Options market making using a stochastic control approach

Aigars Rosenberg

May 10, 2022

Supervised by Professor Zdzislaw Brzezniak

MSc Financial Engineering

University of York

Department of Economics and Related Studies

31,000 words

Abstract

In this paper, we are investigating the market making problem under a stochastic control framework by first deriving the results obtained in seminal work by [AS] for stock market making, which was then expanded by [ElAA3] into an options market making setting. We will focus on the derivation of analytic expressions for optimal controls using dynamic programming principle and their approximation in the case of risk-averse options market makers. We will use the same simplifying assumptions to prove these results, where we consider the market maker as an agent who seeks to maximise the expected utility of terminal wealth. Later we prove derivations in [ElAA3] for approximation of variance of the utility of terminal wealth for a market maker who also seeks to manage the inventory risk. Monte Carlo simulations are used to verify the performance of three different market makers: risk-neutral, risk-averse and zero intelligence as a baseline. In the simulation, we assume the stock price to follow the Heston model for which we briefly investigate some important results concerning option valuation. The aim is to show that stochastic control is a viable method of solving the market making problem under simplifying assumptions.

Contents

Introduction

The role of market makers in financial markets is to provide liquidity. This is achieved by sending limit orders to both sides of the order book, bid and ask. Spread between the bid and ask quotes is market makers profit that is realised continuously as these orders get executed. Orders are executed according to a priority queue, so orders with the highest priority will be executed first. Order priority is determined by the order matching system with a variety of rules. The most common priority rule is price level and time at which it was received, so the latest orders with the best prices are of the highest priority. We will only be looking at the price level as the variable that influences the execution probability and ignore the time implications. In real life, time is of critical importance in order execution and is viewed in terms of latency. We classify limit orders based on their price level as:

- **Aggressive:** If they cross the spread, in other words, orders are sent at opposite touch price or higher (lower) for ask (bid) orders, respectively. These execute immediately provided there is enough liquidity at limit price.

- **Non-passive:** If they are within the spread, so they become new best bid/ask, tightening the spread. These end up resting on the order book.

- **Passive:** If they do not cross the spread, in other words, orders are sent at touch price or lower (higher) for ask (bid) orders, respectively. These end up resting on the order book.

Touch prices are the best bid (ask) prices and opposite touch prices are best to ask (bid) prices, respectively. Market maker wants his quotes to be as far away from the mid price as possible to maximise earnings from incoming aggressive orders lifting (hitting) his resting ask (bid) quote. But, with the presence of competition for the order flow market makers cannot set the spread too wide as the probability of execution will decrease. So there is a balancing act between the increasing probability of execution and earning a spread. Ideally, market makers want to have their resting orders to be filled simultaneously, which does not happen in real life under continuous time. At the time instants where only one side is filled, the market maker is holding inventory and so is exposed to the risk of change in stock price. In order to manage this risk, market makers have to manage their inventory such that any directional exposure is minimised. Managing risk is even more vital in options market making owing to their non-linear nature, which increases severity of potential losses. To manage this risk, option market makers seek to hedge their exposure to "speed" and "acceleration" of changes in option price with respect to stock price and volatility, in addition to other risks. This problem is the market making problem.

1 Background

1.1 Lévy processes

Reader should proceed to Appendices 5.2 and 5.3 if not familiar with the background of the below results. Also refer to Appendix 5.1 for notations. For the rest of the section we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$.

Definition 1.1. (Lévy process [Ap, Section 1.3, p43]). Let $X = (X(t)), t \geq 0$ be an $\mathbb{R}^d$-valued process, we say that X is a Lévy process if the following conditions are satisfied:

(1) $X(0) = 0$ a.s.

(2) Independent increments: for each $n \in \mathbb{N}, 0 \leq t_1 \leq t_2 \leq ... \leq t_{n+1} < \infty$, $(X(t_{i+1}) - X(t_i)), i \in [1, n]$ are independent

(3) Stationary increments: for all $s \in [0, t)$, $X(t) - X(s)$ is equal in distribution to $X(t - s)$

(4) Stochastic continuity: for each $\epsilon > 0, t \geq 0$, $\lim_{h\to 0} \mathbb{P}(|X(t + h) - X(t)| > \epsilon) = 0$

Definition 1.2. (Brownian motion [BZ, Definition 6.9]). Let $W = (W_t)_{t\geq 0}$ be an $\mathbb{R}$-valued process. We call W standard Brownian motion if

(1) $W_0 = 0$ a.s.

(2) Independent increments: for each $n \in \mathbb{N}, 0 \leq t_1 \leq t_2 \leq ... \leq t_{n+1} < \infty$, $(W_{t_{i+1}} - W_{t_i}), i \in [1, n]$ are independent

(3) Sample paths continuity: $W : [0, \infty) \to \mathbb{R}$ are continious a.s.

(4) for all $s \in [0, t)$, $W_t - W_s \sim N(0, t - s)$, i.e.

$$\mathbb{P}(W_t - W_s = x) = \frac{1}{\sqrt{2\pi(t - s)}} e^{-\frac{x^2}{2(t-s)}} \text{ for all } x \in \mathbb{R} \tag{1.1}$$

Definition 1.3. (Poisson process [Ap, Example 1.3.9]). An $\mathbb{N}$-valued Lévy process $N = (N_t)_{t\geq 0}$ with intensity measure $\lambda \geq 0$ such that for every $t \geq 0$, $N_t \sim Pois(\lambda t)$, i.e.

$$\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \text{ for each } n \in \mathbb{N} \tag{1.2}$$

is a Poisson process. Compensated Poisson process denoted $\tilde{N} = (\tilde{N}_t)_{t\geq 0}$ is defined by

$$\tilde{N}_t = N_t - \lambda t \text{ for } t \geq 0 \tag{1.3}$$

Proposition 1.4. ***(Moments of compensated Poisson process).*** *Let $N = (N_t)_{t\geq 0}$ be a Poisson process with intensity λ, then $\mathbb{E}(\tilde{N}_t) = 0$ and $Var(\tilde{N}_t) = \lambda t$*

Proof. Since $N_t \sim Pois(\lambda t)$, by properties of Poisson distribution

$$\mathbb{E}(N_t) = \lambda t,\ Var(N_t) = \lambda t \tag{1.4}$$

Hence, by linearity of expectation and variance

$$\mathbb{E}(\tilde{N}_t) = \mathbb{E}(N_t - \lambda t) = \mathbb{E}(N_t) - \lambda t = 0 \tag{1.5}$$

$$Var(\tilde{N}_t) = Var(N_t - \lambda t) = Var(N_t) = \lambda t \tag{1.6}$$

□

Proposition 1.5. ***(Compensated Poisson process is a Martingale).*** *Let $\tilde{N} = (\tilde{N}_t)_{t\geq 0}$ be a Compensated Poisson process with intensity λ, then $\tilde{N}$ is a martingale.*

Proof. In $\tilde{N}_t = N_t - \lambda t$ is $(\mathcal{F}_t)$-measurable and so is $(\mathcal{F}_t)$-adapted. Moreover, it is integrable as $\mathbb{E}(\tilde{N}_t) = 0$ for all $t \geq 0$. Fix $0 \leq s \leq t$, then by adaptness and independence

$$\mathbb{E}(\tilde{N}_t|\tilde{N}_s) = \mathbb{E}(N_t - \lambda t|N_s) = \mathbb{E}(N_t|N_s) - \lambda t = N_s + \lambda s \tag{1.7}$$

Therefore, $\tilde{N}$ is a martingale. □

Definition 1.6. (Compound Poisson process [Ap, Example 1.3.10, Proposition 1.3.11], [OS, Example 1.6]). Let $(Z_n)_{n=1}^{\infty}$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}^d$ with common law μ_Z and let $N = (N_t)_{t\geq 0}$ be a Poisson process of intensity λ that is independent of all the Z_i. A compound Poisson process Y is defined as

$$Y_t = \sum_{i=1}^{N_t} Z_i \text{ for each } t \geq 0 \tag{1.8}$$

Proposition 1.7. ***(Compound Poisson process is a Lévy process).*** *Let $Y = (Y_t)_{t\geq 0}$ be a compound Poisson process from Definition 1.6, then Y is a Lévy process.*

Proof. To prove this, we verify that Y satisfies properties in Definition 1.1. Property (1) is satisfied as

$$\mathbb{P}(N_0 = n) = 0 \text{ for each } n \in \mathbb{N} \tag{1.9}$$

implies that

$$Y_0 = \mathbb{P}\left(\sum_{i=1}^{N_0} Z_i = 0\right) = 1 \tag{1.10}$$

Properties (2) and (3) are also satisfied as by Definition 1.6, each Z_i is an element of a sequence of i.i.d. random variables. Independence of each Z_i implies that their sum is also independent. Moreover, for each $n, m \in \mathbb{N}$ with $0 \leq n < m$

$$\textstyle\sum_{i=1}^{m-n} Z_i \quad = \quad Z_1 + Z_2 + ... + Z_{m-n} \tag{1.11}$$

and

$$\begin{aligned} \textstyle\sum_{i=1}^m Z_i - \sum_{i=1}^n Z_i &= Z_1 + Z_2 + ... + Z_m - Z_1 - Z_2 - ... - Z_n \\ &= Z_m + Z_{m-1} + ... + Z_{m-n} \end{aligned} \tag{1.12}$$

Independence and identical distribution of each Z_i implies that

$$\sum_{i=1}^{m-n} Z_i \stackrel{d}{=} \sum_{i=1}^{m} Z_i - \sum_{i=1}^{n} Z_i \text{ for each } n, m \in \mathbb{N},\ 0 \leq n < m \tag{1.13}$$

For property (4) note that

$$\lim_{h \to 0} \mathbb{P}(|Y_{t+h} - Y_t| > 0) = 0 \tag{1.14}$$

This condition implies that

$$\lim_{h \to 0} \mathbb{P}(|Y_{t+h} - Y_t| = 0) = 1 \tag{1.15}$$

where

$$|Y_{t+h} - Y_t| = \left| \sum_{i=1}^{N_{t+h}} Z_i - \sum_{i=1}^{N_t} Z_i \right| \tag{1.16}$$

Since

$$\lim_{h \to 0} \mathbb{P}(N_{t+h} = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t},\ \lim_{h \to 0} \mathbb{P}(N_h = n) = 0 \text{ for each } n \in \mathbb{N} \tag{1.17}$$

We deduce that

$$\lim_{h \to 0} \mathbb{P}\left(N_{t+h} - N_t + N_h = 0\right) = 1 \tag{1.18}$$

Hence

$$\lim_{h \to 0} \mathbb{P}\left(\left| \sum_{i=1}^{N_{t+h}} Z_i - \sum_{i=1}^{N_t} Z_i \right| = 0 \right) = 1 \tag{1.19}$$

Therefore Y is a Lévy process. $\square$

Definition 1.8. (Markov process [Ap, Subsection 3.1.1, p143-144]). Let $X = (X_t)_{t \geq 0}$ be an $(\mathcal{F}_t)$-adapted process. X is a Markov process if for all bounded and Borel measurable functions f

$$\mathbb{E}\left(f(X_t) | \mathcal{F}_s\right) = \mathbb{E}\left(f(X_t) | X_s\right) \text{ for all } 0 \leq s \leq t < \infty \tag{1.20}$$

Proposition 1.9. ***(Brownian motion is a Lévy process).*** *Let $W = (W_t)_{t \geq 0}$ is be a $\mathbb{R}^d$-valued Brownian motion. It is a Lévy process.*

Proof. To prove this proposition, we verify that W satisfies properties in Definition 1.1. Properties (1), (2) and (3) is satisfied immediately by Definition 1.2. It remains to prove property (4), where we need to verify that

$$\lim_{h\to 0}\mathbb{P}(|W_{t+h}-W_t|>\epsilon)=0 \text{ for each } \epsilon>0, t\geq 0 \tag{1.21}$$

This condition implies a convergence in probability from Definition (5.64)

$$\lim_{h\to 0}\mathbb{P}(|W_{t+h}-W_t|=0)=1 \tag{1.22}$$

Note that by Definition 1.2, sample paths of W are continious, which by Definition 5.11 implies that

$$W_t=\lim_{s\uparrow t}W_s=\lim_{s\downarrow t}W_s \tag{1.23}$$

Hence

$$\lim_{h\to 0}W_{t+h}=W_t \Rightarrow \lim_{h\to 0}W_{t+h}-W_t=0 \tag{1.24}$$

and

$$\mathbb{P}(\lim_{h\to 0}W_{t+h}-W_t=0)=1 \tag{1.25}$$

Which is an almost sure convergence from Definition 5.65. By convergence relations from Definition 5.67, almost sure convergence implies convergence in probability. Therefore property (4) is satisfied and W is indeed a Lévy process. □

Definition 1.10. (Càdlàg process [Ap, Chapter 2, p82]). Let $X=(X(t,\omega)), t\geq 0$ be an $\mathbb{R}^d$-valued process. We say that X is a càdlàg process if its sample paths are càdlàg, i.e. a mapping

$$\mathbb{R}^+ \ni t \to X(t,\omega)\in\mathbb{R}^d \text{ for each } \omega\in\Omega \tag{1.26}$$

satisfies properties from Definition 5.11.

Proposition 1.11. ***(Adapted and càdlàg process [Ph, Proposition 1.1.1]).*** *Let $X=(X(t)), t\geq 0$ be an $\mathbb{R}^d$-valued process. If X is $(\mathcal{F}_t)$-adapted and càdlàg, then X is $(\mathcal{F}_t)$-progressively measurable.*

Theorem 1.12. ***(Càdlàg modification [Ap, Theorem 2.1.8]).*** *If $X=(X(t,\omega)), t\geq 0$ is an $\mathbb{R}^d$-valued Lévy process, then there always exists a càdlàg modification to sample paths of X, such that X is still a Lévy process. By càdlàg modification we mean that mapping*

$$\mathbb{R}^+ \ni t \to X(t,\omega)\in\mathbb{R}^d \text{ for each } \omega\in\Omega \tag{1.27}$$

is modified such that it satisfies properties from Definition 5.11.

Definition 1.13. (Jump process [Ap, Chapter 2: Section 2.3]). Let $X=(X(t)), t\geq 0$ be an $\mathbb{R}^d$-valued Lévy process with càdlàg sample paths. A jump process $\Delta X=(\Delta X(t)), t\geq 0$ starting at

$\Delta X(t) = 0$ associated with X is defined by

$$\Delta X(t) = X(t) - X(t-) \text{ for each } t \in (0, T] \tag{1.28}$$

Theorem 1.14. ***(Càdlàg process of finite variation [Ap, Theorem 2.3.14]).*** *Let $X = (X_t)_{t\geq 0}$ be $\mathbb{R}^d$-valued càdlàg process defined on $[0,T]\times\Omega$. If X a process of finite variation then*

$$\sum_{0\leq t\leq T} |\Delta X(t)| \leq V_0^T(X) \text{ for } T > 0 \tag{1.29}$$

Theorem 1.15. ***(Levy process with bounded jumps [Ap, Theorem 2.4.7]).*** *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with associated jump process $\Delta X = (\Delta X_t)_{t\geq 0}$ as in Definition 1.13. We say that X has bounded jumps if there exists a constant $C > 0$ such that*

$$\sup_{t\geq 0} |\Delta X_t| \leq C \text{ for all } t > 0, \omega \in \Omega \tag{1.30}$$

If X has bounded jumps then

$$\mathbb{E}\left(|X_t|^n\right) < \infty \text{ for all } n \in \mathbb{N}, t > 0 \tag{1.31}$$

Definition 1.16. (Lévy measure [Ap, Subsection 1.2.4, p28-29]). Let ν be a measure on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$. We say that it is a Lévy measure if

$$\int_{\mathbb{R}^d\setminus\{0\}} (a \wedge |x|^2)\nu(dx) = \int_{|x|\geq a} \nu(dx) + \int_{|x|<a} |x|^2\nu(dx) < \infty \tag{1.32}$$

where $a > 0$.

Proposition 1.17. ***(Lévy measure is σ-finite measure [Ap, Exercise 1.2.13]).*** *If ν is a Lévy measure on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$, then it is a σ-finite measure.*

Proof. Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathbb{R}^d \setminus \{0\}$ such that $\mathbb{R}^d \setminus \{0\} = \bigcup_{n=1}^\infty B_n$. Since ν is a Lévy measure, we have that

$$\int_{\mathbb{R}^d\setminus\{0\}} (a \wedge |x|^2)\nu(dx) = \int_{\bigcup_{n=1}^\infty B_n} (a \wedge |x|^2)\nu(dx) < \infty \tag{1.33}$$

Hence

$$\int_{B_n} (a \wedge |x|^2)\nu(dx) < \infty \text{ for each } n \in \mathbb{N} \tag{1.34}$$

Then, by Definition 5.21, ν is a σ-finite measure. $\square$

Definition 1.18. (Random measure [Ap, Subsection 2.3.1, p103-104]). A random measure N on $(M, \mathcal{B}(M))$ is a collection of random variables $(N(A,\omega)), A \in \mathcal{B}(M), \omega \in \Omega$ such that

(1) for each $\omega \in \Omega$, $N(\cdot,\omega)$ is a measure on $(M, \mathcal{B}(M))$

(2) for each $A \in \mathcal{B}(M)$, $N(A,\cdot)$ is a random variable

Definition 1.19. (Counting measure [Ap, Section 2.3, p100]). Let $X = (X(t,\omega)), t \geq 0$ be an $\mathbb{R}^d$-valued Lévy process with càdlàg sample paths. Let $\Delta X = (\Delta X(t,\omega)), t \geq 0$ be its associated jump process with $\Delta X(0,\omega) = 0$. We define a counting measure on $(M, \mathcal{B}(M))$ by

$$N(t, A, \omega) = \#\{\Delta X(s,\omega) \in A : 0 < s \leq t\} = \sum_{s \in (0,t]} \mathbf{1}_{\{\Delta X(s,\omega) \in A\}} \in \mathbb{N} \cup \{\infty\} \text{ for } A \in \mathcal{B}(M), \omega \in \Omega \quad (1.35)$$

Proposition 1.20. ***(Counting measure is a measure).*** *Let N be a counting measure on $(M, \mathcal{B}(M))$, then N is a measure.*

Proof. First note that process X from Definition 1.19 is càdlàg, hence left limits $X(s-,\omega) = \lim_{u \uparrow s} X(u,\omega)$ exists at each point $s \in (0,t]$. Now, we verify conditions from Definition 5.16. Non-negativity is satisfied as N takes values in $\mathbb{N} \cup \{\infty\}$. Moreover,

$$N(t, \emptyset, \omega) = \#\{\Delta X(s,\omega) \in \emptyset : 0 < s \leq t\} = 0 \quad (1.36)$$

and so Null-empty set condition is also satisfied. Finally, given any sequence $(B_n)_{n \in \mathbb{N}}$ of pairwise disjoint sets such that $A = \bigcup_{n \in \mathbb{N}} B_n \in \mathcal{B}(M)$ we have that

$$\begin{aligned} N\left(t, \bigcup_{n \in \mathbb{N}} B_n, \omega\right) &= \#\left\{\Delta X(s,\omega) \in \bigcup_{n \in \mathbb{N}} B_n : 0 < s \leq t\right\} \\ &= \sum_{n \in \mathbb{N}} \#\left\{\Delta X(s,\omega) \in B_n : 0 < s \leq t\right\} \text{ for } \omega \in \Omega \end{aligned} \quad (1.37)$$

Therefore, Countable additivity is also satisfied and N is a measure. ☐

Definition 1.21. (Intensity measure [Ap, Section 2.3, p101]) Let N be a counting measure on $(M, \mathcal{B}(M))$, we call ν the intensity measure of N on $(M, \mathcal{B}(M))$ if

$$\nu(A) = \mathbb{E}\left(N(1, A)\right) \text{ for } A \in \mathcal{B}(M) \quad (1.38)$$

Proposition 1.22. ***(Intensity measure is a measure).*** *Let N be a counting measure on $(M, \mathcal{B}(M))$ and let ν be the intensity measure of N on $(M, \mathcal{B}(M))$, then ν is a measure.*

Proof. We verify conditions from Definition 5.16. Non-negativity is satisfied as $N(1, A) \geq 0$ for any $A \in \mathcal{B}(M)$, and so $\nu(A) = \mathbb{E}\left(N(1,A)\right) \geq 0$ for any $A \in \mathcal{B}(M)$. Null-empty set condition is trivially satisfied. Finally, given any sequence $(B_n)_{n \in \mathbb{N}}$ of pairwise disjoint sets such that $A = \bigcup_{n \in \mathbb{N}} B_n \in \mathcal{B}(M)$ we have by linearity of expectation that

$$\begin{aligned} \nu\left(\bigcup_{n \in \mathbb{N}} B_n\right) &= \mathbb{E}\left(N\left(1, \bigcup_{n \in \mathbb{N}} B_n\right)\right) \\ &= \mathbb{E}\left(\#\left\{\Delta X(s) \in \bigcup_{n \in \mathbb{N}} B_n : 0 < s \leq 1\right\}\right) \\ &= \sum_{n \in \mathbb{N}} \mathbb{E}\left(\#\left\{\Delta X(s) \in B_n : 0 < s \leq 1\right\}\right) \\ &= \sum_{n \in \mathbb{N}} \nu\left(B_n\right) \end{aligned} \quad (1.39)$$

Therefore, countable additivity is also satisfied and ν is a measure. ☐

Definition 1.23. (Poisson random measure [IW, Chapter 1, Definition 8.1]). Let N be a random variable defined by map

$$\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^d \setminus \{0\}) \times \Omega \ni (t, A, \omega) \to N(t, A, \omega) \in \mathbb{N} \cup \{\infty\} \tag{1.40}$$

We call N a Poisson random measure on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ with intensity measure ν on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$ if

(1) for each $t \geq 0, \omega \in \Omega$, $N(t, \cdot, \omega)$ is a counting measure on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$

(2) for each $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, $N(\cdot, A, \cdot)$ is a random variable such that $N(\cdot, A, \cdot) \sim Pois(\nu(A))$ where $\nu(A) = \mathbb{E}\left(N(\cdot, A, \cdot)\right)$

(3) for each disjoint $A_1, ..., A_n \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, $N(\cdot, A_1, \cdot), ..., N(\cdot, A_n, \cdot)$ are independent random variables

Theorem 1.24. ***(Existence of Poisson random measure [IW, Theorem 8.1]).*** *Given an intensity measure ν on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$ that is σ-finite, there exists a Poisson random measure N with*

$$\mathbb{E}(N(t, A)) = t\nu(A) \text{ for every } t \geq 0, A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}) \tag{1.41}$$

Remark 1.25. **(Bounded below [Ap, Section 2.3, p101], [Pr, Theorem 34]).** Let $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, then we say A is bounded below if $\bar{A} \notin 0$, i.e there exists $r > 0$ such that

$$B^d(r) \cap A = \emptyset \tag{1.42}$$

where

$$B^d(r) = \{x \in \mathbb{R}^d : |x| < r\} \tag{1.43}$$

Definition 1.26. (Poisson point process [Ap, Subsection 2.3.1, p104-105]). Let $(U, \mathcal{U})$ and $(S, \mathcal{S})$ be two measurable spaces, where $S = \mathbb{R}^+ \times U$ and $\mathcal{S} = \mathcal{B}(\mathbb{R}^+) \times \mathcal{U}$. Let $p = (p(t)), t \geq 0$ be a U-valued $\mathcal{U}$-adapted process such that N is a Poisson random measure on $(S, \mathcal{S})$, where

$$N((0, T] \times A) = \#\{p(t) \in A : 0 < t \leq T\} \text{ for each } T \geq 0, A \in \mathcal{U} \tag{1.44}$$

We say that p is a Poisson point process and N is its associated Poisson random measure.

Proposition 1.27. ***(Jump process is a Poisson point process [Ap, Subsection 2.3.1, p104-105]).*** *Let $X = (X_t)_{t \geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with càdlàg sample paths. Let $\Delta X = (\Delta X_t)_{t \geq 0}$ be its associated jump process with $\Delta X_0 = 0$. Let $(U, \mathcal{U})$ be a measure space from Definition 1.26. If $U = \mathbb{R}^d \setminus \{0\}$, $\mathcal{U} = \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ and $A \in \mathcal{U}$ is bounded below. Then ΔX* is a *Poisson point process and N is its associated Poisson random measure.*

Theorem 1.28. ***(Integer-valued Lévy process [Ap, Theorem 2.2.13]).*** *Suppose we have an integer valued Lévy process $N = (N_t)_{t \geq 0}$, then N is a Poisson process if*

(1) N is increasing a.s.

(2) $\Delta N = (\Delta N_t)_{t\geq 0}$ takes values in $\{0, 1\}$

Proposition 1.29. ***(Collection of Poisson random measures is a Lévy process [Pr, Corollary p.27]).*** *Let $(N(t, A_n))_{n\in\mathbb{N}}$ be a collection of Poisson random measures with $A_1, ..., A_n \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$. Then $(N(t, A_n))_{n\in\mathbb{N}}$ is a Lévy process.*

Lemma 1.30. ***(Integrability of Poisson random measure [Ap, Lemma 2.3.4]).*** *Let $N = (N(t, A)), t \geq 0$ be a Poisson random measure with intensity measure ν on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$. If $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ is bounded below, then*

$$\mathbb{P}\left(N(t, A) < \infty\right) = 1 \textit{ for all } t \geq 0 \tag{1.45}$$

Theorem 1.31. ***(Special case of Poisson random measure [Ap, Theorem 2.3.5]).*** *Let $N = (N(t, A)), t \geq 0$ be a Poisson random measure with intensity measure ν on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$. If $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ is bounded below, then N is a Poisson process with intensity $\nu(A) = \mathbb{E}\left(N(1, A)\right)$.*

Definition 1.32. (Compensated Poisson random measure [Ap, Subsection 2.3.1, p105]). Let $N = (N(t, A)), t \geq 0$ be a Poisson random measure with intensity measure ν on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$. For each $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ bounded below, we define a compensated Poisson random measure by

$$\tilde{N}(t, A) = N(t, A) - t\nu(A) \text{ for each } t \geq 0 \tag{1.46}$$

which is a martingale.

1.2 Stochastic integrals

Definition 1.33. ([IW, Chapter 2: Definition 1.1, Remark 1.1]). Let $\Phi = (\Phi_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued and $(\mathcal{F}_t)$-predictable process, $\Phi \in \mathscr{L}_2$ if

$$\mathbb{E}\left(\int_0^T |\Phi_t|^2 dt\right) < \infty \text{ for every } T > 0 \tag{1.47}$$

Definition 1.34. ([IW, Chapter 2: Definition 1.6, Remark 1.1]). Let $\Phi = (\Phi_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued and $(\mathcal{F}_t)$-predictable process, $\Phi \in \mathscr{L}_2^{loc}$ if

$$\mathbb{P}\left(\int_0^T |\Phi_t|^2 dt < \infty\right) = 1 \text{ for every } T > 0 \tag{1.48}$$

Definition 1.35. ([IW, Chapter 2: Definition 1.3]). Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued process, $X \in \mathscr{M}_2$ if

(1) X is a L^2-martingale with respect to $(\mathcal{F}_t)$

(2) $X_0 = 0$ a.s.

Futhermore, if the above conditions are satisfied then $X \in \mathscr{M}_2^c$ if

(3) Sample paths of X, $t \to X_t$ are continious a.s.

Definition 1.36. ([IW, Chapter 2: Definition 1.8]). Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued process, $X \in \mathscr{M}_2^{loc}$ if

(1) X is a local L^2-martingale with respect to $(\mathcal{F}_t)$

(2) $X_0 = 0$ a.s.

Futhermore, if the above conditions are satisfied then $X \in \mathscr{M}_2^{c,loc}$ if

(3) Sample paths of X, $t \to X_t$ are continious a.s.

Definition 1.37. ([IW, Chapter 2: Definition 2.2]). Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued process such that $X \in \mathscr{M}_2$. Let $\Phi = (\Phi_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-predictable process, $\Phi \in \mathscr{L}_2(X)$ if

$$\mathbb{E}\left(\int_0^T |\Phi_t|^2 d\left\langle X\right\rangle_t\right) < \infty \text{ for every } T > 0 \tag{1.49}$$

where $\left\langle X\right\rangle_t$ is a quadratic variation of process X as in Theorem 5.73.

Definition 1.38. ([IW, Chapter 2: Definition 2.4]). Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued process such that $X \in \mathscr{M}_2^{loc}$. Let $\Phi = (\Phi_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-predictable process, $\Phi \in \mathscr{L}_2^{loc}(X)$ if there exists a sequence of stopping times s_n such that $s_n \uparrow \infty$ a.s. and

$$\mathbb{E}\left(\int_0^{T\wedge s_n} |\Phi_t|^2 d\left\langle X\right\rangle_t\right) < \infty \text{ for every } T > 0, n = 1, 2, ... \tag{1.50}$$

where$\langle X\rangle_t$ is a quadratic variation of process X as in Theorem 5.73. If $\langle X\rangle_t$ is continuous then above is equivalent to

$$\mathbb{P}\left(\int_0^T |\Phi_t|^2 d\langle X\rangle_t < \infty\right) = 1 \text{ for all } T > 0 \tag{1.51}$$

Definition 1.39. (Brownian stochastic integral [IW, Chapter 2: Definition 1.5]). Let $W = (W_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-Brownian motion. Let $\Phi = (\Phi_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued and $(\mathcal{F}_t)$-predictable process such that $\Phi \in \mathscr{L}_2$. We define a process $(I_t(\Phi))_{t\geq 0} \in \mathscr{M}_2^c$ where for each fixed $t > 0$

$$I_t(\Phi) = \int_0^t \Phi_s dW_s \tag{1.52}$$

is a random variable called a stochastic integral with respect to Brownian motion.

Proposition 1.40. ***(Properties of Brownian stochastic integral [IW, Chapter 2: Proposition 1.1]).*** *Let $I_t(\Phi)$ be a stochastic integral with respect to an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-Brownian motion $W = (W_t)_{t\geq 0}$. It has following properties for each $0 \leq s < t$*

(1) $\mathbb{E}\left(I_t(\Phi) - I_s(\Phi) | \mathcal{F}_s\right) = 0$

(2) $\mathbb{E}\left(\left(I_t(\Phi) - I_s(\Phi)\right)^2 \middle| \mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t |\Phi_u|^2 du \middle| \mathcal{F}_s\right)$

Proposition 1.41. ***([IW, Chapter 2: Proposition 2.4]).*** *Let $M \in \mathscr{M}_2$ and $\Phi \in \mathscr{L}_2(M)$ (or $M \in \mathscr{M}_2^{loc}$ and $\Phi \in \mathscr{L}_2^{loc}(M)$). Then $X = I_T^M(\Phi)$ is characterized as the unique $X \in \mathscr{M}_2$ ($X \in \mathscr{M}_2^{loc}$) such that*

$$\langle X, N\rangle_T = \int_0^T \Phi(t) d\langle M, N\rangle_t \textit{ for every } N \in \mathscr{M}_2 \textit{ (or } N \in \mathscr{M}_2^{loc}\textit{), for all } T \geq 0 \tag{1.53}$$

Proposition 1.42. ***([IW, Chapter 2: Proposition 2.5]).*** *With $T > 0$*

(1) Let $X, Y \in \mathscr{M}_2^{loc}$ and $\Phi \in \mathscr{L}_2^{loc} X) \cap \mathscr{L}_2^{loc}(Y)$. Then $\Phi \in \mathscr{L}_2^{loc}(X+Y)$ and

$$\int_0^T \Phi(t) d(X+Y)(t) = \int_0^T \Phi(t) dX(t) + \int_0^T \Phi(t) dY(t) \tag{1.54}$$

(2) Let $X \in \mathscr{M}_2^{loc}$ and $\Phi, \Psi \in \mathscr{L}_2^{loc}(X)$. Then

$$\int_0^T (\Phi + \Psi)(t) dX(t) = \int_0^T \Phi(t) dX(t) + \int_0^T \Psi(t) dX(t) \tag{1.55}$$

(3) Let $X \in \mathscr{M}_2^{loc}$ and $\Phi \in \mathscr{L}_2^{loc}(X)$. Set $Y = I_T^M(\Phi)$ and let $\Psi \in \mathscr{L}_2^{loc}(Y)$. Then $\Phi\Psi \in \mathscr{L}_2^{loc}(X)$

$$\int_0^T (\Phi\Psi)(t) dX(t) = \int_0^T \Psi(t) dY(t) \tag{1.56}$$

Definition 1.43. (Poisson integral [Ap, Subsection 2.3.2, p106]). Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with càdlàg sample paths. Let $\Delta X = (\Delta X_t)_{t\geq 0}$ be its associated jump process with $\Delta X_0 = 0$. Let $N = (N(t, A)), t \geq 0$ be a Poisson random measure on $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ with intensity measure ν and $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ is bounded below. Here for each $t \geq 0, \omega \in \Omega$, N is a counting measure of jumps of X. Suppose $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ is a Borel measurable function. We define an $\mathbb{R}^d$-valued random variable

$$\int_A \Phi(x)N(t,dx) = \sum_{s\in(0,t]} \Phi(\Delta X(s))\mathbf{1}_{\{\Delta X(s)\in A\}} \text{ for each } t>0, \omega\in\Omega \tag{1.57}$$

called a Poisson integral. It is interpreted as a random finite sum, where collection $\left(\int_A \Phi(x)N(t,dx)\right), t \geq 0$ is a stochastic process.

Theorem 1.44. ***(Poisson integral is a Compound Poisson process [Ap, Theorem 2.3.9]).*** *Let all of the assumptions from Definition 1.43 hold, then* $\left(\int_A \Phi(x)N(t,dx)\right), t \geq 0$ *is a compound Poisson process.*

Theorem 1.45. ***(Poisson integrals for L^p integrands [Pr, Theorem 38]).*** *Let* $N = (N(t,A)), t \geq 0$ *be a Poisson random measure on* $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ *with* σ*-finite intensity measure* ν *satisfying Definition 1.16, where* $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ *is bounded below. If* $\Phi(z)$ *is bounded in* L^p*, i.e. satisfy Definition 5.69 for*
(1) $p = 1$*, then*

$$\mathbb{E}\left(\int_A \Phi(x)N(t,dx)\right) = t\int_A \Phi(x)\nu(dx) \tag{1.58}$$

(2) $p = 2$*, then*

$$\mathbb{E}\left(\left|\int_A \Phi(x)N(t,dx) - t\int_A \Phi(x)\nu(dx)\right|^2\right) = t\int_A |\Phi(x)|^2\,\nu(dx) \tag{1.59}$$

Theorem 1.46. ***(Lévy processes on disjoint sets [Pr, Theorem 39]).*** *Let* $X = (X_t)_{t\geq 0}$ *be an* $\mathbb{R}^d$*-valued Lévy process with càdlàg sample paths. Let* $\Delta X = (\Delta X_t)_{t\geq 0}$ *be its associated jump process with* $\Delta X_0 = 0$*. Suppose* $A_1, A_2 \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ *such that* $A_1 \cap A_2 = \emptyset$ *and* A_1, A_2 *are bounded below. Then processes* $Y^1 = (Y^1_t)_{t\geq 0}, Y^2 = (Y^2_t)_{t\geq 0}$ *defined by*

$$Y^1_t = \sum_{s\in(0,t]} \Delta X_s \mathbf{1}_{\{\Delta X_s\in A_1\}} \textit{ for each } t>0 \tag{1.60}$$

$$Y^2_t = \sum_{s\in(0,t]} \Delta X_s \mathbf{1}_{\{\Delta X_s\in A_2\}} \textit{ for each } t>0 \tag{1.61}$$

are independent Lévy processes.

Proposition 1.47. ***(Independence of Brownian motion and Poisson process).*** *Let* $W = (W_t)_{t\geq 0}$ *be an* $\mathbb{R}$*-valued* $(\mathcal{F}_t)$*-Brownian motion and let* $N = (N_t)_{t\geq 0}$ *be a* $(\mathcal{F}_t)$*-Poisson process. Process* W *is independent of* N*.*

Proof. Note that both W, N are Lévy processes by Proposition 1.9 and by Definitions 1.3. Moreover, N is a special case of Poisson random measure for $A \in \mathcal{B}(\mathbb{R}\setminus\{0\})$ bounded below as defined in Theorem (1.31).

By Definition (1.2) sample paths of W are continious, which in turn implies by Definition 5.11 that

$$W_{t-} = \lim_{s \uparrow t} W_s = W_t \text{ for } t \geq 0 \tag{1.62}$$

Hence

$$\Delta W_t = W_t - W_{t-} = 0 \text{ for } t \geq 0 \tag{1.63}$$

Let B be a Borel set such that

$$\{\Delta W_t \in B : 0 < s \leq t\} \tag{1.64}$$

In this case $B \in \mathcal{B}(\{0\})$ and so $A \cap B = \emptyset$. Therefore, by Theorem 1.46 processes W, N are independent. $\square$

Theorem 1.48. ***(Decomposition of Lévy process with bounded jumps [Pr, Theorem 41]).*** *Let $Z = (Z_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with càdlàg sample paths and jumps bounded by a as in Theorem 1.15. Let $\Delta Z = (\Delta Z_t)_{t\geq 0}$ be its associated jump process with $\Delta Z_0 = 0$. Let $N = (N(t, A)), t \geq 0$ be a Poisson random measure on $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ with intensity measure ν and $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$. Here for each $t \geq 0, \omega \in \Omega$, N is a counting measure of jumps of Z. Let $Y = (Y_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued process where each $Y_t = Z_t - \mathbb{E}(Z_t)$. Then*

$$Y_t = Y_t^c + Y_t^d = Y_t^c + \int_{|z|\leq a} z\left(N(t, dz) - t\nu(dz)\right) \text{ for each } t \geq 0 \tag{1.65}$$

where Y_t^c has continious sample paths, Y_t^c, Y_t^d are L^2-martingales and independent Lévy processes.

