbb{P}_{\tilde{\theta}}} + \frac{J}{\sigma^2}\, dM_t \right]\,,$$

where θ^* is an unknown value and $\beta_t = \sigma/(1+t)$ is the learning rate; see estimator (1.9) in Bhudisaksang and Cartea (2021b). This choice of β_t simplifies the calculations in our example.

The investor's performance criterion is

$$V^\nu = \mathbb{E}^{\mathbb{P}} \left[X_T + Q_T\, (S_T + \eta\, Q_T) \right]\,, \tag{4.2.1}$$

where $\eta \geq 0$ is a terminal penalty parameter, and the investor's adaptive robust problem is given by

$$V(t, x, S, \hat{\theta}, q) := \inf_{\nu \in \mathcal{A}} \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[X_T + Q_T\, (S_T + \eta\, Q_T) \right]\,, \tag{4.2.2}$$

where $\hat{\theta}$ is the estimate of the unknown drift parameter θ^*. Here, the uncertainty set for the estimator of θ^* is

$$G(t, \hat{\theta}) = \left[\hat{\theta} - c/\sqrt{1+t}\,,\; \hat{\theta} + c/\sqrt{1+t} \right]\,, \tag{4.2.3}$$

where the uncertainty parameter $c > 0$ is a constant and it is easy to check that the function G satisfies Assumptions 4.1.14 and 4.1.16. The set of admissible strategies is

$$\mathcal{A} = \left\{ \nu = (\nu_t)_{\{0 \le t \le T\}} \mid \nu \text{ is progressively measurable, } \sup_{\mathbb{P} \in \mathcal{P}(0, S_0, \hat{\theta}_0)} \mathbb{E}^{\mathbb{P}} \left[\int_0^T |\nu_t|^2 \, dt \right] < \infty \right\}.$$

Next, we show that Assumption 4.1.4 holds. Use Gronwall's inequality and the definition of the function G to write

$$\sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^T \left(\hat{\theta}_u^2 + S_u^2 \right) du \right] < \infty \,, \tag{4.2.4}$$

for all $t \in [0,T]$ and $x = (S_t, \hat{\theta}_t) \in \mathbb{R}^2$. Let $\nu \in \mathcal{A}$, then

$$\begin{aligned} \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} [X_T] &= \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[x + \int_t^T |S_u + k\, \nu_u| \, |\nu_u| \, du \right] \\ &\le x + \sup_{\mathbb{P} \in \mathcal{P}(t,x)} \left(\mathbb{E}^{\mathbb{P}} \left[\int_t^T S_u^2 \, du \right] \mathbb{E}^{\mathbb{P}} \left[\int_t^T \nu_u^2 \, du \right] + \mathbb{E}^{\mathbb{P}} \left[\int_t^T k\, \nu_u^2 \, du \right] \right) \\ &< \infty \,, \end{aligned} \tag{4.2.5}$$

where the first inequality follows from the Cauchy inequality and the second inequality follows from (4.2.4). Next, apply a similar argument to obtain

$$\sup_{\mathbb{P} \in \mathcal{P}(t,x)} \mathbb{E}^{\mathbb{P}} \left[|Q_T| \, |S_T + \eta\, Q_T| \right] < \infty \,. \tag{4.2.6}$$

Hence, the functions J and w are finite from (4.2.5) and (4.2.6). Now, we show that the function V is finite. Let $(t,x,y) \in [0,T] \times \mathbb{R}^{1+d} \times \mathbb{R}^{\tilde{d}}$, for each $\nu \in \mathcal{A}$, we notice that $V(t,x,y) \le w(t,x,y,\nu) < \infty$. To show that $V(t,x,y) > -\infty$ we consider

$$\begin{aligned} X_T + Q_T\, (S_T + \eta\, Q_T) &\ge x + \int_t^T (S_u + k\, \nu_u)\, \nu_u \, du - \frac{S_T^2}{4\,\eta} \\ &\ge x - \int_t^T \frac{S_u^2}{4\,k} \, du - \frac{S_T^2}{4\,\eta} \,. \end{aligned}$$

Therefore,

$$V(t,x,y) \ge \sup_{\mathbb{P} \in \mathcal{P}(t,x,G)} \mathbb{E}^{\mathbb{P}} \left[x - \int_t^T \frac{S_u^2}{4\,k} \, du - \frac{S_T^2}{4\,\eta} \right] > -\infty \,,$$

where the last inequality follows from (4.2.4). Then, from Theorem 4.1.18, the value function V satisfies the HJBI

$$\partial_t V + \inf_{\nu} \sup_{\tilde{\theta} \in \tau(t,\hat{\theta})} \Big[(S + k\,\nu)\, \nu\, \partial_x V - \nu\, \partial_q V + \tilde{\theta}\, \partial_S V$$

$$+\frac{\sigma^2}{2}\,\partial_{SS}V+\frac{\beta_t\,(\tilde\theta-\hat\theta)}{\sigma^2}\,\partial_{\hat\theta}V+\beta_t\,\partial_{\hat\theta\,S}V+\frac{\beta_t^2}{\sigma^2}\,\partial_{\hat\theta\hat\theta}V\Big]$$
$$+\frac{\lambda}{2}\Big[V(t,x,S,\hat\theta+\beta_t\,\epsilon\,/\sigma^2,q)+V(t,x,S,\hat\theta-\beta_t\,\epsilon\,/\sigma^2,q)-2\,V\Big]=0\,,$$

subject to the terminal condition $V(T,x,S,\hat\theta,q)=x+S\,q+\eta\,q^2$.

The Proposition below shows the optimal acquisition speed.

Proposition 4.2.1. *The optimal speed of trading that solves the agent's acquisition problem, see* (4.2.2)*, is*

$$\nu_t^*=\frac{q}{T-t+k/\eta}+\frac{\partial_q h(t,\hat\theta,q)}{2\,k}\,, \tag{4.2.7}$$

where the function h satisfies

$$\begin{aligned}
&\partial_t h-\frac{1}{4\,k}\left(\partial_q h\right)^2+\left(\hat\theta+\frac{c}{\sqrt{1+t}}\right)q+\frac{c\,\partial_{\hat\theta}h}{\sigma\,(t+1)\,\sqrt{1+t}}\\
&\quad+\frac{\lambda}{2}\Big(h(t,\hat\theta+\beta_t\,\epsilon\,/\sigma^2,q)+\,h(t,\hat\theta-\beta_t\,\epsilon\,/\sigma^2,q)-2\,h\Big)+\frac{\partial_{\hat\theta\hat\theta}h}{(t+1)^2}=0\,,\quad \textit{if } q+\frac{\partial_{\hat\theta}h}{\sigma\,(t+1)}\geq 0\\
&\partial_t h-\frac{1}{4\,k}\left(\partial_q h\right)^2+\left(\hat\theta-\frac{c}{\sqrt{1+t}}\right)q-\frac{c\,\partial_{\hat\theta}h}{\sigma\,(t+1)\,\sqrt{1+t}}\\
&\quad+\frac{\lambda}{2}\Big(h(t,\hat\theta+\beta_t\,\epsilon\,/\sigma^2,q)+\,h(t,\hat\theta-\beta_t\,\epsilon\,/\sigma^2,q)-2\,h\Big)+\frac{\partial_{\hat\theta\hat\theta}h}{(t+1)^2}=0\,,\quad \textit{otherwise}\,,
\end{aligned} \tag{4.2.8}$$

with terminal condition $h(T,\hat\theta,q)=0$. Moreover, the optimal speed of trading of the adaptive agent (i.e., $c=0$) is

$$\nu^*=\frac{q}{T-t+k/\eta}+\frac{(T-t)\,(T-t+2\,k/\eta)\,\hat\theta}{4\,k\,(T-t+k/\eta)} \tag{4.2.9}$$

Proof. It is straightforward to show that the optimal speed of trading in feedback form is given by $\nu^*=(\partial_q V-S)/2\,k$. Next, use the ansatz

$$V(t,x,S,\hat\theta,q)=x+S\,q+\tilde h(t,q,\hat\theta)$$

and reduce the HJBI equation to

$$\begin{aligned}
&\partial_t\tilde h-\frac{1}{4\,k}\left(\partial_q\tilde h\right)^2+\frac{\lambda}{2}\left(\tilde h(t,\hat\theta+\beta_t\,\epsilon\,/\sigma^2,q)+\,\tilde h(t,\hat\theta-\beta_t\,\epsilon\,/\sigma^2,q)-2\,\tilde h\right)\\
&\qquad+\sup_{\tilde\theta\in\tau(t,\hat\theta)}\left[\tilde\theta\,q+\frac{(\tilde\theta-\hat\theta)}{\sigma\,(t+1)}\,\partial_{\hat\theta}\tilde h+\frac{1}{(t+1)^2}\,\partial_{\hat\theta\hat\theta}\tilde h\right]=0\,,
\end{aligned} \tag{4.2.10}$$

where $\tilde{h}(T,\hat{\theta},q) = \eta\, q^2$. The supremum attains at either $\tilde{\theta} = \hat{\theta} + c/\sqrt{1+t}$ or $\tilde{\theta} = \hat{\theta} - c/\sqrt{1+t}$. Thus, we write

$$\begin{aligned}
&\partial_t \tilde{h} - \frac{1}{4\,k}\left(\partial_q \tilde{h}\right)^2 + \left(\hat{\theta} + \frac{c}{\sqrt{1+t}}\right) q + \frac{c\,\partial_{\hat{\theta}}\tilde{h}}{\sigma\,(t+1)\,\sqrt{1+t}} \\
&\quad + \frac{1}{2}\left(\tilde{h}(t,\hat{\theta}+\beta_t\,\epsilon\,/\sigma^2,q) + \tilde{h}(t,\hat{\theta}-\beta_t\,\epsilon\,/\sigma^2,q) - 2\,\tilde{h}\right) + \frac{\partial_{\hat{\theta}\hat{\theta}}\tilde{h}}{(t+1)^2} = 0\,,\text{if } q + \frac{\partial_{\hat{\theta}}\tilde{h}}{\sigma\,(t+1)} \geq 0 \\
&\partial_t \tilde{h} - \frac{1}{4\,k}\left(\partial_q \tilde{h}\right)^2 + \left(\hat{\theta} - \frac{c}{\sqrt{1+t}}\right) q - \frac{c\,\partial_{\hat{\theta}}\tilde{h}}{\sigma\,(t+1)\,\sqrt{1+t}} \\
&\quad + \frac{1}{2}\left(\tilde{h}(t,\hat{\theta}+\beta_t\,\epsilon\,/\sigma^2,q) + \tilde{h}(t,\hat{\theta}-\beta_t\,\epsilon\,/\sigma^2,q) - 2\,\tilde{h}\right) + \frac{\partial_{\hat{\theta}\hat{\theta}}\tilde{h}}{(t+1)^2} = 0\,,\text{otherwise}\,,
\end{aligned} \tag{4.2.11}$$

Substitute

$$\tilde{h}(t,q,\hat{\theta}) = \frac{k\,q^2}{T-t+k/\eta} + h(t,q,\hat{\theta})$$

into the equation above to obtain (4.2.8) and it is straightforward to write the optimal speed of trading as in (4.2.9).

When the uncertainty parameter is $c = 0$, the system of PDEs in (4.2.8) reduces to

$$\partial_t h - \frac{1}{4\,k}\left(\partial_q h\right)^2 + \hat{\theta}\,q + \frac{\partial_{\hat{\theta}\hat{\theta}} h}{(t+1)^2} + \frac{\lambda}{2}\left(h(t,\hat{\theta}+\beta_t\,\epsilon\,/\sigma^2,q) + h(t,\hat{\theta}-\beta_t\,\epsilon\,/\sigma^2,q) - 2\,h\right) = 0 \tag{4.2.12}$$

with $h(T,\hat{\theta},q) = 0$. Next, substitute the ansatz

$$h(t,\hat{\theta},q) = q\,\tilde{h}_1(t,\hat{\theta}) + \tilde{h}_2(t,\hat{\theta})\,,$$

into (4.2.12) to obtain the coupled PDEs:

$$\begin{aligned}
&\partial_t \tilde{h}_1 - \frac{\tilde{h}_1}{T-t+k/\eta} + \frac{\partial_{\hat{\theta}\hat{\theta}}\tilde{h}_1}{(t+1)^2} + \hat{\theta} + \frac{\lambda}{2}\left(\tilde{h}_1(t,\hat{\theta}+\beta_t\,\epsilon\,/\sigma^2) + \tilde{h}_1(t,\hat{\theta}-\beta_t\epsilon\,/\sigma^2) - 2\,\tilde{h}_1\right) = 0, \\
&\partial_t \tilde{h}_2 - \frac{\tilde{h}_1^2}{4\,k} + \frac{\partial_{\hat{\theta}\hat{\theta}}\tilde{h}_2}{(t+1)^2} + \frac{\lambda}{2}\left(\tilde{h}_2(t,\hat{\theta}+\beta_t\,\epsilon\,/\sigma^2) + \tilde{h}_2(t,\hat{\theta}-\beta_t\,\epsilon\,/\sigma^2) - 2\,\tilde{h}_2\right) = 0\,.
\end{aligned}$$

Let $\tilde{h}_1(t,\hat{\theta}) = \hat{\theta}\,f^{\hat{\theta}q}(t) + f^q(t)$ and $\tilde{h}_2(t,\hat{\theta}) = \hat{\theta}^2\,f^{\hat{\theta}\,\hat{\theta}}(t) + \hat{\theta}\,f^{\hat{\theta}}(t) + f^{\theta}(t)$, where all the functions depend only on t, and obtain the ordinary differential equation

$$(T-t+k/\eta)\,\partial_t f^{\hat{\theta}\,q} - f^{\hat{\theta}\,q} = -(T-t+k/\eta)\,.$$

Integrate the equation above from t to T to obtain

$$\begin{aligned}
f^{\hat{\theta}\,q}(t) &= \frac{1}{T-t+k/\eta}\int_t^T (T-s+k/\eta)\,ds \\
&= \frac{(T-t)\,k/\eta + \frac{1}{2}\,(T-t)^2}{T-t+k/\eta}\,.
\end{aligned} \tag{4.2.13}$$

Substitute the expression above into the optimal speed of trading to obtain (4.2.9). □

Note that when the value of the uncertainty parameter c is zero, the optimal strategy does not depend on the size of the jump ϵ, nor on the arrival rate parameter λ. When $c > 0$, we cannot find a closed-form solution for the function h, so we employ the Crank-Nicolson finite-difference method to solve the PDE in (4.2.8) and obtain the optimal speed of trading in (4.2.7). For the set of parameters we study, our numerical solution shows that the term $\partial_q h(t, \hat{\theta}, q)\,/2\,k$ in (4.2.7) is greater (smaller) than the second term on the right-hand side of the optimal speed with $c = 0$ in (4.2.9) when $q > 0$ ($q < 0$). Also, we find that all else being equal, as the bands of the set G widen (i.e., the larger is the value of the uncertainty parameter c), the value $\left|\partial_q h(t, \hat{\theta}, q)\right|$ increases.

Therefore, as the investor is more uncertain about the estimate of the drift of the asset, the speed of trading is adjusted as follows. When the remaining target to purchase is positive (negative), i.e., $q > 0$ ($q < 0$), the investor speeds up the purchase (sales) of shares as the value of the uncertainty parameter c increases. Note that if $q < 0$, the investor holds more shares than the original target Q_0, so the investor must sell the excess shares before the end of the trading horizon.

In other words, as the investor perceives more uncertainty about the estimate of the drift parameter, the conservative (i.e., robust) strategy is to accelerate the purchase of the shares when the target is positive and to accelerate the sales of shares when the target is negative.

4.2.2 Performance of adaptive robust strategies

In this subsection, we compare the performance of the adaptive robust strategies with that of three strategies in which the agent: knows the true value of the drift parameter, employs a wrong value of the drift parameter, and employs a robust strategy. In the robust strategy, the agent uses the framework derived above, but does not learn the value of the unknown parameter. Instead, the agent assumes that the true parameter θ^* lies in the interval $[\underline{\theta}, \overline{\theta}]$ and solves

$$v(t, x, y) = \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}} J(t, x, y, \mathbb{P}, \alpha)\,, \tag{4.2.14}$$

where the set $\mathcal{P}$ contains all probability measures $\mathbb{P}_{\tilde{\theta}}$, such that $\tilde{\theta}_u \in [\underline{\theta}, \overline{\theta}]$ for all $u \in [t, T]$.

4.2.2.1 Adaptive robust optimal acquisition

The terminal time is $T = 20$ minutes and other model parameters are:

$$Q_0 = 10^5,\; X_0 = 0,\; S_0 = 10,\; \theta^* = 0.09,\; \hat{\theta}_0 = -0.03,\; \sigma = 0.2,\; \eta = 10^{-3}\,,$$

and

$$k = 10^{-4},\; c = 0.02\,, \epsilon = 0.1\,,$$

and for the robust strategy, the set of possible values of the parameter θ^* is $[-0.1, 0.2]$.

When the agent believes that θ is the true drift parameter, the optimal speed to trade is given by

$$\nu_t^* = \frac{q}{T - t + k/\eta} + \frac{(T-t)\,(T-t+2\,k/\eta)\,\theta}{4\,(T-t+k/\eta)\,k}\,. \tag{4.2.15}$$

We compare the adaptive robust strategy in (4.2.7) with the following strategies: i) the agent employs (4.2.15) with the true drift $\theta = \theta^*$, ii) the agent employs the robust strategy without learning in (4.2.14), in which case the speed of trading is in (4.2.15) with $\theta = \bar{\theta} = 0.2$, iii) the agent employs (4.2.15) with the wrong drift parameter $\theta = -0.03$.

We discretise the time space into 2,000 time-steps and employ 1,000 simulations to analyse the performance of the four strategies. The left-hand panel of Figure 4.1 shows the mean acquisition cost of the strategies. At the terminal date, the lowest mean cost is that of the strategy with the true drift parameter, followed by the adaptive robust strategy – see third column (for $c = 0.02$) in Table 4.1 below.

The right-hand panel of Figure 4.1 shows the standard deviation of the acquisition costs. For most of the trading horizon, the adaptive strategy shows the highest value of the standard deviation of the acquisition costs – the standard deviation peaks half-way through and then declines. This sharp increase, followed by a sharp decrease, in the standard deviation of the acquisition costs results from the strategy learning the correct value of the parameter. At every time-step, the adaptive strategy refines the estimate of the drift parameter, thus the speed of trading readjusts, so the costs are more volatile during the first half of the trading horizon.

Figure 4.2 shows the mean target inventory for the four strategies. Recall that the true, wrong, and robust strategies are deterministic, see (4.2.15), so for each strategy the mean of the remaining target is the same as the remaining target for each simulation. Observe that the robust strategy overshoots the target, that is, the robust strategy purchases more than the target and unwinds the extra shares at the end of the trading horizon. It is 'optimal' to speculate because the agent

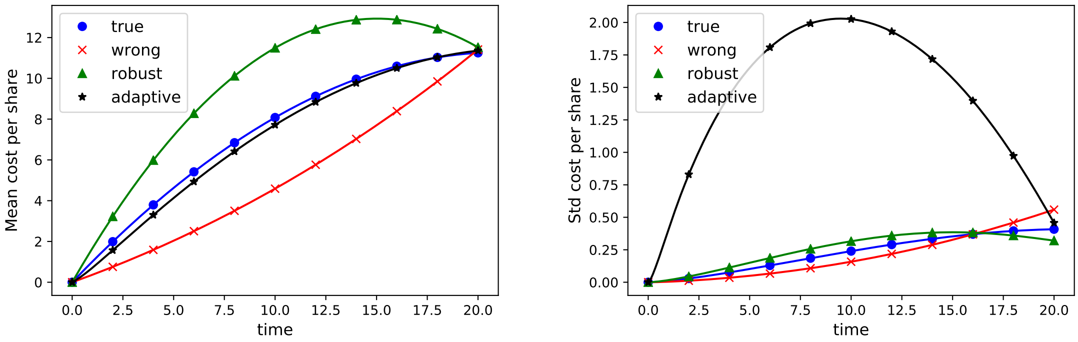

Figure 4.1: Mean and standard deviation of the acquisition costs per share.

expects an increase in the price of the asset – recall that the agent's estimate of the drift parameter is $\bar{\theta} = 0.2$. Finally, although the mean target is non-negative for the adaptive robust strategy, we note that in 300 simulations there was at least one time-step when the target became negative (because the agent's estimate of the drift was high enough to justify speculative purchases).

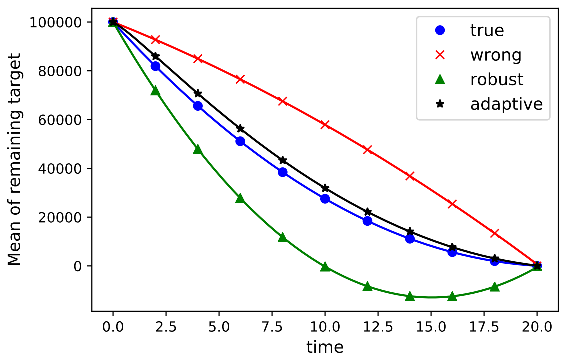

Figure 4.2: Target inventory.

We repeat the analysis for various values of the uncertainty parameter c in the set G (see (4.2.3)) and report the findings in Table 4.1; recall that the case discussed in the figures above is when $c = 0.2$. As the value of the uncertainty parameter c increases, the agent is less certain about the estimate of the drift of the stock price, so the agent speeds up the purchase of the target inventory (see figure (4.2) above), and the performance of the strategy worsens.

We repeat the analysis for various values of the terminal date of the trading horizon; see Table 4.2. Here, the initial acquisition target is $Q_0 = T \times 5 \times 10^3$ to keep

	$c = 0.01$		$c = 0.02$		$c = 0.03$		$c = 0.04$	
θ	mean	std	mean	std	mean	std	mean	std
true	11.26	0.41	11.26	0.41	11.26	0.41	11.26	0.41
false	11.41	0.55	11.41	0.55	11.41	0.55	11.41	0.55
robust	11.55	0.32	11.55	0.32	11.55	0.32	11.55	0.32
adaptive	11.38	0.47	11.38	0.46	11.36	0.47	11.37	0.46

Table 4.1: Performance of strategies.

	$T = 10$		$T = 20$		$T = 30$		$T = 40$	
θ	mean	std	mean	std	mean	std	mean	std
true	10.86	0.32	11.25	0.42	11.55	0.45	11.79	0.48
false	10.85	0.36	11.40	0.58	11.99	0.70	12.62	0.84
robust	10.98	0.28	11.54	0.33	12.11	0.40	12.76	0.56
adaptive	10.89	0.36	11.36	0.48	11.77	0.52	12.13	0.60

Table 4.2: Performance of strategies.

the ratio Q_0/T constant.[3] Our results show that as the terminal date increases, the adaptive robust strategy performs better, relative to the false and robust strategy, because there is more time for the agent to learn the value of the unknown drift parameter. Finally, note that the jump component has a similar effect as that of the volatility because the jump component is symmetric around zero.

4.3 Conclusions and future research

We proposed a continuous-time version of the adaptive robust methodology in Bielecki, Chen, and Cialenco (2017a). In our extension, the underlying process follows a jump-diffusion process with unknown drift and the agent is continuously estimating the drift while making optimal decisions that are time consistent. Our methodology has a balance between making a model robust to misspecification and learning an unknown parameter. Our result is general and the value function is characterised as the solution of a PDE. As an particular example we considered an optimal execution problem when the agent purchases a large amount of shares and her trades walk the limit order book of the exchange. When the agent has enough time to learn the value

[3]Note that for Figures 4.1 and 4.2, the terminal date is $T = 20$. Therefore, the initial acquisition target show in those figures is $Q_0 = 20 \times 5 \times 10^3 = 100000$.

of the unknown parameter, we showed that the adaptive robust strategy performs better (lower average and lower variance of acquisition costs) than when the agent employs a robust strategy or uses the incorrect parameter estimate.

We propose a number of extensions for future research. Assume that the observed process is noisy, e.g., the stock process has a stochastic drift, which follows an Ornstein-Uhlenbeck process with unknown mean. Another extension is to include a penalty term in the expectation of the objective function instead of considering the unknown parameter inside the confidence region. To this end, one can explore the approach in Bion-Nadal (2009) to impose a penalty for choosing an alternative model in the adaptive robust framework.

Chapter 5

Utility maximisation problem under adaptive robust control

In this chapter, we present our ongoing work on the application of the adaptive robust control framework in Chapter 4. Specifically, we focus on the utility maximisation problem under the adaptive robust control framework. Some of our results are still in development.

The outline of this chapter is listed as follows. The first Section discusses how our model relates to the literature and presents the motivation of our analysis. The second Section presents the model in detail and proves the stochastic representation of the value function of the stochastic control problem. The third Section performs an asymptotic analysis of the optimal strategy using the previous stochastic representation theorem that we develop in the previous Section. In the last Section, we provide an implication analysis of each hyper-parameter in the model to understand the impact of the adaptive robust control model on the optimal strategy

5.1 Introduction

The utility maximisation problem is a classical problem in financial mathematics that dates back to Merton (1975). The problem has been extended in multiple directions. In this chapter, we are particularly interested in the relaxation of the assumption of the knowledge of parameters. Here, we assume that the agent *does not* know the true value of the parameter of the underlying processes in the utility maximisation problem, but the agent estimates the parameter dynamically as a new observation arrives. We follow the adaptive-robust framework in Bhudisaksang and Cartea (2021a) which allows the agent to have a time-consistent behavior.