Lemma 1.49. ***(Integrability of Compound Poisson process [Ap, Lemma 2.5.1]).*** *If $Y = (Y_t)_{t\geq 0}$ is an $\mathbb{R}^d$-valued Compound Poisson process, $(Z_n)_{n=1}^{\infty}$ is a sequence of i.i.d. $\mathbb{R}^d$-valued random variables with common law μ_Z and $N = (N_t)_{t\geq 0}$ is a Poisson process of intensity λ that is independent of all the Z_n. Then*

$$\mathbb{E}\left(|Y_t|^n\right) = \mathbb{E}\left(\left|\sum_{m=1}^{N_t} Z_m\right|^n\right) < \infty \text{ for all } t > 0 \text{ and for each } n = 1, 2, ... \tag{1.66}$$

if and only if

$$\mathbb{E}\left(|Z_m|^n\right) < \infty \text{ for each } m = 1, ..., N_t,\ n = 1, 2, ... \tag{1.67}$$

Theorem 1.50. ***(Integrability of Lévy process [Ap, Theorem 2.5.2]).*** *If $X = (X_t)_{t\geq 0}$ is an $\mathbb{R}^d$-valued Lévy process with intensity measure ν. Then*

$$\mathbb{E}\left(|X_t|^n\right) < \infty \text{ for all } t > 0 \text{ and for each } n = 1, 2, ... \tag{1.68}$$

if and only if

$$\int_{|x|\geq a} |x|^n \nu(dx) < \infty \text{ for each } n = 1, 2, ... \tag{1.69}$$

Definition 1.51. (Classes of Predictable integrands [IW, Chapter 2: Section 3]). Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with càdlàg sample paths. Let $\Delta X = (\Delta X_t)_{t\geq 0}$ be its associated jump process with $\Delta X_0 = 0$. Let $N = (N(t, A)), t \geq 0$ be a Poisson random measure on $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ with intensity measure ν and $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ being bounded below. Here for each $t \geq 0, \omega \in \Omega$, N is a counting measure of jumps of X. Let $\Phi : [0, T] \times A \to \mathbb{R}^d$ is an $(\mathcal{F}_t)$-predictable mapping as in Definition 5.52. Then

(1) $\Phi \in \boldsymbol{F}$ if $\mathbb{P}\left(\int_0^T \int_A |\Phi(t, x)| N(dt, dx) < \infty\right) = 1$ for every $T > 0$

(2) $\Phi \in \boldsymbol{F}_1$ if $\mathbb{E}\left(\int_0^T \int_A |\Phi(t, x)|\nu(dx)dt\right) < \infty$ for every $T > 0$

(3) $\Phi \in \boldsymbol{F}_1^{loc}$ if $\mathbb{P}\left(\int_0^T \int_A |\Phi(t, x)|\nu(dx)dt < \infty\right) = 1$ for every $T > 0$

(4) $\Phi \in \boldsymbol{F}_2$ if $\mathbb{E}\left(\int_0^T \int_A |\Phi(t, x)|^2\nu(dx)dt\right) < \infty$ for every $T > 0$

(5) $\Phi \in \boldsymbol{F}_2^{loc}$ if $\mathbb{P}\left(\int_0^T \int_A |\Phi(t, x)|^2\nu(dx)dt < \infty\right) = 1$ for every $T > 0$

Definition 1.52. (Poisson stochastic integral [Ap, Subsection 4.3.2, p230-231]). Let all of the assumptions from Definition 1.51 hold. Let $P = (P_t)_{t\geq 0}$ be a compound Poisson process where each $P_t = \int_A zN(t, dx)$ and Let $\Delta P = (\Delta P_t)_{t\geq 0}$ be its associated jump process with $\Delta P_0 = 0$. Suppose $\Phi : [0, T] \times A \to \mathbb{R}^d \in \boldsymbol{F}$. We define an $\mathbb{R}^d$-valued random variable

$$I_T(\Phi) = \int_0^T \int_A \Phi(t, x)N(dt, dx) = \sum_{t\in[0,T]} \Phi(t, \Delta P_t)\mathbf{1}_{\{\Delta P_t \in A\}} \text{ for } T > 0 \tag{1.70}$$

called a stochastic Poisson integral. It is interpreted as a random finite sum, where collection $(I_T(\Phi)), T \geq 0$ is a stochastic process.

Theorem 1.53. ***(Properties of Poisson stochastic integral [IW, Chapter 2: Section 3], [Ap, Subsection 4.3.2, p230-231]).*** *Let all of the assumptions from Definition 1.52 hold. Let $\Phi \in \boldsymbol{F}$ and set*

$$I_T^M(\Phi) = \int_0^T \int_A \Phi(t, x)\tilde{N}(dt, dx) = \textit{ for } T > 0 \tag{1.71}$$

We then may define

$$I_T^M(\Phi) = \int_0^T \int_A \Phi(t, x)N(dt, dx) - \int_0^T \int_A \Phi(t, x)\nu(dx)dt = \int_0^T \int_A \Phi(t, x)\left(N(dt, dx) - \nu(dx)dt\right) \textit{ for } T > 0 \tag{1.72}$$

given that any of the below integrability (or square integrability) conditions hold:

(1) $\Phi \in \boldsymbol{F}_1$ then $I_T^M(\Phi)$ is a $(\mathcal{F}_t)$-martingale

(2) $\Phi \in \boldsymbol{F}_1^{loc}$ then $I_T^M(\Phi)$ is a local $(\mathcal{F}_t)$-martingale

(3) $\Phi \in \boldsymbol{F}_2$ then $I_T^M(\Phi) \in \mathscr{M}_2$

(4) $\Phi \in \boldsymbol{F}_2^{loc}$ then $I_T^M(\Phi) \in \mathscr{M}_2^{loc}$

Theorem 1.54. ***(Lévy-Itô decomposition [OS, Theorem 1.7], [Ap, Theorem 2.4.16]).*** *If* $X = (X(t)), t \geq 0$ *is an* $\mathbb{R}^d$*-valued Lévy process with Lévy measure* ν*. Then process* X *can be decomposed as*

$$X(t) = bt + \sigma W(t) + \int_{|x|<a} x\tilde{N}(t, dx) + \int_{|x|\geq a} xN(t, dx) \tag{1.73}$$

where $0 \leq n \leq d$, a is an $[0, \infty]$-valued constant, x, b *are an* $\mathbb{R}^d$-valued vectors, σ *is an* $\mathbb{R}^{d\times n}$*-valued matrix,* $W = (W(t)), t \geq 0$ *is an* $\mathbb{R}^n$*-valued Brownian motion,* $N = (N(t, x)), t \geq 0$ *is a Poisson random measure on* $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^d \setminus \{0\})$*,* $\tilde{N} = (\tilde{N}(t, x))), t \geq 0$ *is a Compensated Poisson random measure on* $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^d \setminus \{0\})$*,* ν *is Lévy intensity measure of* N *on* $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$ *and* $W, N, \tilde{N}$ *are independent.*

1.3 Stochastic differential equations

Definition 1.55. (SDEs driven by Lévy processes [Ap, Sections 6.2, p363-365, Sections 6.3, p377-378]). Let $X = (X(t)), t \geq 0$ is an $\mathbb{R}^d$-valued Lévy process from Theorem 1.54. Suppose we have an SDE

$$\begin{cases} dY(t) &= L(Y(t))dX(t) \\ &= b(Y(t))dt + \sigma(Y(t))dW(t) \\ &+ \int_{|x|<a} F(Y(t-),x)\tilde{N}(dt,dx) \\ &+ \int_{|x|\geq a} G(Y(t-),x)N(dt,dx) \\ Y(0) &= y_0 \end{cases} \tag{1.74}$$

Which in scalar form is for each $k = 1, ..., d$

$$\begin{cases} dY^k(t) &= L^k(Y(t))dX^k(t) \\ &= b^k(Y(t))dt + \sum_{i=1}^m \sigma_i^k(Y(t))dW^i(t) \\ &+ \int_{|x|<a} F^k(Y(t-),x)\tilde{N}(dt,dx) \\ &+ \int_{|x|\geq a} G^k(Y(t-),x)N(dt,dx) \\ Y^k(0) &= y_0^k \end{cases} \tag{1.75}$$

With standard initial condition

$$\mathbb{P}\left(Y(0) = y_0\right) = 1 \text{ for } \mathcal{F}_0\text{-measurable } \mathbb{R}^d\text{-valued random variable } y_0 \tag{1.76}$$

We assume that x is an $\mathbb{R}^d$-valued vector, a is an $[0,\infty]$-valued constant, $W = (W(t,\omega)), t \geq 0$ is an $\mathbb{R}^m$-valued $(\mathcal{F}_t)$-Brownian motion on $\mathbb{R}^+ \times \Omega$, $N = (N(t,x,\omega)), t \geq 0$ is an $(\mathcal{F}_t)$-Poisson random measure on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^n \setminus \{0\}) \times \Omega$, $\tilde{N} = (\tilde{N}(t,x,\omega))), t \geq 0$ is an $(\mathcal{F}_t)$-Compensated Poisson random measure on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^n \setminus \{0\}) \times \Omega$, ν is a Lévy intensity measure of N, processes $W, N, \tilde{N}$ are mutually independent and independent of $\mathcal{F}_0$. Moreover, for all $i = 1, ..., n$, $k = 1, ..., d$, $0 \leq n \leq d$, functions $b^k : \mathbb{R}^d \to \mathbb{R}$, $\sigma_i^k : \mathbb{R}^d \to \mathbb{R}$, $F^k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $G^k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ are Borel measurable. We call (1.74) an SDEs driven by an $\mathbb{R}^n$-valued Lévy process $X = (X(t)), t \geq 0$. The solution to (1.74) is an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-adapted and càdlàg process $Y = (Y(t)), t \geq 0$ with standard initial condition (1.76) which satisfies the following stochastic integral equation almost surely

$$\begin{aligned} Y(t) &= y_0 + \int_0^t L(Y(s))dX(s) \\ &= y_0 + \int_0^t b(Y(s))ds + \int_0^t \sigma(Y(s))dW(s) \\ &+ \int_0^t \int_{|x|<a} F(Y(s-),x)\tilde{N}(ds,dx) \\ &+ \int_0^t \int_{|x|\geq a} G(Y(s-),x)N(ds,dx) \text{ for all } t \geq 0. \end{aligned} \tag{1.77}$$

Remark 1.56. Solution to (1.90) depends on the initial condition 1.92. The following result provides sufficient conditions for existence and uniqueness of the solution to (1.74) satisfying (1.77).

Theorem 1.57. ***(Existence and uniqueness of Lévy SDE solution [Ap, Section 6.2, p365-366], [Ap, Theorem 6.2.3], [OS, Theorem 1.19]).*** *We impose following assumptions on functions b, σ, F, G from (1.74)*

(1) Linear growth condition: there exists $C_1 > 0$ such that for all $y \in \mathbb{R}^d$

$$\begin{array}{lll} & \sum_{k=1}^d |(\sigma(y)\sigma(y)^T)_k^k| + \sum_{k=1}^d |b^k(y)|^2 & \\ + & \sum_{k=1}^d \int_{|x|<a} |F^k(y,x)|^2 \nu(dx) & \\ + & \sum_{k=1}^d \int_{|x|\geq a} |G^k(y,x)|^2 \nu(dx) & \leq \quad C_1(1+|y|^2) \end{array} \tag{1.78}$$

(2) Lipschitz condition: there exists $C_2 > 0$ such that for all $y_1, y_2 \in \mathbb{R}^d$

$$\begin{array}{lll} & \sum_{k=1}^d \sum_{i=1}^m |\sigma_i^k(y_1) - \sigma_i^k(y_2)|^2 + \sum_{k=1}^d |b^k(y_1) - b^k(y_2)|^2 & \\ + & \sum_{k=1}^d \int_{|x|<a} |F^k(y_1,x) - F^k(y_2,x)|^2 \nu(dx) & \\ + & \sum_{k=1}^d \int_{|x|\geq a} |G^k(y_1,x) - G^k(y_2,x)|^2 \nu(dx) & \leq \quad C_2|y_1 - y_2|^2 \end{array} \tag{1.79}$$

If conditions (1) and (2) are satisfied, then there exists an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-adapted and càdlàg process $Y = (Y(t)), t \geq 0$ with standard initial condition (1.76) which satisfies (1.77) and is a unique solution to (1.74) with following properties

(1) Pathwise uniqueness

(2) Square-integrability, i.e.

$$\mathbb{E}\left(|Y(t)|^2\right) < \infty \textit{ for all } t \geq 0 \tag{1.80}$$

Definition 1.58. (Pathwise uniqueness [IW, Chapter 4: Definition 1.5]). We say that the pathwise uniqueness of solutions holds if whenever $Y_1 = (Y_1(t)), t \geq 0$ and $Y_2 = (Y_2(t)), t \geq 0$ are any two solutions to (1.74) defined on same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with same filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, same $\mathbb{R}^m$-valued $(\mathcal{F}_t)$-Brownian motion $W = (W(t,\omega)), t \geq 0$ on $\mathbb{R}^+ \times \Omega$, same $\mathbb{N}^n$-valued $(\mathcal{F}_t)$-Poisson random measure $N = (N(t,x,\omega)), t \geq 0$ on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^n \setminus \{0\}) \times \Omega$ and same $\mathbb{N}^n$-valued $(\mathcal{F}_t)$-Compensated Poisson random measure $\tilde{N} = (\tilde{N}(t,x,\omega))), t \geq 0$ on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^n \setminus \{0\}) \times \Omega$, such that

$$\mathbb{P}\left(Y_1(0) = Y_2(0)\right) = 1 \tag{1.81}$$

then

$$\mathbb{P}\left(Y_1(t) = Y_2(t)\right) = 1 \text{ for all } t \geq 0 \tag{1.82}$$

Corollary 1.59. ***(Quadratic growth of expectation [Ap, Corollary 6.2.4]).*** *If $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-adapted and càdlàg process $Y = (Y(t)), t \geq 0$ with standard initial condition (1.76) satisfies (1.77) and is* *a solution to (1.74). Then there exists a constant* $C(t)$ *such that*

$$\mathbb{E}\left(|Y(t)|^2\right) \leq C(t)\left(1+\mathbb{E}\left(|y_0|^2\right)\right) \text{ for each } t \geq 0 \tag{1.83}$$

Theorem 1.60. ***(Markov property of Lévy SDE solution [Ap, Theorem 6.4.5], [Ok, Theorem 7.1.2]).*** *If* $\mathbb{R}^d$*-valued* $(\mathcal{F}_t)$*-adapted and càdlàg process* $Y=(Y(t)), t \geq 0$ *with standard initial condition (1.76) satisfies (1.77) and is a solution to (1.74). Then* Y *is a Markov process as defined in Definition 1.8, i.e. for all bounded and Borel measurable functions* $f:\mathbb{R}^d \to \mathbb{R}$ *following holds*

$$\mathbb{E}\left(f\left(Y(t)\right)|\,\mathcal{F}_s\right) = \mathbb{E}\left(f\left(Y(t)\right)|\,Y(s)\right) \text{ for all } 0 \leq s \leq t < \infty \tag{1.84}$$

Theorem 1.61. ***(Multi-dimensional Itô formula [OS, Theorem 1.16], [Bj, Proposition 4.18]).*** *Let* $Y=(Y(t)), t\geq 0$ *be an* $\mathbb{R}^d$*-valued Lévy process with càdlàg modification of the form*

$$\begin{aligned} Y(t) &= Y(0) + \int_0^t b(s,Y(s))ds + \int_0^t \sigma(s,Y(s))dW(s) \\ &+ \int_0^t \int_{|x|<a} F(s,x,Y(s))\tilde{N}(ds,dx) \\ &+ \int_0^t \int_{|x|\geq a} G(s,x,Y(s))N(ds,dx) \end{aligned} \tag{1.85}$$

where for each $k=1,...,d$

$$\begin{aligned} Y^k(t) &= Y^k(0) + \int_0^t b^k(s,Y(s))ds + \sum_{i=1}^m \int_0^t \sigma_i^k(s,Y(s))dW_i(s) \\ &+ \sum_{j=1}^n \int_0^t \int_{|x_j|<a_j} F_j^k(s,x_j,Y(s))\tilde{N}_j(ds,dx_j) \\ &+ \sum_{j=1}^n \int_0^t \int_{|x_j|\geq a_j} G_j^k(s,x_j,Y(s))N_j(ds,dx_j) \end{aligned} \tag{1.86}$$

where $b^k:[0,t]\times\mathbb{R}^d\to\mathbb{R}$, $\sigma_i^k:[0,t]\times\mathbb{R}^d\to\mathbb{R}$, $F_j^k:[0,t]\times\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}$, $G_j^k:[0,t]\times\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}$ *are* $(\mathcal{F}_t)$*-adapted processes such that the integrals exist. Let* $f\in C^2([0,t]\times\mathbb{R}^d)$, *then*

$$
\begin{aligned}
&f(t,Y(t)) = f(0,Y(0))\\
&+\int_0^t \left[\frac{\partial f}{\partial s}(s,Y(s)) + \sum_{k=1}^d b^k(s,Y(s))\frac{\partial f}{\partial Y_k}(s,Y(s)) + \frac{1}{2}\sum_{k,l=1}^d (\sigma(s,Y(s))\rho\sigma(s,Y(s))^T)_l^k \frac{\partial^2 f}{\partial Y_k \partial Y_l}(s,Y(s))\right] ds\\
&\qquad + \int_0^t \left[\sum_{k=1}^d \sum_{i=1}^m \sigma_i^k(s,Y(s))\frac{\partial f}{\partial Y_k}(s,Y(s))\right] dW(s)\\
&+\sum_{j=1}^n \int_0^t \int_{|x_j|<a_j} \left[f(s,Y(s-)+F_j(s,Y(s-),x_j)) - f(s,Y(s-)) - \sum_{k=1}^d F_j^k(s,Y(s-),x_j)\frac{\partial f}{\partial Y_k}(s,Y(s-))\right]\nu_j(dx_j)ds\\
&\qquad + \sum_{j=1}^n \int_0^t \int_{|x_j|<a_j} \left[f(s,Y(s-)+F_j(s,Y(s-),x_j)) - f(s,Y(s-))\right]\tilde{N}_j(ds,dx_j)\\
&\qquad + \sum_{j=1}^n \int_0^t \int_{|x_j|\geq a_j} \left[f(s,Y(s-)+G_j(s,Y(s-),x_j)) - f(s,Y(s-))\right]N_j(ds,dx_j) \qquad (1.87)
\end{aligned}
$$

where ρ is $m \times m$-matrix with values in $[-1,1]$ coming from $\rho_j^i dt = \langle dW_i, dW_j\rangle_t$ for $i,j = 1,...,m$ with $\rho_i^i = 1$ for $i = 1,...,m$.

Lemma 1.62. ***(Integration by parts [OS, Lemma 3.6]).*** *Suppose we have two $\mathbb{R}^d$-valued processes $Y^1 = (Y_t^1)_{t\geq 0}$ and $Y^2 = (Y_t^2)_{t\geq 0}$ which satisfy following equation*

$$
\begin{cases} dY^l(t) &= b^l(t)dt + \sigma^l(t)dW(t) + \int_{\mathbb{R}^d} F^l(t,x)\tilde{N}(t,dx) \text{ for } l = 1,2 \\ Y^l(0) &= y^l \in \mathbb{R}^d \text{ for } l = 1,2 \end{cases} \qquad (1.88)
$$

where $b^l \in \mathbb{R}^d$, $\sigma^l \in \mathbb{R}^{d\times m}$ and $F^l \in \mathbb{R}^{d\times n}$. Moreover assume that $\mathbb{E}\left(Y^l(t)^2\right) < \infty$ for $l = 1,2$. Then

$$
\begin{aligned}
\mathbb{E}\left(Y^1(t)Y^2(t)\right) &= y^1y^2 + \mathbb{E}\Big(\int_0^t Y^1(s)dY^2(s) + \int_0^t Y^2(s)dY^1(s) + \int_0^t \sum_{i=1}^m (\sigma^1(s)\sigma^2(s)^T)_i^i ds\\
&+ \int_0^t \Big(\sum_{j=1}^n \Big(\sum_{k=1}^d \int_{\mathbb{R}} F_{kj}^1(s,x_j)F_{kj}^2(s,x_j)\Big)\nu_j(dx_j)\Big)ds\Big)
\end{aligned} \qquad (1.89)
$$

1.4 Controlled state process

In this and following subsections we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_s)_{s\in[t,T]}$ and $0 \leq t \leq T < \infty$

Definition 1.63. (Control process [OS, Section 3.1, p39]). Suppose we have a set $U \subset \mathbb{R}^d$. We call U a control space and function u a control process, where we assume that

(1) $u : [t, T] \times \Omega \to U$ is a stochastic process

(2) u is $(\mathcal{F}_s)$-adapted and càdlàg

Definition 1.64. (Admissible control process [FS, Definition 2.1]). Suppose that $u = (u(s)), s \in [t, T]$ is a U-valued control process defined on $[t, T] \times \Omega$. We say it is admissible if

(1) u is $(\mathcal{F}_s)$-progressively measurable

(2) $\mathbb{E}\left(\int_t^T |u(s)|^2 ds\right) < \infty$

We denote $\mathcal{U}$ by the set containing every admissible control u.

Proposition 1.65. ***(Bounded control process [FS, Definition 2.1]).*** *Let $u = (u(s)), s \in [t, T]$ is a U-valued control process. If set set U is bounded, then assumption (2) from Definition 1.64 is satisfied.*

Definition 1.66. (Controlled state SDE [Ok, Subsection 11.1, p225-226], [Ph, Section 1.3, p22]). Consider the following modification of (1.74)

$$\begin{cases} dY_u(T) &= L(Y_u(T), u(T))dX(T) \\ &= b(Y_u(T), u(T))dT + \sigma(Y_u(T), u(T))dW(T) \\ &+ \int_{|x|<a} F(Y_u(T-), x, u(T-))\tilde{N}(dT, dx) \\ &+ \int_{|x|\geq a} G(Y_u(T-), x, u(T-))N(dT, dx) \\ Y_u(t) &= y \end{cases} \tag{1.90}$$

Which in scalar form is for each $k = 1, ..., d$

$$\begin{cases} dY_u^k(T) &= L^k(Y_u(T))dX^k(T) \\ &= b^k(Y_u(T), u(T))dT + \sum_{i=1}^m \sigma_i^k(Y_u(T), u(T))dW^i(T) \\ &+ \int_{|x|<a} F^k(Y_u(T-), x, u(T-))\tilde{N}(dT, dx) \\ &+ \int_{|x|\geq a} G^k(Y_u(T-), x, u(T-))N(dT, dx) \\ Y_u^k(t) &= y \end{cases} \tag{1.91}$$

With initial condition

$$\mathbb{P}\left(Y_u(t) = y\right) = 1 \text{ for } \mathcal{F}_t\text{-measurable } \mathbb{R}^d\text{-valued random variable } y \tag{1.92}$$

We assume that control process $u : [t,T]\times\Omega \to U \subset \mathbb{R}^d$ is $(\mathcal{F}_s)$-progressively measurable, x is an $\mathbb{R}^d$-valued vector, a is an $[0,\infty]$-valued constant, $W = (W(s,\omega)), s \in [t,T]$ is an $\mathbb{R}^m$-valued $(\mathcal{F}_s)$-Brownian motion on $\mathbb{R}^+ \times \Omega$, $N = (N(t,x,\omega)), s \in [t,T]$ is an $(\mathcal{F}_s)$-Poisson random measure on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^n \setminus \{0\}) \times \Omega$, $\tilde{N} = (\tilde{N}(t,x,\omega))), s \in [t,T]$ is an $(\mathcal{F}_s)$-Compensated Poisson random measure on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^n \setminus \{0\}) \times \Omega$, ν is a Lévy intensity measure of N, processes $W, N, \tilde{N}$ are mutually independent and independent of $\mathcal{F}_0$. Moreover, for all $i = 1,...,n,\ k = 1,...,d,\ 0 \leq n \leq d$, functions $b^k : \mathbb{R}^d \times U \to \mathbb{R}$, $\sigma_i^k : \mathbb{R}^d \times U \to \mathbb{R}$, $F^k : \mathbb{R}^d \times \mathbb{R}^d \times U \to \mathbb{R}$, $G^k : \mathbb{R}^d \times \mathbb{R}^d \times U \to \mathbb{R}$ are Borel measurable. We call (1.90) a controlled state SDE. The solution to (1.90) is an $\mathbb{R}^d$-valued $(\mathcal{F}_s)$-progressively measurable process $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$ with standard initial condition (1.92) which satisfies the following stochastic integral equation almost surely

$$\begin{array}{rcl} Y_u^{t,y}(T) & = & y + \int_t^T L(Y_u^{t,y}(s))dX(s) \\ & = & y + \int_t^T b(Y_u^{t,y}(s), u(s))ds + \int_t^T \sigma(Y_u^{t,y}(s), u(s))dW(s) \\ & + & \int_t^T \int_{|x|<a} F(Y_u^{t,y}(s-), x, u(s-))\tilde{N}(ds,dx) \\ & + & \int_t^T \int_{|x|\geq a} G(Y_u^{t,y}(s-), x, u(s-))N(ds,dx) \text{ for all } T \geq t \end{array} \tag{1.93}$$

We call $Y^{t,y}$ a controlled state process.

Remark 1.67. Solution to (1.90) depends on both the initial condition (1.92) and the control process u. The following result provides sufficient conditions for existence and uniqueness of the solution to (1.90) satisfying (1.93).

Theorem 1.68. ***(Existence and uniqueness of controlled state SDE solution [FS, Section IV.2, p152-153], [Ph, Section 1.3, p23], [OS, Theorem 1.19], [Ph, Theorem 1.3.15]).*** *We impose following assumptions on functions b, σ, F, G from 1.90*
(1) Linear growth condition: there exists $C_1 > 0$ such that for all $y \in \mathbb{R}^d, \vartheta \in U \subset \mathbb{R}^d$

$$\begin{array}{rll} & \sum_{k=1}^d |(\sigma(y,\vartheta)\sigma(y,\vartheta)^T)_k^k| + \sum_{k=1}^d |b^k(y,\vartheta)|^2 & \\ + & \sum_{k=1}^d \int_{|x|<a} |F^k(y,x,\vartheta)|^2 \nu(dx) & \\ + & \sum_{k=1}^d \int_{|x|\geq a} |G^k(y,x,\vartheta)|^2 \nu(dx) & \leq \quad C_1(1+|y|^2) \end{array} \tag{1.94}$$

(2) Lipschitz condition: there exists $C_2 > 0$ such that for all $y_1, y_2 \in \mathbb{R}^d, \vartheta \in U \subset \mathbb{R}^d$

$$\begin{array}{rll} & \sum_{k=1}^d \sum_{i=1}^m |\sigma_i^k(y_1,\vartheta) - \sigma_i^k(y_2,\vartheta)|^2 + \sum_{k=1}^d |b^k(y_1,\vartheta) - b^k(y_2,\vartheta)|^2 & \\ + & \sum_{k=1}^d \int_{|x|<a} |F^k(y_1,x,\vartheta) - F^k(y_2,x,\vartheta)|^2 \nu(dx) & \\ + & \sum_{k=1}^d \int_{|x|\geq a} |G^k(y_1,x,\vartheta) - G^k(y_2,x,\vartheta)|^2 \nu(dx) & \leq \quad C_2|y_1 - y_2|^2 \end{array} \tag{1.95}$$

If control u is admissible, conditions (1) and (2) are satisfied, then there exists an $\mathbb{R}^d$-valued $(\mathcal{F}_s)$-progressively measurable process $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$ with initial condition (1.92) which satisfies

(1.93) and is a solution to (1.90). Moreover, it satisfies properties in Theorems 1.57 and 1.60.

Definition 1.69. (Conditional expectation notation). Let $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$ be an $\mathbb{R}^d$-valued controlled state process from Definition (1.90). Going forward we will use the following notation

$$\mathbb{E}\left(Y_u^{t,y}(T)\right) = \mathbb{E}_{t,y}\left(Y^u(T)\right) = \mathbb{E}\left(Y^u(T) \middle| Y^u(t) = y\right) \tag{1.96}$$

1.5 Stochastic control

In this and following subsections we assume that $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$ is an $\mathbb{R}^d$-valued controlled state process from Definition 1.66.

Definition 1.70. (Stochastic control problem [Ph, Chapter 3: Section 3.2]). Let $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$ be an $\mathbb{R}^d$-valued controlled state process from Definition (1.90). Let $u = (u(t)), t \geq 0$ be an admissible U-valued control process and $\mathcal{U}$ be a set from 1.64. Let $f : [t,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$ be a Borel measurable functions. We define a finite horizon stochastic control problem by

$$J(t,y,u) = \mathbb{E}\left(\int_t^T f(s, Y_u^{t,y}(s), u(s))ds + g(Y_u^{t,y}(T))\right) \text{ for all } (t,y,u) \in [0,T] \times \mathbb{R}^d \times \mathcal{U} \tag{1.97}$$

The stochastic control problem is to maximise J over admissible control processes $u \in \mathcal{U}$, i.e.

$$v(t,y) = \sup_{u \in \mathcal{U}} J(t,y,u) \geq J(t,y,u) \tag{1.98}$$

We call f a running gain function, g a terminal gain function, $J : [t,T] \times \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}$ a gain function, $v : [t,T] \times \mathbb{R}^d \to \mathbb{R}$ a value function. If $g(y) \equiv 0$ then the problem is said to be in Lagrange form, while if $f(s,y,u) \equiv 0$, the problem is said to in Mayer form.

Definition 1.71. (Optimal control process [Ph, Section 3.2, p38-39]). Let $J : [t,T] \times \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}$ be a gain function from Definition 1.70 and let $\hat{u} = (\hat{u}(t)), t \geq 0$ be an admissible U-valued control process. Suppose that $\mathbb{R}^d$-valued process $Y_{\hat{u}}^{t,y} = (Y_{\hat{u}}^{t,y}(s)), s \in [t,T]$ is the unique solution to (1.90) such that

$$v(t,y) = J(t,y,\hat{u}) = \mathbb{E}\left(\int_t^T f(s, Y_{\hat{u}}^{t,y}(s), \hat{u}(s))ds + g(Y_{\hat{u}}^{t,y}(T))\right) \tag{1.99}$$

In this case we call $Y_{\hat{u}}^{t,y}$ optimally controlled state process and say that $\hat{u}$ is an optimal control process given initial condition $(t,y) \in [0,T) \times \mathbb{R}^d$.

Definition 1.72. (Quadratic growth [FS, Section IV.2, p154]). Let $f : [t,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$ be a Borel measurable functions. We impose following assumptions:

(1) There exists $C_1 > 0$ such that

$$|f(t,y,\vartheta)| \leq C_1(1 + |y|^2 + |\vartheta|^2) \text{ for all } (t,y,\vartheta) \in [0,T] \times \mathbb{R}^d \times U \tag{1.100}$$

(2) There exists $C_2 > 0$ such that

$$|g(y)| \leq C_2(1 + |y|^2) \text{ for all } y \in \mathbb{R}^d \tag{1.101}$$

Proposition 1.73. ***(Integrability of gain function).*** *Let* $\mathcal{U}$ be a set from *Definition 1.64. If* f, g *satisfy assumptions from Definition 1.72 and* $u = (u(s)), s \in [t,T]$ *is an admissible* U*-valued control process, then* *the gain function* $J : [t,T] \times \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}$ *from Definition 1.70, is well defined, i.e.*

$$\mathbb{E}\left(\int_t^T f(s, Y_u^{t,y}(s), u(s))ds + g(Y_u^{t,y}(T))\right) < \infty \text{ for all } (t,y,u) \in [0,T] \times \mathbb{R}^d \times \mathcal{U} \tag{1.102}$$

Proof. Let us choose and fix $(t,y,u) \in [0,T] \times \mathbb{R}^d \times \mathcal{U}$ Firstly, from Assumption (2) in Definition 1.72 we have that

$$\left|g(Y_u^{t,y}(T))\right| \le C_2(1 + |Y_u^{t,y}(T)|^2) \tag{1.103}$$

Hence, by monotinicity of expectation

$$\mathbb{E}\left(\left|g(Y_u^{t,y}(T))\right|\right) \le \mathbb{E}\left(C_2\left(1 + \left|Y_u^{t,y}(T)\right|^2\right)\right) = C_2\left(1 + \mathbb{E}\left(\left|Y_u^{t,y}(T)\right|^2\right)\right) \tag{1.104}$$

Since $Y^{t,y}$ is a unique solution to (1.90), by Theorem 1.68 we have that

$$\mathbb{E}\left(\left|Y_u^{t,y}(T)\right|^2\right) < \infty \tag{1.105}$$

Hence

$$\mathbb{E}\left(\left|g(Y_u^{t,y}(T))\right|\right) \le C_2\left(1 + \mathbb{E}\left(\left|Y_u^{t,y}(T)\right|^2\right)\right) < \infty \tag{1.106}$$

Secondly, by Fubini Theorem 5.19

$$\mathbb{E}\left(\int_t^T \left|f(s, Y_u^{t,y}(s), u(s))\right| ds\right) = \int_t^T \mathbb{E}\left(\left|f(s, Y_u^{t,y}(s), u(s))\right|\right) ds \tag{1.107}$$

by linearity of integral and expectation

$$\int_t^T \mathbb{E}\left(C_1\left(1 + \left|Y_u^{t,y}(s)\right|^2 + |u(s)|^2\right)\right) ds = C_1\left(1 + \int_t^T \mathbb{E}\left(\left|Y_u^{t,y}(s)\right|^2\right) ds + \int_t^T \mathbb{E}\left(|u(s)|^2\right) ds\right) \tag{1.108}$$

from Assumption (2) in Definition 1.72 we have that

$$\left|f(s, Y_u^{t,y}(s), u(s))\right| \le C_1(1 + |Y_u^{t,y}(s)|^2 + |u(s)|^2) \text{ for all } s \in [0,T] \tag{1.109}$$

Hence, by monotinicity of expectation

$$\int_t^T \mathbb{E}\left(\left|f(s, Y_u^{t,y}(s), u(s))\right|\right) ds \le C_1\left(1 + \int_t^T \mathbb{E}\left(\left|Y_u^{t,y}(s)\right|^2\right) ds + \int_t^T \mathbb{E}\left(|u(s)|^2\right) ds\right) \tag{1.110}$$

Since controls are admissible, from assumption (2) of Definition 1.64, we have that

$$\mathbb{E}\left(\int_t^T |u(s)|^2 ds\right) < \infty \tag{1.111}$$

Hence, by Fubini Theorem 5.19

$$\mathbb{E}\left(\int_t^T |u(s)|^2 ds\right) = \int_t^T \mathbb{E}\left(|u(s)|^2\right) ds < \infty \tag{1.112}$$

Combining (1.105) and (1.112) it follows that

$$\int_t^T \mathbb{E}\left(\left|f(s, Y_u^{t,y}(s), u(s))\right|\right) ds \le C_1\left(1 + \int_t^T \mathbb{E}\left(\left|Y_u^{t,y}(s)\right|^2\right) ds + \int_t^T \mathbb{E}\left(|u(s)|^2\right) ds\right) < \infty \tag{1.113}$$

Therefore, by linearity of expectation

$$\mathbb{E}\left(\int_t^T f(s, Y_u^{t,y}(s), u(s))ds\right) + \mathbb{E}\left(g(Y_u^{t,y}(T))\right) = \mathbb{E}\left(\int_t^T f(s, Y_u^{t,y}(s), u(s))ds + g(Y_u^{t,y}(T))\right) < \infty \tag{1.114}$$

□

Definition 1.74. (Integrable running gain function [Ph, Section 3.2, p38]). Let $f : [t,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ be a running gain function from Definition 1.70. Let $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$ be an $\mathbb{R}^d$-valued controlled state process from Definition 1.66, let $\mathcal{U}$ be a set from Definition 1.64. Let $u = (u(s)), s \in [t,T]$ be an admissible U-valued control process. For $(t,y) \in [0,T] \times \mathbb{R}^d$, we denote $\mathcal{U}(t,y)$ by the set of all $u \in \mathcal{U}$ such that

$$\mathbb{E}\left(\int_t^T \left|f(s, Y_u^{t,y}(s), u(s))\right| ds\right) < \infty \tag{1.115}$$

where we assume that $\mathcal{U}(t,y)$ is non-empty for all $(t,y) \in [0,T] \times \mathbb{R}^d$.

Proposition 1.75. ***(Integrability of running gain function [Ph, Remark 3.2.1]).*** *Let $f : [t,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ be a running gain function from Definition 1.70. Let $\mathcal{U}$ be a set from Definition 1.64 and let $\mathcal{U}(t,y)$ be a set from Definition 1.74. If f satisfies assumption (1) from Definition 1.72, then for any constant control ϑ*

$$\mathbb{E}\left(\int_t^T \left|f(s, Y_\vartheta^{t,y}(s), \vartheta)\right| ds\right) < \infty \text{ for all } (t,y,\vartheta) \in [0,T] \times \mathbb{R}^d \times U \tag{1.116}$$

Moreover, for any admissible U-valued control process $u = (u(s)), s \in [t,T]$

$$\mathbb{E}\left(\int_t^T \left|f(s, Y_u^{t,y}(s), u(s))\right| ds\right) < \infty \text{ for all } (t,y,u) \in [0,T] \times \mathbb{R}^d \times \mathcal{U} \tag{1.117}$$

Proof. Proof of (1.117) and (1.116) follows from the second part of the proof of Proposition 1.73. □

Definition 1.76. (Deterministic control process [Ok, Section 11.1, p227]). A deterministic (or an open loop) control process U-valued control is a function

$$u : [0,T] \to U \tag{1.118}$$

Definition 1.77. (Feedback control process [Ok, Section 11.1, p227]). A feedback control process (closed loop controls) is

$$u(s,\omega) \text{ is } \sigma\left(Y_u^{t,y}(r,\omega) : r \le s\right)\text{-measurable for all } s \in [t,T] \tag{1.119}$$

Definition 1.78. (Markov control process [Ok, Section 11.1, p227]). A Markov control process is

$$u(s,\omega) = \underline{u}(s, Y_{\underline{u}}^{t,y}(s,\omega)) \text{ for all } s \in [t,T] \tag{1.120}$$

for some Borel measurable function $\underline{u} : [t,T] \times \mathbb{R}^d \to U$.

Definition 1.79. (Regular boundary points [Ok, Definition 9.2.8]). Let $\mathcal{S} = [0,T) \times \mathbb{R}^d$, $\bar{\mathcal{S}} = [0,T] \times \mathbb{R}^d$ and $\partial\mathcal{S} = \{T\} \times \mathbb{R}^d$. Suppose we have an $\mathbb{R}^d$-valued $(\mathcal{F}_s)$-adapted process $Y^{t,y} = (Y^{t,y}(s,\omega)), s \in [t,T]$ with initial condition $\mathbb{P}(Y(t) = y) = 1$ for $y \in \mathbb{R}^d$. Let

$$\tau_{\mathcal{S}}^{t,y}(\omega) = \inf\left\{s > t : Y^{t,y}(s,\omega) \notin \mathcal{S}\right\} \tag{1.121}$$

(1) A point $(t,y) \in \partial\mathcal{S}$ is called regular with respect to $Y^{t,y}$ if

$$\mathbb{P}\left(\tau_{\mathcal{S}}^{t,y}(\omega) = t\right) = 1 \text{ for all } (t,y) \in \partial\mathcal{S} \tag{1.122}$$

(2) A point $(t,y) \in \partial\mathcal{S}$ is called irregular with respect to $Y^{t,y}$ if

$$\mathbb{P}\left(\tau_{\mathcal{S}}^{t,y}(\omega) = t\right) = 0 \text{ for all } (t,y) \in \partial\mathcal{S} \tag{1.123}$$

Theorem 1.80. ***(Sufficiency of Markov controls [Ok, Theorem 11.2.3]).*** *Let $\mathcal{T}_{t,T}$ be a collection of $[t,T]$-valued $(\mathcal{F}_s)$-stopping times, let $\underline{u} = (\underline{u}(s, Y_{\underline{u}}^{t,y}(s,\omega))), s \in [t,T]$ be a Markov control process as in Definition 1.78 and let $u = (u(s,\omega)), s \in [t,T]$ be an $(\mathcal{F}_s)$-adapted control process as in Definition 1.63. Suppose we have two value functions denoted by v_M, v_A and defined as*

$$v_M(t,y) = \sup_{\underline{u}\in\mathcal{U}} J(t,y,\underline{u}), \;\; v_A(t,y) = sup_{u\in\mathcal{U}} J(t,y,u) \tag{1.124}$$

Assume that there exists an optimal Markov control process $\hat{\underline{u}} = (\hat{\underline{u}}(s, Y_{\hat{\underline{u}}}^{t,y}(s,\omega)), s \in [t,T]$ satisfying

$$v_M(t,y) = J(t,y,\hat{\underline{u}}) \text{ for all } (t,y) \in \mathcal{S} \tag{1.125}$$

such that all the boundary points $(t,y) \in \partial\mathcal{S}$ are regular with respect to $Y_{\hat{\underline{u}}}^{t,y}$ and value function $v_M \in C^2(\mathcal{S}) \cap C^0(\bar{\mathcal{S}})$ satisfies

$$\mathbb{E}_{t,y}\left(|v_M(\theta, Y^u(\theta))| + \int_t^\theta |\mathcal{J}(v_M,u)|\, ds\right) < \infty \text{ for all } (t,y) \in \mathcal{S}, \;\; \theta \in \mathcal{T}_{t,T}, \;\; u \in (\mathcal{F}_s)\text{-adapted controls} \tag{1.126}$$

With $\mathcal{J}$ from (1.152). Then

$$v_M(t,y) = v_A(t,y) \text{ for all } (t,y) \in \mathcal{S} \tag{1.127}$$

Remark 1.81. Under assumptions from Theorem 1.80 control process $\underline{u}$ of Markov type from Definition 1.78 coincides with an arbitrary $(\mathcal{F}_s)$-adapted control process u. Interpretation of this result is that one can obtain as good performance with Markov controls as with any $(\mathcal{F}_s)$-adapted controls.

1.6 Dynamic programming principle

Theorem 1.82. ***(Dynamic programming equation [Ph, Theorem 3.3.1]).*** *Let $\mathcal{T}_{t,T}$ be a collection of $[t,T]$-valued $(\mathcal{F}_s)$-stopping times, let $\mathcal{U}(t,y)$ be a set from Definition 1.74. Assume that f satisfies assumption (1) from Definition 1.72. Then the following equation holds for all $(t,y) \in [0,T] \times \mathbb{R}^d$*

$$\begin{aligned} v(t,y) &= \sup_{u \in \mathcal{U}(t,y)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + v(\theta, Y_u^{t,y}(\theta))\right) \\ &= \sup_{u \in \mathcal{U}(t,y)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + v(\theta, Y_u^{t,y}(\theta))\right) \end{aligned} \tag{1.128}$$

Proof. We prove this theorem by using same methods as in proof of [Ph, Theorem 3.3.1]. Let $\theta \in \mathcal{T}_{t,T}$ and suppose we have two $\mathbb{R}^d$-valued solutions to (1.90), which satisfy Theorem 1.68. We will denote these by $Y_u^{t,y} = (Y_u^{t,y}(s)), s \in [t,T]$, $Y_u^{\theta, Y_u^{t,y}(\theta)} = (Y_u^{\theta, Y_u^{t,y}(\theta)}(s)), s \in [\theta, T]$ and define them as

$$Y_u^{t,y}(s) = y + \int_t^s L(Y_u^{t,y}(r), u(r))dX(r) \text{ for } s \in [t,T] \tag{1.129}$$

and

$$Y_u^{\theta, Y_u^{t,y}(\theta)}(s) = Y_u^{t,y}(\theta) + \int_\theta^s L(Y_u^{\theta, Y_u^{t,y}(\theta)}(r), u(r))dX(r) \text{ for } s \geq \theta \in [t,T] \tag{1.130}$$

Then by Theorem 1.68 these solutions will satisfy Definition 1.58

$$\begin{aligned} Y_u^{\theta, Y_u^{t,y}(\theta)}(s) &= Y_u^{t,y}(s) \\ &= y + \int_t^s L(Y_u^{t,y}(r), u(r))dX(r) \\ &= y + \int_t^\theta L(Y_u^{t,y}(r), u(r))dX(r) + \int_\theta^s L(Y_u^{t,y}(r), u(r))dX(r) \text{ for } s \geq \theta \in [t,T] \end{aligned} \tag{1.131}$$

Firstly, we start with the gain function from Definition 1.70

$$J(t,y,u) = \mathbb{E}\left(\int_t^T f(s, Y_u^{t,y}(s), u(s))ds + g(Y_u^{t,y}(T))\right) \tag{1.132}$$

by Definition 1.69 we can write it as

$$J(t,y,u) = \mathbb{E}_{t,y}\left(\int_t^T f(s, Y^u(s), u(s))ds + g(Y^u(T))\right) \tag{1.133}$$

by tower rule we get

$$J(t,y,u) = \mathbb{E}_{t,y}\left(\mathbb{E}_{\theta, Y^{t,y}(\theta)}\left(\int_t^T f(s, Y^u(s), u(s))ds + g(Y^u(T))\right)\right) \tag{1.134}$$

using (1.131), we split the integral into $\mathcal{F}_\theta$-measurable and $\mathcal{F}_\theta$-independent parts

$$J(t,y,u) = \mathbb{E}_{t,y}\left(\mathbb{E}_{\theta, Y^{t,y}(\theta)}\left(\int_t^\theta f(s, Y^u(s), u(s))ds + \int_\theta^T f(s, Y^u(s), u(s))ds + g(Y^u(T))\right)\right) \tag{1.135}$$

taking out what is known

$$J(t,y,u) = \mathbb{E}_{t,y}\left(\int_t^\theta f(s, Y^u(s), u(s))ds + \mathbb{E}_{\theta, Y^{t,y}(\theta)}\left(\int_\theta^T f(s, Y^u(s), u(s))ds + g(Y^u(T))\right)\right) \tag{1.136}$$

by Definition 1.69 we rewrite expectations as

$$J(t,y,u) = \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + \mathbb{E}\left(\int_\theta^T f(s, Y_u^{\theta, Y_u^{t,y}(\theta)}(s), u(s))ds + g(Y_u^{\theta, Y_u^{t,y}(\theta)}(T))\right)\right) \quad (1.137)$$

by Definition 1.70 of the gain function

$$J(t,y,u) = \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + J(\theta, Y_u^{t,y}(\theta), u)\right) \quad (1.138)$$

by Proposition 1.75, $u \in \mathcal{U}(t,y)$. Next, by Definition 1.70 of a value function we have the relation

$$J(\theta, Y_u^{t,y}(\theta), u) \leq \sup_{u \in \mathcal{U}(t,y)} J(\theta, Y_u^{t,y}(\theta), u) = v(\theta, Y_u^{t,y}(\theta)) \quad (1.139)$$

Using this relation we obtain

$$\begin{aligned} J(t,y,u) &\leq \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + v(\theta, Y_u^{t,y}(\theta))\right) \\ &\leq \sup_{u \in \mathcal{U}(t,y)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + v(\theta, Y_u^{t,y}(\theta))\right) \end{aligned} \quad (1.140)$$

Maximising u over $\mathcal{U}(t,y)$ yields

$$v(t,y) \leq \sup_{u \in \mathcal{U}(t,y)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + v(\theta, Y_u^{t,y}(\theta))\right) \quad (1.141)$$

Secondly, let $\epsilon > 0$, then by Definition 1.70 of the value function there exists a control process $u^\epsilon = (u^\epsilon(s)), \theta \leq s \leq T$ such that $u^\epsilon \in \mathcal{U}(\theta, Y^{t,y}(\theta))$

$$J(\theta, Y_{u^\epsilon}^{t,y}(\theta), u^\epsilon) \geq v(\theta, Y_{u^\epsilon}^{t,y}(\theta)) - \epsilon = \sup_{u^\epsilon \in \mathcal{U}(\theta, Y_{u^\epsilon}^{t,y}(\theta))} J(\theta, Y_{u^\epsilon}^{t,y}(\theta), u^\epsilon) - \epsilon \quad (1.142)$$

Let $\tilde{u} = (\tilde{u}(s)), s \in [t,T]$ be a control process defined as

$$\tilde{u}(s) = \begin{cases} u(s) & \text{if} \quad s \in [t, \theta] \\ u^\epsilon(s) & \text{if} \quad s \in [\theta, T] \end{cases} \quad (1.143)$$

It can be shown that process $\tilde{u}$ is $(\mathcal{F}_s)$-progressively measurable and so is $\tilde{u} \in \mathcal{U}(t,y)$. By same arguments as before, we get

$$J(t,y,\tilde{u}) = \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + J(\theta, Y_{u^\epsilon}^{t,y}(\theta), u^\epsilon)\right) \quad (1.144)$$

Using (1.142) we obtain

$$\begin{aligned} v(t,y) &\geq J(t,y,\tilde{u}) \\ &\geq \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s, Y_u^{t,y}(s), u(s))ds + v(\theta, Y_{u^\epsilon}^{t,y}(\theta))\right) - \epsilon \end{aligned} \quad (1.145)$$

Since $u(\theta) = u^\epsilon(\theta)$ from (1.143) we have that

$$v(\theta, Y_u^{t,y}(\theta)) = v(\theta, Y_{u^\epsilon}^{t,y}(\theta)) - \epsilon \quad (1.146)$$

Hence

$$\sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s,Y_u^{t,y}(s),u(s))ds + v(\theta,Y_{u^\epsilon}^{t,y}(\theta))\right) - \epsilon = \sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s,Y_u^{t,y}(s),u(s))ds + v(\theta,Y_u^{t,y}(\theta))\right) \tag{1.147}$$

Maximising u over $\mathcal{U}(t,y)$ we get

$$v(t,y) \geq \sup_{u\in\mathcal{U}(t,y)} \sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}\left(\int_t^\theta f(s,Y_u^{t,y}(s),u(s))ds + v(\theta,Y_u^{t,y}(\theta))\right) \tag{1.148}$$

Combining (1.141) with 1.148 we obtain (1.128) as required. □

Remark 1.83. **(Equivalent dynamic programming equation [Ph, Remark 3.3.3]).** Let all of the assumptions from Theorem 1.82 hold. There exists an equivalent equation to (1.128), where for all $(t,y)\in[0,T]\times\mathbb{R}^d$, $u\in\mathcal{U}(t,y)$, $\theta\in\mathcal{T}_{t,T}$

$$v(t,y) \geq \mathbb{E}\left(\int_t^\theta f(s,Y_u^{t,y}(s),u(s))ds + v(\theta,Y_u^{t,y}(\theta))\right) \tag{1.149}$$

Proposition 1.84. ***(Hamilton-Jacobi-Bellman equation [Ph, Subsection 3.4.1, p43]).*** *Let* $\mathcal{S} = [0,T)\times\mathbb{R}^d$*,* $\bar{\mathcal{S}} = [0,T]\times\mathbb{R}^d$ and $\partial\mathcal{S} = \{T\}\times\mathbb{R}^d$*. Let* $U\subset\mathbb{R}^d$ *and suppose we have a constant control* $\vartheta\in U$*. If* $v\in C^2(\bar{\mathcal{S}})$ *then we can obtain following PDE*

$$\begin{cases} \frac{\partial v}{\partial t}(t,y) + \mathcal{H} = 0 \text{ for all } (t,y)\in\mathcal{S} \\ v(t,y) = g(y) \text{ for all } (t,y)\in\partial\mathcal{S} \end{cases} \tag{1.150}$$

where $v(T,\cdot) = g$ *is regular terminal condition. We call it a Hamilton-Jacobi-Bellman (HJB) equation, where* $\mathcal{H}$ *is called Hamiltonian and is defined by*

$$\mathcal{H} = \sup_{\vartheta\in U} \mathcal{J}(v,\vartheta) \tag{1.151}$$

With

$$\begin{aligned} \mathcal{J}(v,\vartheta) &= \textstyle\sum_{k=1}^d b^k(y,\vartheta)\frac{\partial^2 v}{\partial y_k}(t,y) + \frac{1}{2}\sum_{k,l=1}^d(\sigma(y,\vartheta)\sigma(y,\vartheta)^T)_l^k\frac{\partial^2 v}{\partial y_k\partial y_l}(t,y) \\ &+ \textstyle\int_{|x|<a}\left[v(t,y+F(y,x,\vartheta)) - v(t,y) - \sum_{k=1}^d F^k(y,x,\vartheta)\frac{\partial v}{\partial y_k}(t,y)\right]\nu(dx) \\ &+ \textstyle\int_{|x|\geq a}\left[v(t,y+G(y,x,\vartheta)) - v(t,y)\right]\nu(dx) \end{aligned} \tag{1.152}$$

Remark 1.85. **([Ph, Chapter 3: Section 3.4]).** Interpretation of the HJB equation is as follows, as $\theta\to t$ in Theorem 1.82, the dynamics of the value function v approach that of (1.150). In other words, (1.150) is the infinitesimal version of (1.128).

Theorem 1.86. ***(Verification [Ph, Theorem 3.5.2]).*** *Let* $\mathcal{S} = [0,T)\times\mathbb{R}^d$*,* $\bar{\mathcal{S}} = [0,T]\times\mathbb{R}^d$ and $\partial\mathcal{S} = \{T\}\times\mathbb{R}^d$*. Let* $U\subset\mathbb{R}^d$ *and suppose we have a constant control* $\vartheta\in U$*. Let* $w:\bar{\mathcal{S}}\to\mathbb{R}$ *be a function such that* $w\in C^2(\mathcal{S})\cap C^0(\bar{\mathcal{S}})$ *and* w *satisfies Quadratic growth assumption (2) from Definition 1.72. Using definitions from the Proposition 1.84:*

(1) Suppose that

$$\begin{aligned} \frac{\partial w}{\partial t}(t,y) + \sup_{\vartheta \in U} \mathcal{J}(w,\vartheta) &\leq 0 \quad \textit{for } (t,y) \in \mathcal{S} \\ w(t,y) &\geq g(y) \quad \textit{for } (t,y) \in \partial\mathcal{S} \end{aligned} \tag{1.153}$$

Then

$$w(t,y) \geq v(t,y) \textit{ for all } (t,y) \in \bar{\mathcal{S}} \tag{1.154}$$

(2) Suppose further that $w(T,\cdot) = g$, and there exists a Borel measurable function $\underline{\hat{u}} : \mathcal{S} \to U$ such that

$$\frac{\partial v}{\partial t}(t,y) + \mathcal{J}(v,\underline{\hat{u}}) = \frac{\partial v}{\partial t}(t,y) + \sup_{\vartheta \in U} \mathcal{J}(v,\vartheta) = 0 \tag{1.155}$$

And the modification of SDE (1.90) of the form

$$\begin{cases} dY_{\underline{\hat{u}}}(T) &= b(Y_{\underline{\hat{u}}}(T), \underline{\hat{u}}(T, Y_{\underline{\hat{u}}}(T)))dT \\ &+ \sigma(Y_{\underline{\hat{u}}}(T), \underline{\hat{u}}(T, Y_{\underline{\hat{u}}}(T)))dW(T) \\ &+ \int_{|x|<a} F(Y_{\underline{\hat{u}}}(T-), x, \underline{\hat{u}}(T-, Y_{\underline{\hat{u}}}(T-)))\tilde{N}(dT,dx) \\ &+ \int_{|x|\geq a} G(Y_{\underline{\hat{u}}}(T-), x, \underline{\hat{u}}(T-, Y_{\underline{\hat{u}}}(T-)))N(dT,dx) \\ Y_{\underline{\hat{u}}}(t) &= y \end{cases} \tag{1.156}$$

with initial condition (1.92) has the unique solution (1.93), denoted by $Y^{t,y}_{\underline{\hat{u}}} = (Y^{t,y}_{\underline{\hat{u}}}(s)), s \in [t,T]$, where the process $\underline{\hat{u}} = (\underline{\hat{u}}(s, Y^{t,y}_{\underline{\hat{u}}}(s))), s \in [t,T]$ is such that $\underline{\hat{u}} \in \mathcal{U}(t,y)$. Then

$$v(t,y) = w(t,y) \textit{ for } (t,y) \in \bar{\mathcal{S}} \tag{1.157}$$

Moreover, $\underline{\hat{u}}$ is an optimal Markov control process.

Remark 1.87. Verification Theorem 1.86 gives necessary conditions for constant control ϑ to coincide with optimal Markov control process $\underline{\hat{u}}$, which is sufficient from Theorem 1.80 and as discussed in Remark 1.81.

1.7 Partial differential equations

Definition 1.88. (Classes of PDEs [Ol, Section 4.4, p171-172, Definition 4.12]). Let $u : [0,T] \times \mathbb{R} \to \mathbb{R}$ and suppose we have a linear second-order PDE

$$A\frac{\partial^2 u(t,x)}{\partial t^2} + B\frac{\partial^2 u(t,x)}{\partial t \partial x} + C\frac{\partial^2 u(t,x)}{\partial x^2} + D\frac{\partial u(t,x)}{\partial t} + E\frac{\partial u(t,x)}{\partial x} + Fu(t,x) = G \tag{1.158}$$

where A, B, C, D, E, F are some constants or functions $[0,T] \times \mathbb{R} \to \mathbb{R}$. If $G \equiv 0$, then (1.158) is called homogeneous, otherwise (1.158) is called inhomogeneous. Assuming that at least one of $A, B, C \neq 0$ we can classify (1.158) on the basis of the value of its discriminant, i.e.

$$\Delta = B^2 - 4AC \tag{1.159}$$

where (1.158) at point $(t,x) \in [0,T] \times \mathbb{R}$ is called

(1) Hyperbolic if $\Delta(t,x) > 0$

(2) Parabolic if $\Delta(t,x) = 0,\ A^2 + B^2 + C^2 \neq 0$

(3) Elliptic if $\Delta(t,x) < 0$

Definition 1.89. (Cauchy problem [Fr, Chapter 1, Section 7, p25]). Let $\mathcal{S} = [0,T) \times \mathbb{R}^d$, $\bar{\mathcal{S}} = [0,T] \times \mathbb{R}^d$ and $\partial\mathcal{S} = \{T\} \times \mathbb{R}^d$. Let $u : \bar{\mathcal{S}} \to \mathbb{R}$, $f : \bar{\mathcal{S}} \to \mathbb{R}$ and $\phi : \mathbb{R}^d \to \mathbb{R}$ be continuous functions. Suppose we have a parabolic PDE with terminal condition

$$\begin{cases} \frac{\partial u(t,x)}{\partial t} + \sum_{i,j=1}^d a^i_j(t,x)\frac{\partial^2 u(t,x)}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(t,x)\frac{\partial u(t,x)}{\partial x_i} - c(t,x)u(t,x) + f(t,x) = 0 \text{ for all } (t,x) \in \mathcal{S} \\ u(t,x) = \phi(x) \text{ for all } (t,x) \in \partial\mathcal{S} \end{cases} \tag{1.160}$$

where $a^i_j : \bar{\mathcal{S}} \to \mathbb{R}$, $b_i : \bar{\mathcal{S}} \to \mathbb{R}$ and $c : \bar{\mathcal{S}} \to \mathbb{R}$. Moreover, f, ϕ satisfy boundedness conditions, i.e. there exist a constant C such that

$$\begin{aligned} |f(t,x)| &\leq Ce^{h|x|^2} \text{ for all } (t,x) \in \mathcal{S} \\ |\phi(x)| &\leq Ce^{h|x|^2} \text{ for all } x \in \mathbb{R}^d \end{aligned} \tag{1.161}$$

with $h \in (0, \frac{\lambda_0}{4T})$ and $\lambda_0 > 0$. We define a solution to (1.160) by

$$u(t,x) = \int_{\mathbb{R}^d} \Gamma(t,x;\xi,0)\phi(\xi)d\xi - \int_0^t \int_{\mathbb{R}^d} \Gamma(t,x;\xi,\tau)f(\xi,\tau)d\xi d\tau \tag{1.162}$$

where Γ is called a fundamental solution to 1.160 without terminal condition and is defined by

$$\Gamma(t,x;\xi,\tau) = \frac{1}{(2\sqrt{\pi})^d}(t-\tau)^{-d/2}\exp\left(-\frac{\sum_{i=1}^d (x_i - \xi)^2}{4(t-\tau)}\right) \text{ for } t > \tau \tag{1.163}$$

for every fixed (ξ, τ).

Theorem 1.90. ***(Existence and uniqueness of solution to Cauchy problem [Fr, Chapter 1, Theorems 12, 16]).*** *We impose following assumptions on (1.160):*

(1) Uniform parabolicity: there exist $\lambda_0, \lambda_1 > 0$ *such that for any* $\mathbb{R}^d$*-valued vector* ξ

$$\lambda_0|\xi|^2 \leq \sum_{i,j=1}^{d} a^i_j(t,x)\xi_i\xi_j \leq \lambda_1|\xi|^2 \text{ for all } (t,x) \in \bar{\mathcal{S}} \tag{1.164}$$

(2) Hölder continuity: there exists a constant C, $\alpha \in (0,1)$ *such that for all* $(x,t) \in \bar{\mathcal{S}}$, $(x_0,t_0) \in \bar{\mathcal{S}}$

$$\begin{aligned} |a^i_j(x,t) - a^i_j(x_0,t_0)| &\leq C(|x-x_0|^\alpha + |t-t_0|^{\alpha/2}) \\ |b_i(x,t) - b_i(x_0,t)| &\leq C|x-x_0|^\alpha \\ |c(x,t) - c(x_0,t)| &\leq C|x-x_0|^\alpha \end{aligned} \tag{1.165}$$

(3) Functions a^i_j, $\frac{\partial a^i_j}{\partial x_i}$, $\frac{\partial^2 a^i_j}{\partial x_i \partial x_j}$, b_i, $\frac{\partial b_i}{\partial x_i}$, c *are bounded continious functions on* $\bar{\mathcal{S}}$*, satisfying uniform Hölder condition: there exists a constant* C, $\alpha \in (0,1)$ *such that for all* $x, y \in \mathbb{R}^d$, $t \in [0,T]$

$$|f(x,t) - f(y,t)| \leq C|x-y|^\alpha \tag{1.166}$$

If assumptions are satisfied, then (1.162) is the solution to (1.160) and there exists a constant C *such that*

$$|u(t,x)| \leq C|x|^2 \text{ for all } (t,x) \in \bar{\mathcal{S}} \tag{1.167}$$

Moreover, solution (1.162) is unique, in the sense that there is at most one solution to (1.160) and there exists a constant $C > 0$ *such that*

$$\int_0^T \int_{\mathbb{R}^d} |u(t,x)| e^{-C|x|^2} dx dt < \infty \tag{1.168}$$

Theorem 1.91. ***(Feynman-Kac representation [Ph, Theorem 1.3.17]).*** *Suppose* u *is the solution to (1.160) such that* $u \in C^2(\mathcal{S}) \cup C^0(\bar{\mathcal{S}})$ *and* $\frac{\partial u(t,x)}{\partial x_1}, ..., \frac{\partial u(t,x)}{\partial x_d}$ *are bounded. Then* u *admits the representation*

$$u(t,x) = \mathbb{E}_{t,x}\left(\int_t^T \exp\left(-\int_t^s c(u,X_u)du\right) f(s,X_s)ds + \exp\left(-\int_t^T c(u,X_u)du\right)\phi(X_T)\right) \text{ for all } (t,x) \in \bar{\mathcal{S}} \tag{1.169}$$

where functions $f : \bar{\mathcal{S}} \to \mathbb{R}$, $\phi : \mathbb{R}^d \to \mathbb{R}$ *and* $c : \bar{\mathcal{S}} \to \mathbb{R}$ *are Borel measurable. Moreover,* $X = (X_s)_{s \in [t,T]}$ *is a* $\mathbb{R}^d$*-valued process, called Ito process and defined as*

$$X_s = x + \int_t^s \sum_{i=1}^{d} b_i(u,X_u)du + \int_t^s \sum_{i,j=1}^{d} \sigma^i_j(u,X_u)dW^j_u \text{ for all } s \in [t,T] \tag{1.170}$$

where x *is* $\mathcal{F}_t$*-measurable* $\mathbb{R}^d$*-valued random variable such that* $\mathbb{P}(X_t = x) = 1$, $W = (W_s)_{s \in [t,T]}$ *is an* $\mathbb{R}^d$*-valued* $(\mathcal{F}_s)$*-Brownian motion,* $b_i : \bar{\mathcal{S}} \to \mathbb{R}$, $\sigma^i_j : \bar{\mathcal{S}} \to \mathbb{R}$ *are Borel measurable functions and* $a^i_j(t,x) = (\sigma(t,x)\sigma(t,x)^T)^i_j$*. Conversly, if* u *admits (1.169), then it is a solution to (1.160).*

1.8 Asymptotic expansion and Fourier transform

In this subsection we assume $0 < \epsilon \ll 1$, which we call a small parameter.

Definition 1.92. (Order notation [Mu, Section 1.1, p2-3]). Let D be any set and let $f, g : D \to \mathbb{C}$ such that $\lim_{x \to x_0} f(x) = f(x_0)$, $\lim_{x \to x_0} g(x) = g(x_0)$ for $x, x_0 \in D$ exists. Then we say $f(x) = O(g(x))$ as $x \to x_0$ if there exists $C, \delta > 0$ such that

$$|f(x)| \leq C\,|g(x)| \text{ whenever } 0 < |x - x_0| < \delta \tag{1.171}$$

And $f(x) = o(g(x))$ as $x \to x_0$ if for every $C > 0$ there exists $\delta > 0$ such that

$$|f(x)| \leq C\,|g(x)| \text{ whenever } 0 < |x - x_0| < \delta \tag{1.172}$$

Condition $0 < |x - x_0| < \delta$ is called δ-neighborhood of x_0.

Proposition 1.93. ***(Implications of order [Mu, Section 1.1, p3]).*** *Let f, g be functions from Definition 1.92. Given that $g(x) \neq 0$ in δ-neighborhood of x_0, except at x_0 itself:*
(1) $f(x) = O(g(x))$ implies that f/g is bounded
(2) $f(x) = o(g(x))$ implies that $f/g \to 0$ as $x \to x_0$

Remark 1.94. **(Order [Mu, Section 1.1, p3]).** O-order describes the asymptotic (limiting) behaviour of f as $x \to x_0$, for example, if $f(x) \to 0$ as $x \to 0$, O-order tells us how fast $f(x) \to 0$. On the other hand, o-order merely confirms that $f(x) \to 0$. If $f(x), g(x) = x^2$ then $f(x) = O(g(x))$, i.e. f grows not faster then g, i.e. $f \leq g$. Whereas if $f(x) = x^2, g(x) = x^3$ then $f(x) = o(g(x))$, i.e. f growths strictly slower then g, i.e. $f < g$, in which case we also write $f \ll g$.

Definition 1.95. (Asymptotic equivalence [Mu, Section 1.1, p4]). Let f, g be functions from Definition 1.92. If f, g are such that $lim_{x \to x_0} f/g = 1$, then we say f is asymptotically equivalent to g under limit $x \to x_0$, where we write

$$f(x) \sim g(x) \text{ as } x \to x_0 \text{ if } \lim_{x \to x_0} \frac{f(x)}{g(x)} = 1 \tag{1.173}$$

Definition 1.96. (Asymptotic sequence [Mu, Section 1.2, p11]). Let $(g_n(x))_{n=0}^{\infty}$ be a sequence of functions defined on some set D. It is called asymptotic as $x \to x_0$ if

$$g_{n+1}(x) = o(g_n(x)) \text{ as } x \to x_0, \text{ for all } n \in \mathbb{N} \cup \{\infty\} \tag{1.174}$$

Definition 1.97. (Asymptotic series [Mu, Section 1.2, p12-13]). Let $(g_n(x))_{n=0}^{\infty}$ be an asymptotic sequence as $x \to x_0$ defined on some set D. Let S_N denote the partial sum defined by $S_N(x) = \sum_{n=0}^{N} a_n g_n(x)$ where a_n are constant. We say that S_N is asymptotic expansion of f, if for each $N \in \mathbb{N} \cup \{\infty\}$

$$f(x) - S_N(x) = R_N(x) \text{ as } x \to x_0 \tag{1.175}$$

where $R_N(x) = O(g_{N+1}(x)) = o(g_N(x))$. We call $\sum_{n=0}^{\infty} a_n g_n(x)$ asymptotic series where

$$f(x) \sim \sum_{n=0}^{\infty} a_n g_n(x) \text{ as } x \to x_0 \tag{1.176}$$

Note that (1.176), does not imply that series $\sum_{n=0}^{\infty} a_n g_n(x)$ is convergent, it only implies (1.175) for all $N \in \mathbb{N} \cup \{\infty\}$.

Remark 1.98. **(Convergent versus Asymptotic series [Mu, Section 1.2, p12-13]).** Suppose we have asymptotic series $\sum_{n=0}^{\infty} a_n g_n(x)$, then

(1) Convergence looks at behaviour of $S_N(x)$ in $|f(x) - S_N(x)| \to 0$ as $N \to \infty$ with x fixed

(2) Asymptoticity looks at behaviour of $S_N(x)$ in $|f(x) - S_N(x)| \to 0$ as $x \to x_0$ with N fixed

Implication is that asymptotic series does not need to be convergent and convergent series does not need to be asymptotic

Proposition. ***(Uniquness of asymptotic sequence [Mu, Section 1.2, p13]).*** *If (1.175) exists for a given f and asymptotic sequence $(g_n(x))_{n=0}^{\infty}$, then it is unique. In this case a_n can be determined uniquely by successive limits:*

$$a_1 = \lim_{x \to x_0} \frac{f(x)}{g_1(x)},\ a_2 = \lim_{x \to x_0} \frac{f(x) - a_1 g_1(x)}{g_2(x)},\ a_N = \lim_{x \to x_0} \frac{1}{g_N(x)} \left(f(x) - \sum_{n=1}^{N-1} a_n g_n(x) \right) \tag{1.177}$$

Definition 1.99. (Power series [Mu, Section 1.2, p13]). Let $((x - x_0)^n)_{n=0}^{\infty}$ be an asymptotic sequence as $x \to x_0$. We call this sequence a power sequence with asymptotic series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ called power series where

$$f(x) \sim \sum_{n=0}^{\infty} a_n (x - x_0)^n \text{ as } x \to x_0 \tag{1.178}$$

Definition 1.100. (Analytic function [Mu, Section 1.1, p5]). If power series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ is convergent to $f(x)$ for x in neighborhood of x_0, then we call f analytic function as $x \to x_0$. In other words, with $S_N(x) = \sum_{n=0}^{N} a_n (x - x_0)^n$ we have

$$|f(x) - S_N(x)| \to 0 \text{ as } x \to x_0, N \to \infty \tag{1.179}$$

In other words, by Remark 1.98, the series is both asymptotic and convergent.

Definition 1.101. (Taylor series [Gr, Chapter 3: Subsection 13.5, p 632]). Let $((x - x_0)^n)_{n=0}^{\infty}$ be a power sequence as $x \to x_0$. Suppose that $f \in C^{\infty}(\mathbb{R})$. Consider modification of power series (1.178) with $a_n = \frac{f^{(n)}(x_0)}{n!}$ where

$$f(x) \sim \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \text{ as } x \to x_0 \tag{1.180}$$

we call this series Taylor series. Here $f^{(n)}(x_0)$ is n-th derivative of f with respect to x evaluated at x_0.

Definition 1.102. (Perturbation series [Na, Section 1.6, p16-17]). Let $(g_n(\epsilon))_{n=0}^{\infty}$ be an asymptotic sequence as $\epsilon \to 0$ and let partial sum be defined as $S_N(x,\epsilon) = \sum_{n=0}^{N} f_n(x) g_n(\epsilon)$, where $g_n(\epsilon) = \epsilon^n$, a_n is function of $x \in \mathbb{R}^d$, f is a function of small parameter ϵ and $x \in \mathbb{R}^d$. Suppose that S_N is asymptotic expansion of f, i.e.

$$f(x,\epsilon) - S_N(x,\epsilon) = R_N(x,\epsilon) \text{ as } x \to x_0 \tag{1.181}$$

Note that contrary to [Na, Section 1.6, p16-17] we refer to $\sum_{n=0}^{\infty} f_n(x) g_n(\epsilon)$ as perturbation series, where

$$f(x,\epsilon) \sim \sum_{n=0}^{\infty} a_n(x) g_n(\epsilon) \text{ as } \epsilon \to 0 \tag{1.182}$$

Definition 1.103. (Uniform versus nonuniform expansion [Na, Section 1.6, p16-17]). The expansion (1.181) is said to be uniformly valid (or regular perturbation) if

$$R_N(x,\epsilon) = O(g_{N+1}(\epsilon)) \text{ as } \epsilon \to 0 \tag{1.183}$$

which will hold if $f_{n+1}(x)$ is no more singular then $f_n(x)$, i.e. for all $n = 0, 1, ..., N$

$$a_n(x) g_n(\epsilon) > a_{n+1}(x) g_{n+1}(\epsilon) \tag{1.184}$$

Whereas straightforward expansion is said to be nonuniformly valid (or singular perturbation) if $f_{n+1}(x)$ is more singular then $f_n(x)$, i.e. for all $n = 0, 1, ..., N$

$$a_n(x) g_n(\epsilon) \leq a_{n+1}(x) g_{n+1}(\epsilon) \tag{1.185}$$

Definition 1.104. (Perturbation problem [Na, Section 1.1, p1-2], [Mu, Section 7.1, p138-139]). Suppose we have a problem L with boundary conditions B

$$L(f, x, \epsilon) = 0, \;\; B(f, \epsilon) = 0 \tag{1.186}$$

where $f = f(x,\epsilon)$ is dependent variable, x is an independent variable, and L, B are operations such as derivatives or integrals of f. Suppose that we guess the solution as perturbation series

$$f(x,\epsilon) \sim \sum_{n=0}^{\infty} a_n(x) \epsilon^n \text{ as } \epsilon \to 0 \tag{1.187}$$

If $f(x,\epsilon)$ is analytic as $\epsilon \to 0$, (1.187) expansion is uniformly valid and it is a regular perturbation problem. This can be solved successivly by means of substitution of $f_0, f_1, ...,$ into (1.186), where $f_0(x) = \lim_{\epsilon \to 0} f(x,\epsilon)$ satisfies

$$L(f_0, x, 0) = 0, \;\; B(f_0, 0) = 0 \tag{1.188}$$

Otherwise if $f(x,\epsilon)$ is not analytic as $\epsilon \to 0$, (1.187) expansion is nonuniformly valid and it is a singular perturbation problem, for which we cannot use the above method to derive the solution to (1.187).

Remark 1.105. Perturbation methods concern with finding an approximate solution to a problem by means of using an exact solution to a related solvable (unperturbed) problem and its perturbation. The solution is expressed as a power series of ϵ, i.e. (1.182), where leading term is the exact solution of solvable (unperturbed) problem (i.e. problem with $\epsilon = 0$) and all the higher order terms are successive solutions to a perturbed problem (i.e. problem with $\epsilon > 0$). Truncating series for sufficiently large N will produce an approximate solution to the problem.

Remark 1.106. **(Sources of non-uniformity [Na, Chapter 2: Subsection 2.2-2.5]).** Consider perturbation problem from Definition 1.104. Sources of nonuniformity in straightforward expansion (1.187) include

(1) Small parameter multiplying highest derivative: setting $\epsilon = 0$ reduces order of the problem, solutions will not be able to satisfy all boundaries.

(2) Change of PDE class: setting $\epsilon = 0$ changes PDE class, for example from elliptic to parabolic, see Definition 1.88.

(3) Presence of singularities: one of the solutions in expansion (1.187) is singular, for example $f_0(x)$, for some values of x.

(4) Coordinate change: setting $\epsilon = 0$ changes boundaries or problem from inhomogeneous to homogeneous, see Definition 1.88.

Definition 1.107. (Fourier integral [Gr, Section 17.109, p913-915]). Let f be a function from Theorem 1.108. Let

$$a(u) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x)\cos(ux)dx,\ b(u) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x)\sin(ux)dx \tag{1.189}$$

Then

$$f(x) = \int_{0}^{\infty} \left[a(u)\cos(ux) + b(u)\sin(ux)\right] du \tag{1.190}$$

where the right hand side of (1.190) is called fourier integral of f.