First, we discuss some of the related papers which work on the relaxed utility maximisation problem. Bordigoni, Matoussi, and Schweizer (2007) consider a utility maximisation problem under drift uncertainty where they use a martingale method to find the optimal strategy. In the recent paper, Neufeld and Nutz (2018) develop a generalised result on a utility maximisation problem under parameter uncertainty in drift, volatility, and jump terms and Liang and Ma (2018) consider a utility maximisation problem where the parameter lies in a time-vary interval. These mentioned papers primarily focus on the aspect of the parameter uncertainty of the parameter in the utility maximisation problem and do not allow the estimation of the parameter in the model.

In another strand of the literature, Bismuth, Guéant, and Pu (2019) consider the utility maximisation problem which embeds both the Bayesian online update of the parameter and the dynamic programming principle to obtain the partial differential equation of the value function. De Franco, Nicolle, and Pham (2019) study the Markowitz portfolio selection problem with an unknown drift vector in the multidimensional framework. The extant literature considers either robustness of the problem or the learning part of the problem, but not both. In this chapter, we use Bhudisaksang and Cartea (2021a) to study the utility maximisation problem with these two features simultaneously.

In particular, we consider the utility maximisation problem under the adaptive-robust framework where the utility function is negative exponential function and the drift parameter of the stock process is unknown to the agent. In other word, the stock processes S satisfies

$$\begin{aligned} dS_t^1 &= \theta^{*,1}\, S_t^1\, dt + \sigma_{11}\, S_t^1\, dW_t^1 + \sigma_{12}\, S_t^1\, dW_t^m\,, \\ dS_t^2 &= \theta^{*,2}\, S_t^2\, dt + \sigma_{21}\, S_t^2\, dW_t^2 + \rho\, \sigma_{22}\, S_t^2\, dW_t^m\,, \end{aligned} \tag{5.1.1}$$

where the agent does not know the parameters $\theta^{*,1}$ and $\theta^{*,2}$. We follow the steps in Bhudisaksang and Cartea (2021a) to set up the adaptive-robust control problem. Then, we solve the corresponding PDE by showing the stochastic representation of the value function. Although this stochastic representation does not have a closed-form solution, the value function and the optimal strategy are simplified to a deterministic integral. This allows us to perform numerical analysis to study the implication and the performance of the optimal strategy.

We list our key contributions. First, we apply a newly developed "adaptive-robust framework" in Bhudisaksang and Cartea (2021b) to a classic utility maximisation problem. Second, we solve the corresponding PDE by showing that the value function has a stochastic representation and can be evaluated as a deterministic integral. Moreover, the optimal strategies can be calculated numerically. This calculation bypasses the use of numerical methods to solve PDEs directly, which is usually not feasible in a high dimensional problem. Third, we build our understanding on the optimal strategy with various assumptions about the volatility of the underlying assets and show numerically that the performance of the adaptive-robust strategy is superior to that of the other common strategies.

The remainder of this chapter is organised as follows. In Section 2, we present our the utility maximisation problem under adaptive-robust framework and also show the main results of this chapter. In particular, we show that the value function of the problem has as a stochastic representation, which allows a feasible numerical method. Section 3 gives us a simple understanding of the optimal strategy for various assumptions on the value of the volatility of the underlying process. Section 4 studies the asymptotic analysis of the adaptive-robust strategy under the independent assumption on the volatility of the assets. Section 5 studies the numerical result of the adaptive-robust framework under the same independent assumption. We present the result in two parts. The first part discusses the intuition and implication of the adaptive-robust strategy that we obtain numerically. The second part gives a comparison of the performance of the adaptive-robust strategy with difference strategies.

5.2 Portfolio investment strategy

The agent considers a dynamic optimal asset allocation problem where shortselling is allowed. The agent trades in two risky assets and in a risk-free asset to maximise the expected utility of terminal wealth. Let r denote the risk-free rate and let S_t^1 and S_t^2 denote the prices of the risky assets at time t, which, under the probability measure $\mathbb{P}^*$, satisfy the SDEs

$$\begin{aligned} dS_t^1 &= \theta^{*,1}\, S_t^1\, dt + \sigma_{11}\, S_t^1\, dW_t^1 + \sigma_{12}\, S_t^1\, dW_t^m\,, \\ dS_t^2 &= \theta^{*,2}\, S_t^2\, dt + \sigma_{21}\, S_t^2\, dW_t^2 + \rho\, \sigma_{22}\, S_t^2\, dW_t^m\,, \end{aligned} \tag{5.2.1}$$

The parameter ρ represents how market risk affects the innovations of the price of asset 2.

The agent knows that the dynamics of the prices are given by (5.2.1) and knows the value of the volatility parameters $\sigma_{ij} > 0$ for $i, j \in \{1, 2\}$ and of ρ. However, the agent but does not know the values of the drift parameters $\theta^{*,1}$ and $\theta^{*,2}$, and we denote $\theta^* := [\theta^{*,1} \quad \theta^{*,2}]^\intercal$, where $^\intercal$ is the transpose operator.

The agent's controlled wealth process is denoted by X_t^α and satisfies the SDE

$$dX_t^\alpha = (\alpha_t^\intercal\,(\theta^* - \mathbf{1}\,r) + r\,X_t^\alpha)\,dt + \alpha_t^1\,\sigma_{11}\,dW_t^1 + \alpha_t^1\,\sigma_{12}\,dW_t^m + \alpha_t^2\,\sigma_{21}\,dW_t^2 + \alpha_t^2\,\rho\,\sigma_{22}\,dW_t^m\,, \tag{5.2.2}$$

where the control α_t^i denotes the amount of money invested in asset i, $\alpha_t = [\alpha_t^1 \quad \alpha_t^2]^\intercal$, and $\mathbf{1} = [1 \quad 1]^\intercal$.

The agent acknowledges that her model of price dynamics in (5.2.1) is incorrectly specified because she does not know the true value of the drift parameters. Hence, the agent employs the 'adaptive-robust' discrete-time framework proposed by Bielecki, Chen, Cialenco, et al. (2017), extended to continuous-time in Bhudisaksang and Cartea (2021b), to derive an investment strategy. The key advantage of the adaptive-robust approach is that the agent uses the evolution of the underlying stochastic price process to continuously update the estimates of the unknown parameters while ensuring that the decisions are robust to model misspesification.

The agent is risk-averse and derives utility from wealth. Her preferences are described by the continuous utility function $U(x)$, which is increasing and concave and bounded from above. The agent's adaptive-robust problem is given by the value function

$$v(t,x,S,\hat{\theta}) = \sup_{\alpha\in\mathcal{A}}\ \inf_{\mathbb{P}\in\mathcal{P}(t,x,S,\hat{\theta},G)} \mathbb{E}^{\mathbb{P}}\left[U(X_T^\alpha)\right]\,, \tag{5.2.3}$$

where $X_t = x$, $\mathcal{A}$ is the set of admissible control processes (square integrable), and $T > 0$ denotes the terminal date of the agent's investment horizon. Here, $\mathbb{E}^{\mathbb{P}}[\,\cdot\,]$ is the expectation operator under the measure $\mathbb{P}$, which is a probability measure in the set of alternative measures $\mathcal{P}(t,x,S,\hat{\theta},G)$ considered by the investor. In particular, the function $G(t,\hat{\theta})$ specifies the investor's model uncertainty that stems from the estimation process $\hat{\theta}$ of the unknown drift parameters.

Next, we describe the drift estimator process for stock i for $i \in \{1,2\}$. First, observe that over the time-step $\Delta t > 0$ the expected return of the stock is given by

$$\mathbb{E}^{\mathbb{P}^*}\left[\frac{S_{t+\Delta t}^i - S_t^i}{S_t^i}\right] = \theta^{*,i}\,\Delta\,t\,. \tag{5.2.4}$$

Hence, for an integrable deterministic function $g(t)$ heuristically, one can write

$$\mathbb{E}^{\mathbb{P}^*}\left[\int_0^t \frac{g(u)}{S_u^i}\,dS_u^i\right] = \theta^{*,i}\int_0^t g(u)\,du\,, \tag{5.2.5}$$

and from the expectation above, it is straightforward to see that

$$\frac{1}{\int_0^t g(u)\,du}\int_0^t \frac{g(u)}{S_u^i}\,dS_u^i \tag{5.2.6}$$

is an estimator of the drift parameter $\theta^{*,i}$. Here, we choose $g(u) = L/(1+u)$, where $L > 0$ is a learning rate parameter, and write

$$\hat{\theta}_t^i = \frac{L}{(1+t)^L} \int_0^t \frac{(1+u)^{L-1}}{S_u^i} \, dS_u^i \tag{5.2.7}$$

for $i = 1, 2$, where $\hat{\theta}$ denotes the estimate of the drift. The choice of the function $g(u)$ allows the estimator in (5.2.7) to converge to $\theta^{*,i}$ in probability. In Appendix C.1, we show that as $t \to \infty$ the estimators in (5.2.7) converge to $\theta^{*,i}$, in probability for all $i = 1, 2$.

Then, by Itô's Lemma, the estimators of the drift parameters in (5.2.7), under the true probability $\mathbb{P}^*$, are

$$\begin{aligned} d\hat{\theta}_t^1 &= \beta_t \left((\theta^{*,1} - \hat{\theta}_t^1)\, dt + \sigma_{11}\, dW_t^1 + \sigma_{12}\, dW_t^m \right) , \\ d\hat{\theta}_t^2 &= \beta_t \left((\theta^{*,2} - \hat{\theta}_t^2)\, dt + \sigma_{21}\, dW_t^2 + \rho\, \sigma_{22}\, dW_t^m \right) , \end{aligned} \tag{5.2.8}$$

where the learning rate is

$$\beta_t = \frac{L}{1+t} . \tag{5.2.9}$$

Therefore, one can implement (5.2.8) because it only depends on the realisation of the stock price processes S as in (5.2.7).

Next, to determine the set of alternative measures $\mathcal{P}(t, x, S, \hat{\theta}, G)$, the agent considers the function

$$G(t, \hat{\theta}) = \left\{ (\tilde{\theta}_1, \tilde{\theta}_2) \in \mathbb{R}^2 \;\middle|\; \hat{\theta}_i - c/\sqrt{1+t} \le \tilde{\theta}_i \le \hat{\theta}_i + c/\sqrt{1+t} \quad \text{for } i = 1, 2 \right\} , \tag{5.2.10}$$

which represents the model uncertainty that stems from not knowing the true value of the drift parameter and $c \ge 0$ is the uncertainty parameter.

To solve the investment problem, the agent derives a Hamilton–Jacobi–Bellman–Isaacs (HJBI) equation to characterise the value function v. This characterisation is done under the measure $\tilde{\mathbb{P}}$ because the agent searches over all strategies and outcomes that depend on all the models determined by the choice of $\tilde{\theta}$ in (5.2.10). This approach requires to write the estimator process in (5.2.8) under the measure $\tilde{\mathbb{P}}$, so we proceed as follows.

For each alternative measure $\tilde{\mathbb{P}}$ considered by the agent, there exists a progressively measurable process $\tilde{\theta}$ such that $\tilde{\theta}_u \in G(t, \hat{\theta}_u)$ for all $u \in [t, T]$ and

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}^*} = \exp\left(\sum_{i=1}^{2} - \int_0^t \frac{\theta^{*,i} - \tilde{\theta}_s^i}{\sigma_{i\,1}} \, dW^i - \frac{1}{2} \sum_{i=1}^{2} \int_0^t \left(\frac{\theta^{*,i} - \tilde{\theta}_s^i}{\sigma_{i\,1}} \right)^2 ds \right) . \tag{5.2.11}$$

Then, by Girsanov's theorem, the process

$$d\tilde{W}_t^i := dW_t^i + \frac{\theta_t^{*,i} - \tilde{\theta}_t^i}{\sigma_{i\,1}}\, dt\,, \qquad \text{for} \quad i = 1,\, 2\,,$$

is a Brownian motion under the probability measure $\tilde{\mathbb{P}}$; see Papaioannou (2012). Therefore, under the probability measure $\tilde{\mathbb{P}}$, the estimators in (5.2.8) are given by

$$\begin{aligned} d\hat{\theta}_t^1 &= \beta_t \left((\tilde{\theta}_t^1 - \hat{\theta}_t^1)\, dt + \sigma_{11}\, d\tilde{W}_t^1 + \sigma_{12}\, dW_t^m \right)\,, \\ d\hat{\theta}_t^2 &= \beta_t \left((\tilde{\theta}_t^2 - \hat{\theta}_t^2)\, dt + \sigma_{21}\, d\tilde{W}_t^2 + \rho\, \sigma_{22}\, dW_t^m \right)\,, \end{aligned} \tag{5.2.12}$$

where $\tilde{W}^1$, $\tilde{W}^2$, and W^m are independent Brownian motions. Thus, equipped with (5.2.12), the value function in (5.2.3) is the solution of an HJBI for the alternative models $\tilde{\mathbb{P}} \in \mathcal{P}(t, x, S, \hat{\theta}, G)$, which includes the continuous-time updates of the estimated drift parameters.

Note that for a fixed time t, when the value of the uncertainty parameter c in the function $G(t, \theta)$ is high (resp. low), the agent is less (resp. more) certain about the estimate of the drift. On the other hand, for a fixed value of $c > 0$, as time evolves, the intervals in (5.2.10) shrink because the agent feels more confident about the estimate of the drift of prices as more data are used to compute each update of $\hat{\theta}_t$.

In contrast, when the agent fully trusts the estimator of the drift parameter, she fixes the value of the uncertainty parameter c to zero, so the function $G(t, \hat{\theta})$ is the set $\{(\hat{\theta}_1, \hat{\theta}_2)\}$. In this case, the agent fully commits to the value of each update of the estimate $\hat{\theta}_t^i$ in the model of price dynamics (5.2.1). Consequently, the agent's optimisation problem only considers the 'adaptive' part of (5.2.3) (there is no robustness actions because the agent fully trusts estimates), so the optimal investment strategy is the solution to the adaptive control problem

$$v(t, x, S, \hat{\theta}) = \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}^{\hat{\mathbb{P}}} \left[U(X_T^\alpha) \right]\,, \tag{5.2.13}$$

where the probability measure $\hat{\mathbb{P}}$ is defined as

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}^*} = \exp\left(\sum_{i=1}^{2} - \int_0^t \frac{\theta^{*,i} - \hat{\theta}_s^i}{\sigma_{i\,1}}\, dW^i - \frac{1}{2} \sum_{i=1}^{2} \int_0^t \left(\frac{\theta^{*,i} - \hat{\theta}_s^i}{\sigma_{i\,1}} \right)^2 ds \right)\,. \tag{5.2.14}$$

The investment problem in (5.2.13) is only "adaptive" because at every point in time, the agent searches for optimal strategies that use the continuous updates of the estimates of the drift parameter of the stock prices.

In the ideal case in which the agent knows the true value of the drift parameters in (5.2.1), the investment problem in (5.2.3) reduces to the classical Merton problem

$$v(t, x, S, \hat{\theta}) = \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}^{\mathbb{P}^*} \left[U(X_T^\alpha) \right] , \tag{5.2.15}$$

where $\mathbb{P}^*$ is the true probability measure. It is straightforward (see e.g., Cartea, Jaimungal, and Penalva (2015)) to show that the Merton optimal investment strategy is

$$\alpha_t^{*,M} = \frac{\Sigma^{-1} \, (\theta^* - \mathbf{1}\, r)}{\gamma \, \exp(r\,(T-t))} \,. \tag{5.2.16}$$

In the general case, when the uncertainty parameter $c > 0$, the agent solves the adaptive-robust problem in (5.2.3). The value function v is finite (see Appendix C.2), and by Proposition 2.19 in Bhudisaksang and Cartea (2021a), one can show that the value function satisfies the HJBI equation

$$\begin{aligned} 0 = \partial_t v + r\, x\, \partial_x v &+ \tfrac{1}{2}\, \beta_t^2\, \sigma_1^2\, \partial_{\hat{\theta}_1 \hat{\theta}_1} v + \tfrac{1}{2}\, \beta_t^2\, \sigma_2^2\, \partial_{\hat{\theta}_2 \hat{\theta}_2} v + \beta_t^2\, \rho\, \sigma_c^2\, \partial_{\hat{\theta}_1 \hat{\theta}_2} v \\ &+ \sup_{\alpha} \inf_{\tilde{\theta} \in G(t, \hat{\theta})} \left\{ \beta_t\, (\tilde{\theta} - \hat{\theta})^\intercal\, \partial_{\hat{\theta}} v + \alpha^\intercal\, (\tilde{\theta} - \mathbf{1}\, r)\, \partial_x v + \tfrac{1}{2}\, \alpha^\intercal\, \Sigma\, \alpha\; \partial_{xx} v + \beta_t\, \alpha^\intercal\, \Sigma\, \partial_{x\hat{\theta}}\, v \right\} , \end{aligned} \tag{5.2.17}$$

where

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\, \sigma_c^2 \\ \rho\, \sigma_c^2 & \sigma_2^2 \end{bmatrix} , \qquad \text{with} \qquad \sigma_1^2 = \sigma_{11}^2 + \sigma_{12}^2, \quad \sigma_2^2 = \sigma_{21}^2 + \rho^2\, \sigma_{22}^2, \quad \sigma_c^2 = \sigma_{12}\, \sigma_{22} \,, \tag{5.2.18}$$

is the variance-covariance matrix of the returns of assets S^1 and S^2 and recall that $\rho = \{-1,\, 0,\, 1\}$. Note that the correlation between the prices of the two risky assets is $\rho\, \sigma_c^2 / \sqrt{\sigma_1\, \sigma_2}$ and that the HJBI does not depend on the true value of the drift parameter θ^*. To simplify notation, we define

$$\mathbf{1}^{+,+} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} , \quad \mathbf{1}^{-,-} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} , \quad \mathbf{1}^{+,-} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} , \quad \mathbf{1}^{-,+} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} , \quad \mathbf{1}^{+} = 1\,, \quad \mathbf{1}^{-} = -1\,.$$

The investment strategy is a function of many ingredients of the model, including the value of the parameters; e.g., the market risk parameter ρ, the value of the uncertainty parameter c, and the estimate of the drift parameter $\hat{\theta}$. Thus, to streamline the discussion, we delineate nine non-overlapping investment regions in $\mathbb{R}^2$ which we use below to derive the agent's investment strategy.

Let

$$A^{i,j} := \left\{ \hat{\theta} \in \mathbb{R}^2 \mid \Sigma^{-1} \left(\hat{\theta} - \mathbf{1}\, r + \mathbf{1}^{-i,-j}\, c/\sqrt{1+t} \right) \in \mathbb{R}^{i,j} \right\} \qquad \text{for} \quad i, j \in [+,-] \,.$$

Denote by $A^{0,-}$ the region where $\hat{\theta}_2 - r + c/\sqrt{1+t} \le 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat{\theta}_2 - r + c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat{\theta}_1 \le \frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat{\theta}_2 - r + c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,.$$

Denote by $A^{0,+}$ the region where $\hat{\theta}_2 - r - c/\sqrt{1+t} \ge 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat{\theta}_2 - r - c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat{\theta}_1 \le \frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat{\theta}_2 - r - c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,.$$

Denote by $A^{+,0}$ the region where $\hat{\theta}_1 - r - c/\sqrt{1+t} \ge 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat{\theta}_1 - r - c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat{\theta}_2 \le \frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat{\theta}_1 - r - c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,.$$

Denote by $A^{-,0}$ the region where $\hat{\theta}_1 - r + c/\sqrt{1+t} \le 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat{\theta}_1 - r + c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat{\theta}_2 \le \frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat{\theta}_1 - r + c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,.$$

Denote by $A^{0,0}$ the region where

$$r - c/\sqrt{1+t} \le \hat{\theta}_1 \le r + c/\sqrt{1+t} \quad \text{and} \quad r - c/\sqrt{1+t} \le \hat{\theta}_2 \le r + c/\sqrt{1+t}\,.$$

Finally, define the function

$$F(t,\hat{\theta}) = \begin{cases} \frac{1}{2}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{-i,-j}\,\frac{c}{\sqrt{1+t}}\right)^{\intercal}\Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{-i,-j}\,\frac{c}{\sqrt{1+t}}\right), & \hat{\theta} \in A^{i,j}\,, \\ \frac{1}{2}\left(\hat{\theta}_1 - r + \mathbf{1}^{-i}\,\frac{c}{\sqrt{1+t}}\right)\frac{1}{\sigma_1^2}\left(\hat{\theta}_1 - r + \mathbf{1}^{-i}\,\frac{c}{\sqrt{1+t}}\right), & \hat{\theta} \in A^{i,0}\,, \\ \frac{1}{2}\left(\hat{\theta}_2 - r + \mathbf{1}^{-i}\,\frac{c}{\sqrt{1+t}}\right)\frac{1}{\sigma_2^2}\left(\hat{\theta}_2 - r + \mathbf{1}^{-i}\,\frac{c}{\sqrt{1+t}}\right), & \hat{\theta} \in A^{0,i}\,, \\ 0\,, & \hat{\theta} \in A^{0,0}\,, \end{cases} \tag{5.2.19}$$

where $i,j \in [+,-]$ and which we use below to write the optimal investment strategy of the agent.

Proposition 5.2.1. ***Adaptive-robust strategy.*** *Let $U(x) = -e^{-\gamma x}$ be the agent's utility function, where $\gamma > 0$ is the risk-aversion parameter. Let the stock prices satisfy the SDEs in (5.2.1), the learning rate parameter $L > 0$, and investment horizon $T > 0$. The adaptive-robust investment strategy, in feedback form, for the investor's* *problem in (5.2.3) is*

$$\alpha_t^* = \begin{cases} \dfrac{\Sigma^{-1}\,\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{-i,-j}\,c/\sqrt{1+t}\right) - \beta_t\,\partial_{\hat{\theta}}\bar{H}(t,\hat{\theta})}{\gamma\;\exp(r\,(T-t))}\,, & \hat{\theta} \in A^{i,j}\,,\; i,j \in [+,-]\,, \\ \dfrac{\left[\hat{\theta}_1 - r + \mathbf{1}^{-i}\,c/\sqrt{1+t} \quad 0\right]^\intercal - \sigma_1^2\,\beta_t\,\partial_{\hat{\theta}}\bar{H}(t,\hat{\theta})}{\gamma\,\sigma_1^2\;\exp(r\,(T-t))}\,, & \hat{\theta} \in A^{i,0}\,,\; i \in [+,-]\,, \\ \dfrac{\left[0 \quad \hat{\theta}_2 - r + \mathbf{1}^{-i}\,c/\sqrt{1+t}\right]^\intercal - \sigma_2^2\,\beta_t\,\partial_{\hat{\theta}}\bar{H}(t,\hat{\theta})}{\gamma\,\sigma_2^2\;\exp(r\,(T-t))}\,, & \hat{\theta} \in A^{0,i}\,,\; i \in [+,-]\,, \\ \dfrac{-\beta_t\,\partial_{\hat{\theta}}\bar{H}(t,\hat{\theta})}{\gamma\;\exp(r\,(T-t))}\,, & \hat{\theta} \in A^{0,0}\,, \end{cases} \tag{5.2.20}$$

where $\hat{\theta} = [\hat{\theta}_1 \quad \hat{\theta}_2]^\intercal$ is the value of the drift estimator at time t. The function $\bar{H}$ has the stochastic representation

$$\bar{H}(t,\hat{\theta}) = \mathbb{E}^{\mathbb{P}^*}\left[\int_t^T F(u, Z_u)\,du \;\middle|\; Z_t = \hat{\theta}\right]\,, \tag{5.2.21}$$

with F in (5.2.19), and the stochastic process $Z = [Z^1 \quad Z^2]^\intercal$, with $Z_t = \hat{\theta}$, follows

$$\begin{aligned} dZ_u^1 =& \beta_t\,\left((r - Z_u^1)\,du + \sigma_{11}\;dW_u^1 + \sigma_{12}\;dW_u^m\right)\,, \\ dZ_u^2 =& \beta_t\,\left((r - Z_u^2)\,du + \sigma_{21}\;dW_u^2 + \rho\,\sigma_{22}\;dW_u^m\right)\,, \end{aligned} \tag{5.2.22}$$

where W^1, W^2, and W^m are mutually independent Brownian motions under the probability measure $\mathbb{P}^*$ (see price dynamics in (5.2.1)), and recall that β_t is the learning rate in (5.2.9). Conditional on $Z_t = \hat{\theta}$, the process $Z_u \in \mathbb{R}^2$ is normally distributed under the probability measure $\mathbb{P}^*$. For all values of the learning rate parameter L, the mean of Z_u^i is

$$\hat{\theta}_i\,\left(\frac{1+t}{1+u}\right)^L + r\,\left(1 - \frac{(1+t)^L}{(1+u)^L}\right)\,. \tag{5.2.23}$$

On the other hand, when $L \neq 1/2$, the co-variance matrix of Z_u is

$$\Sigma\,\int_t^u L^2\,\frac{(1+v)^{2\,L-2}}{(1+u)^{2\,L}}\,dv = \Sigma\,\frac{L^2}{2\,L-1}\,\left(\frac{(1+u)^{2\,L-1}}{(1+u)^{2\,L}} - \frac{(1+t)^{2\,L-1}}{(1+u)^{2\,L}}\right)\,, \tag{5.2.24}$$

and when $L = 1/2$, the co-variance matrix of Z_u is

$$\Sigma\,\int_t^u L^2\,\frac{(1+v)^{2\,L-2}}{(1+u)^{2\,L}}\,dv = \frac{\Sigma}{4}\,\left(\frac{\log(1+u) - \log(1+t)}{1+u}\right)\,, \tag{5.2.25}$$

where

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\,\sigma_c^2 \\ \rho\,\sigma_c^2 & \sigma_2^2 \end{bmatrix}\,, \qquad \text{with} \qquad \sigma_1^2 = \sigma_{11}^2 + \sigma_{12}^2, \quad \sigma_2^2 = \sigma_{21}^2 + \rho^2\,\sigma_{22}^2, \quad \sigma_c^2 = \sigma_{12}\,\sigma_{22}\,, \tag{5.2.26}$$

Proof. Recall that the value function v satisfies the HJBI in (5.2.17), subject to the terminal condition $v(T,x,\hat{\theta}) = -e^{-\gamma x}$. Substitute the ansatz

$$v(t,x,\hat{\theta}) = -\,U(t,\hat{\theta})\,\exp\Big(-\gamma\,x\,\exp\big(r\,(T-t)\big)\Big)$$

into the HJBI in (5.2.17) to write

$$\begin{aligned} 0 = \partial_t U &+ \tfrac{1}{2}\,\beta_t^2\,\sigma_1^2\,\partial_{\hat{\theta}_1\hat{\theta}_1}U + \tfrac{1}{2}\,\beta_t^2\,\sigma_2^2\,\partial_{\hat{\theta}_2\hat{\theta}_2}U + \,\beta_t^2\,\rho\,\sigma_c^2\,\partial_{\hat{\theta}_1\hat{\theta}_2}U \\ &+ \inf_{\alpha}\sup_{\tilde{\theta}\in G(t,\hat{\theta})}\Big\{\beta_t\,(\tilde{\theta}-\hat{\theta})^\intercal\,\partial_{\hat{\theta}}U - \alpha^\intercal\,(\tilde{\theta}-\mathbf{1}\,r)\,U + \tfrac{1}{2}\,\alpha^\intercal\,\Sigma\,\alpha\,U - \beta_t\,\alpha^\intercal\,\Sigma\,\partial_{\hat{\theta}}\,U\Big\} \end{aligned} \tag{5.2.27}$$

with terminal condition $U(T,\hat{\theta}) = 1$.