Theorem 1.108. ***(Fourier integral [Gr, Theorem 17.9.1, p915])***. *Let f be defined on $\mathbb{R}$ and let $f, \frac{df(x)}{dx}$ be piecewise continious on every finite interval $[-\ell, \ell]$, i.e. for arbitrarily large ℓ. Suppose that $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$. Then Fourier integral of f converges to $f(x)$ at every point x at which f is continious, and to the mean value $(f(x+) + f(x-))/2$ at every point x at which f is discontinious. Refer to Definition 5.11 for the meaning of $f(x+)$ and $f(x-)$.*

Definition 1.109. (Fourier transform [Gr, Section 17.10, p921; Theorem 17.9.1]). Fourier transform of function f is defined by

$$F\{f(x)\} = \hat{f}(u) = \int_{-\infty}^{\infty} e^{-iux} f(x) dx \tag{1.191}$$

Inverse Fourier transform of function f is defined by

$$F^{-1}\left\{\hat{f}(u)\right\} = f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{iux}\hat{f}(u) du \tag{1.192}$$

where integrals of (1.191) and (1.192) are Fourier integrals.

Proposition 1.110. ***(Fourier transform of sign function [Ro, Chapter 3, p65]).*** *Let sgn be a sign function from Definition 5.7. Then sgn has following Fourier integral representation*

$$sgn(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin(ux)}{u} du \tag{1.193}$$

Proof. Note that

$$sgn(x) = \lim_{a\to 0} e^{-a|x|} sgn(x) \tag{1.194}$$

Let $H(x) = \mathbf{1}_{\{x>0\}}(x)$, then

$$sgn(x) = H(x) - H(-x) \tag{1.195}$$

and

$$sgn(x) = \lim_{a\to 0}\left(e^{-ax}H(x) - e^{ax}H(-x)\right) \tag{1.196}$$

By Definition 1.109 of Fourier transform

$$\begin{aligned} F\{sgn(x)\} &= \int_{-\infty}^{\infty} e^{-iux} sgn(x) dx \\ &= \int_{-\infty}^{\infty} \lim_{a\to 0}\left(e^{-ax}H(x) - e^{at}H(-x)\right) e^{-iux} dx \end{aligned} \tag{1.197}$$

By dominated convergence Theorem 5.62

$$\begin{aligned} F\{sgn(x)\} &= \lim_{a\to 0} \int_{-\infty}^{\infty} \left(e^{-ax}H(x) - e^{ax}H(-x)\right) e^{-iux} dx \\ &= \lim_{a\to 0}\left(\int_{-\infty}^{\infty} e^{-x(a+iu)}H(t)dt - \int_{-\infty}^{\infty} e^{x(a-iu)}H(-x)dx\right) \\ &= \lim_{a\to 0}\left(\int_{-\infty}^{\infty} e^{-x(a+iu)}\mathbf{1}_{\{x>0\}}(x)dt - \int_{-\infty}^{\infty} e^{x(a-iu)}\mathbf{1}_{\{x>0\}}(-x)dx\right) \\ &= \lim_{a\to 0}\left(\int_{0}^{\infty} e^{-x(a+iu)}dt - \int_{-\infty}^{0} e^{x(a-iu)}dx\right) \\ &= \lim_{a\to 0}\left(\int_{0}^{\infty} e^{-x(a+iu)}dt - \int_{0}^{\infty} e^{-x(a-iu)}dx\right) \\ &= \lim_{a\to 0}\left(\left.\frac{e^{-x(a+iu)}}{-(a+iu)}\right|_0^{\infty} - \left.\frac{e^{-x(a-iu)}}{-(a-iu)}\right|_0^{\infty}\right) \\ &= \lim_{a\to 0}\left(\frac{1}{a+iu} - \frac{1}{a-iu}\right) \\ &= \frac{2}{iu} \\ &= \widehat{sgn}(u) \end{aligned} \tag{1.198}$$

By Definition 1.109 of inverse Fourier transform

$$\begin{aligned}
F^{-1}\left\{\widehat{sgn}(u)\right\} &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} e^{iux}\widehat{sgn}(u)du \\
&= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} e^{iux}\tfrac{2}{iu}du \\
&= \tfrac{1}{\pi}\int_{-\infty}^{\infty} \tfrac{e^{iux}}{iu}du
\end{aligned} \tag{1.199}$$

By Euler's formula from Definition 5.6 and integral properties of odd functions from Proposition 5.10

$$\begin{aligned}
F^{-1}\left\{\widehat{sgn}(u)\right\} &= \tfrac{1}{\pi}\int_{-\infty}^{\infty}\left(\tfrac{1}{i}\tfrac{\cos(ux)}{iu} + \tfrac{\sin(ux)}{u}\right)du \\
&= \tfrac{1}{\pi}\int_{-\infty}^{\infty}\tfrac{1}{i}\tfrac{\cos(ux)}{u}du + \tfrac{1}{\pi}\int_{-\infty}^{\infty}\tfrac{\sin(ux)}{u}du \\
&= \tfrac{1}{\pi}\int_{-\infty}^{\infty}\tfrac{\sin(ux)}{u}du \\
&= sgn(x)
\end{aligned} \tag{1.200}$$

□

2 Market making in stock

2.1 Notations

$0 \leq t < T < \infty$

$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_s)_{s \in [t,T]}$

$W = (W_s)_{s \in [t,T]}$ is an $\mathbb{R}$-valued $(\mathcal{F}_s)$-Brownian motion

$N^a = (N^a_s)_{s \in [t,T]}$ is an $\mathbb{N}$-valued $(\mathcal{F}_s)$-Poisson processes with intensity λ^a

$N^b = (N^b_s)_{s \in [t,T]}$ is an $\mathbb{N}$-valued $(\mathcal{F}_s)$-Poisson processes with intensity λ^b

N^a, N^b are independent

$\mathcal{S} = [t, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{N}$

$\bar{\mathcal{S}} = [t, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{N}$

$\partial\mathcal{S} = \{T\} \times \mathbb{R} \times \mathbb{R} \times \mathbb{N}$

2.2 Assumptions

Definition 2.1. (Spot price process [AS, Section 2.1]). Let $S = (S_s)_{s \in [t,T]}$ be an $\mathbb{R}$-valued continious process defined by

$$S_T = s + \sigma dW_T \tag{2.1}$$

We call it a spot price process and interpret it as a mid price of a stock at time T traded by the market maker.

Definition 2.2. (Inventory process [AS, Section 2.4]). Let $Q = (Q_s)_{s \in [t,T]}$ be an $\mathbb{N}$-valued piecewise constant and càdlàg process defined by

$$Q_T = q + N^b_T - N^a_T \tag{2.2}$$

We call it an inventory process and interpret it as an amount of stock held in inventory of the market maker at time T. Processes N^a, N^b correspond to an amount of stock bought and sold by the market maker up to time t.

Definition 2.3. (Limit orders [AS, Section 2.4]). We call p^b, p^a bid and ask prices. We call δ^a, δ^b, a distances (or spreads) of bid and ask price from the mid price at time t, which is $S_t = s$ and defined by

$$\delta^b = s - p^b, \;\; \delta^a = p^a - s \tag{2.3}$$

We interpret bid (ask) price as the price at which a limit order is sent by the market maker to sell (buy) a stock.

Definition 2.4. (Intensity functions [AS, Section 2.4]). Intensities λ^a, λ^b of Poisson process N^a, N^b correspond to the probability of execution of the limit order posted by the market maker. Probability of execution of limit order depends on the quotes distances δ^a, δ^b from the mid-price S. As spreads δ^a, δ^b increase, limit orders posted by market maker get deeper into the order book, so it becomes less likely that incoming order will reach them, i.e. probability of execution falls. Hence, intensities λ^a, λ^b are functions of δ^a, δ^b, i.e. $\lambda^a = \lambda^a(\delta^a), \lambda^b = \lambda^b(\delta^b)$

Proposition 2.5. ***(Intensities as decreasing functions of spreads [AS, Section 2.4]).*** *In the view of Definition 2.4, intensities $\lambda^a = \lambda^a(\delta^a), \lambda^b = \lambda^b(\delta^b)$ are decreasing functions of spreads δ^a, δ^b in classical sense, i.e.*
(1) if $\delta^a \geq \delta^b$ then $\lambda^a(\delta^a) \leq \lambda^b(\delta^b)$
(2) if $\delta^a \leq \delta^b$ then $\lambda^a(\delta^a) \geq \lambda^b(\delta^b)$

Definition 2.6. (Cash process [AS, Section 2.4]). Let $X = (X_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process assumed to satisfy following SDE

$$\begin{cases} dX_T = p^a dN_T^a - p^b dN_T^b \\ X_t = x \text{ a.s.} \end{cases} \tag{2.4}$$

which we call cash process.

Definition 2.7. (Utility function [AS, Section 2.4]). Let g be a function taking values in $\mathbb{R}$ defined by

$$g(s, x, q) = -e^{-\gamma(x+qs)} \text{ for } (s, x, q) \in \mathbb{R} \times \mathbb{R} \times \mathbb{N} \tag{2.5}$$

We call it a utility function where constant γ is called a risk-aversion coefficient. Coefficient γ controls the market maker's utility profile through function g, where for $\gamma = 0$ there is no relationship between wealth and utility (market maker is risk-neutral), for $\gamma < 0$ there is positive relationship between wealth and utility (market maker is risk-seeking) and for $\gamma > 0$ there is negative relationship between wealth and utility (market maker is risk-averse).

Definition 2.8. (Value function [AS, Section 3.1]). Let g be a utility function from Definition 2.7. Let $U = \mathbb{R}^+ \times \mathbb{R}^+$ be the set of values for controls (δ^a, δ^b). The stochastic control problem for risk-neutral market maker is to maximise expectation of utility of terminal wealth over controls (δ^a, δ^b). The value function for this problem is defined by

$$v(t, s, x, q) = \max_{(\delta^a,\delta^b)\in U} \mathbb{E}_{t,s,x,q}\left(g(S_T, X_T^{(\delta^a,\delta^b)}, Q_T^{(\delta^a,\delta^b)})\right) \tag{2.6}$$

Definition 2.9. (Market impact [AS, Section 2.4]). Suppose that $\delta \in \{\delta^a, \delta^b\}$ and let Q we the size of the market order, then Q limit orders with best prices will get executed. This causes a temporary market impact as market order crosses the spread and consumes liquidity provided by resting orders on the other side of the book. We denote highest (lowest) price for market order to buy (sell) as p^Q and so market impact from market order Q is defined by

$$\Delta p = p^Q - s \tag{2.7}$$

where s is a spot price of the stock. If market makers limit orders are within range of this market order, i.e.

$$\delta \leq \Delta p \tag{2.8}$$

then these orders will get executed.

2.3 State process

Proposition 2.10. ***(Controlled wealth process).*** *Refer to Subsection 2.1 for notations and Subsection 2.2 for assumptions. Consider the following SDE*

$$\begin{cases} dY_\vartheta(T) & = & bdW_T + G(\vartheta)dN_T \\ Y_\vartheta(t) & = & y \end{cases} \tag{2.9}$$

With initial condition

$$\mathbb{P}\left(Y_\vartheta(t) = y\right) = 1 \textit{ for } \mathcal{F}_t\textit{-measurable } (\mathbb{R}\times\mathbb{R}\times\mathbb{N})\textit{-valued random variable } y \tag{2.10}$$

Where

$$Y(T) = \begin{bmatrix} S_T \\ X_T \\ Q_T \end{bmatrix},\ y = \begin{bmatrix} s \\ x \\ q \end{bmatrix},\ b = \begin{bmatrix} \sigma \\ 0 \\ 0 \end{bmatrix},\ G(\vartheta) = \begin{bmatrix} 0 & 0 \\ p^a & -p^b \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ s+\delta^a & -s+\delta^b \\ -1 & 1 \end{bmatrix} \tag{2.11}$$

$$N_T = \begin{bmatrix} N_T^a \\ N_T^b \end{bmatrix},\ \lambda(\vartheta) = \begin{bmatrix} \lambda^a(\delta^a) \\ \lambda^b(\delta^b) \end{bmatrix},\ \vartheta = \begin{bmatrix} \delta^a \\ \delta^b \end{bmatrix},\ \tilde{N}_T = N_T - \lambda(\vartheta) \tag{2.12}$$

Suppose that control process ϑ is constant, with $\vartheta \in U \subset \mathbb{R}^2$ where set U is bounded. Then 2.9 has a unique solution $Y_\vartheta^{t,y} = (Y_\vartheta^{t,y}(s)), s \in [t,T]$ with initial condition (2.10) satisfying the following stochastic integral equation almost surely

$$Y_\vartheta^{t,y}(T) = y + \int_t^T bdW_s + \int_t^T G(\vartheta)dN_s \tag{2.13}$$

Proof. Suppose $N^A = (N(t,A)), t \geq 0$ and $N^B = (N(t,B)), t \geq 0$ are Poisson random measures over disjoint Borel sets A and B, both bounded below. Then by Theorem 1.46, N^A and N^B are independent. By Theorem 1.31, if A is a Borel set bounded below then N^A is a Poisson process. Hence, both N^A, N^B are in fact, special cases of Poisson random measures, namely Poisson processes. We can interpret their stochastic integrals as

$$\int_t^T \int_A N(ds,dx) = \int_t^T dN_s^a,\ \int_t^T \int_B N(ds,dx) = \int_t^T dN_s^b \tag{2.14}$$

Moreover, by Definition 1.32

$$\int_t^T d\tilde{N}_s^a = \int_t^T dN_s^a - \int_t^T \lambda^a(\delta^a)ds,\ \int_t^T d\tilde{N}_s^b = \int_t^T dN_s^b - \int_t^T \lambda^b(\delta^b)ds \tag{2.15}$$

With the assumption that intensities $\lambda^a(\delta^a), \lambda^b(\delta^b)$ are decreasing functions of controls δ^a, δ^b in classical sense, together with the assumption that set U is bounded, allows us to imply that $\lambda^a(\delta^a), \lambda^b(\delta^b)$ are also bounded and so

$$\lambda^a(\delta^a) < \infty,\ \lambda^b(\delta^b) < \infty \text{ for any } (\delta^a, \delta^b) \in U \tag{2.16}$$

Next, we check if ϑ is admissible control from Definition 1.64. Assumption (2) is satisfied by Proposition 1.65 as set U is bounded. Since ϑ is a constant, it is trivially $(\mathcal{F}_s)$-measurable and so is $(\mathcal{F}_s)$-progressively measurable. Therefore ϑ is admissible control. Now we check if assumptions on coefficients of 2.9 from Theorem 1.68 are satisfied. First note that vector a is constant and matrix G depends on ϑ. Since U is bounded we have that

$$\left|s+\delta^a-1\right|^2 \lambda^a(\delta^a)+\left|-s+\delta^b+1\right|^2 \lambda^b(\delta^b)<\infty \text{ for any } (\delta^a,\delta^b)\in U \tag{2.17}$$

Hence, both Linear growth and Lipschitz conditions are satisfied, which means that $Y_{\vartheta}^{t,y}$ satisfying (2.13) is the unique solution. □

Proposition 2.11. ***(Order arrival intensity [AS, Section 2.5]).*** *We use assumptions from Definition 2.9. Let random variable Q be a size of market order with the distribution f^Q . Let Δp be a market impact, let $\delta \in \{\delta^a, \delta^b\}$ and $\lambda \in \{\lambda^a, \lambda^b\}$. Let $\mathbb{P}(\Delta p \geq \delta)$ be the probability of execution of market makers limit order. Suppose that the distribution f^Q obeys a Power-law, i.e.*

$$f^Q(x) \propto x^{-1-\alpha} \textit{ for } \alpha>0 \tag{2.18}$$

And suppose that market impact Δp is directly proportional to log of size of market order $ln(Q)$, i.e.

$$\Delta p \propto \ln(Q) \tag{2.19}$$

Let Λ be a constant and denote the arrival rate (frequency) of market orders Q. Then intensity λ defined by

$$\lambda(\delta)=\Lambda\mathbb{P}(\Delta p \geq \delta) \;\; \textit{for } \delta \geq 0 \tag{2.20}$$

is of the form

$$\lambda(\delta)=Ae^{-k\delta} \tag{2.21}$$

Proof. Since Δp is directly proportional to $ln(Q)$, there exists $c>0$ such that

$$\Delta p = c \;\ln(Q) \tag{2.22}$$

Moreover, since Q has density f^Q we have that

$$\begin{aligned}
\lambda(\delta) &= \Lambda\mathbb{P}\left(\Delta p \geq \delta\right) \\
&= \Lambda\mathbb{P}\left(c\ln(Q) \geq \delta\right) \\
&= \Lambda\mathbb{P}\left(\ln(Q) \geq \tfrac{1}{c}\delta\right) \\
&= \Lambda\mathbb{P}\left(Q \geq \exp(\tfrac{1}{c}\delta)\right) \\
&= \Lambda\int_{\exp(\frac{1}{c}\delta)}^{\infty} f^Q(x)dx \\
&= \Lambda\int_{\exp(\frac{1}{c}\delta)}^{\infty} x^{-1-\alpha}dx \\
&= \tfrac{\Lambda}{\alpha}e^{-\alpha\frac{1}{c}\delta} \\
&= Ae^{-k\delta}
\end{aligned} \tag{2.23}$$

where $A = \frac{\Lambda}{\alpha}$ and $k = \alpha\frac{1}{c}$.

□

2.4 HJB equation

Remark 2.12. **(Dynamic programming equation).** Let $Y^{t,y}_\vartheta = (Y^{t,y}_\vartheta(s)), s \in [t,T]$ be a unique solution to (2.9) as defined in Proposition 2.10. Let v be a value function from Definition 2.8, which we will rewrite in matrix form as

$$v(t,y) = \max_{\vartheta \in U} \mathbb{E}\left(g(Y^{t,y}_\vartheta(T))\right) \tag{2.24}$$

By equivalent formulation of dynamic programming equation from Remark 1.83 where for all $(t,y) \in [0,T] \times \mathbb{R}^d$, $u \in \mathcal{U}(t,y)$, $\theta \in \mathcal{T}_{t,T}$ we have

$$v(t,y) \geq \mathbb{E}\left(\int_t^\theta f(s, Y^{t,y}_u(s), u(s))ds + v(\theta, Y^{t,y}_u(\theta))\right) \tag{2.25}$$

Since controls ϑ are constant, by Proposition 1.75, ϑ belongs to $\mathcal{U}(t,y)$. Letting $T = \theta$ and noting that we do not have function f in our stochastic control problem, we deduce that v satisfies

$$v(t,y) \geq \mathbb{E}\left(v(T, Y^{t,y}_\vartheta(T))\right) \tag{2.26}$$

Note that by conditional expectation notations from Definition 1.69 we have

$$\mathbb{E}\left(v(T, Y^{t,y}_\vartheta(T))\right) = \mathbb{E}_{t,y}\left(v(T, Y^\vartheta(T))\right) \tag{2.27}$$

Proposition 2.13. ***(Derivation of HJB equation [AS, Section 3.1, Equation 13]).*** *Refer to Subsection 2.1 for notations and Subsection 2.2 for assumptions. Let $Y^{t,y}_\vartheta = (Y^{t,y}_\vartheta(s)), s \in [t,T]$ be a unique solution to (2.9) as defined in Proposition 2.10. Then value function v satisfies following PDE*

$$\begin{cases} \frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial s^2} + \mathcal{H} = 0 \text{ for all } (t,s,x,q) \in \mathcal{S} \\ v(t,s,x,q) = g(s,x,q) \text{ for all } (t,s,x,q) \in \partial\mathcal{S} \end{cases} \tag{2.28}$$

where

$$\begin{aligned} \mathcal{H} &= \max_{\delta^a} \lambda^a(\delta^a)\left[v(t,s,x+s+\delta^a,q-1) - v(t,s,x,q)\right] \\ &+ \max_{\delta^b} \lambda^b(\delta^b)\left[v(t,s,x-s+\delta^b,q+1) - v(t,s,x,q)\right] \end{aligned} \tag{2.29}$$

Proof. Generalising results obtained in [Ph, Subsection 3.4.1, p43]. Let $t + \tau = T$ in (2.26) then

$$v(t,y) \geq \mathbb{E}_{t,y}\left(v(t+h, Y^\vartheta(t+h))\right) \tag{2.30}$$

First notice that a function

$$v(t,y) = g(y) = -e^{-\gamma y} \text{ belongs to } C^2(\bar{\mathcal{S}}) \tag{2.31}$$

So we apply the Ito Lemma from Theorem 1.61 to function v over time interval $t, t+\tau$

$$
\begin{aligned}
v(t+\tau, Y_\vartheta^{t,y}(t+\tau)) &= v(t,y) \\
&+ \int_t^{t+\tau} \left[\frac{\partial v}{\partial u} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial S^2}\right] du + \int_t^{t+\tau} \left[\sigma \frac{\partial v}{\partial S}\right] dW_u \\
&+ \int_t^{t+\tau} \left[v(u, S_u, X_u + s + \delta^a, Q_u - 1) - v(u, S_u, X_u, Q_u)\right] dN_u^a \\
&+ \int_t^{t+\tau} \left[v(u, S_u, X_u - s + \delta^b, Q_u + 1) - v(u, S_u, X_u, Q_u)\right] dN_u^b
\end{aligned}
\tag{2.32}
$$

By Definition 1.32 we get

$$
\begin{aligned}
v(t+\tau, Y_\vartheta^{t,y}(t+\tau)) &= v(t,y) \\
&+ \int_t^{t+\tau} \left[\frac{\partial v}{\partial u} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial S^2}\right] du + \int_t^{t+\tau} \left[\sigma \frac{\partial v}{\partial S}\right] dW_u \\
&+ \int_t^{t+\tau} \left[v(u, S_u, X_u + s + \delta^a, Q_u - 1) - v(u, S_u, X_u, Q_u)\right] dN_u^a \\
&+ \int_t^{t+\tau} \left[v(u, S_u, X_u - s + \delta^b, Q_u + 1) - v(u, S_u, X_u, Q_u)\right] dN_u^b \\
&+ \int_t^{t+\tau} \left[v(u, S_u, X_u + s + \delta^a, Q_u - 1) - v(u, S_u, X_u, Q_u)\right] \lambda^a(\delta^a) du \\
&+ \int_t^{t+\tau} \left[v(u, S_u, X_u - s + \delta^b, Q_u + 1) - v(u, S_u, X_u, Q_u)\right] \lambda^b(\delta^b) du
\end{aligned}
\tag{2.33}
$$

Recall the notation from Definition 1.69 where

$$
\mathbb{E}\left(Y_\vartheta^{t,y}(T)\right) = \mathbb{E}_{t,y}\left(Y^\vartheta(T)\right) = \mathbb{E}\left(Y^\vartheta(T) \middle| Y^\vartheta(t) = y\right)
\tag{2.34}
$$

Next, we substitute the above into (2.26), then by linearity of expectation we have

$$
\begin{aligned}
v(t,y) &\geq v(t,y) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[\frac{\partial v}{\partial u} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial S^2}\right] du\right) + \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[\sigma \frac{\partial v}{\partial S}\right] dW_u\right) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[v(u, S_u, X_u + s + \delta^a, Q_u - 1) - v(u, S_u, X_u, Q_u)\right] dN_u^a\right) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[v(u, S_u, X_u - s + \delta^b, Q_u + 1) - v(u, S_u, X_u, Q_u)\right] dN_u^b\right) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[v(u, S_u, X_u + s + \delta^a, Q_u - 1) - v(u, S_u, X_u, Q_u)\right] \lambda^a(\delta^a) du\right) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[v(u, S_u, X_u - s + \delta^b, Q_u + 1) - v(u, S_u, X_u, Q_u)\right] \lambda^b(\delta^b) du\right)
\end{aligned}
\tag{2.35}
$$

Brownian motion W is a martingale by Definition 1.39 and Compensated Poisson processes $\tilde{N}^a, \tilde{N}^b$ are also martingales by Proposition 1.5 so we get

$$
\begin{aligned}
v(t,y) &\geq v(t,y) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[\frac{\partial v}{\partial u} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial S^2}\right] du\right) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[v(u, S_u, X_u + s + \delta^a, Q_u - 1) - v(u, S_u, X_u, Q_u)\right] \lambda^a(\delta^a) du\right) \\
&+ \mathbb{E}_{t,s,x,q}\left(\int_t^{t+\tau} \left[v(u, S_u, X_u - s + \delta^b, Q_u + 1) - v(u, S_u, X_u, Q_u)\right] \lambda^b(\delta^b) du\right)
\end{aligned}
\tag{2.36}
$$

by mean value Theorem 5.4

$$
\begin{aligned}
v(t,y) &\geq v(t,y) \\
&+ \tau \mathbb{E}_{t,s,x,q}\left(\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial S^2}\right) \\
&+ \tau \mathbb{E}_{t,s,x,q}\left(\left[v(t,S_t,X_t+s+\delta^a,Q_t-1) - v(u,S_t,X_t,Q_t)\right]\lambda^a(\delta^a)\right) \\
&+ \tau \mathbb{E}_{t,s,x,q}\left(\left[v(t,S_t,X_t-s+\delta^b,Q_t+1) - v(u,S_t,X_t,Q_t)\right]\lambda^b(\delta^b)\right)
\end{aligned} \tag{2.37}
$$

The initial condition (2.10) implies $\mathbb{P}\left(\mathbb{E}_{t,y}\left(Y^{\vartheta}(t)\right) = y\right) = 1$, hence

$$
\begin{aligned}
v(t,y) &\geq v(t,y) \\
&+ \tau\left(\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial s^2}\right) \\
&+ \tau\lambda^a(\delta^a)\left(v(t,s,x+s+\delta^a,q-1) - v(t,s,x,q)\right) \\
&+ \tau\lambda^b(\delta^b)\left(v(t,s,x-s+\delta^b,q+1) - v(t,s,x,q)\right)
\end{aligned} \tag{2.38}
$$

Substracting $v(t,y)$ from both sides, diving by τ and taking limit as $\tau \to 0$

$$
\begin{aligned}
0 &\geq \frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial s^2} \\
&+ \lambda^a(\delta^a)\left(v(t,s,x+s+\delta^a,q-1) - v(t,s,x,q)\right) \\
&+ \lambda^b(\delta^b)\left(v(t,s,x-s+\delta^b,q+1) - v(t,s,x,q)\right)
\end{aligned} \tag{2.39}
$$

Letby mean value Theorem 5.4

$$
\begin{aligned}
\mathcal{J}(t,s,\vartheta) &= \lambda^a(\delta^a)\left(v(t,s,x+s+\delta^a,q-1) - v(t,s,x,q)\right) \\
&+ \lambda^b(\delta^b)\left(v(t,s,x-s+\delta^b,q+1) - v(t,s,x,q)\right)
\end{aligned} \tag{2.40}
$$

then

$$
0 \geq \frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial s^2} + \mathcal{J}(t,s,\vartheta) \tag{2.41}
$$

Now suppose that there exists $\hat{\vartheta}$ such that

$$
v(t,y) = \mathbb{E}_{t,y}\left(v(t+\tau, Y^{\hat{\vartheta}}(t+\tau))\right) \tag{2.42}
$$

Then by same arguments as above we obtain equality

$$
0 = \frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial s^2} + \mathcal{J}(t,s,\hat{\vartheta}) \tag{2.43}
$$

It follows that by Definitions 1.70 and 1.71 we have

$$
\mathcal{J}(t,s,\hat{\vartheta}) = \sup_{\vartheta \in U} \mathcal{J}(t,s,y,\vartheta) \geq \mathcal{J}(t,s,\vartheta) \tag{2.44}
$$

This suggests that v should satisfy

$$
\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial s^2} + \mathcal{H} = 0 \tag{2.45}
$$

where by boundedness of U

$$\mathcal{H} = \sup_{\vartheta \in U} \mathcal{J}(t,s,y,\vartheta) = \max_{\vartheta \in U} \mathcal{J}(t,s,y,\vartheta) \tag{2.46}$$

From Proposition 1.84 regular terminal condition associated to PDE (2.45) is

$$v(T,s,x,q) = g(s,x,q) \text{ for all } y \in \mathbb{R}^3 \tag{2.47}$$

□

Remark 2.14. Let v be the value function satisfying (2.28). As it depends on t, s, x, q, we could simplify (2.28) such that dependence of v on x is eliminated, i.e. factor out x from v. The following result covers this.

Proposition 2.15. ***(Simplification of HJB equation [AS, Section 3.1, Equation 15]).*** *Let v be the value function satisfying (2.28). Then (2.28) can be simplified in the sense of removing dependence of v on variable x, by defining a new function w such that w satisfies*

$$\begin{cases} \frac{\partial w}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 w}{\partial s^2} - \frac{1}{2}\sigma^2 \gamma \left(\frac{\partial w}{\partial s}\right)^2 + \mathcal{H} = 0 \text{ for all } t \in [0,T),\ s,q \in \mathbb{R} \\ w(T,s,q) = qs \text{ for all } s,q \in \mathbb{R} \end{cases} \tag{2.48}$$

where $\mathcal{H}$ is defined by

$$\begin{aligned} \mathcal{H} &= \max_{\delta^a} \frac{\lambda^a(\delta^a)}{\gamma}\left[1 - e^{-\gamma(s+\delta^a - r^a(s,q,t))}\right] \\ &+ \max_{\delta^b} \frac{\lambda^b(\delta^b)}{\gamma}\left[1 - e^{\gamma(s-\delta^b - r^b(s,q,t))}\right] \end{aligned} \tag{2.49}$$

and

$$r^b(s,q,t) = w(s,q+1,t) - w(s,q,t),\ \ r^a(s,q,t) = w(s,q,t) - w(s,q-1,t) \tag{2.50}$$

Proof. Note that function w in (2.48) does not depend on x, so we need to factor it out of v. For this we will use separation of variables with a solution (ansatz) to (2.28) proposed in [AS, Section 3.1, Equation 14] of the form

$$v(t,s,x,q) = -e^{-\gamma x} e^{-\gamma w(t,s,q)} = z(x, w(t,s,q)) \tag{2.51}$$

By direct substitution into (2.28)

$$\begin{aligned} 0 = &\ \frac{\partial z(x,w(t,s,q))}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 z(x,w(t,s,q))}{\partial s^2} \\ + &\ \max_{\delta^a} \lambda^a(\delta^a)\left[z(x+s+\delta^a, w(t,s,q-1)) - z(x,w(t,s,q))\right] \\ + &\ \max_{\delta^b} \lambda^b(\delta^b)\left[z(x-s+\delta^b, w(t,s,q+1)) - z(x,w(t,s,q))\right] \end{aligned} \tag{2.52}$$

By the chain rule

$$\begin{aligned}
0 = \quad & \frac{\partial z(x,w(t,s,q))}{\partial w(t,s,q)}\frac{\partial w(t,s,q)}{\partial t} + \frac{1}{2}\sigma^2\left[\frac{\partial z(x,w(t,s,q))}{\partial w(t,s,q)}\frac{\partial^2 w(t,s,q)}{\partial s^2} + \frac{\partial^2 z(x,w(t,s,q))}{\partial w(t,s,q)^2}\left(\frac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\
+ \quad & \max_{\delta^a}\lambda^a(\delta^a)\left[z(x+s+\delta^a, w(t,s,q-1)) - z(x, w(t,s,q))\right] \\
+ \quad & \max_{\delta^b}\lambda^b(\delta^b)\left[z(x-s+\delta^b, w(t,s,q+1)) - z(x, w(t,s,q))\right]
\end{aligned} \tag{2.53}$$

By exponential rule the partial derivatives are

$$\begin{aligned}
\frac{\partial z(x,w(t,s,q))}{\partial w(t,s,q)} &= \gamma e^{-\gamma x}e^{-\gamma w(t,s,q)} = \gamma(-z(x,w(t,s,q))) \\
\frac{\partial^2 z(x,w(t,s,q))}{\partial w(t,s,q)^2} &= -\gamma^2 e^{-\gamma x}e^{-\gamma w(t,s,q)} = -\gamma^2(-z(x,w(t,s,q)))
\end{aligned} \tag{2.54}$$

Substituting these back into equation

$$\begin{aligned}
0 = \quad & \gamma\left(-z(x,w(t,s,q))\right)\frac{\partial w(t,s,q)}{\partial t} \\
+ \quad & \frac{1}{2}\sigma^2\left[\gamma\left(-z(x,w(t,s,q))\right)\frac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma^2\left(-z(x,w(t,s,q))\right)\left(\frac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\
+ \quad & \max_{\delta^a}\lambda^a(\delta^a)\left[z(x+s+\delta^a, w(t,s,q-1)) - z(x, w(t,s,q))\right] \\
+ \quad & \max_{\delta^b}\lambda^b(\delta^b)\left[z(x-s+\delta^b, w(t,s,q+1)) - z(x, w(t,s,q))\right]
\end{aligned} \tag{2.55}$$

Dividing by $\gamma\left(-z(x,w(t,s,q))\right)$

$$\begin{aligned}
0 = \quad & \frac{\partial w(t,s,q)}{\partial t} + \frac{1}{2}\sigma^2\left[\frac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma\left(\frac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\
+ \quad & \max_{\delta^a}\frac{\lambda^a(\delta^a)}{\gamma(-z(x,w(t,s,q)))}\left[z(x+s+\delta^a, w(t,s,q-1)) - z(x, w(t,s,q))\right] \\
+ \quad & \max_{\delta^b}\frac{\lambda^b(\delta^b)}{\gamma(-z(x,w(t,s,q)))}\left[z(x-s+\delta^b, w(t,s,q+1)) - z(x, w(t,s,q))\right]
\end{aligned} \tag{2.56}$$

Moving $\left(-z(x,w(t,s,q))\right)^{-1}$ inside the bracket

$$\begin{aligned}
0 = \quad & \frac{\partial w(t,s,q)}{\partial t} + \frac{1}{2}\sigma^2\left[\frac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma\left(\frac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\
+ \quad & \max_{\delta^a}\frac{\lambda^a(\delta^a)}{\gamma}\left[\frac{z(x,w(t,s,q))}{z(x,w(t,s,q))} - \frac{z(x+s+\delta^a,w(t,s,q-1))}{z(x,w(t,s,q))}\right] \\
+ \quad & \max_{\delta^b}\frac{\lambda^b(\delta^b)}{\gamma}\left[\frac{z(x,w(t,s,q))}{z(x,w(t,s,q))} - \frac{z(x-s+\delta^b,w(t,s,q+1))}{z(x,w(t,s,q))}\right]
\end{aligned} \tag{2.57}$$

Then

$$\begin{aligned}
0 = \quad & \frac{\partial w(t,s,q)}{\partial t} + \frac{1}{2}\sigma^2\left[\frac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma\left(\frac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\
+ \quad & \max_{\delta^a}\frac{\lambda^a(\delta^a)}{\gamma}\left[1 - \frac{z(x+s+\delta^a,w(t,s,q-1))}{z(x,w(t,s,q))}\right] \\
+ \quad & \max_{\delta^b}\frac{\lambda^b(\delta^b)}{\gamma}\left[1 - \frac{z(x-s+\delta^b,w(t,s,q+1))}{z(x,w(t,s,q))}\right]
\end{aligned} \tag{2.58}$$

Substituting for $z(x, w(t,s,q)) = -e^{-\gamma x} e^{-\gamma\theta(t,s,q)}$

$$\begin{aligned} 0 = \quad & \tfrac{\partial w(t,s,q)}{\partial t} + \tfrac{1}{2}\sigma^2 \left[\tfrac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma \left(\tfrac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\ + \quad & \max_{\delta^a} \tfrac{\lambda^a(\delta^a)}{\gamma} \left[1 - \tfrac{-e^{-\gamma(x+s+\delta^a)} e^{-\gamma w(t,s,q-1)}}{-e^{-\gamma x} e^{-\gamma w(t,s,q)}}\right] \\ + \quad & \max_{\delta^b} \tfrac{\lambda^b(\delta^b)}{\gamma} \left[1 - \tfrac{-e^{-\gamma(x-s+\delta^b)} e^{-\gamma w(t,s,q+1)}}{-e^{-\gamma x} e^{-\gamma w(t,s,q)}}\right] \end{aligned} \tag{2.59}$$

Then

$$\begin{aligned} 0 = \quad & \tfrac{\partial w(t,s,q)}{\partial t} + \tfrac{1}{2}\sigma^2 \left[\tfrac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma \left(\tfrac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\ + \quad & \max_{\delta^a} \tfrac{\lambda^a(\delta^a)}{\gamma} \left[1 - e^{-\gamma(s+\delta^a - w(t,s,q) + w(t,s,q-1))}\right] \\ + \quad & \max_{\delta^b} \tfrac{\lambda^b(\delta^b)}{\gamma} \left[1 - e^{\gamma(s-\delta^b - w(t,s,q+1) + w(t,s,q))}\right] \end{aligned} \tag{2.60}$$

Substituting for r^a, r^b from (2.50) we get the required result

$$\begin{aligned} 0 = \quad & \tfrac{\partial w(t,s,q)}{\partial t} + \tfrac{1}{2}\sigma^2 \left[\tfrac{\partial^2 w(t,s,q)}{\partial s^2} - \gamma \left(\tfrac{\partial w(t,s,q)}{\partial s}\right)^2\right] \\ + \quad & \max_{\delta^a} \tfrac{\lambda^a(\delta^a)}{\gamma} \left[1 - e^{-\gamma(s+\delta^a - r^a(t,s,q))}\right] \\ + \quad & \max_{\delta^b} \tfrac{\lambda^b(\delta^b)}{\gamma} \left[1 - e^{\gamma(s-\delta^b - r^b(t,s,q))}\right] \end{aligned} \tag{2.61}$$

□

Proposition 2.16. ***(Optimal controls [AS, Section 3.1]).*** *Given PDE (2.48), optimal controls $\hat{\delta}^a$, $\hat{\delta}^b$ satisfy*

$$r^a - s = \hat{\delta}^a - \frac{1}{\gamma} \ln \left\{1 - \gamma \frac{\lambda^a(\hat{\delta}^a)}{(\partial\lambda^a/\partial\hat{\delta}^a)(\hat{\delta}^a)}\right\}, \; s - r^b = \hat{\delta}^b - \frac{1}{\gamma} \ln \left\{1 - \gamma \frac{\lambda^a(\hat{\delta}^a)}{(\partial\lambda^a/\partial\hat{\delta}^a)(\hat{\delta}^a)}\right\} \tag{2.62}$$

Proof. We suppose that maximum is attained for some $\delta^a = \hat{\delta}^a$ when

$$\nabla \left(\lambda^a(\delta^a) \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right]\right)(\delta^a) = 0 \tag{2.63}$$

Hence, by product and exponential rules, we have

$$\begin{aligned} 0 \;&=\; \tfrac{\partial}{\partial\delta^a} \left\{\tfrac{\lambda^a(\delta^a)}{\gamma} \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right]\right\} \\ &=\; \tfrac{1}{\gamma} \tfrac{\partial}{\partial\delta^a} \left\{\lambda^a(\delta^a) \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right]\right\} \\ &=\; \tfrac{1}{\gamma} \left(\tfrac{\partial}{\partial\delta^a} \left\{\lambda^a(\delta^a)\right\} \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right] + \tfrac{\partial}{\partial\delta^a} \left\{1 - e^{-\gamma(s+\delta^a - r^a)}\right\} \lambda^a(\delta^a)\right) \\ &=\; \tfrac{1}{\gamma} \left(\tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a} \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right] + \left(\tfrac{\partial}{\partial\delta^a}\{1\} - \tfrac{\partial}{\partial\delta^a} \left\{e^{-\gamma(s+\delta^a - r^a)}\right\}\right) \lambda^a(\delta^a)\right) \\ &=\; \tfrac{1}{\gamma} \left(\tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a} \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right] + e^{-\gamma(s+\delta^a - r^a)} \tfrac{\partial}{\partial\delta^a} \left\{\gamma\left(s+\delta^a - r^a\right)\right\} \lambda^a(\delta^a)\right) \\ &=\; \tfrac{1}{\gamma} \left(\tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a} \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right] + e^{-\gamma(s+\delta^a - r^a)} \gamma\lambda^a(\delta^a)\right) \\ &=\; \tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a} \left[1 - e^{-\gamma(s+\delta^a - r^a)}\right] + e^{-\gamma(s+\delta^a - r^a)} \gamma\lambda^a(\delta^a) \\ &=\; 1 - e^{-\gamma(s+\delta^a - r^a)} + e^{-\gamma(s+\delta^a - r^a)} \gamma \tfrac{\lambda^a(\delta^a)}{(\partial\lambda^a/\partial\delta^a)(\delta^a)} \\ &=\; e^{\gamma(s+\delta^a - r^a)} - 1 + \gamma \tfrac{\lambda^a(\delta^a)}{(\partial\lambda^a/\partial\delta^a)(\delta^a)} \end{aligned} \tag{2.64}$$

Therefore, we get a relation for optimal ask distance $\hat{\delta}^a$

$$\begin{aligned}
e^{\gamma\left(s+\hat{\delta}^a-r^a\right)} &= 1-\gamma\tfrac{\lambda^a(\hat{\delta}^a)}{(\partial\lambda^a/\partial\hat{\delta}^a)(\hat{\delta}^a)} \\
\gamma\left(s+\hat{\delta}^a-r^a\right) &= \ln\left(1-\gamma\tfrac{\lambda^a(\hat{\delta}^a)}{(\partial\lambda^a/\partial\hat{\delta}^a)(\hat{\delta}^a)}\right) \\
s+\hat{\delta}^a-r^a &= \tfrac{1}{\gamma}\ln\left(1-\gamma\tfrac{\lambda^a(\hat{\delta}^a)}{(\partial\lambda^a/\partial\hat{\delta}^a)(\hat{\delta}^a)}\right) \\
r^a-s &= \hat{\delta}^a-\tfrac{1}{\gamma}\ln\left(1-\gamma\tfrac{\lambda^a(\hat{\delta}^a)}{(\partial\lambda^a/\partial\hat{\delta}^a)(\hat{\delta}^a)}\right)
\end{aligned} \tag{2.65}$$

By same arguments we can obtain $\hat{\delta}^b$. $\square$

Proposition 2.17. ***(HJB equation with exponential arrival rates [AS, Section 3.2]).*** *Suppose that $\lambda^a(\hat{\delta}^a)=Ae^{-k\hat{\delta}^a}$ and $\lambda^b(\hat{\delta}^b)=Ae^{-k\hat{\delta}^b}$. Then (2.48) can be expressed as*

$$\begin{cases} \frac{\partial w}{\partial t}(t,s,q)+\frac{1}{2}\sigma^2\frac{\partial^2 w}{\partial s^2}(t,s,q)-\frac{1}{2}\sigma^2\gamma\left(\frac{\partial w}{\partial s}(t,s,q)\right)^2+\mathcal{H}=0 \text{ for all } t\in[0,T),\ s,q\in\mathbb{R} \\ w(T,s,q)=qs \text{ for all } s,q\in\mathbb{R} \end{cases} \tag{2.66}$$

where

$$\mathcal{H}=\frac{A}{k+\gamma}\left(2-k\left(\frac{2}{\gamma}\ln\left\{1+\frac{\gamma}{k}\right\}+2w(s,q,t)-w(s,q-1,t)-w(s,q+1,t)\right)\right) \tag{2.67}$$

Proof. First, we take partial derivatives in (2.62)

$$\frac{\partial\lambda^a(\hat{\delta}^a)}{\partial\hat{\delta}^a}=-Ake^{-k\hat{\delta}^a},\ \frac{\partial\lambda^b(\hat{\delta}^b)}{\partial\hat{\delta}^b}=-Ake^{-k\hat{\delta}^b} \tag{2.68}$$

Substituting these into (2.62) we get

$$\begin{aligned}
r^a &= s+\hat{\delta}^a-\tfrac{1}{\gamma}\ln\left\{1-\gamma\tfrac{Ae^{-k\hat{\delta}^a}}{-Ake^{-k\hat{\delta}^a}}\right\}=s+\hat{\delta}^a-\tfrac{1}{\gamma}\ln\left\{1+\tfrac{\gamma}{k}\right\} \\
r^b &= s-\hat{\delta}^b+\tfrac{1}{\gamma}\ln\left\{1-\gamma\tfrac{Ae^{-k\hat{\delta}^b}}{-Ake^{-k\hat{\delta}^b}}\right\}=s-\hat{\delta}^b+\tfrac{1}{\gamma}\ln\left\{1+\tfrac{\gamma}{k}\right\}
\end{aligned} \tag{2.69}$$

Substituting these into (2.49) we get

$$\begin{aligned}
\mathcal{H} &= \tfrac{Ae^{-k\hat{\delta}^a}}{\gamma}\left[1-\exp\left(-\gamma(s+\hat{\delta}^a-s-\hat{\delta}^a+\tfrac{1}{\gamma}\ln\left\{1+\tfrac{\gamma}{k}\right\}\right)\right] \\
&+ \tfrac{Ae^{-k\hat{\delta}^b}}{\gamma}\left[1-\exp\left(\gamma(s-\hat{\delta}^b-s+\hat{\delta}^b-\tfrac{1}{\gamma}\ln\left\{1+\tfrac{\gamma}{k}\right\}\right)\right] \\
&= \tfrac{Ae^{-k\hat{\delta}^a}}{\gamma}\left[1-\exp\left(-\ln\left\{1+\tfrac{\gamma}{k}\right\}\right)\right] \\
&+ \tfrac{Ae^{-k\hat{\delta}^b}}{\gamma}\left[1-\exp\left(-\ln\left\{1+\tfrac{\gamma}{k}\right\}\right)\right] \\
&= \tfrac{A}{k+\gamma}\left(e^{-k\hat{\delta}^a}+e^{-k\hat{\delta}^b}\right)
\end{aligned} \tag{2.70}$$

We then apply asymptotic expansion to $e^{-k\hat{\delta}^a}, e^{-k\hat{\delta}^b}$ in variable k by Definition 1.101 and obtain 1st order Taylor series which linearly approximates $e^{-k\hat{\delta}^a}, e^{-k\hat{\delta}^b}$

$$e^{-k\hat{\delta}^a}\approx\sum_{n=0}^{1}\frac{1}{n!}(-k)^n(e^{-k\hat{\delta}^a})^{(n)}=1-k\hat{\delta}^a \text{ as } q\to 0 \tag{2.71}$$

and

$$e^{-k\hat{\delta}^b}\approx\sum_{n=0}^{1}\frac{1}{n!}(-k)^n(e^{-k\hat{\delta}^b})^{(n)}=1-k\hat{\delta}^b \text{ as } q\to 0 \tag{2.72}$$

Hence, first order linear approximation of (2.70) is

$$\mathcal{H}=\frac{A}{k+\gamma}\left(1-k\hat{\delta}^a+1-k\hat{\delta}^b\right)=\frac{A}{k+\gamma}\left(-k\left(\hat{\delta}^a+\hat{\delta}^b\right)\right) \tag{2.73}$$

Using (2.50) relation we have that

$$\hat{\delta}^a+\hat{\delta}^b=\frac{2}{\gamma}\ln\left\{1+\frac{\gamma}{k}\right\}+2w(s,q,t)-w(s,q-1,t)-w(s,q+1,t) \tag{2.74}$$

Combining this with linear approximation we get

$$\mathcal{H}=\frac{A}{k+\gamma}\left(2-k\left(\frac{2}{\gamma}\ln\left\{1+\frac{\gamma}{k}\right\}+2w(s,q,t)-w(s,q-1,t)-w(s,q+1,t)\right)\right) \tag{2.75}$$

□

Proposition 2.18. ***(Asymptotic expansion and approximation [AS, Section 3.2]).*** *Let w be a function from 2.48. We can approximate it by*

$$w(q,s,t)\approx w^{(0)}(t)+qw^{(1)}(s,t)+\frac{1}{2}q^2w^{(2)}(s,t) \tag{2.76}$$

Proof. First note that terms q,s are separable in w, as $w(q,s,t)=qs$. Hence, we can apply asymptotic expansion of w in variable q from Definition 1.101 and obtain Taylor series which approximates w

$$w(q,s,t)=\sum_{n=0}^{\infty}\frac{1}{n!}q^nw^{(n)}(s,t) \text{ as } q\to 0 \tag{2.77}$$

Suppose we take 3rd order approximation of w, i.e.

$$w(q,s,t)\approx w^{(0)}(s,t)+qw^{(1)}(s,t)+\frac{1}{2}q^2w^{(2)}(s,t) \tag{2.78}$$

By Definition 1.101, $w^{(0)}(s,t)$ is 0-th derivative of w evaluated at $q=0$, so

$$w^{(0)}(s,t)=\left.w(q,s,t)\right|_{q=0}=w(0,s,t) \tag{2.79}$$

Note that s is related to q only via relationship qs. It follows that

$$v(t,s,x,0)=-e^{-\gamma x}e^{-\gamma w^{(0)}(s,t)}=v(x,w^{(0)}(s,t)) \tag{2.80}$$

is independent of s. Hence, we can write $w^{(0)}(s,t)=w^{(0)}(t)$ and required result follows. □

Proposition 2.19. ***(Solution [AS, Section 3.2]).*** *Let all assumptions from Propositions 2.18 and 2.17 hold. Solution to HJB equation (2.66) is*

$$\begin{aligned} w(q,s,t) &= \tfrac{A}{k+\gamma}\left(\left(2-\tfrac{2k}{\gamma}\ln\left\{1+\tfrac{\gamma}{k}\right\}\right)(T-t)-\tfrac{k\gamma\sigma^2}{2}(T-t)^2\right)\\ &+ qs-\tfrac{1}{2}q^2\gamma\sigma^2(T-t) \end{aligned} \tag{2.81}$$

where optimal spread is

$$\hat{\delta}^a + \hat{\delta}^b = \gamma\sigma^2(T-t) + \frac{2}{\gamma} ln\left\{1+\frac{\gamma}{k}\right\} \tag{2.82}$$

Proof. We use similar methods as in [AS, Section 3.2], in the sense that we will use approximation (2.76) to obtain 3 PDEs based on power of q. We solve these successively to obtain 3 solutions that we substitute back into (2.76). Firstly, we substitute approximation (2.76) into PDE (2.66)

$$\begin{array}{rcl} 0 & = & \mathcal{H} \\ & - & \frac{1}{2}\sigma^2\gamma\left(\frac{\partial}{\partial s}\left[w^{(0)}(t) + qw^{(1)}(s,t) + \frac{1}{2}q^2 w^{(2)}(s,t)\right]\right)^2 \\ & + & \frac{1}{2}\sigma^2\frac{\partial^2}{\partial s^2}\left[w^{(0)}(t) + qw^{(1)}(s,t) + \frac{1}{2}q^2 w^{(2)}(s,t)\right] \\ & + & \frac{\partial}{\partial t}\left[w^{(0)}(t) + qw^{(1)}(s,t) + \frac{1}{2}q^2 w^{(2)}(s,t)\right] \end{array} \tag{2.83}$$

Taking partial derivatives

$$\begin{array}{rcl} 0 & = & \mathcal{H} \\ & - & \frac{1}{2}\sigma^2\gamma\left(q\frac{\partial w^{(1)}(s,t)}{\partial s} + \frac{1}{2}q^2\frac{\partial w^{(2)}(s,t)}{\partial s}\right)^2 \\ & + & \frac{1}{2}\sigma^2\left(q\frac{\partial^2 w^{(1)}(s,t)}{\partial s^2} + \frac{1}{2}q^2\frac{\partial^2 w^{(2)}(s,t)}{\partial s^2}\right) \\ & + & \frac{dw^{(0)}(t)}{dt} + q\frac{\partial w^{(1)}(s,t)}{\partial t} + \frac{1}{2}q^2\frac{\partial w^{(2)}(s,t)}{\partial t} \end{array} \tag{2.84}$$

where

$$\begin{array}{rcl} \mathcal{H} & = & \frac{A}{k+\gamma}\left(2 - k\left(\frac{2}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\} + 2w(s,q,t) - w(s,q-1,t) - w(s,q+1,t)\right)\right) \\ \frac{1}{k}\left(2 - \mathcal{H}\left(\frac{A}{k+\gamma}\right)^{-1}\right) & = & \frac{2}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\} + 2w(s,q,t) - w(s,q-1,t) - w(s,q+1,t) \\ & = & \frac{2}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\} + 2w^{(0)}(t) + 2qw^{(1)}(s,t) + q^2 w^{(2)}(s,t) \\ & - & w^{(0)}(t) - qw^{(1)}(s,t) + w^{(1)}(s,t) - \frac{1}{2}q^2 w^{(2)}(s,t) - \frac{1}{2}w^{(2)}(s,t) + qw^{(2)}(s,t) \\ & - & w^{(0)}(t) - qw^{(1)}(s,t) - w^{(1)}(s,t) - \frac{1}{2}q^2 w^{(2)}(s,t) - \frac{1}{2}w^{(2)}(s,t) - qw^{(2)}(s,t) \\ & = & \frac{2}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\} - w^{(2)}(s,t) \\ \mathcal{H} & = & \frac{A}{k+\gamma}\left(2 - \frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\} + kw^{(2)}(s,t)\right) \end{array} \tag{2.85}$$

Hence, we obtain a PDE

$$\left\{\begin{array}{l} \frac{dw^{(0)}(t)}{dt} + q\frac{\partial w^{(1)}(s,t)}{\partial t} + \frac{1}{2}q^2\frac{\partial w^{(2)}(s,t)}{\partial t} + \frac{1}{2}\sigma^2\left(q\frac{\partial^2 w^{(1)}(s,t)}{\partial s^2} + \frac{1}{2}q^2\frac{\partial^2 w^{(2)}(s,t)}{\partial s^2}\right) \\ -\frac{1}{2}\sigma^2\gamma\left(q\frac{\partial w^{(1)}(s,t)}{\partial s} + \frac{1}{2}q^2\frac{\partial w^{(2)}(s,t)}{\partial s}\right)^2 + \frac{A}{k+\gamma}\left(2 - \frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\} + kw^{(2)}(s,t)\right) = 0 \\ w(T,s,q) = w^{(0)}(T) + qw^{(1)}(s,T) + \frac{1}{2}q^2 w^{(2)}(s,T) \end{array}\right. \tag{2.86}$$

Since $w(0,s,T) = w^{(0)}(T) = 0$ by proof of Proposition 2.18, the terminal condition becomes

$$w(q,s,T) = qw^{(1)}(s,T) + \frac{1}{2}q^2 w^{(2)}(s,T) = qs \tag{2.87}$$

will hold if

$$w^{(2)}(s,T)=0,\ w^{(1)}(s,T)=s \tag{2.88}$$

By perturbation methods from Definition 1.104, we split PDE (2.86) on basis of powers of q

$$\begin{array}{ll} q^1 & \begin{cases} \frac{\partial w^{(1)}(s,t)}{\partial t}+\frac{1}{2}\sigma^2\frac{\partial^2 w^{(1)}(s,t)}{\partial s^2}=0 \\ w^{(1)}(s,T)=s \end{cases} \\ q^2 & \begin{cases} \frac{\partial w^{(2)}(s,t)}{\partial t}+\frac{1}{2}\sigma^2\frac{\partial^2 w^{(2)}(s,t)}{\partial s^2}-\sigma^2\gamma\left(\frac{\partial w^{(1)}(s,t)}{\partial s}\right)^2=0 \\ w^{(2)}(s,T)=0 \end{cases} \\ q^0 & \begin{cases} \frac{dw^{(0)}(t)}{dt}+\frac{A}{k+\gamma}\left(2-\frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\}+kw^{(2)}(s,t)\right)=0 \\ w^{(0)}(T)=0 \end{cases} \end{array} \tag{2.89}$$

PDE for q^1 of 2.89, is solved with $w^{(1)}(s,t)=s$, as $\frac{\partial s}{\partial t}=0$ and $\frac{\partial^2 s}{\partial s^2}=0$. By Theorem 1.90 the solution is unique, since $\frac{1}{2}\sigma^2$ is a constant, assumptions are trivially satisfied. Next, we solve PDE for q^2, where we directly substitute $w^{(1)}(s,t)=s$ to obtain

$$\frac{\partial w^{(2)}(s,t)}{\partial t}+\frac{1}{2}\sigma^2\frac{\partial^2 w^{(2)}(s,t)}{\partial s^2}=\sigma^2\gamma \tag{2.90}$$

which is solved with $w^{(2)}(s,t)=-\gamma\sigma^2(T-t)$, as with such $w^{(2)}(s,t)$, $\frac{\partial w^{(2)}(s,t)}{\partial t}=\sigma^2\gamma$ and $\frac{\partial^2 w^{(2)}(s,t)}{\partial s^2}=0$. Since $\sigma^2\gamma$ is a constant, it is trivially bounded and solution is unique by Theorem 1.90. Finally, we solve PDE for q^0, where by direct substitution of $w^{(2)}(s,t)=-\gamma\sigma^2(T-t)$ to obtain an ODE

$$\frac{dw^{(0)}(t)}{dt}=-\frac{A}{k+\gamma}\left(2-\frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\}-k\gamma\sigma^2(T-t)\right) \tag{2.91}$$

which is solved with $w^{(0)}(t)=\frac{A}{k+\gamma}\left(\left(2-\frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\}\right)(T-t)-\frac{k\gamma\sigma^2}{2}(T-t)^2\right)$, as as with such $w^{(0)}(t)$,

$$\begin{aligned} \frac{dw^{(0)}(t)}{dt} &= \frac{A}{k+\gamma}\left(\left(2-\frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\}\right)\frac{d}{dt}\left[(T-t)\right]-\frac{k\gamma\sigma^2}{2}\frac{d}{dt}\left[(T-t)^2\right]\right) \\ &= \frac{A}{k+\gamma}\left(-2+\frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\}+k\gamma\sigma^2(T-t)\right) \end{aligned} \tag{2.92}$$

Combining solutions for (2.89) with approximation (2.76) we obtain

$$\begin{aligned} w(q,s,t) &= w^{(0)}(t)+qw^{(1)}(s,t)+\frac{1}{2}q^2w^{(2)}(s,t) \\ &= \frac{A}{k+\gamma}\left(\left(2-\frac{2k}{\gamma}ln\left\{1+\frac{\gamma}{k}\right\}\right)(T-t)-\frac{k\gamma\sigma^2}{2}(T-t)^2\right) \\ &+ qs-\frac{1}{2}q^2\gamma\sigma^2(T-t) \end{aligned} \tag{2.93}$$

which is solution to (2.66) as required. Using (2.50) and solutions to (2.89) PDEs, we get

$$\begin{aligned}
r^a &= w(s,q,t) - w(s,q-1,t) \\
&= w^{(0)}(t) + qw^{(1)}(s,t) + \tfrac{1}{2}q^2 w^{(2)}(s,t) - w^{(0)}(t) - (q-1)w^{(1)}(s,t) - \tfrac{1}{2}(q-1)^2 w^{(2)}(s,t) \\
&= \tfrac{1}{2}q^2 w^{(2)}(s,t) + w^{(1)}(s,t) - (\tfrac{1}{2}q^2 + \tfrac{1}{2} - q)w^{(2)}(s,t) \\
&= \tfrac{1}{2}q^2 w^{(2)}(s,t) + w^{(1)}(s,t) - \tfrac{1}{2}q^2 w^{(2)}(s,t) - \tfrac{1}{2}w^{(2)}(s,t) + qw^{(2)}(s,t) \\
&= w^{(1)}(s,t) - \tfrac{1}{2}w^{(2)}(s,t) + qw^{(2)}(s,t) \\
&= w^{(1)}(s,t) - (\tfrac{1}{2} - q)w^{(2)}(s,t) \\
&= s + (\tfrac{1}{2} - q)\gamma\sigma^2(T-t)
\end{aligned} \tag{2.94}$$

Recall from (2.69) that optimal spreads $\hat{\delta}^a, \hat{\delta}^b$ are given by relations

$$\hat{\delta}^a = r^a - s + \frac{1}{\gamma} ln \left\{1 + \frac{\gamma}{k}\right\}, \ \hat{\delta}^b = s - r^b + \frac{1}{\gamma} ln \left\{1 + \frac{\gamma}{k}\right\} \tag{2.95}$$

Therefore

$$\begin{aligned}
\hat{\delta}^a &= r^a - s + \tfrac{1}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\} \\
&= s + (\tfrac{1}{2} - q)\gamma\sigma^2(T-t) - s + \tfrac{1}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\} \\
&= (\tfrac{1}{2} - q)\gamma\sigma^2(T-t) + \tfrac{1}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\}
\end{aligned} \tag{2.96}$$

and by the same methods

$$\begin{aligned}
\hat{\delta}^b &= s - r^b + \tfrac{1}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\} \\
&= (q + \tfrac{1}{2})\gamma\sigma^2(T-t) + \tfrac{1}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\}
\end{aligned} \tag{2.97}$$

Combined bid and ask spread is

$$\begin{aligned}
\hat{\delta}^a + \hat{\delta}^b &= \tfrac{1}{2}(1-2q)\gamma\sigma^2(T-t) + \tfrac{1}{2}(2q+1)\gamma\sigma^2(T-t) + \tfrac{2}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\} \\
&= \gamma\sigma^2(T-t) + \tfrac{2}{\gamma} ln \left\{1 + \tfrac{\gamma}{k}\right\}
\end{aligned} \tag{2.98}$$

$\square$

3 Options market making

3.1 Notations

$0 \le t < T < \infty$

$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_s)_{s\in[t,T]}$

$\mathbb{Q}$ is probability measure equivalent to $\mathbb{P}$

$W^{(1)} = (W_s^{(1)})_{s\in[t,T]}$ is an $\mathbb{R}$-valued $(\mathbb{P}, \mathbb{F})$-Brownian motion

$W^{(2)} = (W_s^{(2)})_{s\in[t,T]}$ is an $\mathbb{R}$-valued $(\mathbb{P}, \mathbb{F})$-Brownian motion

$\langle W^{(1)}, W^{(2)} \rangle_t = \rho_R dt$

$\widetilde{W}^{(1)} = (\widetilde{W}_s^{(1)})_{s\in[t,T]}$ is an $\mathbb{R}$-valued $(\mathbb{Q}, \mathbb{F})$-Brownian motion

$\widetilde{W}^{(2)} = (\widetilde{W}_s^{(2)})_{s\in[t,T]}$ is an $\mathbb{R}$-valued $(\mathbb{Q}, \mathbb{F})$-Brownian motion

$\left\langle \widetilde{W}^{(1)}, \widetilde{W}^{(2)} \right\rangle_t = \rho_I dt$

$N^a = (N_s^a)_{s\in[t,T]}$ is an $\mathbb{N}$-valued $(\mathcal{F}_s)$-Poisson processes with intensity λ^a

$N^b = (N_s^b)_{s\in[t,T]}$ is an $\mathbb{N}$-valued $(\mathcal{F}_s)$-Poisson processes with intensity λ^b

N^a, N^b are independent

$\mathcal{S} = [t, T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$

$\bar{\mathcal{S}} = [t, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$

$\partial\mathcal{S} = \{T\} \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$

3.2 Assumptions

Definition 3.1. (Spot price process [ElAA3, Section 2]). Let $S = (S_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process, which is assumed to satisfy following SDEs under:

(1) $\mathbb{P}$ measure:

$$\begin{cases} dS_T &= S_T \mu dT + S_T \sigma(Y_T) dW_T^{(1)} \\ S_t &= s \text{ a.s.} \end{cases} \tag{3.1}$$

(2) $\mathbb{Q}$ measure:

$$\begin{cases} dS_T &= S_T r dT + S_T \sigma(Y_T) d\widetilde{W}_T^{(1)} \\ S_t &= s \text{ a.s.} \end{cases} \tag{3.2}$$

We call it a spot price process and interpret it as a mid price of a stock at time T traded by the market maker. We assume that the stock has infinite liquidity, meaning that mid proce of the stock is same as the best bid and ask.

Definition 3.2. (Variance process [ElAA3, Section 2]). Let $Y = (Y_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process called variance process and assumed to satisfy following SDEs under:

(1) $\mathbb{P}$ measure:

$$\begin{cases} dY_T &= a_R(Y_T)dT + b_R(Y_T)dW_T^{(2)} \\ Y_t &= y \text{ a.s.} \end{cases} \tag{3.3}$$

where function a_R, b_R satisfy sufficient conditions for the existence of unique solutions to (3.3) satisfying

$$\mathbb{E}^{\mathbb{P}}\left(\int_0^T \sigma(Y_t)dt\right) < \infty,\ \mathbb{E}^{\mathbb{P}}\left(\int_0^T (a_R^2(Y_T) + b_R^2(Y_T))dt\right) < \infty \tag{3.4}$$

(2) $\mathbb{Q}$ measure:

$$\begin{cases} dY_T &= a_I(Y_T)dT + b_I(Y_T)d\widetilde{W}_T^{(2)} \\ Y_t &= y \text{ a.s.} \end{cases} \tag{3.5}$$

where function a_I, b_I satisfy sufficient conditions for the existence of unique solutions to 3.5 satisfying

$$\mathbb{E}^{\mathbb{Q}}\left(\int_0^T \sigma(Y_t)dt\right) < \infty,\ \mathbb{E}^{\mathbb{Q}}\left(\int_0^T (a_I^2(Y_T) + b_I^2(Y_T))dt\right) < \infty \tag{3.6}$$

Definition 3.3. (Option price [ElAA3, Section 2]). Let $h : \mathbb{R}^+ \to \mathbb{R}^+$ be a Borel measurable function, which we call payoff function. We define the option mid-price at time t expiring at time T under

(1) $\mathbb{P}$ measure:

$$C_{\mathbb{P}}(t, S_t, Y_t) = \mathbb{E}^{\mathbb{P}}\left(h(S_T)|\,\mathcal{F}_t\right) \text{ for } t \leq T \tag{3.7}$$

(2) $\mathbb{Q}$ measure:

$$C_{\mathbb{Q}}(t, S_t, Y_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left(h(S_T)|\,\mathcal{F}_t\right) \text{ for } t \leq T \tag{3.8}$$

Definition 3.4. (Bid/Ask price [ElAA3, Section 2]). Let $\delta^a, \delta^b > 0$ denote distances (or spreads) of bid and ask price from the option price $C_{\mathbb{Q}}(t, S_t, Y_t)$ and let C_t^b, C_t^a denote bid and ask prices, there are defined by

$$\begin{aligned} C_t^a &= C_{\mathbb{Q}}(t, S_t, Y_t) + \delta^a \text{ for } t \leq T \\ C_t^b &= C_{\mathbb{Q}}(t, S_t, Y_t) - \delta^b \text{ for } t \leq T \end{aligned} \tag{3.9}$$

We interpret bid (ask) price as the price at which a limit order is sent by the market maker to sell (buy) the option.

Definition 3.5. (Option inventory process [ElAA3, Section 2]). Let $Q_1 = (Q_{1,s})_{s\in[t,T]}$ be an $\mathbb{N}$-valued piecewise constant and càdlàg process assumed to satisfy following SDE

$$\begin{cases} dQ_{1,T} &= dN_T^b - dN_T^a \\ Q_{1,t} &= q_1 \text{ a.s.} \end{cases} \tag{3.10}$$

We call it an option inventory process and interpret it as an amount of options held in inventory of the market maker at time T.

Definition 3.6. (Stock inventory process [ElAA3, Section 2]). Let $Q_2 = (Q_{2,s})_{s\in[t,T]}$ be an $\mathbb{N}$-valued piecewise constant and càdlàg process assumed to satisfy following SDE

$$\begin{cases} dQ_{2,T} &= -\Delta_T dQ_{1,T} - Q_{1,T} d\Delta_T - d\left\langle Q_1, \Delta\right\rangle_T \\ &= -\Delta_T dN_T^b + \Delta_T dN_T^a - Q_{1,T} d\Delta_T \\ Q_{2,t} &= q_2 \quad \text{a.s.} \end{cases} \tag{3.11}$$

We call it an stock inventory process and interpret it as an amount of stock held in inventory of the market maker at time T. Note that at time $s \in [t, T]$, process $Q_{2,s}$ depends on process $Q_{1,s}$ via Δ_s, this corresponds to delta-hedging.

Definition 3.7. (Delta hedging [ElAA3, Section 2]). Let $\Delta_t = \Delta(t, S_t, Y_t)$ denote delta of the option, defined as

$$\Delta_t = \frac{\partial C_{\mathbb{Q}}(t, S_t, Y_t)}{\partial S_t} \text{ for } t \leq T \tag{3.12}$$

We interpret this as sensitivity of change in price of the option with respect to change in price of the stock. This will be used for delta-hedging purpuses, whereby market maker will trade in stock to ensure that following identity holds

$$Q_{1,t}\Delta_t + Q_{2,t} = 0 \text{ for } t \leq T \tag{3.13}$$

With the assumption of infinite liquidity in stock, delta hedging is free, i.e. market maker will not incur costs for crossing the spread to buy (or sell) stock whenever $Q_{1,t}\Delta_t > Q_{2,t}$ (or $Q_{1,t}\Delta_t < Q_{2,t}$).

Definition 3.8. (Cash process [ElAA3, Section 2]). Let $X = (X_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process assumed to satisfy following SDE

$$\begin{cases} dX_T = (C_{\mathbb{Q}}(T, S_T, Y_T) + \delta^a)dN_T^a - (C_{\mathbb{Q}}(T, S_T, Y_T) - \delta^b)dN_T^b + Q_{2,T}dS_T \\ X_t = x \quad \text{a.s.} \end{cases} \tag{3.14}$$

We call it cash process and interpret it as the market value of combined (stock and option) inventory held by market maker at time T.

Definition 3.9. (Order arrival intensities [ElAA3, Section 3]). Intensities λ^a, λ^b of Poisson process N^a, N^b correspond to the probability of execution of the limit order posted by the market maker to buy (or sell) the option at C_t^a (or C_t^b) price. These are defined by

$$\lambda^a(\delta^a) = A\left(B + (\delta^a)^{1/\beta}\right)^{-\gamma}, \; \lambda^b(\delta^b) = A\left(B + (\delta^b)^{1/\beta}\right)^{-\gamma} \text{ for all } \delta^a, \delta^b \geq 0 \tag{3.15}$$

where $A, B, \beta > 0$ and $\gamma > 1$ is parameter characterising the market impact function, i.e. in the case where $\beta = \frac{1}{2}$ we mean square-root market impact, whereas if $\beta = 1$ we mean linear market impact.

Definition 3.10. (Utility function [ElAA3, Section 4]). Let h be a payoff function from Definiton 3.3. Let g be a function taking values in $\mathbb{R}$ defined by

$$g(s, y, q_1, x) = x + q_1 h(s) \text{ for } (s, y, q_1, x) \in \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R} \tag{3.16}$$

We call it utility function.