Next, we discuss the inf-sup solutions to determine the optimal investment strategy. Note that the term inside the sup sign depends on $\tilde{\theta}^\intercal\,(\beta_t\,\partial_{\hat{\theta}}U - \alpha\,U)$ where $\tilde{\theta}\in G(t,\hat{\theta})$. Therefore, we consider the value of $\beta_t\,\partial_{\hat{\theta}}U - \alpha\,U$ in different investment regions of $\mathbb{R}^2$, see part 1 in Appendix C.3 for the completed calculation.

Then, we solve the inf-sup problem in (5.2.27) for each $\hat{\theta}$ in the nine non-overlapping investment regions of $\mathbb{R}^2$ discussed above. Let $U(t,\hat{\theta}) = \exp(-H(t,\hat{\theta}))$ and let $C(t,\hat{\theta})$ denote the inf-sup term in (5.2.27). Then, the optimal investment strategy depends on nine regions of $\mathbb{R}^2$ (see Appendix C.3 part 2 and 3 for more details). Moreover, the optimal investment strategy satisfies (5.2.20) and

$$C(t,\hat{\theta}) = -U\,\Big(\beta_t\,(\mathbf{1}\,r-\hat{\theta})^\intercal\,\partial_{\hat{\theta}}H + F(t,\hat{\theta})\Big)\;, \tag{5.2.28}$$

for all t and $\hat{\theta}$. Next, for $\hat{\theta}\in\mathbb{R}^2$ we use (5.2.27), and recall the terminal condition $H(T,\hat{\theta}) = 0$, to write

$$\partial_t H + \tfrac{1}{2}\,\beta_t^2\,\sigma_1^2\,\partial_{\hat{\theta}_1\,\hat{\theta}_1}H + \tfrac{1}{2}\,\beta_t^2\,\sigma_2^2\,\partial_{\hat{\theta}_2\,\hat{\theta}_2}H + \,\beta_t^2\,\rho\,\sigma_c^2\,\partial_{\hat{\theta}_1\,\hat{\theta}_2}H + \beta_t\,(\mathbf{1}\,r-\hat{\theta})^\intercal\;\partial_{\hat{\theta}}H + F(t,\hat{\theta}) = 0\,, \tag{5.2.29}$$

where F is in (5.2.20).

To show that the functions H and $\bar{H}$ in (5.2.21) are equal, we define the stochastic process $Z = [Z^1 \quad Z^2]^\intercal$, which follows

$$\begin{aligned} dZ_u^1 &= \beta_t\,\big((r-Z_u^1)\,du + \sigma_{11}\;dW_u^1 + \sigma_{12}\;dW_u^m\big)\;, \\ dZ_u^2 &= \beta_t\,\big((r-Z_u^2)\,du + \sigma_{21}\;dW_u^2 + \rho\,\sigma_{22}\;dW_u^m\big)\;, \end{aligned}$$

with $Z_t = \hat{\theta}$, where W^1, W^2, and W^m are independent Brownian motions under the probability measure $\mathbb{P}^*$. From the PDE in (5.2.29) and the terminal condition

$H(T, \hat{\theta}) = 0$, the function H has the stochastic representation in (5.2.21) because the process

$$H(t, Z_t) + \int_t^T F(u, Z_u)\, du\,,$$

is a martingale process, by Itô's Lemma. Next, we show that the random variable Z_u for $t \leq u \leq T$ is normally distributed. Let $I_1 = 1$ and $I_2 = \rho$, apply Itô's Lemma to $\exp(\int_t^u \beta_s\, ds)\, Z_u$ and for $i = 1, 2$, we write

$$\begin{aligned}
\exp\left(\int_t^u \beta_s\, ds\right) Z_u^i &= Z_t^i + \int_t^u \exp\left(\int_t^v \beta_s\, ds\right) dZ_v^i + \int_t^u \beta_v \exp\left(\int_t^v \beta_s\, ds\right) Z_v^i\, dv \\
&= Z_t^i + \int_t^u r\, \beta_v \exp\left(\int_t^v \beta_s\, ds\right) dv + \int_t^u \sigma_{i\,1}\, \beta_v \exp\left(\int_t^v \beta_s\, ds\right) dW_v^i \\
&\quad + \int_t^u I_i\, \sigma_{i\,2}\, \beta_v \exp\left(\int_t^v \beta_s\, ds\right) dW_v^m\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
Z_u^i &= Z_t^i \exp\left(-\int_t^u \beta_s\, ds\right) + \int_t^u r\, \beta_v \exp\left(-\int_v^u \beta_s\, ds\right) dv \\
&\quad + \int_t^u \sigma_{i\,1}\, \beta_v \exp\left(-\int_v^u \beta_s\, ds\right) dW_v^i + \int_t^u I_i\, \sigma_{i\,2}\, \beta_v \exp\left(-\int_v^u \beta_s\, ds\right) dW_v^m \\
&= \hat{\theta}_i \left(\frac{1+t}{1+u}\right)^L \\
&\quad + L\left(\int_t^u r\, \frac{(1+v)^{L-1}}{(1+u)^L}\, dv + \int_t^u \sigma_{i\,1}\, \frac{(1+v)^{L-1}}{(1+u)^L}\, dW_v^i + \int_t^u I_i\, \sigma_{i\,2}\, \frac{(1+v)^{L-1}}{(1+u)^L}\, dW_v^m\right).
\end{aligned}$$

Thus, it is easy to see that the mean Z_u^i satisfies (5.2.23) and the co-variance matrix Z_u satisfies (5.2.24) when $L \neq 1/2$ and (5.2.25) when $L = 1/2$. $\square$

In the next Proposition, we derive the optimal investment strategy α^* of the problem in (5.2.13) in closed-form when the value of the uncertainty parameter c is zero. Recall that in this case the investor, at every point in time t, uses the estimate of the drift parameter $\hat{\theta}_t$; i.e., the investor uses the adaptive strategy.

Proposition 5.2.2. ***Adaptive strategy.*** *Let the prices of the asset satisfy the SDEs in (5.2.1), r is the risk-free rate, $L > 0$ the learning rate parameter, $T > 0$ the investment horizon, the uncertainty parameter $c = 0$, and $\rho \in \{-1, 0, 1\}$. The adaptive investment strategy for the investor's problem (5.2.13) is*

$$\alpha_t^{*,A} = \frac{\Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\, r\right)}{\gamma\, \exp(r\,(T-t))} - \frac{\beta_t\, \partial_{\hat{\theta}} \bar{H}^0(t, \hat{\theta})}{\gamma\, \exp(r\,(T-t))}\,, \tag{5.2.30}$$

where $\bar{H}^0$ has the stochastic representation

$$\bar{H}^0(t,\hat{\theta}) = \mathbb{E}^{\mathbb{P}^*}\left[\int_t^T \frac{1}{2}\,(Z_u - \mathbf{1}\,r)^\intercal\,\Sigma^{-1}\,(Z_u - \mathbf{1}\,r)\,du\right], \tag{5.2.31}$$

with Z defined in (5.2.22)*, and*

$$\partial_{\hat{\theta}}\bar{H}^0(t,\hat{\theta}) = \Sigma^{-1}\,(\hat{\theta} - \mathbf{1}\,r)\int_t^T \left(\frac{1+t}{1+u}\right)^{2\,L} du\,. \tag{5.2.32}$$

For a proof see the appendix C.4.

From Proposition 5.2.1, the optimal strategy of the adaptive robust model in (5.2.20) can be written as the sum of two components. The first term on the right-hand side of (5.2.20) is a Merton-style investment strategy with a truncation that depends on the sign of $\hat{\theta} - r$ and the value of the uncertainty parameter c. The second term on the right-hand side is an adjustment that results from the integral of the estimate of the risk-premium with a truncation ($\bar{H}^c$). Similarly, from Proposition 5.2.2, we have that the optimal strategy of the adaptive model in (5.2.30) can be written as the sum of two components. The first term on the right-hand side of (5.2.30) is a Merton-style investment strategy. The second term on the right-hand side is an adjustment that results from the integral of the estimate of the risk-premium ($\bar{H}^0$).

Next, we provide some terminologies and observations regarding the optimal strategies in (5.2.20) and (5.2.30). We see from (5.2.20) and (5.2.30), that the adaptive strategy is compensated by the partial derivative of the integral of the variance of the estimated risk-premium with a truncation for the adaptive-robust strategy and without a truncation for the adaptive strategy. In other words, the optimal strategy is compensated by the sensitivity of the parameter $\hat{\theta}$ on the integral of the variance of the estimator with or without truncation. Also, we refer to the term $\beta_t\,\partial_{\hat{\theta}}\bar{H}^0$ as "compensated parameter estimation term" because the term only appears when the agent simultaneously estimates the value of the unknown parameter and solves the optimal control problem. When the agent accounts for parameter uncertainty, the term $\beta_t\,\partial_{\hat{\theta}}\bar{H} - \beta_t\,\partial_{\hat{\theta}}\bar{H}^0$ depends on c, which accounts for parameter uncertainty of the problem and we refer to this term "compensated parameter uncertainty term".

We return to these points above when we study the case when the value of the market risk direction ρ is zero.

5.2.1 Connection to the related literature

Here, we discuss some related literature that study the utility maximisation problem under relaxed assumptions, which produce similar results and obtain a similar second component. This will help us to understand the second component in (5.2.20) and (5.2.30) intuitively.

We remark that the function $\bar{H}^0$ can be understood as the integration of the variance of the estimated risk-premium term and the function $\bar{H}$ can be understood as the integration of the variance of the estimated risk-premium term with a truncation.

The second term in (5.2.30) has been studied extensively, and is know as the "intertemporal hedging term". Kim and Omberg (1996) study the utility maximisation problem when the risk premium of stock process is specified by the Ornstein–Uhlenbeck process (OU-process) with a known coefficient. The optimal strategy for the problem consists of two components where the first is the usual optimal strategy for the Merton problem and the second is the hedging term against the risk-premium uncertainty. Moreover, it is possible for the optimal strategy to short a stock, although the current risk-premium of the stock is positive, see Figure 4 and Section 3 in Kim and Omberg (1996) for more details. Xia (2001) extends the model to account for the parameter learning of the OU process leading to the additional term for the optimal strategy, which accounts for the parameter estimation part. Moreover, in Liu and Muhle-Karbe (2013) where they use the model in Kim and Omberg (1996), the intertemporal hedging term of the predictable return is the first order partial derivative of the total variance of the stochastic mean of the stock prices and therefore, the intertemporal hedging term in general accounts for the sensitivity of the risk-premium of the stock process. Note that all these paper are slightly working on difference setups to our problem as their utility function is in the HARA form and the risk-premium is not a constant but an OU process.

Next, we show that the second component of the optimal strategy in Kim and Omberg (1996) can be written as a partial derivative of the function that has a stochastic representation as in our optimal strategy in (5.2.30).

Lemma 5.2.3. *Following the model in Kim and Omberg (1996) with the utility function $U(x) = -\exp(-\gamma x)$ and $\rho_{mx} = 1$ (the other parameters and processes follow Kim and Omberg (1996)), we can write the second component of the optimal strategy as a partial derivative of a function with a stochastic representation. In particular,*

$$\hat{H}(x,t) = \mathbb{E}^{\mathbb{P}}\left[\int_t^T Z_s^2\, ds \mid Z_t = x\right] \tag{5.2.33}$$

where
$$dZ_s = \left(-Z_s\,\sigma_X + \lambda_X\,(\bar{X} - Z_s)\right)\,ds + \sigma_X\,dW_s \tag{5.2.34}$$
and $Z_t = x$*. Thus, the optimal strategy is*
$$y_*(W,X,\tau) = \frac{X - \sigma_X\,\partial_X \hat{H}}{\gamma\,\sigma^2\,e^{r\,(T-t)}}\,, \tag{5.2.35}$$
where $\tau = T - t$ *and note that here the term* X *is the risk-premium of the risky asset in this model.*

For a proof see the appendix C.5.

As we see from the above Lemma, the function $\hat{H}$ in (5.2.33) is the integration of the variance of the risk-premium of the stock process. This is similar to our result in Proposition 5.2.2. The optimal strategy has the second component that depends on the first derivative of the function H. This implies that the optimal strategy adjusts its strategy of the sensitive of the integration of the variance of the risk-premium of the stock process.

Next, we provide the intuition of the second term. From figure 4 in Kim and Omberg (1996), it is possible for the nonmyopic investor to short the risky asset, although the risk-premium is positive depending on the risk-aversion coefficient γ. Similar to our result, when the parameter L is less than 1, it can lead to a situation where the "estimated" risk-premium is positive, but the optimal strategy is negative (the investor shorts the risky asset). Although in Kim and Omberg (1996) refers the second component as " hedging against risk-premium uncertainty" or "speculating on risk-premium uncertainty", we prefer to call this term "compensated parameter estimation term".

5.2.2 Implication of the adaptive strategy under different value of L

In this subsection, we discuss the behaviour of the adaptive strategy when the correlation between two stocks is zero (i.e., $\rho = 0$). We rewrite the adaptive strategy in (5.2.30) as
$$\alpha_t^{ad,i} = \frac{(\hat{\theta}_i - r)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))} \times A(t,T,L)\,, \tag{5.2.36}$$
where
$$A(t,T,L) = \begin{cases} 1 - \frac{L}{2\,L-1}\left(1 - \left(\frac{1+t}{1+T}\right)^{2\,L-1}\right) & \text{if} \quad L \neq 1/2\,, \\ 1 - \frac{1}{2}\,\left(\log(1+T) - \log(1+t)\right) & \text{if} \quad L = 1/2\,. \end{cases} \tag{5.2.37}$$

Recall that the first term on the right-hand side of the adaptive strategy in (5.2.36) is the Merton-style strategy. The term $A(t,T,L)$ is an 'adjustment' that results from adapting the investment strategy to the updates in the estimate of the drift parameters. To understand the intuition of the adaptive strategy we focus on the adjustment term as a function of $T-t$ and of the value of the learning rate parameter L.

The Proposition below shows the upper and lower bounds of the adjustment term for $L>0$ and all $t \leq T$.

Proposition 5.2.4. *Let the learning rate parameter $L>0$, the horizon of the investment $T>0$, and market risk parameter $\rho=0$. Then, for all T and L, the adjustment term $A(t,T,L)$ is increasing in t, and obeys the bounds*

$$A(0,T,L) \leq A(t,T,L) \leq A(T,T,L) = 1\,, \tag{5.2.38}$$

where

$$A(0,T,L) = \begin{cases} \frac{L-1}{2\,L-1} + \frac{L}{2\,L-1}\left(\frac{1}{1+T}\right)^{2\,L-1} & \text{if} \quad L \neq 1/2\,, \\ 1-\frac{1}{2}\log(1+T) & \text{if} \quad L = 1/2\,. \end{cases} \tag{5.2.39}$$

For a proof see the appendix C.6.

The lower bound (5.2.39) could be positive or negative. The next Proposition shows that for $L \geq 1$, $A(t,T,L)$ is always positive and when $L<1$, there exists a time $t^L < T$ such that $A(t^L,T,L)=0$.

Proposition 5.2.5. *Let the learning rate parameter $L \geq 1$, the horizon of the investment $T>0$, and market risk parameter $\rho=0$. Then, the adjustment term $A(t,T,L)$ is always positive and when $L<1$, there exists*

$$t^L = \begin{cases} (1+T)\,\left(\frac{1-L}{L}\right)^{1/(2\,L-1)} - 1 & \text{if} \quad L \neq 1/2\,, \\ (1+T)^{1/2} - 1 & \text{if} \quad L = 1/2\,, \end{cases} \tag{5.2.40}$$

such that $A(t^L,T,L)=0$. In other words, if $t>t^L$, then $A(t,T,L)>0$ and if $t<t^L$, then $A(t,T,L)<0$.

For a proof see the appendix C.7.

Therefore, when the learning rate parameter $L \geq 1$, the sign of the holdings in the stock are the same for the adaptive and the Merton-style strategies. For example, if $L \geq 1$ and $\hat{\theta}_i < r$ both the Merton-style and the adaptive strategies hold a short position in stock i for $i=1,2$. On the other hand, when $L<1$, the sign of the holdings in the stock held by the adaptive strategy could be the same or of opposite sign than that of the holdings of the Merton-style strategy.

Next, we consider the property of $A(t,T,L)$ with respect to L, which helps us to analysis the optimal strategy in more details. For $L \neq 1/2$, we have that

$$\frac{\partial A}{\partial L} = \frac{1}{(2\,L-1)^2}\left(1-\left(\frac{1+t}{1+T}\right)^{2\,L-1}\right) + \frac{L}{(2\,L-1)}\,(2\,L-1)\left(\frac{(1+t)^{2\,L-2}}{(1+T)^{2\,L-1}}\right). \tag{5.2.41}$$

Note that the second term in (5.2.41) is always greater than 0 regardless of the value of L. Note that if $L > 1/2$, the first term in (5.2.41) is greater than 0. Note that if $L < 1/2$, the first term in (5.2.41) can be less than 0. This implies that given the same value of t and T, when $L > 1/2$, the adjustment term is increasing as the value of L increases.

The intuition of the adaptive strategy is that the sensitivity of the integral of the variance of the estimated risk-premium has an important effect on the investment strategy; the higher the sensitivity of the integral of the variance of the estimated risk-premium, the more hedging of the strategy becomes. When $L > 1/2$, as the value of L increases, the sensitivity of the integral of the variance of the estimated risk-premium decreases because the investor acknowledges that the investor learns the risk-premium very quickly (large L) and therefore, the integral of the variance of the estimated risk-premium and its sensitivity is smaller because the investor has already learn quickly. For example, assume that $\hat{\theta} > r$. There are two forces at play in the adaptive strategy. The investor goes long the stock because the expected return from holding it exceeds the risk free rate. On the other hand, because new information arrives and the drift parameter is constantly updated leading to the risk of the risk-premium term, the sensitivity of the integral of the variance of the estimated risk-premium reduces the amount of stock the investor would have purchased (i.e., the adjustment parameter $A(t,T,L) < 1$ for $t < T$). Recall that the investor's utility function is increasing and concave in wealth. Thus, all else being equal, if the sensitivity of the integral of the variance of the estimated risk-premium increases, the quantity held in the stock also decreases. Thus, in a similar way, an increase in the variance of the estimator of the drift process reduces the holdings in the stock.

Moreover, the choice of the learning parameter L is crucial. From Proposition 5.2.5, when $L \geq 1$, the adjustment term is always greater than 0 regardless of the values of t and T. Therefore, under this condition, the adaptive strategy always has the same sign as that of the Merton-style strategy because the adjustment term is always greater than 0.

In contrast, when $L < 1$, the adjustment term can be greater than 0 or less than 0 depending on the value of t. Therefore, under this condition, the adaptive strategy can have the same or opposite sign as the Merton investment strategy in (5.2.16).

5.3 Implication of the optimal strategy

In this Section, we study the optimal investment strategy in (5.2.3) when the correlation between the prices of two risky assets in (5.2.1) is $\rho = 0$, $\rho > 0$, $\rho < 0$. Here, we explore the effect of the difference assumption on ρ to the optimal strategy of the adaptive-robust agent. Thus, we study the effect of the uncertainty coefficient c to the optimal strategy under the different assumptions of ρ.

5.3.1 Zero correlation ($\rho = 0$)

In this subsection, we assume that the correlation between S^1 and S^2 is zero, i.e, $\rho = 0$. Thus, the price processes S^1 and S^2 are independent and we have the following result.

Figure 5.1 plots the nine investment regions in $\mathbb{R}^2$, where the center region is $A^{0,0}$ and the upper right most region is $A^{+,+}$ and the remaining regions are $A^{+,0}$, $A^{+,-}$, $A^{0,-}$, $A^{-,-}$, $A^{-,0}$, $A^{-,+}$, $A^{0,+}$ in clockwise order. We use $r = 0.4$, $c = 0.2$, and $t = 3$.

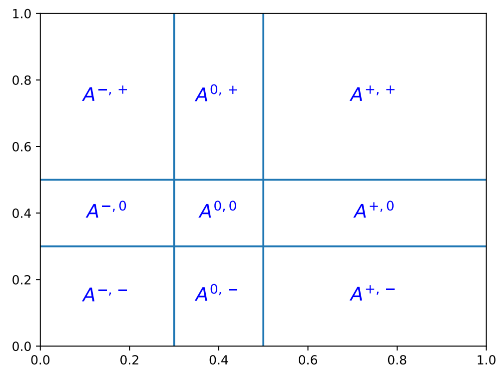

Figure 5.1: Investment regions when $\rho = 0$. Other parameters are $r = 0.4$, $c = 0.2$, and $t = 3$.

In Figure 5.1, the regions $A^{+,+} = [r+c/\sqrt{1+t}, \infty] \times [r+c/\sqrt{1+t}, \infty]$ and $A^{-,-} = [-\infty, r - c/\sqrt{1+t}] \times [-\infty, r - c/\sqrt{1+t}]$. Consider the case $\hat{\theta} \in A^{+,+}$, therefore the values of $\hat{\theta}_1$ and $\hat{\theta}_2$ are greater than $r + c/\sqrt{1+t}$. Thus, the agent acquires a long position in both of the stocks because the estimate of the drift parameters are greater than $r + c/\sqrt{1+t}$, where c is the uncertainty coefficient. Similarly for $\hat{\theta} \in A^{-,-}$, the agent acquires a short position in both of the stocks since both of estimators are less than $r - c/\sqrt{1+t}$ where c is the uncertainty coefficient.

Consider the case $\hat{\theta} \in A^{+,0}$. The agent mainly acquires a long position only on the first stock because $\hat{\theta}_1$ is greater than $r + (c/\sqrt{1+t})$ and $\hat{\theta}_2$ is in between $[r - (c/\sqrt{1+t}), r + (c/\sqrt{1+t})]$. This is due to the value $\hat{\theta}_2$ is not larger than r nor lesser than r enough to account for the uncertainty coefficient c.

Consider the case $\hat{\theta} \in A^{-,0}$, the agent mainly acquires a short position only on the first stock because $\hat{\theta}_1$ is less than $r - (c/\sqrt{1+t})$ and the $\hat{\theta}_2$ is in between $[r - (c/\sqrt{1+t}), r + (c/\sqrt{1+t})]$. This is due to the value $\hat{\theta}_2$ is not larger than r nor lesser than r enough to account for the uncertainty coefficient c.

In contrast, consider the case $\hat{\theta} \in A^{-,+}$. The agent takes a short position in the first stock and takes a long position in the second stock because the value $\hat{\theta}_1$ is less than $r - c/\sqrt{1+t}$ and the value $\hat{\theta}_2$ is greater than $r + c/\sqrt{1+t}$. Similarly for the investment region $A^{+,-}$.

5.3.2 Positive correlation case ($\rho > 0$)

In this subsection, we consider the case where the correlation between the prices of the risky assets is positive i.e, $\rho > 0$. Thus, the process S^1 and S^2 are positive correlated and we have the following immediate result.

Figures 5.2 plots the nine regions in $\mathbb{R}^2$, where the center region is $A^{0,0}$ and the upper right most region is $A^{+,+}$ and the remaining regions are $A^{+,0}$, $A^{+,-}$, $A^{0,-}$, $A^{-,-}$, $A^{-,0}$, $A^{-,+}$, $A^{0,+}$ in clockwise order.

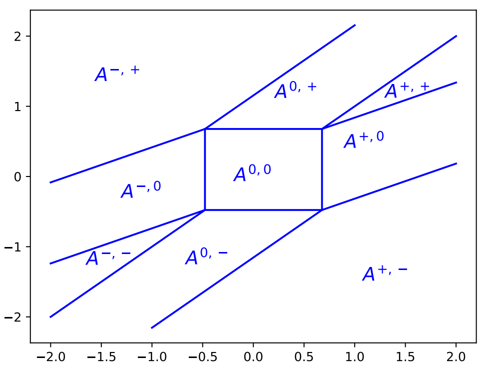

Figure 5.2: Regions for when $\rho > 0$.

In this case, the stock prices have a noise path that is positively correlated since $\rho > 0$. Figure 5.2 shows that the region $A^{+,+} \subset [r + c/\sqrt{1+t}, \infty] \times [r + c/\sqrt{1+t}, \infty]$ which implies that there exits $\hat{\theta} \in [r+c/\sqrt{1+t}, \infty] \times [r+c/\sqrt{1+t}, \infty]$ which is not in $A^{+,+}$. Note that this is due to the positive correlation of these two risky assets makes the diversification in a long-long position becomes less efficient in the risk-adjusted return. Therefore, it is possible for the agent to acquire a long position in only one asset even though both of $\hat{\theta}_1$ and $\hat{\theta}_2$ are greater $r + (c/\sqrt{1+t})$. Similarly, the region $A^{-,-} \subset [-\infty, r - c/\sqrt{1+t}] \times [-\infty, r - c/\sqrt{1+t}]$. Therefore, it is possible for the agent to acquire a short position in only one asset, although both of $\hat{\theta}_1$ and $\hat{\theta}_2$ are less than $r - c/\sqrt{1+t}$.

Moreover, we consider the case where $\hat{\theta} \in A^{+,0}$, we notice that there exists $\hat{\theta} \in \mathbb{R}^2$ such that $\hat{\theta} \in A^{+,0}$ and $\hat{\theta} \in [r + c/\sqrt{1+t}, \infty] \times [r + c/\sqrt{1+t}, \infty]$. Although both of the $\hat{\theta}_1$ and $\hat{\theta}_2$ are greater $r + (c/\sqrt{1+t})$, the positive correlation of these two risky assets makes the diversification in a long-long position becomes less efficient in the risk-adjusted return. Therefore, the agent only acquires a long position in the first asset.

Thus, when $\hat{\theta} \in A^{-,0}$, we notice that there exists $\hat{\theta} \in \mathbb{R}^2$ such that $\hat{\theta} \in A^{-,0}$ and $\hat{\theta} \in [-\infty, r - c/\sqrt{1+t}] \times [-\infty, r - c/\sqrt{1+t}]$. Although both of $\hat{\theta}_1$ and $\hat{\theta}_2$ are less than $r - (c/\sqrt{1+t})$, the positive correlation of these two risky assets makes the diversification in a short-short position becomes less efficient in the risk-adjusted return. Therefore, the agent only acquires a short position in the first asset. Similarly, these implications also hold when $\hat{\theta} \in A^{0,+}$ or $\hat{\theta} \in A^{0,-}$.

In contrast, consider the case where $\hat{\theta} \in A^{-,+}$. We notice that there exists $\hat{\theta} \in \mathbb{R}^2$ such that $\hat{\theta} \in A^{-,+}$ and $\hat{\theta}_1 > r - c/\sqrt{1+t}$ or $\hat{\theta}_2 < r + c/\sqrt{1+t}$. The agent acquires a short position in the first asset and a long position in the second asset even though the estimator can be $\hat{\theta}_1 > r - c/\sqrt{1+t}$ or $\hat{\theta}_2 < r + c/\sqrt{1+t}$, this is because taking an short-long position for these two stocks do increases the risk-adjusted return for the agent. This is due to the positive correlation between the prices of the risky assets. Similarly for the regions $A^{+,-}$.

5.3.3 Negative correlation case ($\rho < 0$)

In this subsection, we consider the case where the correlation between the prices of the risky assets is negative i.e, $\rho < 0$. Thus, the process S^1 and S^2 are negative correlated and we have the following immediate result.

Figure 5.3 shows the regions. There are nine regions in $\mathbb{R}^2$ space where the center region is $A^{0,0}$ and the upper right most region is $A^{+,+}$ and the remaining regions are $A^{+,0}, A^{+,-}, A^{0,-}, A^{-,-}, A^{-,0}, A^{-,+}, A^{0,+}$ in clockwise order.

In this case, the stock prices have a noise path that is negatively correlated since $\rho < 0$. Figure 5.3 shows that the region $[r + c/\sqrt{1+t}, \infty] \times [r + c/\sqrt{1+t}, \infty] \subset A^{+,+}$ which implies $\hat{\theta} \in A^{+,+}$ that even though either of $\hat{\theta}_1$ or $\hat{\theta}_2$ are not greater $r + (c/\sqrt{1+t})$. Note that this is due to the negative correlation of these two risky assets makes the diversification in the long-long position becomes more efficient in the risk-adjusted return. Therefore, it is possible for the agent to acquire a long position in both of assets.

Similarly, the region $[-\infty, r - c/\sqrt{1+t}] \times [-\infty, r - c/\sqrt{1+t}] \subset A^{-,-}$. Therefore, it is possible for the agent to acquire a short position in both risky assets, although either of $\hat{\theta}_1$ or $\hat{\theta}_2$ are greater than $r - c/\sqrt{1+t}$.

Moreover, we consider the case where $\hat{\theta} \in A^{+,0}$, we notice that there exists $\hat{\theta} \in \mathbb{R}^2$ such that $\hat{\theta} \in A^{+,0}$ and $\hat{\theta} \in [r + c/\sqrt{1+t}, \infty] \times [-\infty, r - c/\sqrt{1+t}]$. Although $\hat{\theta}_1$ is greater than $r + (c/\sqrt{1+t})$ and $\hat{\theta}_2$ is less than $r - (c/\sqrt{1+t})$, the negative correlation of these two risky assets makes the diversification in a long-short position becomes less efficient in the risk-adjusted return. Therefore, the agent only acquires a long position in the first asset.

Thus, when $\hat{\theta} \in A^{-,0}$, we notice that there exists $\hat{\theta} \in \mathbb{R}^2$ such that $\hat{\theta} \in A^{-,0}$ and $\hat{\theta} \in [-\infty, r - c/\sqrt{1+t}] \times [r + c/\sqrt{1+t}, \infty]$. Although $\hat{\theta}_1$ is less than $r - (c/\sqrt{1+t})$ and $\hat{\theta}_2$ is greater than $r + (c/\sqrt{1+t})$, the positive correlation of these two stock makes the diversification in a short-long position becomes less efficient in the risk-adjusted return. Therefore, the agent only acquires a short position in the first asset.

Similarly, these implications also hold when $\hat{\theta} \in A^{0,+}$ or $\hat{\theta} \in A^{0,-}$.

In contrast, consider the case where $\hat{\theta} \in A^{-,+}$. We notice that there exists $\hat{\theta} \in \mathbb{R}^2$ such that $\hat{\theta} \notin A^{-,+}$ and $\hat{\theta}_1 < r - c/\sqrt{1+t}$ and $\hat{\theta}_2 > r + c/\sqrt{1+t}$. This implies that although $\hat{\theta}_1 < r - c/\sqrt{1+t}$ and $\hat{\theta}_2 > r + c/\sqrt{1+t}$, the agent may only acquires a short position in the first asset or only acquires a long position in the second asset, but not both. The negative correlation makes an short-long position for these two risky assets becomes less efficient in the risk-adjusted return for the agent. Similarly for the regions $A^{+,-}$.

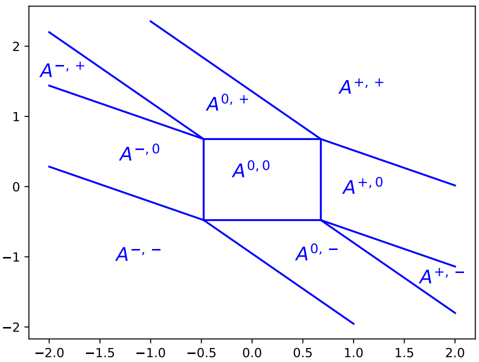

Figure 5.3: Regions for when $\rho < 0$.

5.4 Asymptotic analysis when c is close to zero for zero correlated case $\rho = 0$

Here, we consider the optimal strategy in Proposition 5.2.1 when the uncertainty parameter c is close to zero. We also explore the effect of the learning rate parameter L on the optimal strategy for both the adaptive strategy and the adaptive-robust strategy. For simplicity, we assume that the correlation between the prices of the stocks is zero (i.e, $\rho = 0$). The optimal strategy in (5.2.20) is not available in closed-form because we cannot find an explicit solution for the stochastic representation of H in (5.2.21). Therefore, here we discuss the optimal strategies when the value of the parameter c is small and study the asymptotic formulation of the optimal strategy in (5.2.20).

First, we state some corollaries which are direct results from Propositions 5.2.1 and 5.2.2. The first corollary states the optimal strategy of the adaptive-robust control problem in (5.2.3) when $\rho = 0$ and $c > 0$.

Corollary 5.4.1. *Let the uncertainty parameter $c > 0$ and $\rho = 0$. The optimal adaptive-robust investment strategy is*

$$\alpha_t^{adr,i} = \begin{cases} \dfrac{(\hat{\theta}_i - r + c/\sqrt{1+t}) - \beta_t\, \sigma_i^2\, \partial_{\hat{\theta}_i} \bar{H}(t,\hat{\theta})}{\gamma\, \sigma_i^2\, \exp(r\,(T-t))}\,, & \hat{\theta}_i - r \leq -c/\sqrt{1+t}\,, \\ \dfrac{(\hat{\theta}_i - r - c/\sqrt{1+t}) - \beta_t\, \sigma_i^2\, \partial_{\hat{\theta}_i} \bar{H}(t,\hat{\theta})}{\gamma\, \sigma_i^2\, \exp(r\,(T-t))}\,, & \hat{\theta}_i - r \geq c/\sqrt{1+t}\,, \\ \dfrac{-\beta_t\; \partial_{\hat{\theta}_i} \bar{H}(t,\hat{\theta})}{\gamma\; \exp(r\,(T-t))}\,, & |\, \hat{\theta}_i - r \,| < c/\sqrt{1+t}\,, \end{cases} \tag{5.4.1}$$

where $i = 1, 2$ and we call this optimal strategy "the adaptive-robust strategy" and the function $\bar{H}$ in (5.2.21) equals

$$\bar{H}(t,\hat{\theta}) = \sum_{i=1}^{2} \frac{1}{2\,\sigma_i^2}\,\mathbb{E}^{\mathbb{P}}\left[\int_t^T \left(Z_u^i - r - c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i - r >\ c/\sqrt{1+u}\}}\, du \ \middle|\ Z_t = \hat{\theta}\right]$$
$$+ \frac{1}{2\,\sigma_i^2}\,\mathbb{E}^{\mathbb{P}}\left[\int_t^T \left(Z_u^i - r + c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i - r <\ -c/\sqrt{1+u}\}}\, du \ \middle|\ Z_t = \hat{\theta}\right]. \tag{5.4.2}$$

Note that (5.4.2) has the property that when $c = 0$, the function $\bar{H}(t,\hat{\theta})$ is greater than the function $\bar{H}(t,\hat{\theta})$ when $c > 0$. Also, we notice that the function $\bar{H}$ is an integral of the estimated variance with the truncation when $c > 0$ and is an integral of the estimated variance when $c = 0$.

The second corollary states the optimal strategy of the adaptive-robust control problem in (5.2.3) when $\rho = 0$ and $c = 0$. Recall the function A in (5.2.37)

Corollary 5.4.2. *Let $c = 0$ and $\rho = 0$. The adaptive optimal investment strategy is*

$$\alpha_t^{ad,i} = \frac{(\hat{\theta}_i - r)\left(1 - \frac{L}{t+1}\int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du\right)}{\gamma\,\sigma_i^2 \exp(r\,(T-t))} = \frac{(\hat{\theta}_i - r)}{\gamma\,\sigma_i^2 \exp(r\,(T-t))}\, A(t,T,L)\,. \tag{5.4.3}$$

where $i = 1, 2$ and we call this optimal strategy "the adaptive strategy".

Before we proceed to the main result, we provide some preliminary results which need for the calculation of (5.4.2). We denote by $\phi(x)$ and $\Phi(x)$ the density function and cumulative distribution of a standard normal variable, respectively.

Lemma 5.4.3. *Let the random variable Y be normally distributed with mean μ and variance σ^2, i.e., $Y \sim \mathcal{N}(\mu,\sigma^2)$, then*

$$\mathbb{E}^{\mathbb{P}}\left[Y^2\,\mathbf{1}_{\{Y>0\}}\right] = \left(\mu^2 + \sigma^2\right)\left(1 - \Phi(-\mu/\sigma)\right) + \frac{\mu\,\sigma}{\sqrt{2\,\pi}} \exp\left(-\mu^2/2\,\sigma^2\right). \tag{5.4.4}$$

For a proof, see Appendix C.8. Note that from (5.4.4), we also have that

$$\mathbb{E}^{\mathbb{P}}\left[Y^2\,\mathbf{1}_{\{Y<0\}}\right] = \left(\mu^2 + \sigma^2\right)\Phi(-\mu/\sigma) - \frac{\mu\,\sigma}{\sqrt{2\,\pi}} \exp\left(-\mu^2/2\,\sigma^2\right). \tag{5.4.5}$$

Define the function h as follows:

$$h(u,\hat{\theta}_i) := \sum_{i=1}^{2} \mathbb{E}^{\mathbb{P}}\left[\left(Z_u^i - r - c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i - r >\ c/\sqrt{1+u}\}}\right]$$
$$+\,\mathbb{E}^{\mathbb{P}}\left[\left(Z_u^i - r + c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i - r < -c/\sqrt{1+u}\}}\right]. \tag{5.4.6}$$

Note that from (5.4.2), we have that

$$\bar{H}(t,\hat{\theta}) = \sum_{i=1}^{2} \frac{1}{2\,\sigma_i^2} \int_t^T h(u,\hat{\theta}_i)\,du\,. \tag{5.4.7}$$

Next, we calculate the gradient $\partial_{\hat{\theta}_i} h$ in the next Lemma.

Lemma 5.4.4. ***Properties of the function*** h**.** *Let the function h be as in (5.4.6). Then, h is $C^{1,2}([0,T]\times\mathbb{R})$ and the gradient $\partial_{\hat{\theta}_i} h$ is given by*

$$\partial_{\hat{\theta}_i} h = \left[2f(1-\Phi(-f/\boldsymbol{\sigma}_i)) + 2\boldsymbol{\sigma}_i\phi(-f/\boldsymbol{\sigma}_i) + 2\tilde{f}\Phi\left(-\tilde{f}/\boldsymbol{\sigma}_i\right) - 2\Sigma_i\phi\left(-\tilde{f}/\boldsymbol{\sigma}_i\right)\right]\partial_{\hat{\theta}_i} f, \tag{5.4.8}$$

where

$$f(u,\hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - \frac{c}{\sqrt{1+u}}\,, \qquad \tilde{f}(u,\hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + \frac{c}{\sqrt{1+u}}\,,$$

and when $L \neq 1/2$,

$$\boldsymbol{\sigma}_i(u) = \sqrt{\frac{L^2\,\sigma_i^2}{2\,L-1}\left(\frac{(1+u)^{2\,L-1}}{(1+u)^{2\,L}} - \frac{(1+t)^{2\,L-1}}{(1+u)^{2\,L}}\right)}\,,$$

and when $L = 1/2$,

$$\boldsymbol{\sigma}_i(u) = \frac{\sigma_i^2}{4}\left(\frac{\log(1+u) - \log(1+t)}{1+u}\right)\,.$$

For a proof, see Appendix C.9.

Note that the function $\boldsymbol{\sigma}_i$ depends on the value of t. The next Lemma shows the asymptotic formula of the function h when the value of the parameter c is small.

Lemma 5.4.5. ***Asymptotic analysis of the function*** h**.** *Let the function h be as in (5.4.6). Then, there exists a constant $\delta_{t,\hat{\theta}_i}$ and an integrable function $R(u,\hat{\theta}_i)$ with respect to u such that*

$$\left|\partial_{\hat{\theta}_i} h(u,\hat{\theta}_i) - 2\,(\hat{\theta}_i - r) - \frac{2\,c}{\sqrt{1+u}}\left[\Phi\left(-(\hat{\theta}_i - r)/\boldsymbol{\sigma}_i(u)\right) - \Phi\left((\hat{\theta}_i - r)/\boldsymbol{\sigma}_i(u)\right)\right]\right| \le c^2\,R(u,\hat{\theta}_i)\,, \tag{5.4.9}$$

for all $u \in [t,T]$ and for all $c < \delta_{t,\hat{\theta}_i}$, where $\delta_{t,\hat{\theta}_i}$ depends on t and $\hat{\theta}_i$.

For a proof, see Appendix C.10.

Note that in the Lemma above, it is necessary to show that the function R is integrable with respect to u because $\partial_{\hat{\theta}_i}\bar{H}$ is the integral of $\partial_{\hat{\theta}_i} h$; see (5.4.7).

Next, we show the asymptotic formula for the partial derivative of H with respect to $\hat{\theta}_i$, which is a direct result from Lemma 5.4.5.

Lemma 5.4.6. ***Properties of the function*** $\bar{H}$. *Let the function* $\bar{H}$ *be as in* (5.4.7). *Then,* $\bar{H}$ *is* $C^{1,2}([0,T]\times\mathbb{R}^2)$ *and the gradient* $\partial_{\hat{\theta}_i}\bar{H}$ *is given by*

$$\partial_{\hat{\theta}_i}\bar{H} = \frac{1}{\sigma_i^2}\int_t^T (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{2L} du - \frac{c}{\sigma_i^2}\left(\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du\right) + \mathcal{O}(c^2)\,, \tag{5.4.10}$$

where

$$\begin{aligned}\boldsymbol{H}^L(t,u,\hat{\theta}_i) = \frac{1}{\sqrt{1+u}}\left(\frac{1+t}{1+u}\right)^{2L}\Bigg[&\Phi\left(-(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{L}/\boldsymbol{\sigma}_i(u)\right)\\ &-\Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{L}/\boldsymbol{\sigma}_i(u)\right)\Bigg].\end{aligned} \tag{5.4.11}$$

Proof. Recall the definition of the functions $\bar{H}$ and h and write

$$\partial_{\hat{\theta}_i}\bar{H}(t,\hat{\theta}_i) = \frac{1}{2\,\sigma_i^2}\int_t^T \partial_{\hat{\theta}_i} h(u,\hat{\theta}_i)\,du\,.$$

Therefore, from Lemma 5.4.5, equation (5.4.10) holds because the function $\tilde{R}$ is integrable.

□

Note that the term

$$\left[\Phi\left(-(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{L}/\boldsymbol{\sigma}_i(u)\right) - \Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{L}/\boldsymbol{\sigma}_i(u)\right)\right], \tag{5.4.12}$$

is positive when $\hat{\theta}_i - r < 0$ and negative when $\hat{\theta}_i - r > 0$. Therefore, $\boldsymbol{H}^L(t,u,\hat{\theta}_i)$ is positive when $\hat{\theta}_i - r < 0$ and negative when $\hat{\theta}_i - r > 0$.

Given the asymptotic formula for $\partial_{\hat{\theta}_i}\bar{H}$ when the value of c is close to zero in (5.4.10) we look at three cases for $i = 1, 2$.

When $\hat{\theta}_i > r + c/\sqrt{1+t}$, the optimal investment strategy of the adaptive-robust problem in (5.2.3) is given by

$$\begin{aligned}\alpha_t^{adr,i} &= \frac{(\hat{\theta}_i - r - c/\sqrt{1+t}) - \beta_t\,\sigma_i^2\,\partial_{\hat{\theta}_i}\bar{H}(t,\hat{\theta}_i)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\\ &= \frac{(\hat{\theta}_i - r)\left(1 - \frac{L}{t+1}\int_t^T\left(\frac{1+t}{1+u}\right)^{2L} du\right) - \frac{c}{\sqrt{1+t}} + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,.\end{aligned} \tag{5.4.13}$$

And when $\hat{\theta}_i < r - c/\sqrt{1+t}$, we have that

$$\alpha_t^{adr,i} = \frac{(\hat{\theta}_i - r)\left(1 - \frac{L}{t+1}\int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du\right) + \frac{c}{\sqrt{1+t}} + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}. \tag{5.4.14}$$

And when $|\ \hat{\theta}_i - r\ | < c/\sqrt{1+t}$, we have that

$$\alpha_t^{adr,i} = \frac{(\hat{\theta}_i - r)\left(-\frac{L}{t+1}\int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du\right) + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}. \tag{5.4.15}$$

The agent's adaptive-robust strategy employs the estimate of the drift $\hat{\theta}_i$, which is continuously updated. To understand the intuition of this strategy we compare it to the adaptive strategy in (5.4.3), which we repeat here for convenience:

$$\alpha_t^{ad,i} = \frac{(\hat{\theta}_i - r)\left(1 - \frac{L}{t+1}\int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du\right)}{\gamma\,\sigma_i^2\exp(r\,(T-t))}. \tag{5.4.16}$$

Thus, if the estimate of the drift is such that $\hat{\theta}_i > r + c/\sqrt{1+t}$, then the optimal adaptive-robust strategy in (5.4.16) invests less on the risky asset than in the adaptive strategy because the second term in (5.4.13) is negative. On the other hand, if $\hat{\theta}_i < r - c/\sqrt{1+t}$, then the optimal adaptive-robust strategy in (5.4.16) invests more on the risky asset than in the adaptive strategy in (5.4.16) because the second term in (5.4.14) is positive. When the value of the estimate of the drift parameter is such that $|\ \hat{\theta}_i - r\ | < c/\sqrt{1+t}$, the agent invests a small amount in the risky asset since the estimator does not to far from 0 to be greater or lesser than 0 with certainty.

In the remaining of this subsection, we first analyse the property of the function $\boldsymbol{H}^L$ when the value of the learning rate parameter L is large. Next, we analysis the adaptive strategy in (5.4.16) under the difference values of L. Then, we compare the investment strategy in (5.4.13), (5.4.14) and (5.4.15) with the adaptive strategy in (5.4.16) to understand the intuition of the adaptive-robust strategy.

5.4.1 Properties of the function $\boldsymbol{H}^L$

In this subsection, we explore the properties of the function $\boldsymbol{H}^L$ in (5.4.11), which allow us to obtain the intuition of the adaptive-robust strategy.

First, we show the convergence of the function $L\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du$ with respect to the parameter L, which is useful for the intuition of the adaptive-robust strategy.

Proposition 5.4.7. *For each $\hat{\theta}_i > r\,(< r)$ and t, the function $L\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du$ converges to 0 as L goes to ∞.*

Proof. It suffices to consider the case $\hat{\theta}_i > 0$. Note that when $\hat{\theta}_i > 0$, the function $\boldsymbol{H}^L(t,u,\hat{\theta})$ is negative and thus $L\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du$ is negative. We consider

$$\begin{aligned}
&\lim_{L\to\infty} L\,\boldsymbol{H}^L(t,u,\hat{\theta}_i) \\
&\quad = \lim_{L\to\infty} \frac{L}{\sqrt{1+u}} \left(\frac{1+t}{1+u}\right)^{2\,L} \cdot \lim_{L\to\infty} \left[1 - 2\,\Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{L} / \boldsymbol{\sigma}_i(u)\right)\right] \\
&\quad = 0
\end{aligned} \tag{5.4.17}$$

because both limits converge to 0. Since

$$\left|\frac{L}{\sqrt{1+u}} \left(\frac{1+t}{1+u}\right)^{2\,L} \cdot \left[1 - 2\,\Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{L} / \boldsymbol{\sigma}_i(u)\right)\right]\right| \le \frac{3\,L}{\sqrt{1+u}} \left(\frac{1+t}{1+u}\right)^{2\,L}$$

and note that the function

$$\frac{3\,L}{\sqrt{1+u}} \left(\frac{1+t}{1+u}\right)^{2\,L}$$

is integrable from t to T, therefore by Dominated convergence theorem,

$$L\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du \to 0\,,$$

as $L \to \infty$. □

While in this Section we assume that stock prices are uncorrelated, one can proceed as above to derive the asymptotic investment strategy when prices are correlated.

5.4.2 Implication of the adaptive-robust strategy

Here, we explore the formulation of the adaptive-robust strategy and provide an intuition of the strategy in (5.4.13), (5.4.14), and (5.4.15) as a function of the learning rate parameter L. We denote $\Delta\alpha_t^i := \alpha_t^{adr,i} - \alpha_t^{ad,i}$. Then, when $\hat{\theta}_i > r + c/\sqrt{1+t}$, we have that

$$\Delta\alpha_t^i = \frac{-\frac{c}{\sqrt{1+t}} + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,. \tag{5.4.18}$$

Since $\boldsymbol{H}^L(t,u,\hat{\theta}_i)$ is negative when $\hat{\theta}_i > r + c/\sqrt{1+t}$, then $\Delta\alpha_t^i$ is negative. Thus, compared with the adaptive strategy, the adaptive-robust strategy has a less long position on the risky asset when $\hat{\theta}_i > r + c/\sqrt{1+t}$. Intuitively, this is because the agent adjusts the strategy to compensate the positive estimation of the estimator.