Definition 3.11. (Value function for risk-neutral market maker [ElAA3, Section 4]). Let g be a utility function from Definition 3.10. Let $U = \mathbb{R}^+ \times \mathbb{R}^+$ be the set of values for controls (δ^a, δ^b). The stochastic control problem for risk-neutral market maker is to maximise expectation of utility of terminal wealth over controls (δ^a, δ^b). The value function for this problem is defined by

$$v(t, s, y, q_1, x) = \sup_{(\delta^a, \delta^b) \in U} \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x} \left(g(S_T, Y_T, Q^{(\delta^a,\delta^b)}_{1,T}, X^{(\delta^a,\delta^b)}_T) \right) \tag{3.17}$$

Definition 3.12. (Value function for risk-averse market maker [ElAA3, Section 5]). Let g be a utility function from Definition 3.10. Let $U = \mathbb{R}^+ \times \mathbb{R}^+$ be the set of values for controls (δ^a, δ^b). The stochastic control problem for risk-averse market maker is to maximise expectation of utility of terminal wealth penalised by variance of utility of terminal wealth over controls (δ^a, δ^b). The value function for this problem is defined by

$$\begin{aligned} v(t, s, y, q_1, x) \quad &= \quad \sup_{(\delta^a, \delta^b) \in U} \Big[\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x} \left(g(S_T, Y_T, Q^{(\delta^a,\delta^b)}_{1,T}, X^{(\delta^a,\delta^b)}_T) \right) \\ &- \quad Var^{\mathbb{P}}_{t,s,y,q_1,x} \left(g(S_T, Y_T, Q^{(\delta^a,\delta^b)}_{1,T}, X^{(\delta^a,\delta^b)}_T) \right) \Big] \end{aligned} \tag{3.18}$$

3.3 State process

Proposition 3.13. ***(Controlled wealth process [ElAA3, Sections 2,4]).*** *Refer to Subsection 3.1 for notations and Subsection 3.2 for assumptions. Consider the following SDE*

$$\begin{cases} dZ_\vartheta(T) &= a(Z_\vartheta(T))dT + b(Z_\vartheta(T))dW_T + G(T, Z_\vartheta(T), \vartheta)dN_T \\ Z_\vartheta(t) &= z \end{cases} \tag{3.19}$$

With initial condition

$$\mathbb{P}\left(Z_\vartheta(t) = z\right) = 1 \text{ for } \mathcal{F}_t\text{-measurable } (\mathbb{R}\times\mathbb{R}\times\mathbb{N}\times\mathbb{R})\text{-valued random variable } z \tag{3.20}$$

where

$$Z_\vartheta(T) = \begin{bmatrix} S_T \\ Y_T \\ Q_{1,T} \\ X_T \end{bmatrix}, \; z = \begin{bmatrix} s \\ y \\ q_1 \\ x \end{bmatrix}, \; a(Z_\vartheta(T)) = \begin{bmatrix} S_T\mu \\ a_R(Y_T) \\ 0 \\ Q_{2,T}S_T\mu \end{bmatrix}, \; b(Z_\vartheta(T)) = \begin{bmatrix} S_T\sigma(Y_T) & 0 \\ 0 & b_R(Y_T) \\ 0 & 0 \\ Q_{2,T}S_T\sigma(Y_T) & 0 \end{bmatrix} \tag{3.21}$$

$$G(T, Z_\vartheta(T), \vartheta) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 1 \\ C_{\mathbb{Q}}(T, S_T, Y_T) + \delta^a & -C_{\mathbb{Q}}(T, S_T, Y_T) + \delta^b \end{bmatrix} \tag{3.22}$$

$$W_T = \begin{bmatrix} W_T^{(1)} \\ W_T^{(2)} \end{bmatrix}, N_T = \begin{bmatrix} N_T^a \\ N_T^b \end{bmatrix}, \; \lambda(\vartheta) = \begin{bmatrix} \lambda^a(\delta^a) \\ \lambda^b(\delta^b) \end{bmatrix}, \; \vartheta = \begin{bmatrix} \delta^a \\ \delta^b \end{bmatrix}, \; \tilde{N}_T = N_T - \lambda(\vartheta) \tag{3.23}$$

Suppose that control process ϑ is constant, with $\vartheta \in U \subset \mathbb{R}^2$ where set U is bounded. Then 2.9 has a unique solution $Z_\vartheta^{t,z} = (Z_\vartheta^{t,z}(s)), s \in [t, T]$ with initial condition (2.10) satisfying the following stochastic integral equation almost surely

$$Z_\vartheta^{t,z}(T) = z + \int_t^T a(Z_\vartheta(s))ds + \int_t^T b(Z_\vartheta(s))dW_s + \int_t^T G(s, Z_\vartheta(s), \vartheta)dN_s \tag{3.24}$$

Proof. By same arguments as in proof of Proposition 2.10 we treat Poisson processes N^a, N^b as the special case of disjoint Poisson random measures. In [ElAA3, Sections 4] set $U = \mathbb{R}^+ \times \mathbb{R}^+$ is defined as set of admissible values for controls ϑ, which we will interpret as ϑ being bounded as in Proposition 1.65. Since ϑ is a constant, it is trivially $(\mathcal{F}_s)$-measurable and so is $(\mathcal{F}_s)$-progressively measurable. Hence, ϑ is admissible control from Definition 1.64. With the assumption that intensities $\lambda^a(\delta^a), \lambda^b(\delta^b)$ are decreasing functions of controls δ^a, δ^b in classical sense, together with the assumption that set U is bounded, allows us to imply that $\lambda^a(\delta^a), \lambda^b(\delta^b)$ are also bounded and so

$$\lambda^a(\delta^a) < \infty,\ \lambda^b(\delta^b) < \infty \text{ for any } (\delta^a, \delta^b) \in U \tag{3.25}$$

Now we check if assumptions on coefficients of 3.19 from Theorem 1.68 are satisfied. Note that from Definition 3.2 we assume that functions a_R, b_R satisfy sufficient conditions for existence and uniqueness of solution to SDE (3.19). Moreover, from Definition 3.2, it is assumed that that

$$\mathbb{E}^{\mathbb{P}}\left(\int_0^T \sigma(Y_t)dt\right) < \infty,\ \mathbb{E}^{\mathbb{P}}\left(\int_0^T (a_R^2(Y_T) + b_R^2(Y_T))dt\right) < \infty \tag{3.26}$$

Recall from Theorem 1.68 that Lipshitz condition is satisfied if there exists $C_2 > 0$ for $z_1, z_2 \in \mathbb{R}^4$ and $\vartheta \in U$

$$\begin{aligned} |f(z_1,\vartheta) - f(z_2,\vartheta)|^2 &\leq C_2 |z_1 - z_2|^2 \\ |f(z_1,\vartheta) - f(z_2,\vartheta)| &\leq \sqrt{C_2}\, |z_1 - z_2| \\ \tfrac{|f(z_1,\vartheta) - f(z_2,\vartheta)|}{|z_1 - z_2|} &\leq \sqrt{C_2} \\ \left|\tfrac{f(z_1,\vartheta) - f(z_2,\vartheta)}{z_1 - z_2}\right| &\leq \sqrt{C_2} \end{aligned} \tag{3.27}$$

By mean value Theorem 5.4, there exists $z \in [z_1, z_2]$ such that

$$\left|\frac{\partial f(z,\vartheta)}{\partial z}\right| \leq \sqrt{C} \tag{3.28}$$

We can show that if Lipshitz condition is satisfied, then so is the linear growth. Set $z_2 = 0, z_1 = z$, then by triangle inequality we have

$$\begin{aligned} |f(z,\vartheta) - f(0,\vartheta)| &\leq \sqrt{C_2}\, |z| \\ |f(z,\vartheta)| &\leq |f(0,\vartheta)| + \sqrt{C_2}\, |z| \\ |f(z,\vartheta)|^2 &\leq |f(0,\vartheta)|^2 + C_2 |z|^2 \end{aligned} \tag{3.29}$$

Assuming that $f(0,\vartheta)$ is a constant, then $|f(0,\vartheta)|^2 = C_2 = C_1 > 0$, so we have

$$|f(z,\vartheta)|^2 \leq C_1\left(1 + |z|^2\right) \tag{3.30}$$

Which is exactly linear growth condition from Theorem 1.68. Therefore, it is sufficient for us to verify (3.28) for existence and uniqueness of solution to SDE (3.19).

$$\frac{\partial}{\partial s}[s\mu] = \mu,\ \frac{\partial}{\partial q_2}[q_2 s\mu] = s\mu,\ \frac{\partial}{\partial s}[s\sigma(y)] = \sigma(y),\ \frac{\partial}{\partial q_2}[q_2 s\sigma(y)] = s\sigma(y) \tag{3.31}$$

By boundedness assumptions on σ, a_R, b_R we deduce that functions a, b from (3.19) satisfy conditions for existence and uniqueness of solution to SDE (3.19). Note that function G depends on option price $C_{\mathbb{Q}}$ and controls δ^a, δ^b. Recall from Definition 3.3 that

$$C_{\mathbb{Q}}(t, S_t, Y_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left(h(S_T)\right| \mathcal{F}_t) \text{ for } t \leq T \tag{3.32}$$

where $h > 0$ and is continious. We argue that for each fixed $t \leq T$, there exist $K_1, K_2 > 0$ such that

$$\left|\frac{\partial C_{\mathbb{Q}}(t,s,y)}{\partial s}\right| = K_1, \ \left|\frac{\partial C_{\mathbb{Q}}(t,s,y)}{\partial y}\right| = K_2 \tag{3.33}$$

Together with boundedness of $\lambda^a(\delta^a), \lambda^b(\delta^b)$, we deduce that assumptions on function G of 3.19 from Theorem 1.68 are satisfied. This means that $Z_{\vartheta}^{t,z}$ satisfying 3.19 is the unique solution. □

Proposition 3.14. ***(Order arrival intensity [ElAA3, Section 3, Subsection 8.1]).*** *We use assumptions from Definition 3.9. Let random variable Q be a size of market order with the distribution f^Q . Let Δp be a market impact, let $\delta \in \{\delta^a, \delta^b\}$ and $\lambda \in \{\lambda^a, \lambda^b\}$. Let $\mathbb{P}(\Delta p \geq \delta)$ be the probability of execution of market makers limit order. Suppose that the distribution f^Q obeys a Power-law, i.e.*

$$f^Q(x) = \frac{\gamma L^\gamma}{(L+x)^{\gamma+1}} \ \textit{for } \gamma > 0 \tag{3.34}$$

And suppose that market impact Δp is given by

$$\Delta p = KQ^\beta \tag{3.35}$$

Let Λ be a constant and denote the arrival rate (frequency) of market orders Q. Then intensity λ defined by

$$\lambda(\delta) = \Lambda \mathbb{P}(\Delta p \geq \delta) \ \textit{ for } \delta \geq 0 \tag{3.36}$$

is of the form

$$\lambda(\delta) = \frac{A}{(B + \delta^{\frac{1}{\beta}})^\gamma} \tag{3.37}$$

where $A, B, \beta > 0$ and $\gamma > 1$.

Proof. Since $\Delta p = KQ^\beta$ and Q has density f^Q we have that

$$\begin{aligned}
\lambda(\delta) &= \Lambda \mathbb{P}(\Delta p \geq \delta) \\
&= \Lambda \mathbb{P}\left(Q^\beta \geq \tfrac{\delta}{K}\right) \\
&= \Lambda \mathbb{P}\left(Q \geq (\tfrac{\delta}{K})^{1/\beta}\right) \\
&= \Lambda \textstyle\int_{(\frac{\delta}{K})^{1/\beta}}^{\infty} f^Q(x)dx \\
&= \Lambda \textstyle\int_{(\frac{\delta}{K})^{1/\beta}}^{\infty} \frac{\gamma L^\gamma}{(L+x)^{\gamma+1}} dx \\
&= \Lambda \tfrac{L^\gamma}{(L+(\frac{\delta}{K})^{1/\beta})^\gamma} \\
&= \tfrac{\Lambda K^{\gamma/\beta} L^\gamma}{(LK^{1/\beta}+\delta^{1/\beta})^\gamma} \\
&= \tfrac{A}{(B+\delta^{1/\beta})^\gamma}
\end{aligned} \tag{3.38}$$

where $A = \Lambda K^{\gamma/\beta} L^\gamma$ and $B = LK^{1/\beta}$. □

Proposition 3.15. *Function $\lambda(\delta)$ as defined in (3.37) is decreasing function of δ in classical sense.*

Proof. Taking derivative of $\lambda(\delta)$

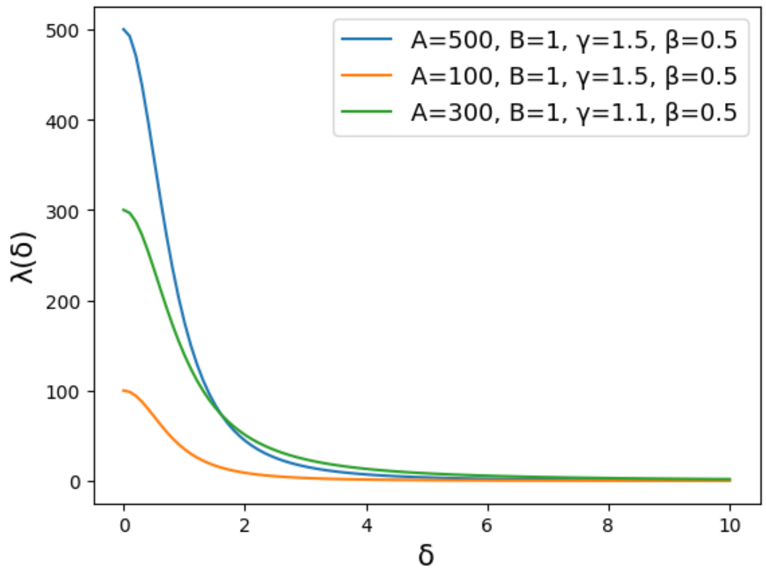
$$\frac{d\lambda(\delta)}{d\delta} = \frac{d}{d\delta}\left[\frac{A}{(B+\delta^{1/\beta})^{\gamma}}\right] = -\frac{A\gamma\delta^{\frac{1}{\beta}-1}\left(\delta^{\frac{1}{\beta}}+B\right)^{-\gamma-1}}{\beta} \tag{3.39}$$

With $A, B, \beta > 0$ and $\gamma > 1$ we can deduce that

$$-\frac{A\gamma\delta^{\frac{1}{\beta}-1}\left(\delta^{\frac{1}{\beta}}+B\right)^{-\gamma-1}}{\beta} < 0 \text{ for } \delta > 0 \tag{3.40}$$

Therefore $\lambda(\delta)$ is decreasing function of δ in classical sense.

Figure 3.1: Order arrival intensity $\lambda(\delta)$ for given spread δ



□

3.4 HJB equation for risk-neutral market maker

Proposition 3.16. ***(Deriving HJB equation [ElAA3, Section 4]).*** *Refer to Subsection 3.1 for notations and Subsection 3.2 for assumptions. Let $Z^{t,z}_{\vartheta} = (Z^{t,z}_{\vartheta}(s)), s \in [t, T]$ be a unique solution to 3.19 as defined in Proposition 3.13. Suppose we have a function $w \in C^2(\mathcal{S}) \cup C^0(\bar{\mathcal{S}})$. Then w satisfies following PDE*

$$\begin{cases} \frac{\partial w}{\partial t} + \mathcal{L}_1 w + \mathcal{L}_2 w + \mathcal{H} = 0 \text{ for all } (t, s, y, q_1, x) \in \mathcal{S} \\ w(t, s, y, q_1, x) = x + q_1 h(s) \text{ for all } (t, s, y, q_1, x) \in \partial\mathcal{S} \end{cases} \tag{3.41}$$

where

$$\begin{aligned} \mathcal{L}_1 w &= \mu s \frac{\partial w}{\partial s} + a_R(y)\frac{\partial w}{\partial y} + \frac{1}{2}\sigma^2(y)s^2\frac{\partial^2 w}{\partial s^2} + \frac{1}{2}b_R^2(y)\frac{\partial^2 w}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\frac{\partial^2 w}{\partial s \partial y} \\ \mathcal{L}_2 w &= q_2 \mu s \frac{\partial w}{\partial x} + \frac{1}{2}q_2^2\sigma^2(y)s^2\frac{\partial^2 w}{\partial x^2} + q_2\sigma^2(y)s^2\frac{\partial^2 w}{\partial x \partial s} + \rho_R b_R(y) q_2 \sigma(y) s \frac{\partial^2 w}{\partial x \partial y} \\ \mathcal{H} &= sup_{\vartheta \in U}\mathcal{J}(t, s, y, \vartheta) \\ \mathcal{J}(t, s, y, \vartheta) &= \lambda^a(\delta^a)\left(w(t, s, y, q_1 - 1, x + C_{\mathbb{Q}}(t, s, y) + \delta^a) - w(t, s, y, q_1, x)\right) \\ &+ \lambda^b(\delta^b)\left(w(t, s, y, q_1 + 1, x - C_{\mathbb{Q}}(t, s, y) + \delta^b) - w(t, s, y, q_1, x)\right) \end{aligned} \tag{3.42}$$

Proof. Generalising results obtained in [Ph, Subsection 3.4.1, p43]. Suppose that w satisfies

$$w(t, z) \geq \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w(t + h, Z^{\vartheta}(t + h))\right) \tag{3.43}$$

Applying the Ito Lemma from Theorem 1.61 to function w, over time interval $t, T = t + \tau$

$$\begin{aligned} w(t + \tau, Z^{t,z}_{\vartheta}(t + \tau)) &= w(t, z) \\ &+ \int_t^{t+\tau}\left[\frac{\partial w}{\partial u} + \mu S_u\frac{\partial w}{\partial S} + a_R(Y_u)\frac{\partial w}{\partial Y} + Q_{2,u}\mu S_u\frac{\partial w}{\partial X}\right] du \\ &+ \int_t^{t+\tau}\frac{1}{2}\left[\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial S^2} + b_R^2(Y_u)\frac{\partial^2 w}{\partial Y^2} + 2\rho_R b_R(Y_u)\sigma(Y_u)S_u\frac{\partial^2 w}{\partial S \partial Y}\right] du \\ &+ \int_t^{t+\tau}\frac{1}{2}\left[Q_{2,u}^2\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial X^2} + 2Q_{2,u}\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial X \partial S} + 2\rho_R b_R(Y_u)Q_{2,u}\sigma(Y_u)S_u\frac{\partial^2 w}{\partial X \partial Y}\right] du \\ &+ \int_t^{t+\tau}\left[\sigma(Y_u)S_u\frac{\partial w}{\partial S} + Q_{2,u}\sigma(Y_u)S_u\frac{\partial w}{\partial X}\right] dW_u^{(1)} + \int_t^{t+\tau} b_R(Y_u)\frac{\partial w}{\partial Y} dW_u^{(2)} \\ &+ \int_t^{t+\tau}\left[w(u, S_u, Y_u, Q_{1,u} - 1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] dN_u^a \\ &+ \int_t^{t+\tau}\left[w(u, S_u, Y_u, Q_{1,u} + 1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] dN_u^b \end{aligned} \tag{3.44}$$

By Definition 1.32 we get

$$
\begin{aligned}
w(t+\tau, Z_{\vartheta}^{t,z}(t+\tau)) &= w(t,z) \\
&+ \int_t^{t+\tau} \left[\tfrac{\partial w}{\partial u} + \mu S_u \tfrac{\partial w}{\partial S} + a_R(Y_u)\tfrac{\partial w}{\partial Y} + Q_{2,u}\mu S_u \tfrac{\partial w}{\partial X}\right] du \\
&+ \int_t^{t+\tau} \tfrac{1}{2}\left[\sigma^2(Y_u) S_u^2 \tfrac{\partial^2 w}{\partial S^2} + b_R^2(Y_u)\tfrac{\partial^2 w}{\partial Y^2} + 2\rho_R b_R(Y_u)\sigma(Y_u) S_u \tfrac{\partial^2 w}{\partial S \partial Y}\right] du \\
&+ \int_t^{t+\tau} \tfrac{1}{2}\left[Q_{2,u}^2 \sigma^2(Y_u) S_u^2 \tfrac{\partial^2 w}{\partial X^2} + 2Q_{2,u}\sigma^2(Y_u) S_u^2 \tfrac{\partial^2 w}{\partial X \partial S} + 2\rho_R b_R(Y_u) Q_{2,u}\sigma(Y_u) S_u \tfrac{\partial^2 w}{\partial X \partial Y}\right] du \\
&+ \int_t^{t+\tau} \left[\sigma(Y_u) S_u \tfrac{\partial w}{\partial S} + Q_{2,u}\sigma(Y_u) S_u \tfrac{\partial w}{\partial X}\right] dW_u^{(1)} + \int_t^{t+\tau} b_R(Y_u)\tfrac{\partial w}{\partial Y} dW_u^{(2)} \\
&+ \int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}-1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] \lambda^a(\delta^a) du \\
&+ \int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}+1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] \lambda^b(\delta^b) du \\
&+ \int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}-1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] d\tilde{N}_u^a \\
&+ \int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}+1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] d\tilde{N}_u^b
\end{aligned}
\tag{3.45}
$$

Recall the notation from Definition 1.69 where

$$
\mathbb{E}\left(Z_{\vartheta}^{t,z}(T)\right) = \mathbb{E}_{t,z}\left(Z^{\vartheta}(T)\right) = \mathbb{E}\left(Z^{\vartheta}(T) \middle| Z^{\vartheta}(t) = z\right) \tag{3.46}
$$

Next, we substitute the above into (3.43), then by linearity of expectation we have

$$
\begin{aligned}
w(t,z) \geq\ & w(t,z) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[\tfrac{\partial w}{\partial u} + \mu S_u \tfrac{\partial w}{\partial S} + a_R(Y_u)\tfrac{\partial w}{\partial Y} + Q_{2,u}\mu S_u \tfrac{\partial w}{\partial X}\right] du\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[\tfrac{1}{2}\sigma^2(Y_u) S_u^2 \tfrac{\partial^2 w}{\partial S^2} + \tfrac{1}{2} b_R^2(Y_u)\tfrac{\partial^2 w}{\partial Y^2} + \rho_R b_R(Y_u)\sigma(Y_u) S_u \tfrac{\partial^2 w}{\partial S \partial Y}\right] du\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[\tfrac{1}{2} Q_{2,u}^2 \sigma^2(Y_u) S_u^2 \tfrac{\partial^2 w}{\partial X^2} + Q_{2,u}\sigma^2(Y_u) S_u^2 \tfrac{\partial^2 w}{\partial X \partial S} + \rho_R b_R(Y_u) Q_{2,u}\sigma(Y_u) S_u \tfrac{\partial^2 w}{\partial X \partial Y}\right] du\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[\sigma(Y_u) S_u \tfrac{\partial w}{\partial S} + Q_{2,u}\sigma(Y_u) S_u \tfrac{\partial w}{\partial X}\right] dW_u^{(1)}\right) + \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} b_R(Y_u)\tfrac{\partial w}{\partial Y} dW_u^{(2)}\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}-1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] \lambda^a(\delta^a) du\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}+1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] \lambda^b(\delta^b) du\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}-1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] d\tilde{N}_u^a\right) \\
+\ & \mathbb{E}_{t,s,y,q_1,x}^{\mathbb{P}}\left(\int_t^{t+\tau} \left[w(u, S_u, Y_u, Q_{1,u}+1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] d\tilde{N}_u^b\right)
\end{aligned}
\tag{3.47}
$$

Brownian motions $W^{(1)}, W^{(2)}$ are martingales by Definition 1.39 and Compensated Poisson processes

$\tilde{N}^a, \tilde{N}^b$ are also martingales by Proposition 1.5 so we get

$$
\begin{aligned}
w(t,z) &\geq w(t,z) \\
&+ \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\int_t^{t+\tau}\left[\tfrac{\partial w}{\partial u} + \mu S_u \tfrac{\partial w}{\partial S} + a_R(Y_u)\tfrac{\partial w}{\partial Y} + Q_{2,u}\mu S_u \tfrac{\partial w}{\partial X}\right] du\right) \\
&+ \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\int_t^{t+\tau}\left[\tfrac{1}{2}\sigma^2(Y_u)S_u^2\tfrac{\partial^2 w}{\partial S^2} + \tfrac{1}{2}b_R^2(Y_u)\tfrac{\partial^2 w}{\partial Y^2} + \rho_R b_R(Y_u)\sigma(Y_u)S_u\tfrac{\partial^2 w}{\partial S\partial Y}\right] du\right) \\
&+ \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\int_t^{t+\tau}\left[\tfrac{1}{2}Q_{2,u}^2\sigma^2(Y_u)S_u^2\tfrac{\partial^2 w}{\partial X^2} + Q_{2,u}\sigma^2(Y_u)S_u^2\tfrac{\partial^2 w}{\partial X\partial S} + \rho_R b_R(Y_u)Q_{2,u}\sigma(Y_u)S_u\tfrac{\partial^2 w}{\partial X\partial Y}\right] du\right) \\
&+ \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\int_t^{t+\tau}\left[w(u,S_u,Y_u,Q_{1,u}-1,X_u+C_{\mathbb{Q}}(u,S_u,Y_u)+\delta^a) - w(u,S_u,Y_u,Q_{1,u},X_u)\right]\lambda^a(\delta^a)du\right) \\
&+ \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\int_t^{t+\tau}\left[w(u,S_u,Y_u,Q_{1,u}+1,X_u-C_{\mathbb{Q}}(u,S_u,Y_u)+\delta^b) - w(u,S_u,Y_u,Q_{1,u},X_u)\right]\lambda^b(\delta^b)du\right)
\end{aligned}
\tag{3.48}
$$

by mean value Theorem 5.4

$$
\begin{aligned}
w(t,z) &\geq w(t,z) \\
&+ \tau\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\tfrac{\partial w}{\partial t} + \mu S_t \tfrac{\partial w}{\partial S} + a_R(Y_t)\tfrac{\partial w}{\partial Y} + Q_{2,t}\mu S_t\tfrac{\partial w}{\partial X}\right) \\
&+ \tau\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\tfrac{1}{2}\sigma^2(Y_t)S_t^2\tfrac{\partial^2 w}{\partial S^2} + \tfrac{1}{2}b_R^2(Y_t)\tfrac{\partial^2 w}{\partial Y^2} + \rho_R b_R(Y_t)\sigma(Y_t)S_t\tfrac{\partial^2 w}{\partial S\partial Y}\right) \\
&+ \tau\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\tfrac{1}{2}Q_{2,t}^2\sigma^2(Y_t)S_t^2\tfrac{\partial^2 w}{\partial X^2} + Q_{2,t}\sigma^2(Y_t)S_t^2\tfrac{\partial^2 w}{\partial X\partial S} + \rho_R b_R(Y_t)Q_{2,t}\sigma(Y_t)S_t\tfrac{\partial^2 w}{\partial X\partial Y}\right) \\
&+ \lambda^a(\delta^a)\tau\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w(t,S_t,Y_t,Q_{1,t}-1,X_t+C_{\mathbb{Q}}(t,S_t,Y_t)+\delta^a) - w(t,S_t,Y_t,Q_{1,t},X_t)\right) \\
&+ \lambda^b(\delta^b)\tau\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w(t,S_t,Y_t,Q_{1,t}+1,X_t-C_{\mathbb{Q}}(t,S_t,Y_t)+\delta^b) - w(t,S_t,Y_t,Q_{1,t},X_t)\right)
\end{aligned}
\tag{3.49}
$$

The initial condition (3.20) implies $\mathbb{P}\left(\mathbb{E}_{t,z}\left(Z^\vartheta(t)\right) = z\right) = 1$, hence

$$
\begin{aligned}
w(t,z) &\geq w(t,z) \\
&+ \tau\left(\tfrac{\partial w}{\partial t} + \mu s\tfrac{\partial w}{\partial s} + a_R(y)\tfrac{\partial w}{\partial y} + q_2\mu s\tfrac{\partial w}{\partial x}\right) \\
&+ \tau\left(\tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial s^2} + \tfrac{1}{2}b_R^2(y)\tfrac{\partial^2 w}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\tfrac{\partial^2 w}{\partial s\partial y}\right) \\
&+ \tau\left(\tfrac{1}{2}q_2^2\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial x^2} + q_2\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial x\partial s} + \rho_R b_R(y)q_2\sigma(y)s\tfrac{\partial^2 w}{\partial x\partial y}\right) \\
&+ \lambda^a(\delta^a)\tau\left(w(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w(t,s,y,q_1,x)\right) \\
&+ \lambda^b(\delta^b)\tau\left(w(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w(t,s,y,q_1,x)\right)
\end{aligned}
\tag{3.50}
$$

Substracting $w(t,z)$ from both sides, dividing by τ and taking limit as $\tau \to 0$

$$
\begin{aligned}
0 &\geq \tfrac{\partial w}{\partial t} + \mu s\tfrac{\partial w}{\partial s} + a_R(y)\tfrac{\partial w}{\partial y} + q_2\mu s\tfrac{\partial w}{\partial x} \\
&+ \tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial s^2} + \tfrac{1}{2}b_R^2(y)\tfrac{\partial^2 w}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\tfrac{\partial^2 w}{\partial s\partial y} \\
&+ \tfrac{1}{2}q_2^2\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial x^2} + q_2\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial x\partial s} + \rho_R b_R(y)q_2\sigma(y)s\tfrac{\partial^2 w}{\partial x\partial y} \\
&+ \lambda^a(\delta^a)\left(w(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w(t,s,y,q_1,x)\right) \\
&+ \lambda^b(\delta^b)\left(w(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w(t,s,y,q_1,x)\right)
\end{aligned}
\tag{3.51}
$$

Using definitions from (3.42), we rewrite previous equation in a compact form

$$
0 \geq \tfrac{\partial w}{\partial t} + \mathcal{L}_1 w + \mathcal{L}_2 w + \mathcal{J}(t,s,y,\vartheta) \tag{3.52}
$$

Now suppose that there exists $\hat{\vartheta}$ such that

$$w(t,z) = \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w(t+\tau, Z^{\hat{\vartheta}}(t+\tau))\right) \tag{3.53}$$

Then by same arguments as above we obtain equality

$$0 = \frac{\partial w}{\partial t} + \mathcal{L}_1 w + \mathcal{L}_2 w + \mathcal{J}(t,s,y,\hat{\vartheta}) \tag{3.54}$$

It follows by Definitions 1.70 and 1.71 that we have

$$\mathcal{J}(t,s,y,\hat{\vartheta}) = \sup_{\vartheta \in U} \mathcal{J}(t,s,y,\vartheta) \geq \mathcal{J}(t,s,y,\vartheta) \tag{3.55}$$

This suggests that w should satisfy

$$\frac{\partial w}{\partial t} + \mathcal{L}_1 w + \mathcal{L}_2 w + \sup_{\vartheta \in U} \mathcal{J}(t,s,y,\vartheta) = 0 \tag{3.56}$$

With regular terminal condition

$$w(T,s,y,q_1,x) = g(s,y,q_1,x) = x + q_1 h(s) \tag{3.57}$$

□

Proposition 3.17. ***(General optimal controls).*** *Let all assumptions from Proposition 3.16 hold. Given the following ansatz to PDE 3.41*

$$w(t,s,y,q_1,x) = x + \theta_0(t,s,y) + q_1\theta_1(t,s,y) \tag{3.58}$$

optimal controls $\hat{\delta}^a_L = \delta^a, \hat{\delta}^b_L = \delta^b$ satisfy

$$\hat{\delta}^a_L = -\frac{\lambda^a(\delta^a)}{\partial\lambda^a(\delta^a)/\partial\delta^a} - M_0(t,s,y),\ \ \hat{\delta}^b_L = -\frac{\lambda^b(\delta^b)}{\partial\lambda^b(\delta^b)/\partial\delta^b} + M_0(t,s,y) \tag{3.59}$$

where

$$M_0(t,s,y) = C_{\mathbb{Q}}(t,s,y) - \theta_1(t,s,y) \tag{3.60}$$

Proof. Note that utility function g from Definition 3.11 is linear and terms $q, h(s)$ in g are separable, while x is not. Hence, we guess that the solution to PDE 3.41 is of the form

$$w(t,s,y,q_1,x) = x + \theta_0(t,s,y) + q_1\theta_1(t,s,y) \tag{3.61}$$

Using this ansatz we get

$$\begin{aligned}
\delta^a + M_0(t,s,y) &= w(t,s,y,q_1-1, x + C_{\mathbb{Q}}(t,s,y) + \delta^a) - w(t,s,y,q_1,x) \\
&= x + C_{\mathbb{Q}}(t,s,y) + \delta^a + \theta_0(t,s,y) + (q_1-1)\theta_1(t,s,y) - x - \theta_0(t,s,y) - q_1\theta_1(t,s,y) \\
&= \delta^a + C_{\mathbb{Q}}(t,s,y) - \theta_1(t,s,y)
\end{aligned} \tag{3.62}$$

Let

$$f_0^a(\delta^a) = \lambda^a(\delta^a)\left(w(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a)-w(t,s,y,q_1,x)\right) \\ = \lambda^a(\delta^a)\left(\delta^a+M_0(t,s,y)\right) \tag{3.63}$$

We are looking for δ^a such that

$$\nabla\left(f_0^a(\delta^a)\right)=0 \tag{3.64}$$

then

$$\frac{\partial}{\partial\delta^a}\left[f_0^a(\delta^a)\right]=\frac{\partial}{\partial\delta^a}\left[\lambda^a(\delta^a)\delta^a\right]+\frac{\partial}{\partial\delta^a}\left[\lambda^a(\delta^a)\right]M_0(t,s,y) \tag{3.65}$$

By product rule

$$\frac{\partial}{\partial\delta^a}\left[\lambda^a(\delta^a)\delta^a\right]=\delta^a\frac{\partial\lambda^a(\delta^a)}{\partial\delta^a}+\lambda^a(\delta^a)\frac{\partial\delta^a}{\partial\delta^a}=\delta^a\frac{\partial\lambda^a(\delta^a)}{\partial\delta^a}+\lambda^a(\delta^a) \tag{3.66}$$

Hence

$$\frac{\partial}{\partial\delta^a}\left[f_0^a(\delta^a)\right]=\delta^a\frac{\partial\lambda^a(\delta^a)}{\partial\delta^a}+\lambda^a(\delta^a)+\frac{\partial\lambda^a(\delta^a)}{\partial\delta^a}M_0(t,s,y)=0 \tag{3.67}$$

We denote the point at which extremum is attained on f_0^a by $\hat{\delta}_L^a$

$$\hat{\delta}_L^a=-\frac{\lambda^a(\hat{\delta}_L^a)}{\partial\lambda^a(\hat{\delta}_L^a)/\partial\hat{\delta}_L^a}-M_0(t,s,y) \tag{3.68}$$

By same methods as above, we obtain $\hat{\delta}_L^b$. In which case f_0^b will be defined by

$$f_0^b(\delta^b) = \lambda^b(\delta^b)\left(w(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b)-w(t,s,y,q_1,x)\right) \\ = \lambda^b(\delta^b)\left(\delta^b-M_0(t,s,y)\right) \tag{3.69}$$

□

Proposition 3.18. ***(Optimal controls [ElAA3, Section 4]).*** *Let all assumptions from Proposition 3.17 hold. Suppose $\beta=0.5$ in (3.37), so order arrival intensity is of the form $\lambda(\delta)=A\left(B+\delta^2\right)^{-\gamma}$. Then optimal controls will be*

$$\hat{\delta}_L^a=\frac{-\gamma M_0+\sqrt{\gamma^2M_0^2+(2\gamma-1)\,B}}{2\gamma-1},\ \hat{\delta}_L^b=\frac{\gamma M_0+\sqrt{\gamma^2M_0^2+(2\gamma-1)\,B}}{2\gamma-1} \tag{3.70}$$

Proof. Using expressions for optimal controls in Proposition 3.17

$$\frac{\partial\lambda^a(\delta^a)}{\partial\delta^a}=\frac{\partial}{\partial\delta^a}\left[\frac{A}{(B+(\delta^a)^2)^\gamma}\right]=A\frac{\partial}{\partial\delta}\left[\frac{1}{(B+(\delta^a)^2)^\gamma}\right]=-A2\gamma\delta^a\left(B+(\delta^a)^2\right)^{-\gamma-1} \tag{3.71}$$

By Proposition 3.14, $\gamma > 1$, so $\gamma - 1 > 0$, then

$$\begin{aligned}
\frac{\partial \lambda^a(\delta^a)}{\partial \delta^a} &= -A2\gamma\delta^a \left(B + (\delta^a)^2\right)^{-\gamma-1} \\
&= -A\frac{2\gamma\delta^a}{(B+(\delta^a)^2)^{\gamma+1}} \\
&= -A\frac{2\gamma\delta^a}{(B+(\delta^a)^2)^{\gamma}(B+(\delta^a)^2)^{1}} \\
&= -\frac{A}{(B+(\delta^a)^2)^{\gamma}}\frac{2\gamma\delta^a}{B+(\delta^a)^2} \\
&= -\lambda^a(\delta^a)\frac{2\gamma\delta^a}{B+(\delta^a)^2}
\end{aligned} \tag{3.72}$$

then

$$\frac{\lambda^a(\delta^a)}{\partial \lambda^a(\delta^a)/\partial \delta^a} = -\frac{\lambda^a(\delta^a)}{\lambda^a(\delta^a)\frac{2\gamma\delta^a}{B+(\delta^a)^2}} = -\frac{B+(\delta^a)^2}{2\gamma\delta^a} = -\left(\frac{B}{2\gamma\delta^a} + \frac{(\delta^a)^2}{2\gamma\delta^a}\right) = -\left(\frac{B}{2\gamma\delta^a} + \frac{\delta^a}{2\gamma}\right) \tag{3.73}$$

Hence optimal spread is

$$\begin{aligned}
\delta^a &= -\frac{\lambda^a(\delta^a)}{\partial \lambda^a(\delta^a)/\partial \delta^a} - M_0(t,s,y) \\
\delta^a &= \frac{B}{2\gamma\delta^a} + \frac{\delta^a}{2\gamma} - M_0(t,s,y) \\
2\gamma(\delta^a)^2 &= B + (\delta^a)^2 - 2\gamma\delta^a M_0(t,s,y) \\
2\gamma(\delta^a)^2 - (\delta^a)^2 &= B - 2\gamma\delta^a M_0(t,s,y) \\
(\delta^a)^2(2\gamma - 1) &= B - 2\gamma\delta^a M_0(t,s,y)
\end{aligned} \tag{3.74}$$

We have a quadratic equation

$$(2\gamma - 1)(\delta^a)^2 + 2\gamma M_0(t,s,y)\delta^a - B = 0 \tag{3.75}$$

Solving for δ^a we have two points

$$\delta^a_+ = \frac{-\gamma M_0(t,s,y) + \sqrt{\gamma^2 M_0^2(t,s,y) + (2\gamma-1)B}}{2\gamma - 1}, \;\; \delta^a_- = \frac{-\gamma M_0(t,s,y) - \sqrt{\gamma^2 M_0^2(t,s,y) + (2\gamma-1)B}}{2\gamma - 1} \tag{3.76}$$

We denote the point at which maximum is attained on f_0^a by $\hat{\delta}^a_L$, since $\delta^a_+ > 0$ we have that

$$\hat{\delta}^a_L = \frac{-\gamma M_0(t,s,y) + \sqrt{\gamma^2 M_0^2(t,s,y) + (2\gamma-1)B}}{2\gamma - 1} \tag{3.77}$$

By same methods as above, we obtain $\hat{\delta}^b_L$. □

Proposition 3.19. ***(Solution [ElAA3, Section 4]).*** *Let all assumptions from Propositions 3.16 and 3.18 hold. Solution to HJB equation (3.41) is*

$$w(t,s,y,q_1,x) = x + \theta_0(t,s,y) + q_1\theta_1(t,s,y) \tag{3.78}$$

where

$$\begin{aligned}\theta_1(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(h(S_T)\right) - \mu\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \Delta_u S_u du\right)\\ \theta_0(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_0(u,S_u,Y_u)du\right)\\ \mathcal{J}_0(t,s,y) &= f_0^a(\hat{\delta}^a_L) + f_0^b(\hat{\delta}^b_L)\\ f_0^a(\delta^a) &= \lambda^a(\delta^a)\left(\delta^a + M_0(t,s,y)\right)\\ f_0^b(\delta^b) &= \lambda^b(\delta^b)\left(\delta^b - M_0(t,s,y)\right)\end{aligned} \tag{3.79}$$

Proof. Using optimal spreads $\hat{\delta}^a_L, \hat{\delta}^b_L$ we get

$$f_0^a(\hat{\delta}^a_L) = \lambda^a(\hat{\delta}^a_L)\left(\frac{M_0(\gamma-1)+\sqrt{\gamma^2M_0^2+(2\gamma-1)\,B}}{2\gamma-1}\right) \tag{3.80}$$

$$f_0^b(\hat{\delta}^b_L) = \lambda^b(\hat{\delta}^b_L)\left(\frac{M_0(1-\gamma)+\sqrt{\gamma^2M_0^2+(2\gamma-1)\,B}}{2\gamma-1}\right) \tag{3.81}$$

By optimality of $\hat{\delta}^a_L, \hat{\delta}^b_L$

$$\mathcal{H} = \sup_{\vartheta\in U}\mathcal{J}(t,s,y,\vartheta) = \mathcal{J}_0(t,s,y) = f_0^a(\hat{\delta}^a_L) + f_0^b(\hat{\delta}^b_L) \tag{3.82}$$

Recall from (3.42) that

$$\mathcal{L}_1 w = \mu s\frac{\partial w}{\partial s} + a_R(y)\frac{\partial w}{\partial y} + \frac{1}{2}\sigma^2(y)s^2\frac{\partial^2 w}{\partial s^2} + \frac{1}{2}b_R^2(y)\frac{\partial^2 w}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\frac{\partial^2 w}{\partial s\partial y} \tag{3.83}$$

Substituting for $w(t,s,y,q_1,x) = x + \theta_0(t,s,y) + q_1\theta_1(t,s,y)$ and taking derivatives we get

$$\begin{aligned}\mathcal{L}_1 w &= \mu s\tfrac{\partial\theta_0}{\partial s} + a_R(y)\tfrac{\partial\theta_0}{\partial y} + \tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2\theta_0}{\partial s^2} + \tfrac{1}{2}b_R^2(y)\tfrac{\partial^2\theta_0}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\tfrac{\partial^2\theta_0}{\partial s\partial y}\\ &+ q_1\mu s\tfrac{\partial\theta_1}{\partial s} + q_1 a_R(y)\tfrac{\partial\theta_1}{\partial y} + q_1\tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2\theta_1}{\partial s^2} + q_1\tfrac{1}{2}b_R^2(y)\tfrac{\partial^2\theta_1}{\partial y^2} + q_1\rho_R b_R(y)\sigma(y)s\tfrac{\partial^2\theta_1}{\partial s\partial y}\end{aligned} \tag{3.84}$$