When $\hat{\theta}_i < r - c/\sqrt{1+t}$, we have that

$$\Delta\alpha_t^i = \frac{\frac{c}{\sqrt{1+t}} + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,. \tag{5.4.19}$$

Since $\boldsymbol{H}^L(t,u,\hat{\theta}_i)$ is positive when $\hat{\theta}_i < r - c/\sqrt{1+t}$, then $\Delta\alpha_t^i$ is positive, which implies that the adaptive-robust strategy has a less short position the risky asset when $\hat{\theta}_i < r - c/\sqrt{1+t}$. Intuitively, this is because the agent adjusts the strategy to compensate the negative estimation of the estimator.

When $\left|\hat{\theta}_i - r\right| < c/\sqrt{1+t}$, the difference between the adaptive-robust strategy and the adaptive strategy is

$$\Delta\alpha_t^i = \frac{-(\hat{\theta}_i - r) + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,. \tag{5.4.20}$$

This implies that when $0 < \hat{\theta}_i - r < c/\sqrt{1+t}$, the term $\Delta\alpha_t^i$ is negative. Therefore, the agent of the adaptive-robust strategy has a less long position than the adaptive strategy. Similarly, when $0 > \hat{\theta}_i - r > -c/\sqrt{1+t}$, the term $\Delta\alpha_t^i$ is positive. Therefore, the agent of the adaptive-robust strategy has a less short position than the adaptive strategy.

In both cases, the adaptive-robust strategy is more conservative in the strategy than the adaptive strategy when the estimator is not far of from the interest rate r.

5.5 Numerical result and its implication

In this Section, we study the numerical result of our utility maximisation under adaptive-robust framework in (5.2.3). From Bhudisaksang and Cartea (2021a), the value function v can be characterised by the equation (5.2.17). In general, the value function v can be solve numerically using the standard numerical PDE techniques. However, in the multidimensional cases, most of numerical schemes become infeasible. Therefore, the simplification of the PDE or the reduction of the dimension of PDE is required.

In this chapter, we approach the numerical method by using the stochastic representation to simplify the numerical method. In particular, instead of solving PDE in (5.2.17) numerically, we only need to evaluate the integral of the term (5.4.8) . This allows us to obtain the numerical solution of the value function v. Note that this simplification is very crucial because the numerical schemes are usually not feasible for this PDE problem.

In this Section, we aim to study two main objectives from the numerical results. First, we would like to understand the implication of the optimal strategy of the adaptive-robust problem under difference values of the uncertainty parameter c. Second, we compare the performance of the adaptive-robust strategies with that of three strategies in which the agent: knows the true value of the drift parameter, employs a wrong value of the drift parameter, and employs a robust strategy. In the robust strategy, the agent uses the framework derived above, but does not learn the value of the unknown parameters.

5.5.1 The implication of the adaptive-robust strategies

In this subsection, we explore the adaptive-robust control strategy of the utility maximisation. We plot the optimal strategy with difference values of c to obtain intuitive view on each hyper-parameters.

Here, we numerically calculate the optimal strategy in (5.2.20). Here, we assume that $T = 5$, $\sigma = 0.3$, $r = 0.01$, and $\tilde{c} = 1$. We plot the following four plots with difference values of $t = 0, 1, 2$ and 4.

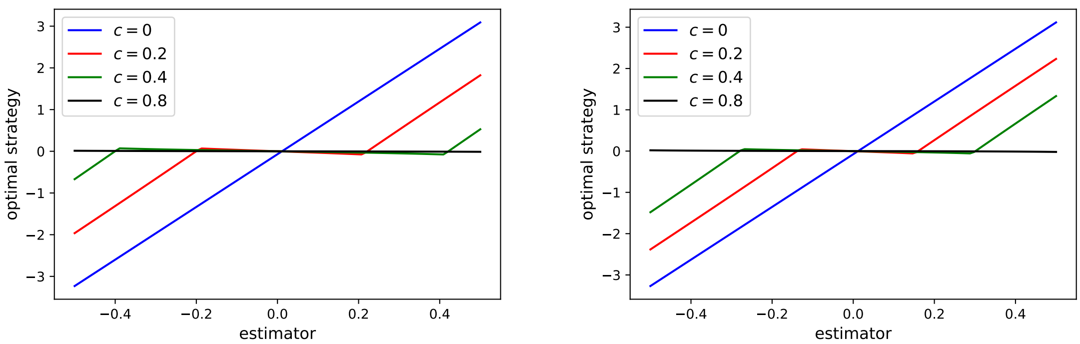

Figure 5.4: The optimal strategy when $t = 0$ and $t = 1$ plot

As expected, we see that when $c = 0$, from the figures 5.4 and 5.5, the optimal strategy is linearly depended on the estimator. When $c > 0$, the agent truncates the optimal strategy to almost zero (there is a small term left) when the estimator is not to far from r. Also, given everything else equals, as c increases, the truncation area is expanding as the agent has a larger value of uncertainty parameter. Moreover, given

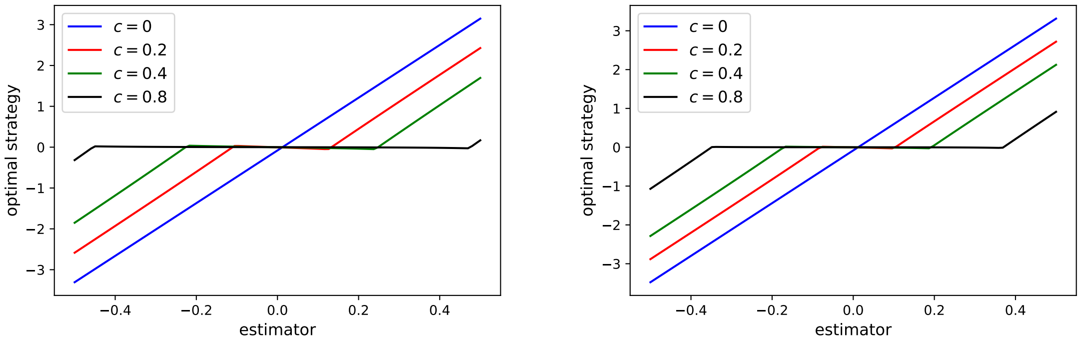

Figure 5.5: The optimal strategy when $t = 2$ and $t = 4$ plot

everything else equals, as t increases, the truncation area is shrinking as the agent has more information to estimate the parameter and thus reduces the uncertainty area.

5.5.2 The performance of the adaptive-robust strategies

In this subsection, we compare the performance of the adaptive-robust strategies with that of three strategies in which the agent: knows the true value of the drift parameter, employs a wrong drift parameter, and employs a robust strategy. In the robust strategy, the agent uses the framework derived above, but does not learn the value of the unknown parameter. Instead, the agent assumes that the true parameter $\theta^* \in [\underline{\theta}, \overline{\theta}]$ where $\underline{\theta}$ is the lowest possible value for the the true parameter and $\overline{\theta}$ is the highest possible value for the the true parameter and solves

$$v(t,x,y) = \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}} J(t,x,y,\mathbb{P},\alpha)\,, \tag{5.5.1}$$

where the set $\mathcal{P}$ contains all probability measure $\mathbb{P}_{\tilde{\theta}}$, such that $\tilde{\theta}_u \in [\underline{\theta}, \overline{\theta}]$ for all $u \in [t, T]$.

Here, we assume that the set of possible values of the parameter $\theta^{*\,1}$ is in $[0.03, 0.20]$, $\theta^{*\,2}$ is in $[0, 0.20]$, the terminal time is $T = 10$ minutes and $L = 1$ and $c = 0.1$. The other model parameters are:

$$X_0 = 1,\ S_0^1 = 10,\ S_0^2 = 5,\ \gamma = 0.5,\ \theta^{*,1} = 0.09,\ \theta^{*,2} = 0.03,\ r = 0.01\,,$$

and

$$\hat{\theta}_0^1 = 0.14,\ \hat{\theta}_0^2 = 0,\ \sigma_1 = 0.1,\ \sigma_2 = 0.3\,.$$

We compare the investment strategy in (5.4.16) with that of an investor who knows the drift parameter θ^* to understand the intuition of the adaptive-robust strategy.

Recall from (5.2.16), the perfect-knowledge strategy is given by

$$\alpha_t^{*,1} = \frac{\theta^{*,1} - r}{\gamma\,\sigma_1^2\,\exp(r\,(T-t))} \quad \text{and} \quad \alpha_t^{*,2} = \frac{\theta^{*,2} - r}{\gamma\,\sigma_2^2\,\exp(r\,(T-t))}\,. \tag{5.5.2}$$

we repeat here for convenience for a general drift parameter θ:

$$\alpha_t^{*,1} = \frac{\theta^{1} - r}{\gamma\,\sigma_1^2\,\exp(r\,(T-t))} \quad \text{and} \quad \alpha_t^{*,2} = \frac{\theta^{2} - r}{\gamma\,\sigma_2^2\,\exp(r\,(T-t))}\,. \tag{5.5.3}$$

As above, we compare the adaptive-robust strategy in (5.4.16) with the following: i) the agent employs (5.5.2) with the true drift $\theta = \theta^*$, ii) the agent employs the robust strategy without learning in (5.5.1), in which case the investment strategy is as in (5.5.3) with $\theta^1 = 0.03$ and $\theta^2 = 0$, iii) the agent employs (5.5.3) with the wrong drift parameters $\theta^1 = 0.15$ and $\theta^2 = 0.17$.

We discretise the time space into 8,000 time steps and employ 1,000 simulations to analyse the performance of the four acquisition strategies. The left panel of Figure 5.6 shows the agent's mean wealth process. The false strategy overestimates the expected growth of the risky asset so it over-invests on the stock, thus the mean value of the portfolio is higher than that of the other three strategies. On the other hand, the right panel of the figure shows that the standard deviation of the wealth for the false strategy is also the highest. Overall, the false strategy is the worst because the agent is risk averse – see Figure 5.7. Finally, note that by the terminal date of the trading horizon, the highest value function is that of the perfect-knowledge strategy followed by that of the adaptive-robust strategy.

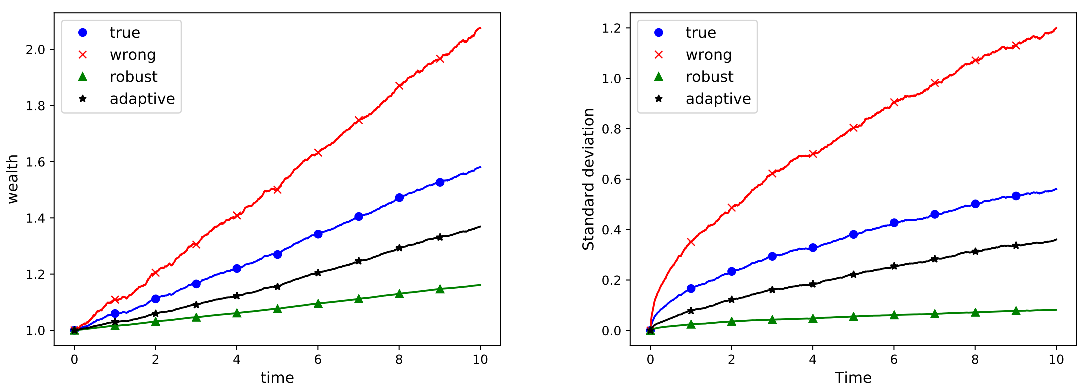

Figure 5.6: mean wealth process and standard deviation of the wealth process plot

From figure (5.6), we see that the wealth process of the correct parameter has a similar performance with the adaptive-robust. Although the false optimal control gives us the higher mean of the wealth process, it has a much higher standard deviation compare to the adaptive-robust portfolio and true portfolio.

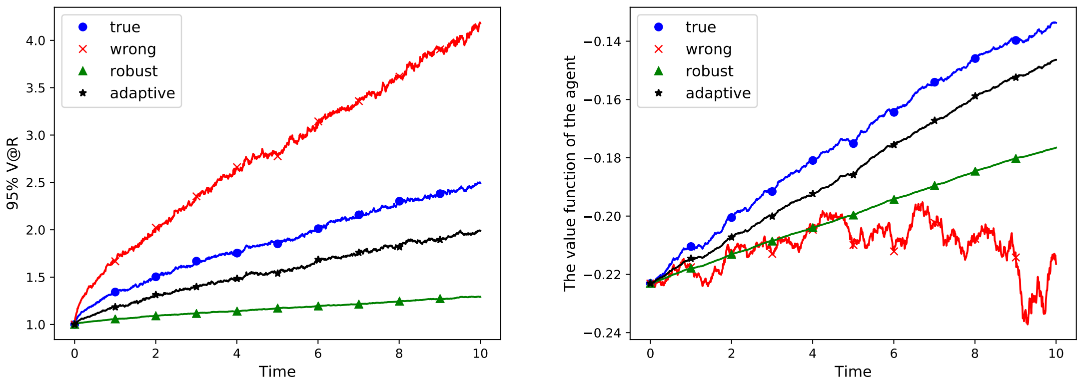

Figure 5.7: 95 percent value at risk and value function plot

From figure (5.7), we see that the wealth process of the false portfolio has a high value at risk. The value function shows that the adaptive-robust control under performance all portfolio when time is close to an initial time since the agent does not have a knowledge much about the unknown parameter. As time increasing, the value function of the adaptive-robust portfolio increases significantly as the adaptive-robust agent has a better estimation of the unknown parameter.

Bibliography

Bannör, Karl et al. (2016). "Parametric model risk and power plant valuation". In: *Energy Economics* 59, pp. 423–434.

Bayraktar, Erhan and Song Yao (2014). "On the robust optimal stopping problem". In: *SIAM Journal on Control and Optimization* 52(5), pp. 3135–3175.

Bergen, V et al. (2018). "Robust multivariate portfolio choice with stochastic covariance in the presence of ambiguity". In: *Quantitative Finance* 18(8), pp. 1265–1294.

Bertsekas, Dimitri and Steven Shreve (1996). "Stochastic Optimal Control: The Discrete-Time Case". In.

Bertsekas, Dimitri and John Tsitsiklis (1996). *Neuro-dynamic Programming*. Athena Scientific.

Bertsekas, Dimitri and John Tsitsiklis (2000). "Gradient convergence in gradient methods with errors". In: *SIAM Journal on Optimization* 10(3), pp. 627–642.

Bhudisaksang, Theerawat and Álvaro Cartea (2021a). "Adaptive Robust Control in Continuous Time". In: *SIAM Journal on Control and Optimization* 59(5), pp. 3912–3945.

Bhudisaksang, Theerawat and Álvaro Cartea (2021b). "Online drift estimation for jump-diffusion processes". In: *Bernoulli* 27(4), pp. 2494–2518.

Biagini, Sara and Mustafa Ç Pınar (2017). "The robust Merton problem of an ambiguity averse investor". In: *Mathematics and Financial Economics* 11(1), pp. 1–24.

Bielecki, Tao Chen, and Igor Cialenco (2017a). "Recursive construction of confidence regions". In: *Electronic Journal of Statistics* 11(2), pp. 4674–4700.

Bielecki, Tao Chen, and Igor Cialenco (2017b). "Recursive construction of confidence regions". In: *Electronic Journal of Statistics* 11(2), pp. 4674–4700.

Bielecki, Tao Chen, Igor Cialenco, et al. (2017). "Adaptive Robust Control Under Model Uncertainty". In: *arXiv preprint arXiv:1706.02227*.

Bion-Nadal, Jocelyne (2009). "Time consistent dynamic risk processes". In: *Stochastic Processes and their Applications* 119(2), pp. 633–654.

Bismuth, Alexis, Olivier Guéant, and Jiang Pu (2019). "Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty". In: *Mathematics and Financial Economics* 13(4), pp. 661–719.

Bordigoni, Giuliana, Anis Matoussi, and Martin Schweizer (2007). "A stochastic control approach to a robust utility maximization problem". In: *Stochastic analysis and applications*. Springer, pp. 125–151.

Cartea, Álvaro, Ryan Donnelly, and Sebastian Jaimungal (2017). "Algorithmic trading with model uncertainty". In: *SIAM Journal on Financial Mathematics* 8(1), pp. 635–671.

Cartea, Álvaro and Sebastian Jaimungal (2017). "Irreversible Investiments and Ambiguity Aversion". In: *International Journal of Theoretical and Applied Finance* 20(07), p. 1750044.

Cartea, Álvaro, Sebastian Jaimungal, and José Penalva (2015). *Algorithmic and high-frequency trading*. Cambridge University Press.

Cartea, Álvaro and Dimitrios Karyampas (2012). "Assessing the Performance of Different Volatility Estimators: A Monte Carlo Analysis". In: *Applied Mathematical Finance* 19(6), pp. 535–552.

Cartea, Álvaro and Leandro Sánchez-Betancourt (2021). "The shadow price of latency: Improving intraday fill ratios in foreign exchange markets". In: *SIAM Journal on Financial Mathematics* 12(1), pp. 254–294.

Chen, Tao and Michael Ludkovski (2019). "A Machine Learning Approach to Adaptive Robust Utility Maximization and Hedging". In: *arXiv preprint arXiv:1912.00244*.

Cohen, Samuel N and Robert James Elliott (2015). *Stochastic calculus and applications*. Vol. 2. Springer.

De Franco, Carmine, Johann Nicolle, and Pham (2019). “Bayesian learning for the Markowitz portfolio selection problem”. In: *International Journal of Theoretical and Applied Finance* 22(07), p. 1950037.

Delyon, Bernard (1996). “General results on the convergence of stochastic algorithms”. In: *IEEE Transactions on Automatic Control* 41(9), pp. 1245–1255.

Denis, Laurent and Magali Kervarec (2013). “Optimal investment under model uncertainty in nondominated models”. In: *SIAM Journal on control and optimization* 51(3), pp. 1803–1822.

Ekren, Ibrahim, Nizar Touzi, and Jianfeng Zhang (2014). “Optimal stopping under nonlinear expectation”. In: *Stochastic Processes and Their Applications* 124(10), pp. 3277–3311.

El Karoui, Nicole and Xiaolu Tan (2013a). “Capacities, measurable selection and dynamic programming Part I: abstract framework”. In: *arXiv preprint arXiv:1310.3363*.

El Karoui, Nicole and Xiaolu Tan (2013b). “Capacities, measurable selection and dynamic programming part ii: application in stochastic control problems”. In: *arXiv preprint arXiv:1310.3364*.

Epstein, Larry G and Shaolin Ji (2014). “Ambiguous volatility, possibility and utility in continuous time”. In: *Journal of Mathematical Economics* 50, pp. 269–282.

Errami, Mohammed, Francesco Russo, and Pierre Vallois (2002). “Itô’s formula for $C^{1,\lambda}$-functions of a càdlàg process and related calculus”. In: *Probability theory and related fields* 122(2), pp. 191–221.

Fadina, Tolulope, Ariel Neufeld, and Thorsten Schmidt (2018). “Affine processes under parameter uncertainty”. In: *arXiv preprint arXiv:1806.02912*.

Föllmer, Hans, Philip Protter, and Albert N Shiryayev (1995). “Quadratic covariation and an extension of Itô’s formula”. In: *Bernoulli*, pp. 149–169.

Gloter, Arnaud, Dasha Loukianova, and Hilmar Mai (2018). “Jump filtering and efficient drift estimation for Lévy-driven SDEs”. In: *The Annals of Statistics* 46(4), pp. 1445–1480.

Hansen, Lars Peter and Thomas J Sargent (2011). “Robustness and ambiguity in continuous time”. In: *Journal of Economic Theory* 146(3), pp. 1195–1223.

Hollender, Julian (2016). *Lévy-Type Processes under Uncertainty and Related Nonlocal Equations*. Technische Universität Dresden.

Ismail, Amine and Pham (2019). “Robust Markowitz mean-variance portfolio selection under ambiguous covariance matrix”. In: *Mathematical Finance* 29(1), pp. 174–207.

Jacod, Jean and Philip Protter (1988). “Time reversal on Lévy processes”. In: *The Annals of Probability* 16(2), pp. 620–641.

Jaimungal, Sebastian and Georg Sigloch (2012). “Incorporating risk and ambiguity aversion into a hybrid model of default”. In: *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 22(1), pp. 57–81.

Kessler, Mathieu (1997). “Estimation of an ergodic diffusion from discrete observations”. In: *Scandinavian Journal of Statistics* 24(2), pp. 211–229.

Kim, Tong Suk and Edward Omberg (1996). “Dynamic nonmyopic portfolio behavior”. In: *The Review of Financial Studies* 9(1), pp. 141–161.

Kubińska, E (2005). “Approximation of Carathéodory functions and multifunctions”. In: *Real Analysis Exchange* 30(1), pp. 351–359.

Kulik, Justin and Veretennikov (2012). “Extended Poisson equation for weakly ergodic Markov processes”. In: *Theory of Probability and Mathematical Statistics* 85, pp. 23–39.

Kunita, Hiroshi (2004). “Representation of martingales with jumps and applications to mathematical finance”. In: *Stochastic Analysis and Related Topics in Kyoto: In honour of Kiyosi Itô*. Mathematical Society of Japan, pp. 209–232.

Kutoyants, Yury A (2013). *Statistical inference for ergodic diffusion processes*. Springer Science & Business Media.

Kwon, Youngmff and Chaniio (1999). “Strong Feller property and irreducibility of diffusions with jumps”. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 67(1-2), pp. 147–157.

Lee, Suzanne S and Per A Mykland (2008). "Jumps in financial markets: A new nonparametric test and jump dynamics". In: *The Review of Financial Studies* 21(6), pp. 2535–2563.

Levanony, David, Adam Shwartz, and Ofer Zeitouni (1994). "Recursive identification in continuous-time stochastic processes". In: *Stochastic Processes and their Applications* 49(2), pp. 245–275.

Liang, Zongxia and Ming Ma (2018). "Robust consumption-investment problem Under CRRA and CARA utilities with time-varying confidence sets". In: *arXiv preprint arXiv:1811.12149*.

Liu, Ren and Johannes Muhle-Karbe (2013). "Portfolio choice with stochastic investment opportunities: A User's guide". In: *arXiv preprint arXiv:1311.1715*.

Lowther, George (2010). "Nondifferentiable functions of one-dimensional semimartingales". In: *The Annals of Probability* 38(1), pp. 76–101.

Masuda, Hiroki (2007). "Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps". In: *Stochastic processes and their applications* 117(1), pp. 35–56.

Masuda, Hiroki (2008). "On stability of diffusions with compound-Poisson jumps". In: *Bulletin of Informatics and Cybernetics* 40, p. 61.

Merton, Robert C (1975). "Optimum consumption and portfolio rules in a continuous-time model". In: *Stochastic Optimization Models in Finance*. Elsevier, pp. 621–661.

Metivier, Michel and Pierre Priouret (1984). "Applications of a Kushner and Clark lemma to general classes of stochastic algorithms". In: *IEEE Transactions on Information Theory* 30(2), pp. 140–151.

Meyn, Sean P and Richard L Tweedie (1993). "Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes". In: *Advances in Applied Probability* 25(3), pp. 518–548.

Neufeld, Ariel and Marcel Nutz (2014). "Measurability of semimartingale characteristics with respect to the probability law". In: *Stochastic Processes and their Applications* 124(11), pp. 3819–3845.

Neufeld, Ariel and Marcel Nutz (2018). "Robust utility maximization with Lévy processes". In: *Mathematical Finance* 28(1), pp. 82–105.

Nowak, Andrzej (2010). "On measurable minimax selectors". In: *Journal of Mathematical Analysis and Applications* 366(1), pp. 385–388.

Nutz, Marcel and Ramon van Handel (2013). "Constructing sublinear expectations on path space". In: *Stochastic processes and their applications* 123(8), pp. 3100–3121.

Papaioannou, Denis (2012). "Applied multidimensional Girsanov theorem". In: *Available at SSRN 1805984*.

Pardoux, Étienne and Veretennikov (2001). "On the Poisson equation and diffusion approximation I". In: *The Annals of Probability* 29(3), pp. 1061–1085.

Pardoux, Étienne and Veretennikov (2003). "On Poisson equation and diffusion approximation II". In: *The Annals of Probability* 31(3), pp. 1166–1192.

Peng, Shige and Xuehong Zhu (2006). "Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations". In: *Stochastic Processes and their Applications* 116(3), pp. 370–380.

Pham (1998). "Optimal stopping of controlled jump diffusion processes: a viscosity solution approach". In: *Journal of Mathematical Systems, Estimation and Control*. Citeseer.

Pham (2009). *Continuous-time stochastic control and optimization with financial applications*. Vol. 61. Springer Science & Business Media.

Pham and Xiaoli Wei (2017). "Dynamic Programming for Optimal Control of Stochastic McKean–Vlasov Dynamics". In: *SIAM Journal on Control and Optimization* 55(2), pp. 1069–1101.

Pham, Xiaoli Wei, and Chao Zhou (2018). "Portfolio diversification and model uncertainty: a robust dynamic mean-variance approach". In: *arXiv preprint arXiv:1809.01464*.

Protter, Philip (1990). *Stochastic integration and differential equations: a new approach*. Springer.

Rao, B.L.S.P (1999). *Statistical inference for diffusion type processes*. Arnold.

Revuz, Daniel (2008). *Markov chains.* Vol. 11. Elsevier.

Sirignano, Justin and Konstantinos Spiliopoulos (2017). “Stochastic gradient descent in continuous time”. In: *SIAM Journal on Financial Mathematics* 8(1), pp. 933–961.

Sirignano, Justin and Konstantinos Spiliopoulos (2020). “Stochastic gradient descent in continuous time: A central limit theorem”. In: *Stochastic Systems* 10(2), pp. 124–151.

Skiadas, Costis (2013). “Smooth ambiguity aversion toward small risks and continuous-time recursive utility”. In: *Journal of Political Economy* 121(4), pp. 775–792.

Soner, H Mete and Nizar Touzi (2002). “Dynamic programming for stochastic target problems and geometric flows”. In: *Journal of the European Mathematical Society* 4(3), pp. 201–236.

Sorensen, Michael (1991). “Likelihood methods for diffusions with jumps”. In: *Statistical inference in stochastic processes.*

Surace, Simone Carlo and Jean-Pascal Pfister (2018). “Online Maximum Likelihood Estimation of the Parameters of Partially Observed Diffusion Processes”. In: *IEEE Transactions on Automatic Control.*

Uehara, Yuma (2018). “Statistical inference for misspecified ergodic Lévy driven stochastic differential equation models”. In: *Stochastic Processes and their Applications.*

Wang, Jian (2010). “Regularity of semigroups generated by Lévy type operators via coupling”. In: *Stochastic Processes and their Applications* 120(9), pp. 1680–1700.

Xia, Yihong (2001). “Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation”. In: *The Journal of Finance* 56(1), pp. 205–246.

Yoshida, Nakahiro (1992). “Estimation for diffusion processes from discrete observation”. In: *Journal of Multivariate Analysis* 41(2), pp. 220–242.

Appendix A

Appendix for Chapter 3

A.1 Additional proof

A.1.1 Proof of Lemma 3.1.9

The existence and uniqueness of the invariant measure π^{θ^*} and the existence of moments of all orders of the process X follow directly from Lemma 2.1 in Gloter, Loukianova, and Mai (2018). By Condition 3.1.3, the Condition (CD3) in Meyn and Tweedie (1993) holds when $V(x) = |x|^p$. Due to Condition 3.1.2 and Theorem 6.1 in Meyn and Tweedie (1993), there exist positive constants C_1 and C_2, such that

$$\left\|\mathbb{P}_{\theta^*,t}(x,\cdot) - \pi^{\theta^*}(\cdot)\right\|_{1+|x|^p} = \left\|\mathbb{P}_{\theta^*,t}(x,\cdot) - \pi^{\theta^*}(\cdot)\right\|_f \leq C_1\, f(x)\, e^{-C_2\, t}\,, \tag{A.1.1}$$

where $f(x) = 1 + |x|^p$. This proves the first part of Lemma 3.1.9. Next, we use the comparison argument on the stochastic differential equation with Condition 3.1.7, see Appendix A.1 in Sirignano and Spiliopoulos (2020) and Theorem 3.1 in Peng and Zhu (2006), to show that the process θ_t has moments of all orders. For the last two inequalities, we notice that

$$X_{t^-}^p = \liminf_{s\to t^-} X_s^p\,. \tag{A.1.2}$$

Then, by using Fatou's lemma,

$$\mathbb{E}_{x,\theta^*}\left[|X_{t^-}|^p\right] \leq \liminf_{s\to t^-} \mathbb{E}_{x,\theta^*}\left[|X_s|^p\right]\,. \tag{A.1.3}$$

Therefore, $\sup_{t>0} \mathbb{E}_{x,\theta^*}\left[|X_{t^-}|^p\right] < \infty$. Similarly, we can show that $\sup_{t>0} \mathbb{E}_{x,\theta^*}\left[|\theta_{t^-}|^p\right] < \infty$.

A.1.2 Proof of Lemma 3.1.13

We recall the notation from Wang (2010). Under the non-degenerate assumption on the diffusion part of the process X in (3.0.1), there exists $\lambda \geq 0$, such that

$\langle \sigma(x)\,\sigma(x)^{\intercal}\,h,\,h\rangle \geq \lambda\,|h|^2$ for all x and h. Denote by σ_λ the unique symmetric nonnegative definite matrix-valued function, such that $\sigma_\lambda^2 = \sigma^2 - \lambda\,I$, where I is the identity matrix. For x and y, define

$$\begin{aligned}
&\alpha(x,y)_\lambda = \tfrac{1}{2}\left\|\sigma_\lambda(x) - \sigma_\lambda(y)\right\|^2, \quad B(x,y,\theta^*) = \langle b(x,\theta^*) - b(y,\theta^*), x-y\rangle, \\
&\left\|\nu(x,\cdot) - \nu(y,\cdot)\right\|_{\mathrm{Var},>1} = \int_{|z|>1} \left|\nu(x,dz) - \nu(y,dz)\right|, \\
&\left\|\tilde{\nu}(x,\cdot) - \tilde{\nu}(y,\cdot)\right\|_{\mathrm{Var},\geq 1} = \frac{1}{2}\int_U \left|\xi(x,z) - \xi(y,z)\right|^2 M(dz)\,.
\end{aligned} \tag{A.1.4}$$

We have that $\left\|\nu(x,\cdot) - \nu(y,\cdot)\right\|_{\mathrm{Var},>1} = 0$ because the Lévy measure $\nu(\cdot,\cdot)$ does not depend on the first variable. Moreover, the value $\left\|\tilde{\nu}(x,\cdot) - \tilde{\nu}(y,\cdot)\right\|_{\mathrm{Var},\geq 1}$ is bounded by $C\,|x-y|^2$ for some positive constant C. Let

$$K(x,y,\theta^*) = \frac{\alpha_\lambda(x,y) + B(x,y,\theta^*) + \left\|\tilde{\nu}(x,\cdot) - \tilde{\nu}(y,\cdot)\right\|_{\mathrm{Var},\geq 1}}{|x-y|^2\,(1+|x-y|)} + \frac{\left\|\nu(x,\cdot) - \nu(y,\cdot)\right\|_{\mathrm{Var},>1}}{|x-y|}. \tag{A.1.5}$$

To apply Theorem A.1.4, it suffices to show that there exist $\delta \in (0,1)$ and a constant $C_\delta \geq 0$, such that for any $x, y \in \mathbb{R}^d$ with $|x-y| \leq \delta$,

$$(K(x,y,\theta^*) + C_\delta)\,|x-y| \leq \frac{2\,\lambda}{1+|x-y|^2}\,. \tag{A.1.6}$$

From Condition 3.1.6 let $\lambda = C_1/2$ where $C_1 \leq \sigma^2(x)$. From Condition 3.1.1, we have that

$$\begin{aligned}
\alpha_\lambda(x,y) = \frac{1}{2}\left\|\sigma_\lambda(x) - \sigma_\lambda(y)\right\|^2 &= \frac{1}{2}\left(\frac{\sigma_\lambda^2(x) - \sigma_\lambda^2(y)}{\sigma_\lambda(x) + \sigma_\lambda(y)}\right)^2 \\
&\leq C_L\left(\frac{\sigma(x) + \sigma(y)}{\sigma_\lambda(x) + \sigma_\lambda(y)}\right)^2 |x-y|^2 \leq \tilde{C}_1\,|x-y|^2\,,
\end{aligned} \tag{A.1.7}$$

where C_L is Lipschitz coefficient of $\sigma(x)$ and for all $|x-y| \leq 1$

$$B(x,y,\theta^*) = \langle b(x,\theta^*) - b(y,\theta^*), x-y\rangle \leq \tilde{C}_2\,|x-y|^2\,, \tag{A.1.8}$$

where $\tilde{C}_1$ and $\tilde{C}_2$ are positive constants that do not depend on x, y. We choose $\delta = \min\{\,2\,\lambda/(C + \tilde{C}_1 + \tilde{C}_2), 1\,\}$ and $C_\delta = 0$, therefore for all x and y, such that $|x-y| \leq \delta$, we have that

$$K(x,y,\theta^*) + C_\delta \leq \frac{C + \tilde{C}_1 + \tilde{C}_2}{1+|x-y|}\,|x-y| \leq \frac{\delta\,\left(C + \tilde{C}_1 + \tilde{C}_2\right)}{1+|x-y|^2} \leq \frac{2\,\lambda}{1+|x-y|^2}\,. \tag{A.1.9}$$

Thus, (A.1.5) holds and we use Theorem A.1.4 to show the locally Lipschitz property of the function $P_t f(x,\theta)$. Let x, y be any two points and define

$$\tilde{f}(u,\theta) := \begin{cases} f(m,\theta), & m \in [x,y]\,, \\ f(y,\theta), & m > y\,, \\ f(x,\theta), & m < x\,. \end{cases}$$

It is not hard to show that the function $\tilde{f}$ is Lipschitz with the constant bounded by $C\,(1+|x|^q+|y|^q)$. The result of the Lemma follows directly from Theorem A.1.4.

Lemma A.1.1. *Suppose that Conditions 3.1.1 to 3.1.6 hold. Then for all $p > 0$ and all t, there exists a constant $K > 0$, such that*

$$\mathbb{E}_x\left[\sup_{0\le s\le t}|X_s|^p\right] \le K\,t^{1/2}\,.$$

Proof. From Lemma 3.1.9, we have

$$\mathbb{E}_x|X_t|^{2p} \le C\,\left(1+|x|^{2p}\right).$$

Let $f(x) = |x|^{2\,p}$ and use Itô's Lemma to write

$$f(X_t) = f(x) + \int_0^t \mathcal{A}f(X_s)\,ds + G_t + J_t\,,$$

where $\mathcal{A}$ is the infinitesimal generator of the process X

$$\begin{aligned} G_t &:= \int_0^t \nabla f(X_s)\,\sigma(X_s)\,dB_s\,, \\ J_t &:= \int_0^t\int \mathbf{1}_{\{|z|\le 1\}}(z)(f(X_{s^-}+\xi(X_{s^-},z)) - f(X_{s^-}))\,\tilde{\mu}(ds,dz) \\ &\quad + \int_0^t\int \mathbf{1}_{\{|z|>1\}}(z)(f(X_{s^-}+\xi(X_{s^-},z)) - f(X_{s^-}))\,\tilde{\mu}(ds,dz) \\ &= J_t^1 + J_t^2, \end{aligned} \tag{A.1.10}$$

where J_t^1 and J_t^2 represent the two double integrals, respectively, in the equation above. Therefore, we have the following bound:

$$\mathbb{E}_x\left[\sup_{0\le s\le t} f(X_s)\right] \le f(x) + \int_0^t \mathbb{E}_x\left[\left|\mathcal{A}f(X_s)\right|\right]\,ds + \mathbb{E}_x\left[\sup_{0\le s\le t} G_t\right] + \mathbb{E}_x\left[\sup_{0\le s\le t} J_t\right]. \tag{A.1.11}$$

Next, consider the second, third, and fourth terms on the right-hand side of (A.1.11). The quantity $\left|\mathcal{A}f(X_s)\right|$ in the second term obeys the upper bound:

$$\left|\mathcal{A}f(X_s)\right| = 2\,p\,|X_s|^{2\,p-2}\,\langle X_s, b(X_s)\rangle + p\,|X_s|^{2\,p-2}\,\mathrm{Tr}\left[\left((p-2)\,[X_s^i X_s^j]_{i,j=1}^d |X_s|^{-2} + I_d\right)\right]\sigma(X_s)^2$$

$$
\begin{aligned}
&+ \int \left(f(X_s + \xi(X_s, z)) - f(X_s) - \nabla f(X_s)\, \xi(X_s, z)\, 1_{|z|\le 1}(z)\right)\, \nu(dz) \\
&\le C_1\, |X_s|^{2\,p}\,,
\end{aligned}
$$

where the last inequality follows from the condition 3.1.3. Therefore, by Lemma 3.1.9, we obtain that

$$
\int_0^t \mathbb{E}_x \left[\left| \mathcal{A} f(X_s) \right| \right]\, ds \le C_1\, t\, (1 + |x|^{2p}) \tag{A.1.12}
$$

To bound the third term, employ the BDG inequality to obtain

$$
\begin{aligned}
\mathbb{E}_x \left[\sup_{0\le s\le t} G_t \right] \le C\, \mathbb{E}_x \left[[G, G]_t^{1/2} \right] \le C\, \mathbb{E}_x \left[[G, G]_t \right]^{1/2} \\
= C\, \mathbb{E}_x \left[\int_0^t |X_s|^{4\,p-4}\; X_s^2\, \sigma^2(X_s)\, ds \right]^{1/2} \\
\le C_2\, \mathbb{E}_x \left[\int_0^t |X_s|^{4\,p-2}\; ds \right]^{1/2} \le C_3\, t^{1/2}\, (1 + |x|^{4\,p-2})^{1/2}\,,
\end{aligned} \tag{A.1.13}
$$

where C_3 is a constant that does not depend on x and t, the fourth inequality follows from the non degenerate assumption in condition 3.1.3 , and the last inequality follows from Lemma 3.1.9.

To bound the forth term, recall that $J_t = J_t^1 + J_t^2$ and apply the BDG inequality to the jump part J_t^1 and write

$$
\begin{aligned}
\mathbb{E}_x \left[\sup_{0\le s\le t} J_t^1 \right] &\le C\, \mathbb{E}_x \left[[J^1, J^1]_t^{1/2} \right] \\
&\le C\, \mathbb{E}_x \left[[J^1, J^1]_t \right]^{1/2} \\
&= C\, \mathbb{E}_x \left[\int_0^t \int \mathbf{1}_{\{|z|\le 1\}}(z)\, \left(f(X_{s^-} + \xi(X_{s^-}, z)) - f(X_{s^-})\right)^2\, \nu(dz)\, ds \right]^{1/2}.
\end{aligned}
$$

For $|z| \le 1$, from the elementary inequality $|x+y|^p \le 2^p \left(|x|^p + |y|^p\right)$, we have that

$$
\begin{aligned}
\left| f(X_{s^-} + \xi(X_{s^-}, z)) - f(X_{s^-}) \right| &\le C \left(|X_{s^-}|^{2\,p} + \left|\xi(X_{s^-}, z)\right|^{2\,p} \right) \\
&\le C_3 \left(|X_{s^-}|^{2\,p} + |X_{s^-}|^{2\,p} \right),
\end{aligned}
$$

where the last inequality follows from Condition 3.1.3. Therefore, based on the last two inequalities we write

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\le s\le t} J_t^1\right] &\le C\,\mathbb{E}_x\left[\int_0^t\int \mathbf{1}_{\{|z|\le 1\}}\,|X_{s-}|^{4p}\,\nu(dz)\,ds\right]^{1/2}\\ &\le C\left(\int_0^t \mathbb{E}_x|X_{s-}|^{4p}\,ds\right)^{1/2} \le C\,t^{1/2}\,(1+|x|^{4p})^{1/2}.\end{aligned}\tag{A.1.14}$$

Similarly, for the jump part J^2, by the BDG inequality, we have that

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\le s\le t} J_t^2\right] &= C\,\mathbb{E}_x\left[\int_0^t\int \mathbf{1}_{\{|z|>1\}}(z)\,(f(X_{s-}+\xi(X_{s-},z))-f(X_{s-}))^2\,\nu(dz)\,ds\right]^{1/2}\\ &\le C\,\mathbb{E}_x\left[\int_0^t\int \mathbf{1}_{\{|z|>1\}}(z)\left(|X_s|^{4p}+|\xi(X_{s-},z)|^{4p}\right)\,\nu(dz)\,ds\right]^{1/2}\\ &\le C\,\mathbb{E}_x\left[\int_0^t\int \mathbf{1}_{\{|z|>1\}}(z)\,|X_s|^{4p}\,(1+|z|^{4p})\,\nu(dz)\,ds\right]^{1/2}\\ &\le C\,t^{1/2}\,(1+|x|^{4p})^{1/2},\end{aligned}\tag{A.1.15}$$

because the value $\int |z|^{4p}\,\nu(dz)$ is finite from Condition 3.1.3.

Therefore, to finalise the proof of the Lemma, combine equations (A.1.12), (A.1.13), (A.1.14), (A.1.15) to show that for all $t>0$,

$$\mathbb{E}_x\left[\sup_{0\le s\le t}|X_s|^p\right] \le \mathbb{E}_x\left[\sup_{0\le s\le t}|X_s|^{2p}\right]^{1/2} \le K\,t^{1/2},$$

for some constant K.

□

Lemma A.1.2. ***Time reversibility****.* *The process (X_t,θ_t) is a time reversible process.*

Proof. First, we recall the result in Jacod and Protter (1988) that Lévy processes are time-reversible. We define the following notation

$$\hat{X}_t = X_{T-t} \quad \text{and} \quad \tilde{X}_t = \hat{X}_{t+}.$$

Then, the result follows directly from the proof of Theorem 4.6 in Errami, Russo, and Vallois (2002). □

Lemma A.1.3. ***Convergence of the function $\tilde{G}$***

$$\lim_{k\to\infty}\int_0^t \nabla_x G_k(X_s,\theta)\,d[X,B]_s = \int_0^t \tilde{G}(X_s,\theta)\,d[X,B]_s\,,\tag{A.1.16}$$

Proof. Let $H(x,\theta) := \lim_{k\to\infty} \nabla_x G_k(x,\theta)$. The function H is a weak derivative with respect to x of the function $G(x,\theta)$. Therefore, for each θ, $H(x,\theta) = \tilde{G}(x,\theta)$ almost everywhere. Moreover, we have that the functions $H(x,\theta)$ and $\nabla_x G_k(x,\theta)$ are uniformly dominated by $C\,(1+|x|^q)$ by an elementary calculation. Next, we apply the dominated convergence theorem to obtain

$$\lim_{k\to\infty} \int_0^t \nabla_x G_k(X_s,\theta)\, d[X,B]_s = \int_0^t H(X_s,\theta)\, d[X,B]_s\,, \tag{A.1.17}$$

$\mathbb{P}$-almost surely because the function $\nabla_\theta G_k$ is dominated by $C\,(1+|x|^q)$.

Now, it remains to show that

$$\int_0^t H(X_s,\theta)\, d[X,B]_s = \int_0^t \tilde{G}(X_s,\theta)\, d[X,B]_s\,.$$

By the Cauchy–Schwarz inequality, we have

$$\begin{aligned} &\mathbb{E}\left[\int_0^t \Big(H(X_s,\theta) - \tilde{G}(X_s,\theta)\Big)\, d[X,B]_s\right]^2 \\ &\qquad \le \mathbb{E}\left[\int_0^t \Big(H(X_s,\theta) - \tilde{G}(X_s,\theta)\Big)^2 d[X,B]_s\right] \mathbb{E}\left[\int_0^t \mathbf{1}_{\{X_t\notin \mathrm{diff}(G)\}}\, d[X,B]_s\right], \end{aligned} \tag{A.1.18}$$

where $\mathbf{1}_{\{x\notin \mathrm{diff}(G)\}}$ is 1 when the function G is differentiable at x and 0 otherwise. Note that the first expectation is finite because the function H and $\tilde{G}$ are dominated by $C\,(1+|x|^q)$ and the second expectation is equal to zero (see Theorem 1.2 in Lowther (2010)). This implies that

$$\int_0^t H(X_s,\theta)\, d[X,B]_s = \int_0^t \tilde{G}(X_s,\theta)\, d[X,B]_s \tag{A.1.19}$$

$\mathbb{P}$-almost surely. □

We use the theorem below (Theorem 2.2 in Wang (2010)) in the proof of Lemma 3.1.13.

Theorem A.1.4. *Suppose Conditions 3.1.1 to 3.1.7 hold. Let the function f be Lipschitz and let $t \ge 0$. If the process X_t follows (3.0.1), then*

$$\big|\mathbb{E}^x f(X_t) - \mathbb{E}^y f(X_t)\big| \le \begin{cases} (2+\delta^{-1})\exp(-C_\delta\, t)\, C(f)\, |x-y|, & |x-y| \le \delta \\ 2\,\delta^{-1}\, \|f\|\, |x-y|, & |x-y| > \delta, \end{cases}$$

where $\mathbb{E}^x f(X_t) = \mathbb{E}\left[f(X_t)\,|\,X_0 = x\right]$, $\|f\|$ is the supremum norm of the function f, and $C(f)$ depends on $\|f\|$ and the Lipschitz constant of f.

Definition A.1.5. Let ν be a (non-negative) measure on a measurable space (X, Σ). The support of ν is a subset of the σ-algebra Σ, which is given by

$$\operatorname{supp}(\nu) = \overline{\{A \in \Sigma \,|\, \nu(A) > 0\}}\,,$$

where $\overline{B}$ is the closure of the set B for $B \in \Sigma$.

Appendix B

Appendix for Chapter 4

B.1 Proofs

B.1.1 Proof of Lemma 4.1.5

We follow the proofs of Neufeld and Nutz (2014) and Fadina, Neufeld, and Schmidt (2018). From theorem 2.6 in Neufeld and Nutz (2014), the set $\mathfrak{P}^{ac}_{sem}(\Omega)$ is Borel measurable in $\mathfrak{P}(\Omega)$. Therefore, the set

$$\{(\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}^{ac}_{sem}(\Omega) \mid \mathbb{P}(\tilde{X}_t = \omega_t) = 1\}\,,$$

is Borel measurable. Recall that the processes $\gamma^{\mathbb{P}}$, a, and $F^{\mathbb{P}}_{\tilde{\omega},t}$ are the drift characteristic, the volatility characteristic, and the jump characteristic, respectively, of the process X_t under the probability measure $\mathbb{P}$. From Theorem 2.6 in Neufeld and Nutz (2014),

$$(\tilde{\omega}, s, \mathbb{P}) \mapsto (\tilde{\omega}, s, \gamma^{\mathbb{P}}_s(\tilde{\omega}), a_s(\tilde{\omega}), F^{\mathbb{P}}_{\tilde{\omega},t})$$

is Borel measurable. From Assumption 4.1.3, the set

$$\{(\tilde{\omega}, s, \mathbb{P}) \in \Omega\times[0, T]\times\mathfrak{P}^{ac}_{sem}(\Omega)|\ s > t, (\gamma^{\mathbb{P}}_s(\tilde{\omega}), a_s(\tilde{\omega}), F^{\mathbb{P}}_{\tilde{\omega},t}) \in \mathbf{b}^*(s, \tilde{\omega}_s)\times a_s(\tilde{\omega}_s)\times L_s(\tilde{\omega}_s)\}$$

is Borel measurable. Therefore, the set

$$\begin{aligned} E := \Big\{&(\omega, t, \mathbb{P}, \tilde{\omega}, s) \in \Omega \times [0, T] \times \mathfrak{P}^{ac}_{sem}(\Omega) \times \Omega \times [0, T]\Big| s > t,\\ &\mathbb{P}(\{\tilde{X}_s = \omega_t\ ,\ 0 \le s \le t\}) = 1, (\gamma^{\mathbb{P}}_s(\tilde{\omega}), a_s(\tilde{\omega}), F^{\mathbb{P}}_{\tilde{\omega},t}) \in \mathbf{b}^*(s, \tilde{\omega}_s) \times a_s(\tilde{\omega}_s) \times L_s(\tilde{\omega}_s)\Big\}\end{aligned}$$

is Borel measurable. Then, by Fubini's Theorem, see Appendix B.2.1 below, we obtain the Borel measurable mapping

$$(\omega, t, \mathbb{P}, \tilde{\omega}) \mapsto \int_0^T \mathbf{1}_E(\omega, t, \mathbb{P}, \tilde{\omega}, s)\, \mathbf{1}_{(t,T]}(s)\, ds\,.$$

The set

$$E^{'} := \left\{ (\omega, t, \mathbb{P}, \tilde{\omega}) \in \Omega \times [0,T] \times \mathfrak{P}^{ac}_{sem}(\Omega) \times \Omega \;\middle|\; \int_0^T \mathbf{1}_E(\omega, t, \mathbb{P}, \tilde{\omega}, s)\, \mathbf{1}_{(t,T]}(s)\, ds = T - t \right\}$$

is Borel measurable because the inverse image of a Borel measurable mapping is also Borel measurable. Therefore, by a monotone class argument, we have that $(\omega, t, \mathbb{P}) \mapsto E^{\mathbb{P}}[\mathbf{1}_{E'}(\omega, t, \mathbb{P}, \cdot)]$ is a Borel measurable function. Therefore, the following result completes the proof:

$$\begin{aligned} &\{(\omega, t, \mathbb{P}) \in \Omega \times [0,T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\} \\ &\qquad\qquad = \left\{ (\omega, t, \mathbb{P}) \in \Omega \times [0,T] \times \mathfrak{P}^{ac}_{sem}(\Omega) \middle|\; E^{\mathbb{P}}[\mathbf{1}_{E'}(\omega, t, \mathbb{P}, \cdot)] = 1 \right\}, \end{aligned}$$

where the set on the right-hand side of the equality above is Borel measurable because $(\omega, t, \mathbb{P}) \mapsto E^{\mathbb{P}}[\mathbf{1}_{E'}(\omega, t, \mathbb{P}, \cdot)]$ is a Borel measurable function.