Let

$$\begin{aligned}\mathcal{L}_1\theta_0 &= \mu s\tfrac{\partial\theta_0}{\partial s} + a_R(y)\tfrac{\partial\theta_0}{\partial y} + \tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2\theta_0}{\partial s^2} + \tfrac{1}{2}b_R^2(y)\tfrac{\partial^2\theta_0}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\tfrac{\partial^2\theta_0}{\partial s\partial y}\\ \mathcal{L}_1\theta_1 &= \mu s\tfrac{\partial\theta_1}{\partial s} + a_R(y)\tfrac{\partial\theta_1}{\partial y} + \tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2\theta_1}{\partial s^2} + \tfrac{1}{2}b_R^2(y)\tfrac{\partial^2\theta_1}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\tfrac{\partial^2\theta_1}{\partial s\partial y}\end{aligned} \tag{3.85}$$

Recall that $Q_2 = (Q_{2,u})_{u\in[t,T]}$ is stock inventory process from Definition 3.6 and $Q_1 = (Q_{1,u})_{u\in[t,T]}$ is option inventory process from Definition 3.5, where $\mathbb{P}(Q_{2,t} = q_2) = 1$ and $\mathbb{P}(Q_{1,t} = q_1) = 1$. There are assumed to satisfy (3.13), i.e

$$Q_{2,t} = -Q_{1,t}\Delta_t \text{ for } t \le T \tag{3.86}$$

where $\Delta_t = \Delta(t,S_t,Y_t)$ is options delta from Definition 3.7 defined by

$$\Delta_t = \frac{\partial C_{\mathbb{Q}}(t,S_t,Y_t)}{\partial S_t} \tag{3.87}$$

Using relation (3.86), we substitute for $q_2 = -q_1\Delta$ to obtain

$$\mathcal{L}_2 w = -q_1\Delta\mu s\frac{\partial w}{\partial x} + \frac{1}{2}q_1^2\Delta^2\sigma^2(y)s^2\frac{\partial^2 w}{\partial x^2} - q_1\Delta\sigma^2(y)s^2\frac{\partial^2 w}{\partial x\partial s} - q_1\Delta\rho_R b_R(y)\sigma(y)s\frac{\partial^2 w}{\partial x\partial y} \tag{3.88}$$

Collecting terms linear in q_1

$$\mathcal{L}_2 w = -q_1 \Delta \mu s \frac{\partial w}{\partial x} - q_1 \Delta \sigma^2(y) s^2 \frac{\partial^2 w}{\partial x \partial s} - q_1 \Delta \rho_R b_R(y) \sigma(y) s \frac{\partial^2 w}{\partial x \partial y} \tag{3.89}$$

Substituting for $w(t, s, y, q_1, x) = x + \theta_0(t, s, y) + q_1 \theta_1(t, s, y)$ and taking derivatives we get

$$\mathcal{L}_2 w = -q_1 \Delta \mu s \tag{3.90}$$

Collecting terms that do not depend on q_1 and that depend on q_1 linearly, we obtain two equations

$$\frac{\partial \theta_0}{\partial t} + \mathcal{L}_1 \theta_0 + \mathcal{J}_0(t, s, y) = 0, \ \frac{\partial \theta_1}{\partial t} + \mathcal{L}_1 \theta_1 - \Delta \mu s = 0 \tag{3.91}$$

Using terminal condition of (3.41)

$$w(T, s, y, q_1, x) = x + \theta_0(T, s, y) + q_1 \theta_1(T, s, y) = x + q_1 h(s) \tag{3.92}$$

we deduce that

$$\theta_0(T, s, y) = 0, \ \theta_1(T, s, y) = h(s) \tag{3.93}$$

By perturbation methods from Definition 1.104, we split PDE (3.41) into two PDEs on basis of powers of q_1

$$\begin{array}{ll} q_1^0 & \begin{cases} \frac{\partial \theta_0}{\partial t} + \mathcal{L}_1 \theta_0 + \mathcal{J}_0(t, s, y) = 0 \\ \theta_0(T, s, y) = 0 \end{cases} \\ q_1^1 & \begin{cases} \frac{\partial \theta_1}{\partial t} + \mathcal{L}_1 \theta_1 - \Delta \mu s = 0 \\ \theta_1(T, s, y) = h(s) \end{cases} \end{array} \tag{3.94}$$

Note that S, Y are both Ito processes, so by Feynman-Kac representation from Theorem 1.91, we get solutions

$$\theta_1(t, s, y) = \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(h(S_T) - \int_t^T \mu \Delta_u S_u du \right) = \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(h(S_T) \right) - \mu \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(\int_t^T \Delta_u S_u du \right) \tag{3.95}$$

$$\theta_0(t, s, y) = \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(\int_t^T \mathcal{J}_0(u, S_u, Y_u) du \right) \tag{3.96}$$

Moreover

$$\begin{aligned} M_0(t, s, y) &= C_{\mathbb{Q}}(t, s, y) - \theta_1(t, s, y) \\ &= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}_{t,s,y} \left(h(S_T) \right) - \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(h(S_T) \right) + \mu \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(\int_t^T \Delta_u S_u du \right) \\ &= C_{\mathbb{Q}}(t, s, y) - C_{\mathbb{P}}(t, s, y) + \mu \mathbb{E}^{\mathbb{P}}_{t,s,y} \left(\int_t^T \Delta_u S_u du \right) \end{aligned} \tag{3.97}$$

Therefore solution to HJB equation (3.41) is

$$w(t, s, y, q_1, x) = x + \theta_0(t, s, y) + q_1 \theta_1(t, s, y) \tag{3.98}$$

□

Proposition 3.20. ***(Verification [ElAA3, Appendix 8.4]).*** *Let $\mathcal{S} = [t,T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$ and $\bar{\mathcal{S}} = [t,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$. Let w be the solution to HJB equation (3.41) defined by (3.78) and let v be the value function from Definition 3.11. Then*

$$w(t,s,y,q_1,x) = v(t,s,y,q_1,x) \text{ for } (t,s,y,q_1,x) \in \bar{\mathcal{S}} \tag{3.99}$$

Proof. To prove this proposition we will use the same methods as in [ElAA3, Appendix 8.4] and as in the proof of [Ph, Theorem 3.5.2]. The proof is split into two parts, where in part 1 we verify that w satisfies quadratic growth condition from Verification Theorem 1.86, while in part 2 we prove the equality (3.99).

Part 1

Recall from proof of Proposition 3.19 that w satisfies

$$\begin{aligned} w(t,s,y,q_1,x) &= x + \theta_0(t,s,y) + q_1\theta_1(t,s,y) \\ &= x + \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_0(u,S_u,Y_u)du\right) + q_1\left(C_{\mathbb{P}}(t,s,y) - \mu\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \Delta_u S_u du\right)\right) \end{aligned} \tag{3.100}$$

where

$$\begin{aligned} \mathcal{J}_0(t,s,y) &= f_0^a(\hat{\delta}_L^a) + f_0^b(\hat{\delta}_L^b) \\ &= \lambda^a(\hat{\delta}_L^a)\left(\hat{\delta}_L^a + M_0(t,s,y)\right) + \lambda^a(\hat{\delta}_L^b)\left(\hat{\delta}_L^b - M_0(t,s,y)\right) \\ &= \lambda^a(\hat{\delta}_L^a)\tfrac{M_0(\gamma-1)+\sqrt{\gamma^2M_0^2+(2\gamma-1)B}}{2\gamma-1} + \lambda^b(\hat{\delta}_L^b)\tfrac{M_0(1-\gamma)+\sqrt{\gamma^2M_0^2+(2\gamma-1)B}}{2\gamma-1} \end{aligned} \tag{3.101}$$

and

$$\lambda^a(\hat{\delta}_L^a) + \lambda^b(\hat{\delta}_L^b) = \frac{A}{\left(B+(\hat{\delta}_L^a)^2\right)^\gamma} + \frac{A}{\left(B+(\hat{\delta}_L^b)^2\right)^\gamma} \tag{3.102}$$

Suppose we have a function $f(x) = \frac{A}{(B+x^2)^\gamma}$, then

$$\frac{d^2f(x)}{dx^2} = \frac{df(x)}{dx}\left[-2A\gamma x(x^2+B)^{-\gamma-1}\right] = 2A\gamma(x^2+B)^{-\gamma-2}\left((2\gamma+1)x^2 - B\right) \tag{3.103}$$

with $A, B > 0$ and $\gamma > 1$, we have that

$$2A\gamma(x^2+B)^{-\gamma-2}\left((2\gamma+1)x^2 - B\right) \leq 0 \text{ for } x > 0 \tag{3.104}$$

Hence, by Theorem 5.3, $\frac{A}{(B+x^2)^\gamma}$ is concave and so f_0^a, f_0^b are concave in $\hat{\delta}_L^a, \hat{\delta}_L^b$. By Theorem 5.2 f_0^a, f_0^b are bounded by first-order Taylor series, so we have as $\hat{\delta}_L^a \to 0$

$$\begin{aligned} f_0^a(\hat{\delta}_L^a) &\leq f_0^a(0) + (\hat{\delta}_L^a - 0)\left.\tfrac{\partial f_0^a(\delta^a)}{\partial\delta^a}\right|_{\delta^a=0} \\ f_0^a(\hat{\delta}_L^a) &\leq \lambda^a(0)M_0 + \hat{\delta}_L^a\left.\delta^a\tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a} + \lambda^a(\delta^a) + \tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a}M_0(t,s,y)\right|_{\delta^a=0} \\ f_0^a(\hat{\delta}_L^a) &\leq \lambda^a(0)M_0 + \hat{\delta}_L^a\left.\delta^a\tfrac{\partial\lambda^a(\delta^a)}{\partial\delta^a} + \lambda^a(\delta^a) + \left(-\lambda^a(\delta^a)\tfrac{2\gamma\delta^a}{B+(\delta^a)^2}\right)M_0(t,s,y)\right|_{\delta^a=0} \\ f_0^a(\hat{\delta}_L^a) &\leq \lambda^a(0)\left(M_0 + \hat{\delta}_L^a\right) \end{aligned} \tag{3.105}$$

and as $\hat{\delta}^b_L \to 0$

$$\begin{array}{rcl}
f^a_0(\hat{\delta}^b_L) & \leq & f^b_0(0) + (\hat{\delta}^b_L - 0)\ \frac{\partial f^b_0(\delta^b)}{\partial \delta^b}\Big|_{\delta^b=0} \\
f^a_0(\hat{\delta}^b_L) & \leq & \lambda^b(0)M_0 + \hat{\delta}^b_L\ \delta^b \frac{\partial \lambda^b(\delta^b)}{\partial \delta^b} + \lambda^b(\delta^b) + \frac{\partial \lambda^b(\delta^b)}{\partial \delta^b} M_0(t,s,y)\Big|_{\delta^b=0} \\
f^a_0(\hat{\delta}^b_L) & \leq & \lambda^b(0)M_0 + \hat{\delta}^b_L\ \delta^b \frac{\partial \lambda^b(\delta^b)}{\partial \delta^b} + \lambda^b(\delta^b) - \left(-\lambda^b(\delta^b)\frac{2\gamma\delta^b}{B+(\delta^b)^2}\right) M_0(t,s,y)\Big|_{\delta^b=0} \\
f^a_0(\hat{\delta}^b_L) & \leq & \lambda^b(0)\left(M_0 + \hat{\delta}^b_L\right)
\end{array} \tag{3.106}$$

hence

$$f^a_0(\hat{\delta}^a_L) + f^b_0(\hat{\delta}^b_L) \leq \lambda^a(0)\left(M_0 + \hat{\delta}^a_L\right) + \lambda^b(0)\left(M_0 + \hat{\delta}^b_L\right) \tag{3.107}$$

note that $\lambda^a(0) = \lambda^b(0) = \frac{A}{B^\gamma}$

$$\begin{array}{rcl}
f^a_0(\hat{\delta}^a_L) + f^b_0(\hat{\delta}^b_L) & \leq & \frac{A}{B^\gamma}\left(2M_0 + \hat{\delta}^a_L + \hat{\delta}^b_L\right) \\
& = & \frac{A}{B^\gamma}\left(2M_0 + \frac{-\gamma M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}{2\gamma-1} + \frac{\gamma M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}{2\gamma-1}\right) \\
& = & \frac{2A}{B^\gamma}\left(M_0 + \frac{\sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}{2\gamma-1}\right) \\
& = & \frac{2A}{B^\gamma}\frac{2M_0\gamma - M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}{2\gamma-1} \\
& = & \frac{2A}{B^\gamma}\frac{(2\gamma-1)M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}{2\gamma-1}
\end{array} \tag{3.108}$$

then by triangle inequality

$$\begin{array}{rcl}
\left|f^a_0(\hat{\delta}^a_L) + f^b_0(\hat{\delta}^b_L)\right| & \leq & \left|\frac{2A}{B^\gamma}\frac{(2\gamma-1)M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}{2\gamma-1}\right| \\
& \leq & \left|\frac{2A}{B^\gamma}\right|\left|\frac{(2\gamma-1)M_0}{2\gamma-1}\right| + \left|\frac{2A}{B^\gamma}\right|\left|\frac{1}{2\gamma-1}\right|\left|\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}\right|
\end{array} \tag{3.109}$$

Suppose we have a function $f(x) = \sqrt{x}$, then

$$\frac{d^2 f(x)}{dx^2} = \frac{df(x)}{dx}\left[\frac{1}{2\sqrt{x}}\right] = -\frac{1}{4x^{3/2}} \leq 0 \text{ for } x > 0 \tag{3.110}$$

Hence, by Theorem 5.3, $\sqrt{x}$ is concave. Since $B > 0$ and $\gamma > 1$, $\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}$ is concave in M_0. By Theorem 5.2, $\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}$ is bounded by first-order Taylor series, so we have as $M_0 \to 0$

$$\begin{array}{rcl}
\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B} & \leq & \sqrt{(2\gamma-1)\,B} + (M_0 - 0)\frac{\partial}{\partial M_0}\left[\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}\right] \\
\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B} & \leq & \sqrt{(2\gamma-1)\,B} + M_0\ \frac{\gamma^2 M_0}{\sqrt{\gamma^2 M_0^2 + (2\gamma-1)B}}\Big|_{M_0=0} \\
\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B} & \leq & \sqrt{(2\gamma-1)\,B}
\end{array} \tag{3.111}$$

Moreover, with $A, B > 0$ and $\gamma > 1$

$$\begin{array}{rcl}
\left|f^a_0(\hat{\delta}^a_L) + f^b_0(\hat{\delta}^b_L)\right| & \leq & \left|\frac{2A}{B^\gamma}\right|\left|\frac{(2\gamma-1)M_0}{2\gamma-1}\right| + \left|\frac{2A}{B^\gamma}\right|\left|\frac{1}{2\gamma-1}\right|\left|\sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}\right| \\
& \leq & \left|\frac{2A}{B^\gamma}\right|\left|\frac{(2\gamma-1)M_0}{2\gamma-1}\right| + \left|\frac{2A}{B^\gamma}\right|\left|\frac{1}{2\gamma-1}\right|\left|\sqrt{(2\gamma-1)\,B}\right| \\
& = & \frac{2A}{B^\gamma}\frac{(2\gamma-1)|M_0| + \sqrt{(2\gamma-1)B}}{2\gamma-1}
\end{array} \tag{3.112}$$

Suppose we have a function $f(x) = |x|$ and $c \in [0,1]$, then by triangle inequality we have

$$\begin{aligned} f((1-c)x+cy) &= |(1-c)x+cy| \\ &\leq |(1-c)|\,|x| + |c|\,|y| \\ &= (1-c)f(x) + cf(y) \text{ for } x, y \in \mathbb{R} \end{aligned} \tag{3.113}$$

Hence, by Definition 5.1, $|x|$ is convex. By convexity of absolute function and Jensen inequality from Theorem 5.47 we have

$$\begin{aligned} \left|\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\mathcal{J}_0(u,S_u,Y_u)\right)\right| &\leq \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\left|\mathcal{J}_0(u,S_u,Y_u)\right|\right) \\ &\leq \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\tfrac{2A}{B^\gamma}\tfrac{(2\gamma-1)|M_0(u,S_u,Y_u)|+\sqrt{(2\gamma-1)B}}{2\gamma-1}\right) \\ &= \tfrac{2A}{B^\gamma}\tfrac{(2\gamma-1)\mathbb{E}^{\mathbb{P}}_{t,s,y}(|M_0(u,S_u,Y_u)|)+\sqrt{(2\gamma-1)B}}{2\gamma-1} \end{aligned} \tag{3.114}$$

Recall from proof of Proposition 3.19 that

$$M_0(t,s,y) = C_{\mathbb{Q}}(t,s,y) - C_{\mathbb{P}}(t,s,y) + \mu\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \Delta_u S_u du\right) \tag{3.115}$$

Recall from Definition 3.7 that delta is defined as

$$\Delta_t = \frac{\partial C(t,s,y)}{\partial s} \tag{3.116}$$

Under assumption that C is a call from Proposition 3.34 we have that

$$C(t,s,y) = S_t P_1 - Ke^{-r\tau}P_0 \tag{3.117}$$

Then using initial condition $\mathbb{P}(S_t = s) = 1$ from (3.20), we get

$$\Delta_t = \frac{\partial C(t,s,y)}{\partial s} = \frac{\partial}{\partial s}\left[S_t P_1 - Ke^{-r\tau}P_0\right] = P_1\frac{\partial S_t}{\partial s} = P_1 \tag{3.118}$$

Note that P_1 is a probability and so $\Delta_t \in [0,1]$. If C is a put, then similarly we have

$$C(t,s,y) = Ke^{-r\tau}P_0 - S_t P_1 \tag{3.119}$$

In this case by same arguments $\Delta_t \in [-1,0]$. Combining both we get

$$|\Delta_t| \leq 1 \tag{3.120}$$

Moreover, if C is either a call or a put, then there exist $C_1 > 0$ such that

$$C_{\mathbb{Q}}(t,s,y) \vee C_{\mathbb{P}}(t,s,y) = C_1(1+s) \tag{3.121}$$

Hence, we can assume that there exist $C_2 > 0$ such that

$$|M_0(u,S_u,Y_u)| \leq C_2(1+S_u) \tag{3.122}$$

By monotinicity of expectation and linearity of expectation

$$\begin{aligned}\left|\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\mathcal{J}_0(u,S_u,Y_u)\right)\right| &\leq \tfrac{2A}{B^\gamma}\tfrac{(2\gamma-1)\mathbb{E}^{\mathbb{P}}_{t,s,y}(C_2(1+S_u))+\sqrt{(2\gamma-1)B}}{2\gamma-1} \\ &= \tfrac{2A}{B^\gamma}\tfrac{(2\gamma-1)\left(C_2+C_2\mathbb{E}^{\mathbb{P}}_{t,s,y}(S_u)\right)+\sqrt{(2\gamma-1)B}}{2\gamma-1}\end{aligned} \tag{3.123}$$

Recall from Definition 3.1 that process S satisfies SDE

$$dS_t = S_t\mu dt + S_t\sigma(Y_t)dW^{(1)}_t \tag{3.124}$$

by Ito Lemma from Theorem 1.61, we have

$$f(u,S_u) = f(t,S_t) + \int_t^u \left[\frac{\partial f}{\partial r} + \mu S_r\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2(Y_r)S_r^2\frac{\partial^2 f}{\partial S^2}\right]dr + \int_t^u \sigma(Y_r)S_r\frac{\partial f}{\partial S}dW^{(1)}_r \tag{3.125}$$

Let $f(t,s) = \ln(s)$, where derivatives are

$$\frac{\partial f}{\partial s} = \frac{1}{s},\ \frac{\partial^2 f}{\partial s^2} = -\frac{1}{s^2},\ \frac{\partial f}{\partial t} = 0 \tag{3.126}$$

Hence

$$\begin{aligned}\ln(S_u) &= \ln(S_t) + \int_t^u\left[\mu - \tfrac{1}{2}\sigma^2(Y_r)\right]dr + \int_t^u \sigma(Y_r)dW^{(1)}_r \\ \ln\left(\tfrac{S_u}{S_t}\right) &= \int_t^u\left[\mu - \tfrac{1}{2}\sigma^2(Y_r)\right]dr + \int_t^u \sigma(Y_r)dW^{(1)}_r \\ \tfrac{S_u}{S_t} &= e^{\int_t^u\left[\mu-\frac{1}{2}\sigma^2(Y_r)\right]dr+\int_t^u\sigma(Y_r)dW^{(1)}_r} \\ S_u &= S_te^{\int_t^u\left[\mu-\frac{1}{2}\sigma^2(Y_r)\right]dr+\int_t^u\sigma(Y_r)dW^{(1)}_r} \\ S_u &= S_te^{\mu(u-t)-\frac{1}{2}\int_t^u\sigma^2(Y_r)dr+\int_t^u\sigma(Y_r)dW^{(1)}_r} \\ S_u &= S_te^{\mu(u-t)}e^{-\frac{1}{2}\int_t^u\sigma^2(Y_r)dr+\int_t^u\sigma(Y_r)dW^{(1)}_r}\end{aligned} \tag{3.127}$$

Since $\mathbb{P}\left(S_t = s\right) = 1$ from initial conditions (3.20), we get

$$\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(S_u\right) = se^{\mu(u-t)}\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(e^{-\frac{1}{2}\int_t^u\sigma^2(Y_r)dr+\int_t^u\sigma(Y_r)dW^{(1)}_r}\right) \tag{3.128}$$

Let

$$X_u = -\frac{1}{2}\int_t^u \sigma^2(Y_r)dr + \int_t^u \sigma(Y_r)dW^{(1)}_r \tag{3.129}$$

by Ito Lemma from Theorem 1.61 we have

$$f(X_u) = f(X_t) + \int_t^u\left[-\frac{1}{2}\sigma^2(Y_r)\frac{\partial f}{\partial X} + \frac{1}{2}\sigma^2(Y_r)\frac{\partial^2 f}{\partial X^2}\right]dr + \int_t^u \sigma(Y_r)\frac{\partial f}{\partial X}dW^{(1)}_r \tag{3.130}$$

Let $f(x) = e^x$, then

$$\begin{aligned}e^{X_u} &= e^0 + \int_t^u\left[-\tfrac{1}{2}\sigma^2(Y_r)e^{X_r} + \tfrac{1}{2}\sigma^2(Y_r)e^{X_r}\right]dr + \int_t^u \sigma(Y_r)e^{X_r}dW^{(1)}_r \\ &= 1 + \int_t^u \sigma(Y_r)e^{X_r}dW^{(1)}_r\end{aligned} \tag{3.131}$$

By martingale property of Brownian integral from Definition 1.39

$$\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(e^{X_u}\right) = \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(1 + \int_t^u \sigma(Y_r)e^{X_r}dW^{(1)}_r\right) = 1 \tag{3.132}$$

Hence

$$\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(S_u\right) = se^{\mu(u-t)} \tag{3.133}$$

Therefore θ_0 is bounded by linear growth in s

$$\begin{aligned}
\left|\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_0(u,S_u,Y_u)du\right)\right| = |\theta_0(t,s,y)| &\leq \int_t^T \left(\tfrac{2A}{B^\gamma}\tfrac{(2\gamma-1)(C_2+C_2se^{\mu(u-t)})+\sqrt{(2\gamma-1)B}}{2\gamma-1}\right)du \\
&= \tfrac{2A}{B^\gamma}\int_t^T\left(C_2+C_2se^{\mu(u-t)}+\tfrac{\sqrt{B}}{\sqrt{2\gamma-1}}\right)du \\
&= \tfrac{2A}{B^\gamma}\left(\left(C_2+\sqrt{\tfrac{B}{2\gamma-1}}\right)\int_t^T du + C_2 s\int_t^T e^{\mu(u-t)}du\right) \\
&= \tfrac{2A}{B^\gamma}\left(C_2+\sqrt{\tfrac{B}{2\gamma-1}}\right)(T-t)+\tfrac{2A}{B^\gamma}C_2\tfrac{e^{\mu(T-t)}-1}{\mu}s
\end{aligned} \tag{3.134}$$

Recall from proof of Proposition 3.19 that θ_1 satisfies

$$\theta_1(t,s,y) = \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(h(S_T)-\int_t^T \mu\Delta_u S_u du\right) = C_{\mathbb{P}}(t,s,y)-\mu\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \Delta_u S_u du\right) \tag{3.135}$$

By mean value Theorem 5.4 and using initial conditions $\mathbb{P}\left(S_t = s\right) = 1$ from (3.20), we have that

$$\begin{aligned}
\theta_1(t,s,y) &= C_{\mathbb{P}}(t,s,y)-\mu\mathbb{E}^{\mathbb{P}}_{t,s,y}\left((T-t)\Delta_t S_t\right) \\
&= C_{\mathbb{P}}(t,s,y)-\mu(T-t)\Delta s
\end{aligned} \tag{3.136}$$

Note that we have previously showed that there exist $C_1 > 0$ such that

$$C_{\mathbb{Q}}(t,s,y)\vee C_{\mathbb{P}}(t,s,y) = C_1(1+s) \tag{3.137}$$

This implies that θ_1 is also bounded by linear growth in s. Next, by triangle inequality we have that

$$|w(t,s,y,q_1,x)| \quad\leq\quad |x|+|\theta_0(t,s,y)|+|q_1|\,|\theta_1(t,s,y)| \tag{3.138}$$

which shows a linear relationship between w and $x, q_1, \theta_0, \theta_1$. Since θ_0, θ_1 are functions of t, s, y, (3.138) implies linear growth, i.e. there exists $C_3 > 0$ such that

$$|w(t,s,y,q_1,x)| \leq C_3\left(1+|s|+|y|+|q_1|+|x|\right) \tag{3.139}$$

Linear growth implies quadratic growth, so there exists $C_4 > 0$ such that

$$|w(t,s,y,q_1,x)| \leq C_4\left(1+|s|^2+|y|^2+|q_1|^2+|x|^2\right) \tag{3.140}$$

Therefore w satisfies quadratic growth condition from Verification Theorem 1.86. Moreover $w \in C^2(\mathcal{S})\cap C^0(\bar{\mathcal{S}})$ by assumption in the Proposition 3.16.

Part 2

Let $Z^{t,z}_{\vartheta} = (Z^{t,z}_{\vartheta}(s)), s \in [t,T]$ be a unique solution to 3.19 as defined in Proposition 3.13. Let $s \in [t,T)$ and τ be a $[t,\infty)$-valued $(\mathcal{F}_s)$-stopping time and recall that $0 \leq t < T < \infty$. Applying the Ito Lemma from Theorem 1.61 to function w, over time interval $t, s \wedge \tau$

$$
\begin{aligned}
& w(s \wedge \tau, Z_{\vartheta}^{t,z}(s \wedge \tau)) \\
= \;& w(t, z) \\
+ \;& \int_t^{s\wedge\tau} \left[\frac{\partial w}{\partial u} + \mu S_u \frac{\partial w}{\partial S} + a_R(Y_u)\frac{\partial w}{\partial Y} + Q_{2,u}\mu S_u \frac{\partial w}{\partial X}\right] du \\
+ \;& \int_t^{s\wedge\tau} \frac{1}{2}\left[\sigma^2(Y_u) S_u^2 \frac{\partial^2 w}{\partial S^2} + b_R^2(Y_u)\frac{\partial^2 w}{\partial Y^2} + 2\rho_R b_R(Y_u)\sigma(Y_u) S_u \frac{\partial^2 w}{\partial S \partial Y}\right] du \\
+ \;& \int_t^{s\wedge\tau} \frac{1}{2}\left[Q_{2,u}^2\sigma^2(Y_u) S_u^2 \frac{\partial^2 w}{\partial X^2} + 2Q_{2,u}\sigma^2(Y_u) S_u^2 \frac{\partial^2 w}{\partial X \partial S} + 2\rho_R b_R(Y_u) Q_{2,u}\sigma(Y_u) S_u \frac{\partial^2 w}{\partial X \partial Y}\right] du \\
+ \;& \int_t^{s\wedge\tau} \left[\sigma(Y_u) S_u \frac{\partial w}{\partial S} + Q_{2,u}\sigma(Y_u) S_u \frac{\partial w}{\partial X}\right] dW_u^{(1)} + \int_t^{s\wedge\tau} b_R(Y_u)\frac{\partial w}{\partial Y} dW_u^{(2)} \\
+ \;& \int_t^{s\wedge\tau} \left[w(u, S_u, Y_u, Q_{1,u} - 1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] dN_u^a \\
+ \;& \int_t^{s\wedge\tau} \left[w(u, S_u, Y_u, Q_{1,u} + 1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] dN_u^b
\end{aligned} \tag{3.141}
$$

By Definition 1.32 we get

$$
\begin{aligned}
& w(s \wedge \tau, Z_{\vartheta}^{t,z}(s \wedge \tau)) \\
= \;& w(t, z) \\
+ \;& \int_t^{s\wedge\tau} \left[\frac{\partial w}{\partial u} + \mu S_u \frac{\partial w}{\partial S} + a_R(Y_u)\frac{\partial w}{\partial Y} + Q_{2,u}\mu S_u \frac{\partial w}{\partial X}\right] du \\
+ \;& \int_t^{s\wedge\tau} \frac{1}{2}\left[\sigma^2(Y_u) S_u^2 \frac{\partial^2 w}{\partial S^2} + b_R^2(Y_u)\frac{\partial^2 w}{\partial Y^2} + 2\rho_R b_R(Y_u)\sigma(Y_u) S_u \frac{\partial^2 w}{\partial S \partial Y}\right] du \\
+ \;& \int_t^{s\wedge\tau} \frac{1}{2}\left[Q_{2,u}^2\sigma^2(Y_u) S_u^2 \frac{\partial^2 w}{\partial X^2} + 2Q_{2,u}\sigma^2(Y_u) S_u^2 \frac{\partial^2 w}{\partial X \partial S} + 2\rho_R b_R(Y_u) Q_{2,u}\sigma(Y_u) S_u \frac{\partial^2 w}{\partial X \partial Y}\right] du \\
+ \;& \int_t^{s\wedge\tau} \left[\sigma(Y_u) S_u \frac{\partial w}{\partial S} + Q_{2,u}\sigma(Y_u) S_u \frac{\partial w}{\partial X}\right] dW_u^{(1)} + \int_t^{s\wedge\tau} b_R(Y_u)\frac{\partial w}{\partial Y} dW_u^{(2)} \\
+ \;& \int_t^{s\wedge\tau} \left[w(u, S_u, Y_u, Q_{1,u} - 1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] \lambda^a(\delta^a) du \\
+ \;& \int_t^{s\wedge\tau} \left[w(u, S_u, Y_u, Q_{1,u} + 1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] \lambda^b(\delta^b) du \\
+ \;& \int_t^{s\wedge\tau} \left[w(u, S_u, Y_u, Q_{1,u} - 1, X_u + C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^a) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] d\tilde{N}_u^a \\
+ \;& \int_t^{s\wedge\tau} \left[w(u, S_u, Y_u, Q_{1,u} + 1, X_u - C_{\mathbb{Q}}(u, S_u, Y_u) + \delta^b) - w(u, S_u, Y_u, Q_{1,u}, X_u)\right] d\tilde{N}_u^b
\end{aligned} \tag{3.142}
$$

First note that $s \wedge \tau$ is also a stopping time, as s is a constant and so it trivially $(\mathcal{F}_s)$-measurable, satisfying Definition 5.54. Let τ_n be the first hitting time from Definition 5.55 defined by

$$
\tau_n = \inf\left\{s \geq t : \{|S_s - s| \geq n\} \wedge \{|Y_s - y| \geq n\} \wedge \{|N_s^a - N_t^a| \geq n\} \wedge \left\{\left|N_s^b - N_t^b\right| \geq n\right\}\right\} \text{ for } n \in \mathbb{N} \tag{3.143}
$$

By Theorem 5.56, τ_n is a stopping time and so we select it to be our stopping time τ, i.e. $\tau_n = \tau$. By Definition 5.55, $\inf\{\emptyset\} = \infty$, in which case we have that $\lim_{n\to\infty} \tau_n = \infty$ by boundedness assumption on intensities $\lambda^a(\delta^a), \lambda^b(\delta^b)$ of N^a, N^b from proof of Proposition 3.11. Since $s \wedge \infty = s$, we therefore have

$$
\lim_{n\to\infty} w(s \wedge \tau_n, Z_{\vartheta}^{t,z}(s \wedge \tau_n)) = w(s, Z_{\vartheta}^{t,z}(s)) \tag{3.144}
$$

This means that sequence $\left(w(s \wedge \tau_n, Z_{\vartheta}^{t,z}(s \wedge \tau_n))\right), n \in \mathbb{N}$ converges to $w(s, Z_{\vartheta}^{t,z}(s))$ pointwise by Definition 5.60. Note that Brownian motions $W^{(1)}, W^{(2)}$ are martingales by Definition 1.39 and Compensated Poisson processes $\tilde{N}^a, \tilde{N}^b$ are also martingales by Proposition 1.5. Hence, stochastic integrals in (3.142) are local $(\mathcal{F}_s)$-martingales for each $n \in \mathbb{N}$ by Definition 5.57. By taking expectations of 3.142 we obtain

$$
\begin{aligned}
& \mathbb{E}\left(w(s\wedge\tau, Z_{\vartheta}^{t,z}(s\wedge\tau))\right) \\
= \;& w(t,z) \\
+ \;& \mathbb{E}\left(\int_t^{s\wedge\tau}\left[\frac{\partial w}{\partial u} + \mu S_u\frac{\partial w}{\partial S} + a_R(Y_u)\frac{\partial w}{\partial Y} + Q_{2,u}\mu S_u\frac{\partial w}{\partial X}\right]du\right) \\
+ \;& \mathbb{E}\left(\int_t^{s\wedge\tau}\left[\frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial S^2} + \frac{1}{2}b_R^2(Y_u)\frac{\partial^2 w}{\partial Y^2} + \rho_R b_R(Y_u)\sigma(Y_u)S_u\frac{\partial^2 w}{\partial S\partial Y}\right]du\right) \\
+ \;& \mathbb{E}\left(\int_t^{s\wedge\tau}\left[\frac{1}{2}Q_{2,u}^2\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial X^2} + Q_{2,u}\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial X\partial S} + \rho_R b_R(Y_u)Q_{2,u}\sigma(Y_u)S_u\frac{\partial^2 w}{\partial X\partial Y}\right]du\right) \\
+ \;& \mathbb{E}\left(\int_t^{s\wedge\tau}\left[w(u,S_u,Y_u,Q_{1,u}-1,X_u+C_{\mathbb{Q}}(u,S_u,Y_u)+\delta^a) - w(u,S_u,Y_u,Q_{1,u},X_u)\right]\lambda^a(\delta^a)du\right) \\
+ \;& \mathbb{E}\left(\int_t^{s\wedge\tau}\left[w(u,S_u,Y_u,Q_{1,u}+1,X_u-C_{\mathbb{Q}}(u,S_u,Y_u)+\delta^b) - w(u,S_u,Y_u,Q_{1,u},X_u)\right]\lambda^b(\delta^b)du\right)
\end{aligned}
\tag{3.145}
$$

Futhermore, quadratic growth condition (3.140) implies that

$$\left|w(T\wedge\tau, Z_{\vartheta}^{t,z}(s\wedge\tau))\right| \leq C_3\left(1 + sup_{s\in[t,T]}\left|Z_{\vartheta}^{t,z}(s)\right|^2\right) \tag{3.146}$$

Since $Z_{\vartheta}^{t,z}$ is defined as the unique solution to 3.19 as defined in Proposition 3.13, by Theorem 1.68 we have that

$$\mathbb{E}\left(\left|Z_{\vartheta}^{t,z}(s)\right|^2\right) < \infty \text{ for all } s \in [t,T] \tag{3.147}$$

Pointwise convergence from (3.144), inequality (3.146) and integrability of (3.147) allows us to apply dominated convergence from Theorem 5.62 onto (3.145), where we get

$$\lim_{n\to\infty}\mathbb{E}\left(w(s\wedge\tau_n, Z_{\vartheta}^{t,z}(s\wedge\tau_n))\right) = \mathbb{E}\left(w(s, Z_{\vartheta}^{t,z}(s))\right) \tag{3.148}$$

Since $w \in C^2(\mathcal{S}) \cap C^0(\bar{\mathcal{S}})$

$$\lim_{s\to T}\mathbb{E}\left(w(s, Z_{\vartheta}^{t,z}(s))\right) = \mathbb{E}\left(w(T, Z_{\vartheta}^{t,z}(T))\right) \tag{3.149}$$

where

$$
\begin{aligned}
& \mathbb{E}\left(w(s\wedge\tau, Z_{\vartheta}^{t,z}(s\wedge\tau))\right) \\
= \;& w(t,z) \\
+ \;& \mathbb{E}\left(\int_t^{T}\left[\frac{\partial w}{\partial u} + \mu S_u\frac{\partial w}{\partial S} + a_R(Y_u)\frac{\partial w}{\partial Y} + Q_{2,u}\mu S_u\frac{\partial w}{\partial X}\right]du\right) \\
+ \;& \mathbb{E}\left(\int_t^{T}\left[\frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial S^2} + \frac{1}{2}b_R^2(Y_u)\frac{\partial^2 w}{\partial Y^2} + \rho_R b_R(Y_u)\sigma(Y_u)S_u\frac{\partial^2 w}{\partial S\partial Y}\right]du\right) \\
+ \;& \mathbb{E}\left(\int_t^{T}\left[\frac{1}{2}Q_{2,u}^2\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial X^2} + Q_{2,u}\sigma^2(Y_u)S_u^2\frac{\partial^2 w}{\partial X\partial S} + \rho_R b_R(Y_u)Q_{2,u}\sigma(Y_u)S_u\frac{\partial^2 w}{\partial X\partial Y}\right]du\right) \\
+ \;& \mathbb{E}\left(\int_t^{T}\left[w(u,S_u,Y_u,Q_{1,u}-1,X_u+C_{\mathbb{Q}}(u,S_u,Y_u)+\delta^a) - w(u,S_u,Y_u,Q_{1,u},X_u)\right]\lambda^a(\delta^a)du\right) \\
+ \;& \mathbb{E}\left(\int_t^{T}\left[w(u,S_u,Y_u,Q_{1,u}+1,X_u-C_{\mathbb{Q}}(u,S_u,Y_u)+\delta^b) - w(u,S_u,Y_u,Q_{1,u},X_u)\right]\lambda^b(\delta^b)du\right)
\end{aligned}
\tag{3.150}
$$

Recall the notation from Definition 1.69 where

$$\mathbb{E}\left(Z_{\vartheta}^{t,z}(T)\right) = \mathbb{E}_{t,z}\left(Z^{\vartheta}(T)\right) = \mathbb{E}\left(Z^{\vartheta}(T)\middle| Z^{\vartheta}(t) = z\right) \tag{3.151}$$

So we rewrite (3.150) as

$$\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w(T, Z_\vartheta(T))\right) = \mathbb{E}\left(w(s\wedge\tau, Z^{t,z}_\vartheta(s\wedge\tau))\right) \tag{3.152}$$

Recall definitions (3.42) from Proposition 3.16

$$\begin{aligned}
\mathcal{L}_1 w &= \mu s\tfrac{\partial w}{\partial s} + a_R(y)\tfrac{\partial w}{\partial y} + \tfrac{1}{2}\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial s^2} + \tfrac{1}{2}b_R^2(y)\tfrac{\partial^2 w}{\partial y^2} + \rho_R b_R(y)\sigma(y)s\tfrac{\partial^2 w}{\partial s\partial y}\\
\mathcal{L}_2 w &= q_2\mu s\tfrac{\partial w}{\partial x} + \tfrac{1}{2}q_2^2\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial x^2} + q_2\sigma^2(y)s^2\tfrac{\partial^2 w}{\partial x\partial s} + \rho_R b_R(y)q_2\sigma(y)s\tfrac{\partial^2 w}{\partial x\partial y}\\
\mathcal{J}(t,s,y,\vartheta) &= \lambda^a(\delta^a)\left(w(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w(t,s,y,q_1,x)\right)\\
&+ \lambda^b(\delta^b)\left(w(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w(t,s,y,q_1,x)\right)
\end{aligned} \tag{3.153}$$

Recall from Definition 3.11 that the objective of the stochastic control problem is given by

$$v(t,s,y,q_1,x) = \sup_{(\delta^a,\delta^b)\in U}\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q^{(\delta^a,\delta^b)}_{1,T}, X^{(\delta^a,\delta^b)}_T)\right) \tag{3.154}$$

Recall that w is the solution to HJB equation (3.41), so we have that

$$\frac{\partial w}{\partial t} + \mathcal{L}_1 w + \mathcal{L}_2 w + \mathcal{J}(t,s,y,\vartheta) \le 0 \tag{3.155}$$

This means that

$$\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q_{1,T}, X_T)\right) \le w(t,s,y,q_1,x) \tag{3.156}$$

and then

$$v(t,s,y,q_1,x) = \sup_{(\delta^a,\delta^b)\in U}\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q^{(\delta^a,\delta^b)}_{1,T}, X^{(\delta^a,\delta^b)}_T)\right) \le w(t,s,y,q_1,x) \tag{3.157}$$

Recall from Proposition 3.18 that $\hat\delta^a_L, \hat\delta^b_L$ denote maximum points of U-valued controls δ^a, δ^b, then

$$\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q^{(\hat\delta^a_L,\hat\delta^b_L)}_{1,T}, X^{(\hat\delta^a_L,\hat\delta^b_L)}_T)\right) = w(t,s,y,q_1,x) \tag{3.158}$$

and then

$$v(t,s,y,q_1,x) = sup_{(\delta^a,\delta^b)\in U}\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q^{(\delta^a,\delta^b)}_{1,T}, X^{(\delta^a,\delta^b)}_T)\right) \ge w(t,s,y,q_1,x) \tag{3.159}$$

Combining (3.157) and (3.159), we obtain the required result

$$w(t,s,y,q_1,x) = v(t,s,y,q_1,x) \text{ for } (t,s,y,q_1,x) \in \bar{\mathcal{S}} \tag{3.160}$$

□

3.5 HJB equation for risk-averse market maker

Remark 3.21. Following proposition makes an assertion that variance of the utility function can be approximated in a way which will simplify the stochastic control problem defined in Definition 3.12. This simplification will allow us to use background results to solve this problem as noted at the end of Remark 3.23.