B.1.2 Proof of Lemma 4.1.6

From the definition of a probability measure $\mathbb{P}^{\tau,\omega}$, we have that for $\mathbb{P}-$ a.s. $\bar{\omega} \in \Omega$:

$$\begin{aligned} &\mathbb{E}^{\mathbb{P}} \left[\int_\tau^T f(s, Y_s, \alpha_s)\, ds + g(Y_T) \;\middle|\; \mathcal{F}_\tau \right] (\bar{\omega}) \\ &\quad = \mathbb{E}^{\mathbb{P}^{\tau(\bar{\omega}),\bar{\omega}}} \left[\int_{\tau(\bar{\omega})}^T f\left(s, (Y_s)^{\tau(\bar{\omega}),\bar{\omega}}, \alpha_s^{\tau(\bar{\omega}),\bar{\omega}}\right) ds + g\left(Y_T^{\tau(\bar{\omega}),\bar{\omega}}\right) \right] \\ &\quad = \mathbb{E}^{\mathbb{P}^{\tau(\bar{\omega}),\bar{\omega}}} \left[\int_{\tau(\bar{\omega})}^T f\left(s, Y_s, \alpha_s^{\tau(\bar{\omega}),\bar{\omega}}\right) ds + g(Y_T) \right], \end{aligned}$$

where the last equation follows because of the path uniqueness of the process Y.

B.1.3 Proof of Lemma 4.1.7

First, due to Assumption 4.1.4, the functions J, w, and v are finite. Next, we check the regularity of the function J. Use a similar argument to that in Lemma 3.3 in Pham and Wei (2017) to show that for a fixed measure $\mathbb{P}$ the function J is continuous with respect to t, x, y, α. Write the function $J(t, x, y, \mathbb{P}, \alpha)$ as

$$J(t, x, y, \mathbb{P}, \alpha) = \int_\Omega \int_t^T f(s, Y_s(\omega), \alpha_s(\omega))\, ds + g(Y_T(\omega))\, \mathbb{P}(d\omega)\,.$$

For fixed t, x, y, α, the mapping $\omega \mapsto Y_s(\omega)$, $\omega \mapsto \alpha_s(\omega)$ is Borel measurable for all $s \in [t, T]$. Apply Corollary 7.29.1 in Bertsekas and Shreve (1996) to show that the function J is Borel measurable with respect to $\mathbb{P}$. Therefore, by the Carathéodory theorem (see Appendix 2.1.5) the function J is Borel measurable. Recall that the set $\{(\omega, t, \mathbb{P}) \in \Omega \times [0,T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\}$ is Borel measurable. Therefore, all conditions in Lemma B.1.1 hold, which completes the proof.

B.1.4 Proof of Lemma B.1.1

Lemma B.1.1. *Let A, X, Y, Z be metrizable separable spaces. Let J be a Borel measurable function on $X \times Y \times Z$ and let $J(x,y,z)$ be a continuous function in the variable y. Define $w(x,y) := \sup_{z \in B(x)} J(x,y,z)$ and $v^*(x) := \inf_{y \in A} w(x,y)$. Assume that the sets $\{(x,z) |\ z \in B(x)\}$ and $\{(x,y) |\ y \in A\}$ are Borel measurable and w and v^* are finite. Then we have the following results:*

- *(a) The function w is upper semianalytic and $\mathcal{L}_X \otimes B_Y$-measurable.*
- *(b) The function v^* is universally measurable, and for any $\epsilon > 0$ there exists a universally measurable ϵ-minimax strategy.*

This Lemma is a modification of Theorem 1 in Nowak (2010). We replace the σ-compact assumption on the set $B(x)$ with the continuity of the function J on the second variable because the σ-compact assumption is not satisfied in our framework.

(a) From a standard measurable selection, the function w is upper semianalytic, and $w(\cdot, y)$ is $\mathcal{L}_X$-measurable. Let

$$f_n(x,y,z,b) := J(x,b,z) + n\, d(y,b), \;\; J_n(x,y,z) := \inf_{b \in A} f_n(x,y,z,b)\,,$$

and define the function w_n as a lower envelope for the function w

$$w_n(x,y) := \sup_{z \in B(x)} J_n(x,y,z)\,,$$

where d is a metric distance in the space A. Note that the function f_n is continuous in the variables y and b, and that the space A is separable. Then, write the function J_n as the infimum of the countable measurable function

$$J_n(x,y,z) = \inf_{k \in \mathbb{N}} f_n(x,y,z,b_k)\,,$$

where $\{b_1, .., b_k, ..\}$ is the countable dense subset in A. Therefore, J_n is Borel measurable and continuous in y. By the DPP, the function w_n is upper semianalytic, and hence $w_n(\cdot, y)$ is $\mathcal{L}_X$-measurable. Now, use the inequality

$$\left|w_n(x,y) - w_n(x,y')\right| \leq \sup_{z \in B(x)} \left|J_n(x,y,z) - J_n(x,y',z)\right|$$

$$\leq \sup_{z \in B(x)} \sup_{b \in A} \left| f_n(x, y, z, b) - f_n(x, y', z, b) \right| \leq n\, d(y, y')$$

to check that the function w_n is continuous in y.

Therefore, by the Carathéodory Theorem 2.1.5, the function w_n is $\mathcal{L}_X \otimes B_Y$-measurable. From the definition of the functions f_n, we have that $J_n \leq J_{n+1}$, $w_n \leq w_{n+1}$, and $w_n \leq w$. Thus, $\lim_{n\to\infty} w_n(x, y) \leq w(x, y)$. Next, we prove that the reverse inequality holds by showing that

$$\lim_{n\to\infty} J_n(x, y, z) = J(x, y, z)\,.$$

Assume there exists x, y, z such that $\lim_{n\to\infty} J_n(x, y, z) < J(x, y, z)$. Therefore, there exists an $\epsilon > 0$ and $N \in \mathbb{N}$ such that $J_n(x, y, z) + \epsilon < J(x, y, z)$ for all $n \geq N$. Hence, there exists a sequence $\tilde{b}_n$ such that

$$J(x, \tilde{b}_n, z) + n\, d(y, \tilde{b}_n) + \frac{\epsilon}{2} < J(x, y, z)\,,$$

for all $n \geq N$. Because the function $J(x, \cdot, z)$ is continuous, the sequence $\tilde{b}_n \in A$ converges to y, and this leads to a contradiction. Therefore, the limit of function J_n is equal to J, and hence

$$\lim_{n\to\infty} \sup_{z \in B(x)} J_n(x, y, z) \geq \sup_{z \in B(x)} \lim_{n\to\infty} J_n(x, y, z) = \sup_{z \in B(x)} J(x, y, z) = w(x, y)\,.$$

Thus, $\lim_{n\to\infty} w_n(x, y) \geq w(x, y)$, and $\lim_{n\to\infty} w_n(x, y) = w(x, y)$, so the function w is also $\mathcal{L}_X \otimes B_Y$-measurable. The rest of the proof follows from Theorem 1 in Nowak (2010).

B.2 Additional definitions and theorems

Theorem B.2.1. *(Fubini's Theorem) Let μ and v be σ-finite outer measures on X and Y respectively. For any non-negative $\mu \times v$-measurable function f,*

$$x \mapsto f(x, y) \;\; \textit{is } \mu - \textit{measurable for } v\textit{-a.e.}\,,$$
$$y \mapsto \int_X f(x, y)\, d\mu(x) \;\; \textit{is } \mu - \textit{measurable}\,.$$

Definition B.2.2. (Hausdorff metric) Let (M, d) be a metric space and let X, Y be non-empty subsets of the space M. The Hausdorff metric is given by

$$d_{\text{haus}}(X, Y) := \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y),\, \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}\,.$$

Definition B.2.3. (Analytic set and analytically measurable function)

(i) Let E be a Borel space, then a subset B is an analytic set in E if there is another Borel space F and a Borel subset $A \subseteq E \times F$ such that $B = \pi_E(A)$. A subset $C \subseteq E$ is co-analytic if its complement C^c is analytic.

(ii) A function $g : E \to \mathbb{R} = \mathbb{R} \cup \infty$ is upper semianalytic if $\{x \in E : g(x) > c\}$ is analytic for every $c \in \mathbb{R}$.

(iii) Let E be a Borel set and $\mathcal{A}(E)$ denote a σ-field generated by all analytic subsets. A function $f : E \to F$, where F is a Borel set, is analytically measurable if $f^{-1}(C) \in \mathcal{A}(E)$ for every $C \in \mathcal{B}(F)$.

Definition B.2.4. Let $(\Omega, \mathcal{F})$ be a measurable space, the universal completion of $\mathcal{F}$ is the σ-field defined as the intersection of $\mathcal{F}^{\mathbb{P}}$ for all probability measures $\mathbb{P} \in \mathcal{P}(\Omega)$ on $(\Omega, \mathcal{F})$ i.e.,

$$\mathcal{F}^U := \bigcap_{\mathbb{P} \in \mathfrak{P}(\Omega)} \mathcal{F}^{\mathbb{P}}.$$

A function f from $(\Omega, \mathcal{F})$ to E is called universally measurable, if $f^{-1}(A) \in \mathcal{F}^U$ for each $A \in \mathcal{B}(E)$.

Definition B.2.5. Let (S, Σ) be a measurable space, and let X and Y be topological spaces. A function $f : S \times X \to Y$ is a **Carathéodory function** if:

1. For each $x \in X$, the function $f^x = f(\cdot, x) : S \to Y$ is $(\Sigma, \mathcal{B}_Y)$-measurable.

2. For each $s \in S$, the function $f_s = f(s, \cdot) : X \to Y$ is continuous.

Theorem B.2.6. *Let (S, Σ) be a measurable space, X a separable metrizable space, and Y a metrizable space. Then every Carathéodory function $f : S \times X \to Y$ is jointly measurable.*

Appendix C

Appendix for Chapter 5

C.1 Proof of the convergence of the estimator in (5.2.8)

Let $I_1 = 1$ and $I_2 = \rho$, apply Itô's Lemma to $\exp(\int_0^t \beta_s\, ds)\, \hat{\theta}_t^i$ and write

$$\begin{aligned}
\exp\left(\int_0^t \beta_s\, ds\right) \hat{\theta}_t^i =& \hat{\theta}_0^i + \int_0^t \exp\left(\int_0^s \beta_u\, du\right) d\hat{\theta}_s^i + \int_0^t \beta_s \exp\left(\int_0^s \beta_u\, du\right) \hat{\theta}_s^i\, ds \\
=& \hat{\theta}_t^i + \int_0^t \theta^{*,i}\, \beta_s \exp\left(\int_0^s \beta_u\, du\right) ds + \int_0^t \sigma_{i\,1}\, \beta_s \exp\left(\int_0^s \beta_u\, du\right) dW_s^i \\
&+ \int_0^t I_i\, \sigma_{i\,2}\, \beta_s \exp\left(\int_0^s \beta_u\, du\right) dW_s^m .
\end{aligned}$$

Therefore,

$$\begin{aligned}
\hat{\theta}_t^i =& \hat{\theta}_0^i \exp\left(-\int_0^t \beta_s\, ds\right) + \int_0^t \theta^{*,i}\, \beta_s \exp\left(-\int_s^t \beta_u\, du\right) ds \\
&+ \int_0^t \sigma_{i\,1}\, \beta_s \exp\left(-\int_s^t \beta_u\, du\right) dW_s^i + \int_0^t I_i\, \sigma_{i\,2}\, \beta_s \exp\left(-\int_s^t \beta_u\, du\right) dW_s^m \\
=& \hat{\theta}_0^i \left(\frac{1}{1+t}\right)^L \\
&+ L \left(\int_0^t \theta^{*,i}\, \frac{(1+s)^{L-1}}{(1+t)^L}\, ds + \int_0^t \sigma_{i\,1}\, \frac{(1+s)^{L-1}}{(1+t)^L}\, dW_s^i + \int_0^t I_i\, \sigma_{i\,2}\, \frac{(1+s)^{L-1}}{(1+t)^L}\, dW_s^m \right) .
\end{aligned}$$

Therefore, for all values of the learning rate parameter L, the mean of $\hat{\theta}_t^i$ is

$$\hat{\theta}_0^i \left(\frac{1}{1+t}\right)^L + \theta^{*,i} \left(1 - \frac{1}{(1+t)^L}\right) ,$$

and when $L \neq 1/2$, the variance of $\hat{\theta}_t^i$ is

$$(\sigma_{i\,1}^2 + I_i^2\,\sigma_{i\,2}^2)\int_0^t L^2\,\frac{1^{2\,L-2}}{(1+s)^{2\,L}}\,ds = (\sigma_{i\,1}^2 + I_i^2\,\sigma_{i\,2}^2)\,\frac{L^2}{2\,L-1}\left(\frac{(1+t)^{2\,L-1}}{(1+t)^{2\,L}} - \frac{(1)^{2\,L-1}}{(1+t)^{2\,L}}\right),$$

and when $L = 1/2$, the variance of $\hat{\theta}_t^i$ is

$$(\sigma_{i\,1}^2 + I_i^2\,\sigma_{i\,2}^2)\int_0^t L^2\,\frac{(1+s)^{2\,L-2}}{(1+t)^{2\,L}}\,ds = \frac{(\sigma_{i\,1}^2 + I_i^2\,\sigma_{i\,2}^2)}{4}\left(\frac{\log(1+t)}{1+t}\right).$$

For each $\epsilon > 0$

$$\mathbb{P}^*\left(\left|\hat{\theta}_t^i - \theta^{*,i}\right| > \epsilon\right) \leq \frac{\mathbb{E}^{\mathbb{P}^*}\left[(\hat{\theta}_t^i - \theta^{*,i})^2\right]}{\epsilon^2} \leq C_\epsilon \max\left\{\frac{\log(1+t)}{1+t}, \frac{1}{1+t}\right\} \tag{C.1.1}$$

where C_ϵ is a constant that depends on ϵ. Then, we have that $\hat{\theta}_t^i$ converges to $\theta^{*,i}$ in probability as t go to infinity.

C.2 Proof of the finiteness of the value function

First, we notice that there exists a constant M such that $M \geq U(X)$ for all $X \in \mathbb{R}$, which implies $v \leq M$. Let $\alpha = 0$ is a constant admissible process, then the random variable $X_T^\alpha = X\,e^{r\,(T-t)}$ because under this control process α the agent invests only the risk-free asset. Then, substitute X_T^α to obtain

$$v(t, X, S, \hat{\theta}) \geq \inf_{\mathbb{P}\in\mathcal{P}(t,x,G)} \mathbb{E}^{\mathbb{P}}\left[-e^{-\gamma\,X\,e^{r\,(T-t)}}\right] = -e^{-\gamma\,X\,e^{r\,(T-t)}}, \tag{C.2.1}$$

since the value inside the expectation is deterministic. Therefore, $v > -\infty$ and implies that v is finite.

C.3 Additional details in the proof of Proposition 5.2.1

C.3.1 First part

1) If $\beta_t\,\partial_{\hat{\theta}}U - \alpha\,U \in \mathbb{R}^{+,+}$ (first quadrant of $\mathbb{R}^2$), then

$$\begin{aligned}&\sup_{\tilde{\theta}\in G(t,\hat{\theta})}\left\{\beta_t\,(\tilde{\theta}-\hat{\theta})\,\partial_{\hat{\theta}}U - \alpha^\intercal\left(\tilde{\theta} - \mathbf{1}\,r\right)U + \tfrac{1}{2}\,\alpha^\intercal\,\Sigma\,\alpha\,U - \beta_t\,\alpha^\intercal\,\Sigma\,\partial_{\hat{\theta}}\,U\right\}\\ &\quad= \beta_t\,c/\sqrt{1+t}\,\mathbf{1}^\intercal\,\partial_{\hat{\theta}}U - \alpha^\intercal\left(\hat{\theta} - \mathbf{1}\,(r - c/\sqrt{1+t})\right)U + \tfrac{1}{2}\,\alpha^\intercal\,\Sigma\,\alpha\,U - \beta_t\,\alpha^\intercal\,\Sigma\,\partial_{\hat{\theta}}\,U\,,\end{aligned} \tag{C.3.1}$$

and the infimum is attained at either

$$\alpha^* = \beta_t \frac{\partial_{\hat\theta} U}{U} \qquad \text{or} \qquad \alpha^* = \Sigma^{-1}\left(\hat\theta - \mathbf{1}\, r + \mathbf{1}^{+,+}\, c/\sqrt{1+t})\right) + \beta_t \frac{\partial_{\hat\theta} U}{U}\,, \tag{C.3.2}$$

or

$$\alpha^{*,1} = \beta_t \frac{\partial_{\hat\theta_1} U}{U} \qquad \text{and} \qquad \alpha^{*,2} = \frac{\hat\theta_2 - r + c/\sqrt{1+t}}{\sigma_2^2} + \beta_t \frac{\partial_{\hat\theta_2} U}{U}\,, \tag{C.3.3}$$

or

$$\alpha^{*,2} = \beta_t \frac{\partial_{\hat\theta_2} U}{U} \qquad \text{and} \qquad \alpha^{*,1} = \frac{\hat\theta_1 - r + c/\sqrt{1+t}}{\sigma_1^2} + \beta_t \frac{\partial_{\hat\theta_1} U}{U}\,. \tag{C.3.4}$$

2) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{-,+}$ (second quadrant of $\mathbb{R}^2$), the infimum is attained at either

$$\alpha^* = \beta_t \frac{\partial_{\hat\theta} U}{U} \qquad \text{or} \qquad \alpha^* = \Sigma^{-1}\left(\hat\theta - \mathbf{1}\, r + \mathbf{1}^{-,+}\, c/\sqrt{1+t}\right) + \beta_t \frac{\partial_{\hat\theta} U}{U}\,, \tag{C.3.5}$$

or

$$\alpha^{*,1} = \beta_t \frac{\partial_{\hat\theta_1} U}{U} \qquad \text{and} \qquad \alpha^{*,2} = \frac{\hat\theta_2 - r + c/\sqrt{1+t}}{\sigma_2^2} + \beta_t \frac{\partial_{\hat\theta_2} U}{U}\,, \tag{C.3.6}$$

or

$$\alpha^{*,2} = \beta_t \frac{\partial_{\hat\theta_2} U}{U} \qquad \text{and} \qquad \alpha^{*,1} = \frac{\hat\theta_1 - r - c/\sqrt{1+t}}{\sigma_1^2} + \beta_t \frac{\partial_{\hat\theta_1} U}{U}\,. \tag{C.3.7}$$

3) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{-,-}$ (third quadrant of $\mathbb{R}^2$), the infimum is attained at either

$$\alpha^* = \beta_t \frac{\partial_{\hat\theta} U}{U} \qquad \text{or} \qquad \alpha^* = \Sigma^{-1}\left(\hat\theta - \mathbf{1}\, r + \mathbf{1}^{-,-}\, c/\sqrt{1+t})\right) + \beta_t \frac{\partial_{\hat\theta} U}{U}\,, \tag{C.3.8}$$

or

$$\alpha^{*,1} = \beta_t \frac{\partial_{\hat\theta_1} U}{U} \qquad \text{and} \qquad \alpha^{*,2} = \frac{\hat\theta_2 - r - c/\sqrt{1+t}}{\sigma_2^2} + \beta_t \frac{\partial_{\hat\theta_2} U}{U}\,, \tag{C.3.9}$$

or

$$\alpha^{*,2} = \beta_t \frac{\partial_{\hat\theta_2} U}{U} \qquad \text{and} \qquad \alpha^{*,1} = \frac{\hat\theta_1 - r - c/\sqrt{1+t}}{\sigma_1^2} + \beta_t \frac{\partial_{\hat\theta_1} U}{U}\,. \tag{C.3.10}$$

4) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{+,-}$ (fourth quadrant of $\mathbb{R}^2$), the infimum is attained at either

$$\alpha^* = \beta_t \frac{\partial_{\hat\theta} U}{U} \qquad \text{or} \qquad \alpha^* = \Sigma^{-1}\left(\hat\theta - \mathbf{1}\, r + \mathbf{1}^{+,-}\, c/\sqrt{1+t})\right) + \beta_t \frac{\partial_{\hat\theta} U}{U}\,, \tag{C.3.11}$$

or

$$\alpha^{*,1} = \beta_t \frac{\partial_{\hat\theta_1} U}{U} \qquad \text{and} \qquad \alpha^{*,2} = \frac{\hat\theta_2 - r - c/\sqrt{1+t}}{\sigma_2^2} + \beta_t \frac{\partial_{\hat\theta_2} U}{U}\,, \tag{C.3.12}$$

or

$$\alpha^{*,2} = \beta_t \frac{\partial_{\hat\theta_1} U}{U} \qquad \text{and} \qquad \alpha^{*,1} = \frac{\hat\theta_1 - r + c/\sqrt{1+t}}{\sigma_1^2} + \beta_t \frac{\partial_{\hat\theta_1} U}{U}\,. \tag{C.3.13}$$

C.3.2 Second part

Consider $\hat\theta \in A^{-,-}$ and we can divide the value α in the following four cases.
Case 1. $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{+,+}$, then the infimum is attained at

$$\alpha^* = \Sigma^{-1}\,\left(\hat\theta - \mathbf{1}\, r + \mathbf{1}^{+,+}\, c/\sqrt{1+t})\right) + \beta_t\, \frac{\partial_{\hat\theta} U}{U}\,,$$

since $\hat\theta \in A^{-,-}$.
Case 2. $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{-,+}$. Note that from $\hat\theta \in A^{-,-}$, implies that

$$\theta_1 - r + \frac{c}{\sqrt{1+t}} \le 0 \quad \text{and} \quad \sigma_2^2\,\left(\hat\theta_1 - r + \frac{c}{\sqrt{1+t}}\right) \le \rho\, \sigma_c^2\,\left(\hat\theta_2 - r + \frac{c}{\sqrt{1+t}}\right)\,.$$

Therefore, the value $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \notin \mathbb{R}^{-,+}$ when

$$\alpha^2 = \beta_t\, \frac{\partial_{\hat\theta_2} U}{U} \qquad \text{and} \qquad \alpha^1 = (\sigma_1^2)^{-1}\,\left(\hat\theta_1 - r - c/\sqrt{1+t})\right) + \beta_t\, \frac{\partial_{\hat\theta_1} U}{U}\,, \quad \text{(C.3.14)}$$

or

$$\alpha = \Sigma^{-1}\,\left(\hat\theta - \mathbf{1}\, r + \mathbf{1}^{-,+}\, c/\sqrt{1+t})\right) + \beta_t\, \frac{\partial_{\hat\theta} U}{U}\,.$$

Therefore, the infimum is possibly attained at

$$\alpha^1 = \beta_t\, \frac{\partial_{\hat\theta_1} U}{U} \qquad \text{and} \qquad \alpha^2 = (\sigma_2^2)^{-1}\,\left(\hat\theta_2 - r + c/\sqrt{1+t})\right) + \beta_t\, \frac{\partial_{\hat\theta_2} U}{U}\,, \quad \text{(C.3.15)}$$

or $\alpha = \beta_t \partial_{\hat\theta} U/U$.
Case 3. $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{+,-}$. Similarly to the Case 2, the infimum is possibly attained at

$$\alpha^2 = \beta_t\, \frac{\partial_{\hat\theta_2} U}{U} \qquad \text{and} \qquad \alpha^1 = (\sigma_1^2)^{-1}\,\left(\hat\theta_1 - r + c/\sqrt{1+t})\right) + \beta_t\, \frac{\partial_{\hat\theta_1} U}{U}\,, \quad \text{(C.3.16)}$$

or $\alpha = \beta_t \partial_{\hat\theta} U/U$.
Case 4. $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{+,-}$. It is easy to check that the infimum is only attained at $\alpha = \beta_t \partial_{\hat\theta} U/U$.

Note that all possible optimal strategies is in the form $K + \beta_t \partial_{\hat\theta} U/U$. Therefore, we substitute into (C.3.1). Therefore, we have that

$$\sup_{\tilde\theta \in G(t,x)} \left\{\beta_t\,(\tilde\theta - \hat\theta)\,\partial_{\hat\theta} U - (K + \beta_t \partial_{\hat\theta} U/U)^\intercal\,\left(\tilde\theta - r\right)\, U\right\} + \frac{1}{2} U\, K^\intercal\, \Sigma\, K - \frac{\beta_t^2}{2\,U}\, \partial_{\hat\theta} U^\intercal\, \Sigma\, \partial_{\hat\theta} U\,, \quad \text{(C.3.17)}$$

and it suffices to not consider the last term which is common to all possible strategy.

Case 1. $K = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)$ and $\beta_t\,\partial_{\hat{\theta}}U - \alpha\,U \in \mathbb{R}^{+,+}$, then (C.3.17) is

$$\begin{aligned}
&\frac{\beta_t\,c}{\sqrt{1+t}}(\partial_{\theta_1}U + \partial_{\theta_2}U) - \frac{\beta_t(\,\partial_\theta U)^\intercal}{U}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)U \\
&-\frac{1}{2}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)^\intercal \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right) \\
&= \beta_t(\,\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) - \frac{1}{2}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)^\intercal \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)
\end{aligned} \tag{C.3.18}$$

Similarly, for Case 2, we have that (C.3.17) is

$$\beta_t(\,\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) - \frac{1}{2}\left(\hat{\theta}_2 - r + \,c/\sqrt{1+t})\right)^\intercal (\sigma_2^2)^{-1}\left(\hat{\theta}_2 - r + c/\sqrt{1+t})\right) \tag{C.3.19}$$

Similarly, for Case 3, we have that (C.3.17) is

$$\beta_t(\,\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) - \frac{1}{2}\left(\hat{\theta}_1 - r + \,c/\sqrt{1+t})\right)^\intercal (\sigma_1^2)^{-1}\left(\hat{\theta}_1 - r + c/\sqrt{1+t})\right) \tag{C.3.20}$$

Similarly, for Case 4, we have that (C.3.17) is

$$\beta_t(\,\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) \tag{C.3.21}$$

Then, it is straight forward to check Case 1 has the minimal value and for the other regions, we follow the same argument and this completes the proof.

C.3.3 Third part

Next, we solve the inf-sup problem in (5.2.27) for each $\hat{\theta}$ in the nine non-overlapping investment regions of $\mathbb{R}^2$ discussed above.

If $\hat{\theta} \in A^{-,-} \subset \mathbb{R}^2$, the optimal investment strategy is

$$\alpha^* = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t}\right) - \beta_t\,\partial_{\hat{\theta}}\,H\,, \tag{C.3.22}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}}H + F(t,\hat{\theta})\right)\,. \tag{C.3.23}$$

If $\hat{\theta} \in A^{+,-} \subset \mathbb{R}^2$, the optimal strategy is

$$\alpha^* = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{-,+}\,c/\sqrt{1+t})\right) - \beta_t\,\partial_{\hat{\theta}}\,H\,, \tag{C.3.24}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}}H + F(t,\hat{\theta})\right)\,. \tag{C.3.25}$$

If $\hat{\theta} \in A^{-,+} \subset \mathbb{R}^2$, the optimal strategy is

$$\alpha^* = \Sigma^{-1} \left(\hat{\theta} - \mathbf{1}\, r + \mathbf{1}^{+,-}\, c/\sqrt{1+t}\right) - \beta_t\, \partial_{\hat{\theta}}\, H\,, \tag{C.3.26}$$

and

$$C(t,\hat{\theta}) = -U\, \left(\beta_t\, (\mathbf{1}\, r - \hat{\theta})^\intercal\, \partial_{\hat{\theta}} H + F(t,\hat{\theta})\right)\,. \tag{C.3.27}$$

If $\hat{\theta} \in A^{+,+} \subset \mathbb{R}^2$, the optimal strategy is

$$\alpha^* = \Sigma^{-1} \left(\hat{\theta} - \mathbf{1}\, r + \mathbf{1}^{-,-}\, c/\sqrt{1+t}\right) - \beta_t\, \partial_{\hat{\theta}}\, H\,, \tag{C.3.28}$$

and

$$C(t,\hat{\theta}) = -U\, \left(\beta_t\, (\mathbf{1}\, r - \hat{\theta})^\intercal\, \partial_{\hat{\theta}} H + F(t,\hat{\theta})\right)\,. \tag{C.3.29}$$

If $\hat{\theta} \in A^{0,-} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = -\beta_t\, \partial_{\hat{\theta}_1}\, H \qquad \text{and} \qquad \alpha^{*,2} = \frac{\hat{\theta}_2 - r + c/\sqrt{1+t}}{\sigma_2^2} - \beta_t\, \partial_{\hat{\theta}_2}\, H\,, \tag{C.3.30}$$

and

$$C(t,\hat{\theta}) = -U\, \left(\beta_t\, (\mathbf{1}\, r - \hat{\theta})^\intercal\, \partial_{\hat{\theta}} H + \frac{1}{2}\, \left(\hat{\theta}_2 - r + c/\sqrt{1+t}\right)\, \frac{1}{\sigma_2^2}\, \left(\hat{\theta}_2 - r + c/\sqrt{1+t}\right)\right)\,. \tag{C.3.31}$$

If $\hat{\theta} \in A^{+,0} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = \frac{\hat{\theta}_1 - r - c/\sqrt{1+t}}{\sigma_1^2} - \beta_t\, \partial_{\hat{\theta}_1}\, H \qquad \text{and} \qquad \alpha^{*,2} = -\beta_t\, \partial_{\hat{\theta}_2}\, H\,, \tag{C.3.32}$$

and

$$C(t,\hat{\theta}) = -U\, \left(\beta_t\, (\mathbf{1}\, r - \hat{\theta})^\intercal\, \partial_{\hat{\theta}} H + \frac{1}{2}\, \left(\hat{\theta}_1 - r - c/\sqrt{1+t}\right)\, \frac{1}{\sigma_1^2}\, \left(\hat{\theta}_1 - r - c/\sqrt{1+t}\right)\right)\,. \tag{C.3.33}$$

If $\hat{\theta} \in A^{0,+} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = -\beta_t\, \partial_{\hat{\theta}_1}\, H \qquad \text{and} \qquad \alpha^{*,2} = \frac{\hat{\theta}_2 - r - c/\sqrt{1+t}}{\sigma_2^2} - \beta_t\, \partial_{\hat{\theta}_2}\, H\,, \tag{C.3.34}$$

and

$$C(t,\hat{\theta}) = -U\, \left(\beta_t\, (\mathbf{1}\, r - \hat{\theta})^\intercal\, \partial_{\hat{\theta}} H + \frac{1}{2}\, \left(\hat{\theta}_2 - r - c/\sqrt{1+t}\right)\, \frac{1}{\sigma_2^2}\, \left(\hat{\theta}_2 - r - c/\sqrt{1+t}\right)\right)\,. \tag{C.3.35}$$

If $\hat{\theta} \in A^{-,0} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = \frac{\hat{\theta}_1 - r + c/\sqrt{1+t}}{\sigma_1^2} - \beta_t\, \partial_{\hat{\theta}_1}\, H \qquad \text{and} \qquad \alpha^{*,2} = -\beta_t\, \partial_{\hat{\theta}_2}\, H\,, \tag{C.3.36}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r-\hat{\theta})^\intercal\,\partial_{\hat{\theta}}H + \frac{1}{2}\left(\hat{\theta}_1 - r + c/\sqrt{1+t}\right)\frac{1}{\sigma_1^2}\left(\hat{\theta}_1 - r + c/\sqrt{1+t}\right)\right). \tag{C.3.37}$$

If $\hat{\theta} \in A^{0,0} \subset \mathbb{R}^2$, then

$$\alpha^* = -\beta_t\,\partial_{\hat{\theta}}\,H\,, \tag{C.3.38}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r-\hat{\theta})^\intercal\,\partial_{\hat{\theta}}H\right). \tag{C.3.39}$$

C.4 Proof of Proposition 5.2.2

Let $c = 0$ in Proposition 5.2.1 to obtain (5.2.30) and (5.2.31). Next, we show (5.2.32). Denote the variance of the random variable X and the covariance between the random variables X and Y by $Var(X)$ and $Cov(X,Y)$, respectively. From (5.2.31), we have that

$$\begin{aligned}\bar{H}^0(t,\hat{\theta}) &= \frac{1}{2\,(\sigma_1^2\,\sigma_2^2 - \rho^2\,\sigma_c^4)} \\ &\quad\times\int_t^T\left(\sigma_2^2\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1-r)^2\right] + \sigma_1^2\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^2-r)^2\right] - 2\rho\sigma_c^2\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1-r)(Z_u^2-r)\right]\right)du\,.\end{aligned}$$

It is straightforward to compute the expectations that appear above:

$$\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1-r)^2\right] = \mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1-r)\right]^2 + Var(Z_u^1) = (\hat{\theta}_1 - r)^2\left(\frac{1+t}{1+u}\right)^{2L} + Var(Z_u^1)\,, \tag{C.4.1}$$

$$\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^2-r)^2\right] = (\hat{\theta}_2 - r)^2\left(\frac{1+t}{1+u}\right)^{2L} + Var(Z_u^2) \tag{C.4.2}$$

and

$$\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1-r)\,(Z_u^2-r)\right] = (\hat{\theta}_1 - r)\,(\hat{\theta}_2 - r)\left(\frac{1+t}{1+u}\right)^{2L} + Cov(Z_u^1, Z_u^2)\,. \tag{C.4.3}$$

The terms $Var(Z_u^1)$, $Var(Z_u^2)$, and $Cov(Z_u^1, Z_u^2)$ do not depend on $\hat{\theta}$; thus, we have that

$$\partial_{\hat{\theta}_1}\bar{H}^0(t,\hat{\theta}) = \frac{\sigma_2^2\,(\hat{\theta}_1 - r) - \rho\,\sigma_c^2\,(\hat{\theta}_2 - r)}{\sigma_1^2\,\sigma_2^2 - \rho^2\,\sigma_c^4}\int_t^T\left(\frac{1+t}{1+u}\right)^{2L}du \tag{C.4.4}$$

and

$$\partial_{\hat{\theta}_2}\bar{H}^0(t,\hat{\theta}) = \frac{-\rho\,\sigma_c^2\,(\hat{\theta}_1 - r) + \sigma_1^2\,(\hat{\theta}_2 - r)}{\sigma_1^2\,\sigma_2^2 - \rho^2\,\sigma_c^4}\int_t^T\left(\frac{1+t}{1+u}\right)^{2L}du\,. \tag{C.4.5}$$

Finally, the expression in (5.2.32) follows from the two equations above.

C.5 Proof of Lemma 5.2.3

Recall that the optimal strategy in Kim and Omberg (1996) can be written as

$$y_*(W,X,\tau) = \left(\frac{\partial_W J}{-\partial_{WW} J}\right)\left(\frac{X(t)}{\sigma}\right) + \left(\frac{\partial_{WX} J}{-\partial_{WW} J}\right) \tag{C.5.1}$$

and the partial differential equation

$$-\,\partial_\tau J + r\,W\,\partial_W J - \frac{y_*^2\,\sigma^2}{2}\,\partial_{W\,W} J - \lambda_X\,(X-\bar{X})\,\partial_X J + \frac{\sigma_X^2}{2}\,\partial_{X\,X} J = 0 \tag{C.5.2}$$

where $J(W,X,0) = U(W) = -\exp(-\gamma\,W)$. Let $J(W,X,\tau) = \Phi(X,\tau)\,U(W\,e^{r\,\tau})$, then by direct substitution, we have that

$$\partial_\tau \Phi - \frac{1}{2}\,\Phi\left(X + \frac{\sigma_X\,\partial_X \Phi}{\Phi}\right)^2 - \lambda_X\,(X-\bar{X})\,\partial_X \Phi + \frac{\sigma_X^2}{2}\,\partial_{XX}\Phi = 0 \tag{C.5.3}$$

We then substitute $\Phi(X,\tau) = \exp(-\hat{H}(x,\tau))$ to obtain

$$\begin{aligned} 0 &= \partial_\tau \hat{H} + \frac{1}{2}\left(X - \sigma_X\,\partial_X \hat{H}\right)^2 + \lambda_X\,(X-\bar{X})\,\partial_X \hat{H} + \frac{1}{2}\,\sigma_X^2\left(\partial_{XX}\hat{H} - (\partial_X \hat{H})^2\right) \\ &= \partial_\tau \hat{H} + \frac{1}{2}\,X^2 + (-\sigma_X\,X + \lambda_X\,(X-\bar{X}))\,\partial_X \hat{H} + \frac{1}{2}\,\sigma_X^2\left(\partial_{XX}\hat{H}\right) \end{aligned} \tag{C.5.4}$$

Therefore, it is easy to see that the function H above is the valid solution. Note that the optimal strategy can be written as

$$y_*(W,X,\tau) = \left(\frac{\partial_W J}{-\partial_{WW} J}\right)\left(\frac{X(t)}{\sigma}\right) + \left(\frac{\partial_{WX} J}{-\partial_{WW} J}\right)\left(\frac{\sigma_X}{\sigma}\right) = \frac{X - \sigma_X\,\partial_X \hat{H}}{\gamma\,e^{r\,(T-t)}\,\sigma}\,. \tag{C.5.5}$$

This completes the proof.

C.6 Proof of Proposition 5.2.4

First, we show that $A(t,T,L)$ is an increasing function on t. Recall that

$$A(t,T,L) = \begin{cases} 1 - \frac{L}{2\,L-1}\left(1 - \left(\frac{1+t}{1+T}\right)^{2\,L-1}\right) & \text{if} \quad L \neq 1/2\,, \\ 1 - \frac{1}{2}\left(\log(1+T) - \log(1+t)\right) & \text{if} \quad L = 1/2\,. \end{cases} \tag{C.6.1}$$

Therefore, by direct computation, the partial derivation of A with respect to t is

$$\partial_t A(t,T,L) = \begin{cases} L\,\frac{(1+t)^{2\,L-2}}{(1+T)^{2\,L-1}} & \text{if} \quad L \neq 1/2\,, \\ \frac{1}{2\,(1+t)} & \text{if} \quad L = 1/2\,. \end{cases} \tag{C.6.2}$$

Therefore, the partial derivative is always positive and the function $A(t, T, L)$ is an increasing on t. Thus, for all $t \leq T$, the adjustment term satisfies $A(0, T, L) \leq A(t, T, L) \leq A(T, T, L) = 1$. Then, by substitution, we obtain the upper and lower bounds.

C.7 Proof of Proposition 5.2.5

Next we discuss the adaptive investment strategy for four ranges of the value of the learning rate L. If $L \geq 1$, it is easy to see that the term $A(0, T, L)$

$$A(0, T, L) = 1 - \frac{L}{2\,L - 1} \left(1 - \left(\frac{1}{1 + T}\right)^{2\,L-1}\right) > 0\,,$$

for $0 \leq t \leq T$ since $L/(2\,L - 1) \leq 1$ and

$$\left(1 - \left(\frac{1 + t}{1 + T}\right)^{2\,L-1}\right) \leq 1\,.$$

Therefore, $A(t, T, L)$ is always positive. When $L \neq 1/2$ and $L < 1$, we have that

$$\left(\frac{1 - L}{L}\right)^{1/(2\,L-1)} < 1\,,$$

since either $(1 - L)/L < 1$ and $1/(2\,L - 1) > 0$ or $(1 - L)/L > 1$ and $1/(2\,L - 1) < 0$. We consider

$$t^L := (1 + T) \left(\frac{1 - L}{L}\right)^{1/(2\,L-1)} - 1\,. \tag{C.7.1}$$

It is easy to check that $A(t^L, T, L) = 0$. When $L = 1/2$, we consider

$$t^L = (1 + T)^{1/2} - 1\,.$$

It is easy to check that $A(t^L, T, L) = 0$. From Proposition 5.2.4, the function $A(t, T, L)$ is an increasing function on t. Then, if $t > t^L$, then $A(t, T, L) > 0$ and if $t < t^L$, then $A(t, T, L) < 0$.

C.8 Proof of Lemma 5.4.3

The second equation follows directly from the first equation. Let $Z \sim \mathcal{N}(0, 1)$ and $Y = \mu + \sigma\, Z$. Write

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[Y^2\, \mathbf{1}_{\{Y>0\}}\right] &= \mathbb{E}^{\mathbb{P}}\left[(\mu + \sigma\, Z)^2\, \mathbf{1}_{\{Z>-\mu/\sigma\}}\right] \\ &= \mu^2\, \big(1 - \Phi(-\mu/\sigma)\big) + 2\,\mu\,\sigma\, \mathbb{E}^{\mathbb{P}}\left[Z\, \mathbf{1}_{\{Z>-\mu/\sigma\}}\right] + \sigma^2\, \mathbb{E}^{\mathbb{P}}\left[Z^2\, \mathbf{1}_{\{Z>-\mu/\sigma\}}\right]\,.\end{aligned} \tag{C.8.1}$$

From the density of the standard normal distribution and straightforward calculations, obtain

$$\mathbb{E}^{\mathbb{P}}\left[Z\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right] = \frac{1}{\sqrt{2\,\pi}}\int_{-\mu/\sigma}^{\infty} z\,e^{-z^2/2}\,dz = \frac{1}{\sqrt{2\,\pi}}\,\exp\left(-\mu^2/2\,\sigma^2\right) \tag{C.8.2}$$

and

$$\mathbb{E}^{\mathbb{P}}\left[Z^2\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right] = \int_{-\mu/\sigma}^{\infty} z^2\,e^{-z^2/2}\,dz = -\frac{1}{\sqrt{2\,\pi}}\,\mu/\sigma\,\exp\left(-\mu^2/2\,\sigma^2\right) + 1 - \Phi(-\mu/\sigma)\,. \tag{C.8.3}$$

Therefore, (C.8.1), (C.8.2), (C.8.3) imply (5.4.4).

C.9 Proof of Lemma 5.4.4

To simplify notation, let $\Sigma_i = \Sigma_i(u)$. From Lemma 5.4.3, we write

$$h_1(u,\hat\theta_i) := (f^2+\Sigma_i^2)\,(1-\Phi(-f/\Sigma_i)) + f\,\Sigma_i\,\phi(-f/\Sigma_i)$$

and compute its gradient with respect to $\hat\theta$:

$$\begin{aligned}
\partial_{\hat\theta_i} h_1 =&\ 2\,f\,\partial_{\hat\theta_i} f\,(1-\Phi(-f/\Sigma_i)) + (f^2+\Sigma_i^2)\,\partial_{\hat\theta_i}\Phi(f/\Sigma_i)\\
&+ \Sigma_i\,\phi(-f/\Sigma_i) + f\,\Sigma_i\,\partial_{\hat\theta_i}\phi(-f/\Sigma_i)\\
=&\ 2\,f\,\partial_{\hat\theta_i} f\,(1-\Phi(-f/\Sigma_i)) + (f^2+\Sigma_i^2)\,\partial_{\hat\theta_i} f/\Sigma_i\,\phi(f/\Sigma_i)\\
&+ \partial_{\hat\theta_i} f\,\Sigma_i\,\phi(-f/\Sigma_i) - (f^2/\Sigma_i)\,\partial_{\hat\theta} f\,\phi(-f/\Sigma_i)\\
=&\ 2\,f\,\partial_{\hat\theta_i} f\,(1-\Phi(-f/\Sigma_i)) + 2\,\partial_{\hat\theta_i} f\,\Sigma_i\,\phi(-f/\Sigma_i)\,,
\end{aligned} \tag{C.9.1}$$

where the equation follows from $\partial_{\hat\theta_i}\Phi(f/\Sigma_i) = 1/\Sigma_i\,\partial_{\hat\theta_i} f\,\phi(f/\Sigma_i)$ and $\partial_{\hat\theta_i}\phi(-f/\Sigma_i) = -f/\Sigma_i^2\,\partial_{\hat\theta_i} f\,\phi(-f/\Sigma_i)$. Similarly, write

$$h_2(u,\hat\theta_i) := (\tilde f^2+\Sigma_i^2)\,\Phi\left(-\tilde f/\Sigma_i\right) - \tilde f\,\Sigma_i\,\Phi\left(-\tilde f/\Sigma_i\right)\,.$$

and compute its gradient with respect to $\hat\theta$:

$$\partial_{\hat\theta} h_2 = 2\,\tilde f\,\partial_{\hat\theta_i}\tilde f\,\Phi\left(-\tilde f/\Sigma_i\right) - 2\,\Sigma_i\,\partial_{\hat\theta_i}\tilde f\,\phi\left(-\tilde f/\Sigma_i\right)\,, \tag{C.9.2}$$

Let $h(u,\hat\theta_i) := h_1(u,\hat\theta_i) + h_2(u,\hat\theta_i)$ and from Lemma 5.4.3, we have that

$$\begin{aligned}
h(u,\hat\theta_i) = \mathbb{E}^{\mathbb{P}}&\left[\left(Z_u^i - r - c/\sqrt{1+u}\right)^2\,\mathbf{1}_{\{Z_u^i>c/\sqrt{1+u}\}}\right]\\
&+ \mathbb{E}^{\mathbb{P}}\left[\left(Z_u^i - r + c/\sqrt{1+u}\right)^2\,\mathbf{1}_{\{Z_u^i<-c/\sqrt{1+u}\}}\right]\,.
\end{aligned}$$

Then, the gradient of h with respect to $\hat\theta_i$, which is given by

$$\partial_{\hat\theta_i} h = \partial_{\hat\theta_i} f\left[2\,f\,(1-\Phi(-f/\Sigma_i)) + 2\,\Sigma_i\,\phi(-f/\Sigma_i) + 2\,\tilde f\Phi\left(-\tilde f/\Sigma_i\right) - 2\,\Sigma_i\,\phi\left(-\tilde f/\Sigma_i\right)\right].$$

C.10 Proof of Lemma 5.4.5

Proof. Recall that

$$f(u, \hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}$$

and

$$\tilde{f}(u, \hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}$$

and $\Sigma_i = \Sigma_i(u)$. We write

$$\partial_{\hat{\theta}_i} h = \partial_{\hat{\theta}_i} f\left[2\, f\, \Phi(f/\Sigma_i) + 2\,\Sigma_i\, \phi(-f/\Sigma_i) + 2\,\tilde{f}\,\Phi\left(-\tilde{f}/\Sigma_i\right) - 2\,\Sigma_i\,\phi\left(-\tilde{f}/\Sigma_i\right)\right]. \tag{C.10.1}$$

By Taylor expansion, we write

$$\begin{aligned}
&2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\right)\Phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}}{\Sigma}\right)\\
&= 2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\right)\cdot\\
&\left[\Phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right) - \frac{c}{\sqrt{1+u}\,\Sigma_i(u)}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right. \qquad \text{(C.10.2)}\\
&\left.+\frac{c^2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}}{\Sigma_i(u)}\right)\right].
\end{aligned}$$

For the value on the right-hand side of (C.10.2), we collect the terms with at least an order of c^2, which is dominated by

$$\tilde{R}_1(u,\hat{\theta}_i) := \frac{2\,\phi\left((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)}{(1+u)\,\Sigma_i(u)} + \max_{0\le c'\le c}\left[\frac{((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u})\,((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u})}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\right.$$
$$\left.\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}}{\Sigma_i(u)}\right)\right]. \tag{C.10.3}$$

Now, we need to choose c such that the function $\tilde{R}_1(u,\hat{\theta}_i)$ is integrable. The first term on the right hand side of (C.10.3) is integrable by the property of the normal density. For the second term, the maximal is attained at either 0, c, or the local minimal which satisfies

$$(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u} = \Sigma_i(u)\,.$$

Note that the following sum is integrable function

$$\left[\frac{((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u})\,((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u})}{2\,\Sigma(u)^3\,(1+u)^{3/2}}\right.$$
$$\left.\cdot\phi\left(\frac{((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u})}{\Sigma(u)}\right)\right]$$
$$+\left[\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L\,((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u})}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)\right],$$

which dominates the second term on the right-hand side of (C.10.3) when the maximal attains either 0 or c. For the local minimal case, we can choose $\delta^1_{t,\hat{\theta}_i}$ small enough such that for all $c < \delta^1_{t,\hat{\theta}_i}$, the local minimal lies in $[0, c]$ for only when $u > t$. Therefore, $\tilde{R}_1$ is integrable.

Next, we apply Taylor expansion around $c = 0$ to the second term on the right-hand side of (C.10.1).

$$
\begin{aligned}
&2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}\right)\Phi\left(-\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}}{\Sigma_i}\right)\\
&= 2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}\right)\\
&\left[\Phi\left(-\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right) - \frac{c}{\sqrt{1+u}\,\Sigma_i(u)}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.\\
&\left.+\frac{c^2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c'/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c'/\sqrt{1+u}}{\Sigma_i(u)}\right)\right].
\end{aligned}
\tag{C.10.4}
$$

Then, we define an integrable function $\tilde{R}_2$ and $\delta^2_{t,\hat{\theta}_i}$ in a similar manner of the equation (C.10.3). Similarly for the the third term and the forth term on the right-hand side of (C.10.1), we write

$$
\begin{aligned}
&2\,\Sigma_i(u)\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}}{\Sigma_i}\right) = 2\,\Sigma_i(u)\\
&\left[\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma(u)}\right) + \frac{c\,(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\sqrt{1+u}\,\Sigma_i(u)^2}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.\\
&\left.+\frac{c^2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right]
\end{aligned}
\tag{C.10.5}
$$

and

$$-2\,\Sigma_i(u)\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}}{\Sigma_i}\right) = -2\,\Sigma_i(u)$$

$$\left[\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right) - \frac{c\,(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\sqrt{1+u}\,\Sigma_i(u)}\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.$$

$$\left.+\frac{c^2\,((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u})}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right],$$

(C.10.6)

and there are integrable functions $\tilde{R}_3$ and $\tilde{R}_4$ and $\delta^3_{t,\hat{\theta}_i}$ and $\delta^4_{t,\hat{\theta}_i}$. By combining (C.10.2), (C.10.4), (C.10.5), and (C.10.6), we obtain

$$\left|\partial_{\hat{\theta}_i} h(u, \hat{\theta}_i) - 2\,(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^{2\,L} - S(t, u, \hat{\theta}_i)\right| \le c^2\,R(u, \hat{\theta}_i)\,, \tag{C.10.7}$$

where

$$S(t, u, \hat{\theta}_i) = \frac{2\,c}{\sqrt{1+u}}\,\left(\frac{1+t}{1+u}\right)^{2\,L}\left[\Phi\left(-(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)\right.$$
$$\left.-\Phi\left((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)\right] \tag{C.10.8}$$

and $R := \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3 + \tilde{R}_4$ and $\delta_{t,\hat{\theta}} = \min\left\{\delta^1_{t,\hat{\theta}_i}, \delta^2_{t,\hat{\theta}_i}, \delta^3_{t,\hat{\theta}_i}, \delta^4_{t,\hat{\theta}_i}\right\}$. □