Proposition 3.22. ***(Variance approximation [ElAA3, Section 5]).*** *Refer to Subsection 3.2 for the assumptions used in this proposition. We also introduce new assumptions. Assume that $\hat{\delta}^a_L, \hat{\delta}^b_L$ are optimal controls from Proposition 3.18. Suppose that $\lambda^a = \lambda^a(\hat{\delta}^a_L), \lambda^b = \lambda^b(\hat{\delta}^b_L)$ are intensities of independent Poisson processes $N^a = (N^a_u)_{u\in[t,T]}, N^b = (N^b_u)_{u\in[t,T]}$. Suppose we have a process $\tilde{X} = (\tilde{X}_u)_{u\in[t,T]}$ which satisfies*

$$\tilde{X}_T = \int_t^T C_{\mathbb{Q}}(u, S_u, Y_u)dN^a_u - \int_t^T C_{\mathbb{Q}}(u, S_u, Y_u)dN^b_u + \int_t^T Q_{2,u}dS_u \tag{3.161}$$

Then

$$Var\left(\tilde{X}_T + Q_{1,T}h(S_T)\middle| \mathcal{F}_t\right) = \mathbb{E}^{\mathbb{P}}\left(\int_t^T (Q^2_{1,u}K_u + L_u)du\middle| \mathcal{F}_t\right) \tag{3.162}$$

where

$$\begin{aligned} L_t &= \left(C^2_{\mathbb{Q}}(t, S_t, Y_t) + \mathbb{E}^{\mathbb{P}}\left(h^2(S_T)\middle| \mathcal{F}_t\right) - 2C_{\mathbb{Q}}(t, S_t, Y_t)C_{\mathbb{P}}(t, S_t, Y_t)\right)(\lambda^a + \lambda^b) \\ K_t &= \Delta^2_t\sigma^2(Y_t)S^2_t \end{aligned} \tag{3.163}$$

Proof. We will use linearity of variance result in this proof, which refers to Proposition 5.46. Also, recall the notation from Definition 1.69 where

$$\mathbb{E}\left(Z^{t,z}_{\vartheta}(T)\right) = \mathbb{E}_{t,z}\left(Z^{\vartheta}(T)\right) = \mathbb{E}\left(Z^{\vartheta}(T)\middle| Z^{\vartheta}(t) = z\right) \tag{3.164}$$

To prove this proposition we start with the assumption that $\tilde{X}_T$ is not independent from $Q_{1,T}h(S_T)$, in which case by linearity of variance

$$\begin{aligned} Var\left(\tilde{X}_T + Q_{1,T}h(S_T)\middle| \mathcal{F}_t\right) &= Var\left(\tilde{X}_T\middle| \mathcal{F}_t\right) + Var\left(Q_{1,T}h(S_T)\middle| \mathcal{F}_t\right) \\ &+ 2Cov\left(\tilde{X}_T, Q_{1,T}h(S_T)\middle| \mathcal{F}_t\right) \end{aligned} \tag{3.165}$$

In order to verify that (3.165) is equal to (3.162) we derive explicit expressions for each term on the right hand side of (3.165) and see if these explicit expressions result in (3.162). We split this procedure in 4 parts, in part 1 deriving explicit expressions for $Var\left(\tilde{X}_T\middle| \mathcal{F}_t\right)$, in part 2 for $Var\left(Q_{1,T}h(S_T)\middle| \mathcal{F}_t\right)$, in part

3 for $Cov\left(\tilde{X}_T, Q_{1,T}h(S_T)\middle|\mathcal{F}_t\right)$ and in part 4 we substitute the explicit expressions for each terms into (3.165) and verify that it is equal to (3.162).

Part 1

Substituting for $\tilde{X}_T$ from (3.161) we get

$$Var\left(\tilde{X}_T\middle|\mathcal{F}_t\right) = Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^a - \int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^b + \int_t^T Q_{2,u}dS_u\middle|\mathcal{F}_t\right) \tag{3.166}$$

Note that $W^{(1)}$ is independent of both N^a, N^b by Proposition 1.47. Therefore by linearity of variance

$$\begin{aligned} Var\left(\tilde{X}_T\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) + Var\left(-\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^b\middle|\mathcal{F}_t\right) \\ &+ Var\left(\int_t^T Q_{2,u}dS_u\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) + Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^b\middle|\mathcal{F}_t\right) \\ &+ Var\left(\int_t^T Q_{2,u}dS_u\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.167}$$

Substituting for dS_u as defined in Definition 3.1 and by linearity of variance, we get

$$\begin{aligned} Var\left(\int_t^T Q_{2,u}dS_u\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T Q_{2,u}\left(\mu S_u du + \sigma(Y_u)S_u dW_u^{(1)}\right)\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T Q_{2,u}\mu S_u du\middle|\mathcal{F}_t\right) + Var\left(\int_t^T Q_{2,u}\sigma(Y_u)S_u dW_u^{(1)}\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.168}$$

Substituting for $Q_{2,u}$ as defined by identity in Definition 3.7

$$\begin{aligned} Var\left(\int_t^T Q_{2,u}dS_u\middle|\mathcal{F}_t\right) &= Var\left(-\int_t^T (Q_{1,u}\Delta_u)\,\mu S_u du\middle|\mathcal{F}_t\right) + Var\left(-\int_t^T (Q_{1,u}\Delta_u)\,\sigma(Y_u)S_u dW_u^{(1)}\middle|\mathcal{F}_t\right) \\ &= \mu^2 Var\left(\int_t^T Q_{1,u}\Delta_u S_u du\middle|\mathcal{F}_t\right) + Var\left(\int_t^T Q_{1,u}\Delta_u\sigma(Y_u)S_u dW_u^{(1)}\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.169}$$

By Proposition 1.40

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}\Delta_u\sigma(Y_u)S_u dW_u^{(1)}\middle|\mathcal{F}_t\right) &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T Q_{1,u}\Delta_u\sigma(Y_u)S_u dW_u^{(1)}\right)^2\middle|\mathcal{F}_t\right) \\ &+ \left(\mathbb{E}^{\mathbb{P}}\left(\int_t^T Q_{1,u}\Delta_u\sigma(Y_u)S_u dW_u^{(1)}\middle|\mathcal{F}_t\right)\right)^2 \\ &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T Q_{1,u}\Delta_u\sigma(Y_u)S_u dW_u^{(1)}\right)^2\middle|\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T Q_{1,u}^2\Delta_u^2\sigma^2(Y_u)S_u^2 du\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.170}$$

Let $T = t+\tau$, then by mean value Theorem 5.4

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}\Delta_u S_u du\middle|\mathcal{F}_t\right) &= Var\left(\int_t^{t+\tau} Q_{1,u}\Delta_u S_u du\middle|\mathcal{F}_t\right) \\ &= Var\left(h\left(Q_{1,t}\Delta_t S_t\right)\middle|\mathcal{F}_t\right) \\ &= \tau^2 Var\left(Q_{1,t}\Delta_t S_t\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.171}$$

The initial condition $\mathbb{P}\left(Z_{\vartheta}^{t,z}(t) = z\right) = 1$ from (3.20) and Definition 5.45, variance of constant random variable is zero, which implies that

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}\Delta_u S_u du \middle| \mathcal{F}_t \right) &= \tau^2 Var\left(Q_{1,t}\Delta_t S_t \middle| \mathcal{F}_t \right) \\ &= \tau^2 Var\left(q_1 \Delta s \right) \\ &= 0 \end{aligned} \tag{3.172}$$

Therefore

$$\begin{aligned} Var\left(\int_t^T Q_{2,u} dS_u \middle| \mathcal{F}_t \right) &= \mu^2 Var\left(\int_t^T Q_{1,u}\Delta_u S_u du \middle| \mathcal{F}_t \right) + Var\left(\int_t^T Q_{1,u}\Delta_u \sigma(Y_u) S_u dW_u^{(1)} \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T Q_{1,u}^2 \Delta_u^2 \sigma^2(Y_u) S_u^2 du \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.173}$$

By Definition 1.32 and by linearity of variance

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) dN_u^a \middle| \mathcal{F}_t \right) &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a + \int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) \\ &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a \middle| \mathcal{F}_t \right) + Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.174}$$

By Theorem 1.45

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a \middle| \mathcal{F}_t \right) &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a \right)^2 \middle| \mathcal{F}_t \right) \\ &+ \left(\mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a \middle| \mathcal{F}_t \right) \right)^2 \\ &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a \right)^2 \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.175}$$

Let $T = t + \tau$, then by mean value Theorem 5.4

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) &= Var\left(\int_t^{t+\tau} C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) \\ &= Var\left(\tau C_{\mathbb{Q}}(t,S_t,Y_t)\lambda^a \middle| \mathcal{F}_t \right) \\ &= \tau^2 Var\left(C_{\mathbb{Q}}(t,S_t,Y_t)\lambda^a \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.176}$$

The initial condition $\mathbb{P}\left(Z_{\vartheta}^{t,z}(t) = z \right) = 1$ from (3.20) and Definition 5.45, variance of constant random variable is zero, which implies that

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) &= \tau^2 Var\left(C_{\mathbb{Q}}(t,S_t,Y_t)\lambda^a \middle| \mathcal{F}_t \right) \\ &= \tau^2 Var\left(C_{\mathbb{Q}}(t,s,y)\lambda^a \right) \\ &= 0 \end{aligned} \tag{3.177}$$

Hence

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) dN_u^a \middle| \mathcal{F}_t \right) &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u) d\tilde{N}_u^a \middle| \mathcal{F}_t \right) + Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u,S_u,Y_u)\lambda^a du \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.178}$$

Similarly

$$\begin{aligned}
Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^b\Big|\,\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)d\tilde{N}_u^b\Big|\,\mathcal{F}_t\right) + Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right)\\
&= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right)
\end{aligned} \tag{3.179}$$

Therefore by linearity of expectation

$$\begin{aligned}
Var\left(\tilde{X}_T\Big|\,\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^a\Big|\,\mathcal{F}_t\right) + Var\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^b\Big|\,\mathcal{F}_t\right)\\
&+ Var\left(\int_t^T Q_{2,u}dS_u\Big|\,\mathcal{F}_t\right)\\
&= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u,S_u,Y_u)\lambda^a du\Big|\,\mathcal{F}_t\right) + \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right)\\
&+ \mathbb{E}^{\mathbb{P}}\left(\int_t^T Q_{1,u}^2\Delta_u^2\sigma^2(Y_u)S_u^2 du\Big|\,\mathcal{F}_t\right)\\
&= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u,S_u,Y_u)\left(\lambda^a+\lambda^b\right)du\Big|\,\mathcal{F}_t\right) + \mathbb{E}^{\mathbb{P}}\left(\int_t^T Q_{1,u}^2\Delta_u^2\sigma^2(Y_u)S_u^2 du\Big|\,\mathcal{F}_t\right)
\end{aligned} \tag{3.180}$$

Part 2

By integration by parts from Lemma 1.62

$$\begin{aligned}
N_T^b C_{\mathbb{P}}(T,S_T,Y_T) &= N_t^b C_{\mathbb{P}}(t,s,y) + \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b + \int_t^T N_u^b dC_{\mathbb{P}}(u,S_u,Y_u)\\
\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b &= N_T^b C_{\mathbb{P}}(T,S_T,Y_T) - N_t^b C_{\mathbb{P}}(t,s,y) - \int_t^T N_u^b dC_{\mathbb{P}}(u,S_u,Y_u)
\end{aligned} \tag{3.181}$$

and

$$\begin{aligned}
N_T^a C_{\mathbb{P}}(T,S_T,Y_T) &= N_t^a C_{\mathbb{P}}(t,s,y) + \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a + \int_t^T N_u^a dC_{\mathbb{P}}(u,S_u,Y_u)\\
\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a &= N_T^a C_{\mathbb{P}}(T,S_T,Y_T) - N_t^a C_{\mathbb{P}}(t,s,y) - \int_t^T N_u^a dC_{\mathbb{P}}(u,S_u,Y_u)
\end{aligned} \tag{3.182}$$

By Definition 3.5, we get

$$\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b - \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a = \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\left(dN_u^b - dN_u^a\right) = \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u} \tag{3.183}$$

and so

$$\begin{aligned}
\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u} &= N_T^b C_{\mathbb{P}}(T,S_T,Y_T) - N_t^b C_{\mathbb{P}}(t,s,y) - \int_t^T N_u^b dC_{\mathbb{P}}(u,S_u,Y_u)\\
&- N_T^a C_{\mathbb{P}}(T,S_T,Y_T) + N_t^a C_{\mathbb{P}}(t,s,y) + \int_t^T N_u^a dC_{\mathbb{P}}(u,S_u,Y_u)\\
&= C_{\mathbb{P}}(T,S_T,Y_T)\left(N_T^b - N_T^a\right) - C_{\mathbb{P}}(t,s,y)\left(N_t^b - N_t^a\right) - \int_t^T \left(N_u^b - N_u^a\right)dC_{\mathbb{P}}(u,S_u,Y_u)\\
&= C_{\mathbb{P}}(T,S_T,Y_T)Q_{1,T} - C_{\mathbb{P}}(t,s,y)Q_{1,t} - \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)
\end{aligned} \tag{3.184}$$

Rearranging and using initial conditions at time t from Definitions 3.5, (3.1) and (3.2), we get

$$Q_{1,T}C_{\mathbb{P}}(T,S_T,Y_T) = q_1 C_{\mathbb{P}}(t,s,y) + \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u} + \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u) \tag{3.185}$$

Process S is driven by Brownian motion $W^{(1)}$, whereas process Q_1 is driven by two Poisson processes

N^a, N^b. Hence, $C_{\mathbb{P}}$ is independent of Q_1 by Proposition 1.47 and by linearity of variance we get

$$\begin{aligned} Var\left(Q_{1,T}C_{\mathbb{P}}(T,S_T,Y_T)\middle|\mathcal{F}_t\right) &= Var\left(q_1 C_{\mathbb{P}}(t,s,y) + \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u} + \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u} + \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\middle|\mathcal{F}_t\right) + Var\left(\int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.186}$$

Substituting for $Q_{1,u}$ as defined by identity in Definition 3.7

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) &= Var\left(-\int_t^T \left(\frac{Q_{2,u}}{\Delta_u}\right)dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T \frac{Q_{2,u}}{\Delta_u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.187}$$

Using Definition (3.3), we apply the Ito Lemma from Theorem 1.61 on function h over time interval t, T

$$h(S_T) = h(s) + \int_t^T \left[\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\right]du + \int_t^u \sigma(Y_u)S_u\frac{\partial h}{\partial S}dW_u^{(1)} \tag{3.188}$$

Note that $h(s) = 0$ as there is no such payoff for option expiring at time T. Applying conditional expectation and by Proposition 1.40, we get

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left(h(S_T)\middle|\mathcal{F}_u\right) &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T \left[\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\right]du + \int_t^u \sigma(Y_u)S_u\frac{\partial h}{\partial S}dW_u^{(1)}\middle|\mathcal{F}_u\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T \left[\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\right]du\middle|\mathcal{F}_u\right) \end{aligned} \tag{3.189}$$

By Fubini Theorem 5.19 and measurability

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left(h(S_T)\middle|\mathcal{F}_u\right) &= \int_t^T \mathbb{E}^{\mathbb{P}}\left(\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\middle|\mathcal{F}_u\right)du \\ &= \int_t^T \mathbb{E}^{\mathbb{P}}\left(\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\right)du \end{aligned} \tag{3.190}$$

Hence

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T \frac{Q_{2,u}}{\Delta_u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T \frac{Q_{2,u}}{\Delta_u}\mathbb{E}^{\mathbb{P}}\left(\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\right)du\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.191}$$

Let $T = t + \tau$, then by mean value Theorem 5.4

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) &= Var\left(\int_t^{t+\tau} \frac{Q_{2,u}}{\Delta_u}\mathbb{E}^{\mathbb{P}}\left(\mu S_u \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_u)S_u^2\frac{\partial^2 h}{\partial S^2}\right)du\middle|\mathcal{F}_t\right) \\ &= Var\left(\tau\frac{Q_{2,t}}{\Delta_t}\mathbb{E}^{\mathbb{P}}\left(\mu S_t \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_t)S_t^2\frac{\partial^2 h}{\partial S^2}\right)\middle|\mathcal{F}_t\right) \\ &= \tau^2 Var\left(\frac{Q_{2,t}}{\Delta_t}\mathbb{E}^{\mathbb{P}}\left(\mu S_t \frac{\partial h}{\partial S} + \frac{1}{2}\sigma^2(Y_t)S_t^2\frac{\partial^2 h}{\partial S^2}\right)\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.192}$$

The initial condition $\mathbb{P}\left(Z_{\vartheta}^{t,z}(t) = z\right) = 1$ from (3.20) and Definition 5.45, variance of constant random variable is zero, which implies that

$$\begin{aligned} Var\left(\int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\middle|\mathcal{F}_t\right) &= \tau^2 Var\left(\frac{q_2}{\Delta}\left(\mu s\frac{\partial h}{\partial s} + \frac{1}{2}\sigma^2(y)s^2\frac{\partial^2 h}{\partial s^2}\right)\middle|\mathcal{F}_t\right) \\ &= 0 \end{aligned} \tag{3.193}$$

Substituting for $dQ_{1,u}$ as defined in Definition 3.5

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\left(dN_u^b-dN_u^a\right)\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b-\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.194}$$

Recall the assumption from the beginning of Subsection 3 that processes N^a, N^b are independent, so by linearity of variance, we get

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b\middle|\mathcal{F}_t\right)+Var\left(-\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b\middle|\mathcal{F}_t\right)+Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.195}$$

By Definition 1.32

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a+\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) \\ &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a\middle|\mathcal{F}_t\right)+Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.196}$$

By Theorem 1.45

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a\middle|\mathcal{F}_t\right) &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a\right)^2\middle|\mathcal{F}_t\right) \\ &+ \left(\mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a\middle|\mathcal{F}_t\right)\right)^2 \\ &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a\right)^2\middle|\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.197}$$

Let $T=t+\tau$, then by mean value Theorem 5.4

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) &= Var\left(\int_t^{t+\tau} C_{\mathbb{P}}(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) \\ &= Var\left(\tau C_{\mathbb{P}}(t,S_t,Y_t)\lambda^a\middle|\mathcal{F}_t\right) \\ &= \tau^2 Var\left(C_{\mathbb{P}}(t,S_t,Y_t)\lambda^a\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.198}$$

The initial condition $\mathbb{P}\left(Z_{\vartheta}^{t,z}(t)=z\right)=1$ from (3.20) and Definition 5.45, variance of constant random variable is zero, which implies that

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) &= \tau^2 Var\left(C_{\mathbb{P}}(t,S_t,Y_t)\lambda^a\middle|\mathcal{F}_t\right) \\ &= \tau^2 Var\left(C_{\mathbb{P}}(t,s,y)\lambda^a\right) \\ &= 0 \end{aligned} \tag{3.199}$$

Hence

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^a\middle|\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^a\middle|\mathcal{F}_t\right)+Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^a du\middle|\mathcal{F}_t\right) \end{aligned} \tag{3.200}$$

Similarly

$$\begin{aligned} Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dN_u^b\Big|\,\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)d\tilde{N}_u^b\Big|\,\mathcal{F}_t\right) + Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right) \end{aligned} \tag{3.201}$$

Therefore

$$\begin{aligned} Var\left(Q_{1,T}h(S_T)|\,\mathcal{F}_t\right) &= Var\left(\int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\Big|\,\mathcal{F}_t\right) + Var\left(\int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\Big|\,\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^a du\Big|\,\mathcal{F}_t\right) + \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^a du + \int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\lambda^b du\Big|\,\mathcal{F}_t\right) \\ &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}}^2(u,S_u,Y_u)\left(\lambda^a+\lambda^b\right)du\Big|\,\mathcal{F}_t\right) \end{aligned} \tag{3.202}$$

Part 3

By linearity of variance

$$\begin{aligned} Cov\left(\tilde{X}_T, Q_{1,T}h(S_T)\Big|\,\mathcal{F}_t\right) &= Cov\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^a, Q_{1,T}h(S_T)\Big|\,\mathcal{F}_t\right) \\ &- Cov\left(\int_t^T C_{\mathbb{Q}}(u,S_u,Y_u)dN_u^b, Q_{1,T}h(S_T)\Big|\,\mathcal{F}_t\right) \\ &+ Cov\left(\int_t^T Q_{2,u}dS_u, Q_{1,T}h(S_T)\Big|\,\mathcal{F}_t\right) \end{aligned} \tag{3.203}$$

Using Definition (3.3) we deduce by measurability

$$h(S_T) = \mathbb{E}^{\mathbb{P}}\left(h(S_T)|\,\mathcal{F}_T\right) = C_{\mathbb{P}}(T,S_T,Y_T) \tag{3.204}$$

Recall that process S is driven by Brownian motion $W^{(1)}$, whereas process Q_1 is driven by two Poisson processes N^a, N^b. Hence, $C_{\mathbb{P}}$ is independent of Q_1 by Proposition 1.47, so by measurability and independence we have

$$Q_{1,T}h(S_T) = Q_{1,T}\mathbb{E}^{\mathbb{P}}\left(h(S_T)|\,\mathcal{F}_T\right) = Q_{1,T}C_{\mathbb{P}}(T,S_T,Y_T) \tag{3.205}$$

Using derived expression for $Q_{1,T}h(S_T)$ from (3.185), note that random variable $q_1C_{\mathbb{P}}(t,s,y)$ is constant so by linearity of variance we obtain

$$\begin{aligned} Cov\left(\int_t^T Q_{2,u}dS_u, Q_{1,T}h(S_T)\Big|\,\mathcal{F}_t\right) &= Cov\left(\int_t^T Q_{2,u}dS_u, \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\Big|\,\mathcal{F}_t\right) \\ &+ Cov\left(\int_t^T Q_{2,u}dS_u, \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\Big|\,\mathcal{F}_t\right) \end{aligned} \tag{3.206}$$

Substituting for dS_u from Definition (3.1), we get

$$\begin{aligned} Cov\left(\int_t^T Q_{2,u}dS_u, Q_{1,T}h(S_T)\Big|\,\mathcal{F}_t\right) &= \mu Cov\left(\int_t^T Q_{2,u}S_u du, \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\Big|\,\mathcal{F}_t\right) \\ &+ Cov\left(\int_t^T Q_{2,u}\sigma(Y_u)S_u dW_u^{(1)}, \int_t^T C_{\mathbb{P}}(u,S_u,Y_u)dQ_{1,u}\Big|\,\mathcal{F}_t\right) \\ &+ \mu Cov\left(\int_t^T Q_{2,u}S_u du, \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\Big|\,\mathcal{F}_t\right) \\ &+ Cov\left(\int_t^T Q_{2,u}\sigma(Y_u)S_u dW_u^{(1)}, \int_t^T Q_{1,u}dC_{\mathbb{P}}(u,S_u,Y_u)\Big|\,\mathcal{F}_t\right) \end{aligned} \tag{3.207}$$

by Definition 3.5 and independence of $W^{(1)}, N^a, N^b$

$$\begin{aligned} Cov\left(\int_t^T Q_{2,u} dS_u, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= \mu Cov\left(\int_t^T Q_{2,u} S_u du, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dQ_{1,u} \middle| \mathcal{F}_t \right) \\ &+ \mu Cov\left(\int_t^T Q_{2,u} S_u du, \int_t^T Q_{1,u} dC_{\mathbb{P}}(u, S_u, Y_u) \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.208}$$

Let $T = t + \tau$, then by mean value Theorem 5.4

$$\begin{aligned} Cov\left(\int_t^T Q_{2,u} dS_u, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= \tau\mu Cov\left(Q_{2,t} S_t, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dQ_{1,u} \middle| \mathcal{F}_t \right) \\ &+ \tau\mu Cov\left(Q_{2,t} S_t, \int_t^T Q_{1,u} dC_{\mathbb{P}}(u, S_u, Y_u) \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.209}$$

The initial condition $\mathbb{P}\left(Z_{\vartheta}^{t,z}(t) = z \right) = 1$ from (3.20) and Definition 5.45, variance of constant random variable is zero, which implies that

$$\begin{aligned} Cov\left(\int_t^T Q_{2,u} dS_u, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= \tau\mu Cov\left(q_2 s, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dQ_{1,u} \middle| \mathcal{F}_t \right) \\ &+ \tau\mu Cov\left(q_2 s, \int_t^T Q_{1,u} dC_{\mathbb{P}}(u, S_u, Y_u) \middle| \mathcal{F}_t \right) \\ &= 0 \end{aligned} \tag{3.210}$$

Therefore

$$\begin{aligned} Cov\left(X_T, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) \\ &- Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^b, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.211}$$

Using (3.185) and linearity of variance

$$\begin{aligned} Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dQ_{1,u} \middle| \mathcal{F}_t \right) \\ &+ Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T Q_{1,u} dC_{\mathbb{P}}(u, S_u, Y_u) \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.212}$$

Substituting for $dQ_{1,u}$ as defined in Definition 3.5

$$\begin{aligned} Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) \left(dN_u^b - dN_u^a \right) \middle| \mathcal{F}_t \right) \\ &+ Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T Q_{1,u} dC_{\mathbb{P}}(u, S_u, Y_u) \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.213}$$

By linearity of variance

$$\begin{aligned} Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) &= Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dN_u^b \middle| \mathcal{F}_t \right) \\ &- Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dN_u^a \middle| \mathcal{F}_t \right) \\ &+ Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T Q_{1,u} dC_{\mathbb{P}}(u, S_u, Y_u) \middle| \mathcal{F}_t \right) \end{aligned} \tag{3.214}$$

by independence of $N^a, N^b, W^{(1)}$ and linearity of variance, we get

$$Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, Q_{1,T} h(S_T) \middle| \mathcal{F}_t \right) = -Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, \int_t^T C_{\mathbb{P}}(u, S_u, Y_u) dN_u^a \middle| \mathcal{F}_t \right) \tag{3.215}$$

By Definition 1.3, with $C_{\mathbb{Q}} = C_{\mathbb{Q}}(u, S_u, Y_u)$ and $C_{\mathbb{P}} = C_{\mathbb{P}}(u, S_u, Y_u)$

$$\begin{aligned}
Cov\left(\int_t^T C_{\mathbb{Q}} dN_u^a, \int_t^T C_{\mathbb{P}} dN_u^a \middle| \mathcal{F}_t\right) &= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{Q}} dN_u^a - \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}} dN_u^a\right)\right)\left(\int_t^T C_{\mathbb{P}} dN_u^a - \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{P}} dN_u^a\right)\right)\middle| \mathcal{F}_t\right) \\
&= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{Q}} dN_u^a - \int_t^T C_{\mathbb{Q}} \lambda^a du\right)\left(\int_t^T C_{\mathbb{P}} dN_u^a - \int_t^T C_{\mathbb{P}} \lambda^a du\right)\middle| \mathcal{F}_t\right) \\
&= \mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{Q}} d\tilde{N}_u^a\right)\left(\int_t^T C_{\mathbb{P}} d\tilde{N}_u^a\right)\middle| \mathcal{F}_t\right)
\end{aligned} \tag{3.216}$$

By Theorem 1.45

$$\mathbb{E}^{\mathbb{P}}\left(\left(\int_t^T C_{\mathbb{Q}} d\tilde{N}_u^a\right)\left(\int_t^T C_{\mathbb{P}} d\tilde{N}_u^a\right)\middle| \mathcal{F}_t\right) = \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \lambda^a du \middle| \mathcal{F}_t\right) \tag{3.217}$$

Hence

$$Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^a, Q_{1,T} h(S_T) \middle| \mathcal{F}_t\right) = -\mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \lambda^a du \middle| \mathcal{F}_t\right) \tag{3.218}$$

Similarly

$$Cov\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) dN_u^b, Q_{1,T} h(S_T) \middle| \mathcal{F}_t\right) = \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \lambda^b du \middle| \mathcal{F}_t\right) \tag{3.219}$$

Therefore, by linearity of expectation

$$\begin{aligned}
Cov\left(\tilde{X}_T, Q_{1,T} h(S_T) \middle| \mathcal{F}_t\right) &= -\mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \lambda^a du \middle| \mathcal{F}_t\right) \\
&- \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \lambda^b du \middle| \mathcal{F}_t\right) \\
&= -\mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \left(\lambda^b + \lambda^b\right) du \middle| \mathcal{F}_t\right)
\end{aligned} \tag{3.220}$$

Part 4

Substituting results obtain from parts 1-3 into (3.165), we get

$$\begin{aligned}
Var\left(\tilde{X}_T + Q_{1,T} h(S_T) \middle| \mathcal{F}_t\right) &= \mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}^2(u, S_u, Y_u) \left(\lambda^a + \lambda^b\right) du \middle| \mathcal{F}_t\right) \\
&+ \mathbb{E}^{\mathbb{P}}\left(\int_t^T Q_{1,u}^2 \Delta_u^2 \sigma^2(Y_u) S_u^2 du \middle| \mathcal{F}_t\right) \\
&+ \mathbb{E}^{\mathbb{P}}\left(\int_t^T \left(h^2(S_T) \middle| \mathcal{F}_u\right) \left(\lambda^a + \lambda^b\right) du \middle| \mathcal{F}_t\right) \\
&- 2\mathbb{E}^{\mathbb{P}}\left(\int_t^T C_{\mathbb{Q}}(u, S_u, Y_u) C_{\mathbb{P}}(u, S_u, Y_u) \left(\lambda^b + \lambda^b\right) du \middle| \mathcal{F}_t\right)
\end{aligned} \tag{3.221}$$

With

$$\begin{aligned}
L_t &= \left(C_{\mathbb{Q}}^2(t, S_t, Y_t) + \mathbb{E}^{\mathbb{P}}\left(h^2(S_T) \middle| \mathcal{F}_t\right) - 2C_{\mathbb{Q}}(t, S_t, Y_t) C_{\mathbb{P}}(t, S_t, Y_t)\right) \left(\lambda^a + \lambda^b\right) \\
K_t &= \Delta_t^2 \sigma^2(Y_t) S_t^2
\end{aligned} \tag{3.222}$$

and by linearity of expectation we obtain the required result

$$Var\left(\tilde{X}_T + Q_{1,T} h(S_T) \middle| \mathcal{F}_t\right) = \mathbb{E}^{\mathbb{P}}\left(\int_t^T \left(Q_{1,u}^2 K_u + L_u\right) du \middle| \mathcal{F}_t\right) \tag{3.223}$$

□

Remark 3.23. **(Value function and variance penalty [ElAA3, Section 5]).** Let us repeat relevant assumptions from Subsection 3.2. Let $X = (X_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process assumed to satisfy

$$\begin{cases} dX_T = (C_{\mathbb{Q}}(T, S_T, Y_T) + \delta^a)dN_T^a - (C_{\mathbb{Q}}(T, S_T, Y_T) - \delta^b)dN_T^b + Q_{2,T}dS_T \\ X_t = x \text{ a.s.} \end{cases} \tag{3.224}$$

The utility function g is defined by

$$g(s, y, q_1, x) = x + q_1 h(s) \tag{3.225}$$

and the value function v is defined by

$$\begin{aligned} v(t, s, y, q_1, x) &= \sup_{(\delta^a,\delta^b)\in U} \Big[\mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q_{1,T}^{(\delta^a,\delta^b)}, X_T^{(\delta^a,\delta^b)})\right) \\ &- \quad Var^{\mathbb{P}}_{t,s,y,q_1,x}\left(g(S_T, Y_T, Q_{1,T}^{(\delta^a,\delta^b)}, X_T^{(\delta^a,\delta^b)})\right)\Big] \end{aligned} \tag{3.226}$$

Let us introduce new assumptions. Assume that process $\tilde{X}$ from Proposition 3.22 approximates process X, i.e.

$$X_T \approx \tilde{X}_T \tag{3.227}$$

Given this approximation, we substitute the variance of g as in (3.162) into (3.226) and define another value function v^ϵ by

$$v^\epsilon(t, s, y, q_1, x) = \sup_{(\delta^a,\delta^b)\in U} \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(X_T + Q_{1,T}h(S_T) - \epsilon\int_t^T (Q_{1,u}^2 K_u + L_u)du\right) \tag{3.228}$$

where we assume ϵ to be a small parameter, i.e. $0 < \epsilon \ll 1$ and drop (δ^a, δ^b) to simplify the notation. Note that as in Definition 1.70, the expectation now contains both the terminal gain function, which in our case is

$$g(Z_\vartheta^{t,z}(T)) = X_T + Q_{1,T}h(S_T) \tag{3.229}$$

and the running gain function, which in our case is

$$f(s, Z_\vartheta^{t,z}(s), \vartheta) = -\epsilon\left(Q_{1,s}^2 K_s + L_s\right) \tag{3.230}$$

Proposition 3.24. ***(Deriving HJB equation [ElAA3, Section 5]).*** *Refer to Subsection 3.1 for notations and Subsection 3.2 for assumptions. Let $Z_\vartheta^{t,z} = (Z_\vartheta^{t,z}(s)), s \in [t, T]$ be a unique solution to 3.19 as defined in Proposition 3.13. Suppose we have a function $w^\epsilon \in C^2(\mathcal{S}) \cup C^0(\bar{\mathcal{S}})$. Then w^ϵ satisfies following PDE*

$$\begin{cases} \frac{\partial w^\epsilon}{\partial t} + \mathcal{L}_1 w^\epsilon + \mathcal{L}_2 w^\epsilon + \mathcal{H}^\epsilon = \epsilon(q_1^2 k + l) \text{ for all } (t, s, y, q_1, x) \in \mathcal{S} \\ w^\epsilon(t, s, y, q_1, x) = x + q_1 h(s) \text{ for all } (t, s, y, q_1, x) \in \partial\mathcal{S} \end{cases} \tag{3.231}$$

where

$$
\begin{aligned}
\mathcal{L}_1 w^\epsilon &= \mu s \tfrac{\partial w^\epsilon}{\partial s} + a_R(y)\tfrac{\partial w^\epsilon}{\partial y} + \tfrac{1}{2}\sigma^2(y) s^2 \tfrac{\partial^2 w^\epsilon}{\partial s^2} + \tfrac{1}{2} b_R^2(y)\tfrac{\partial^2 w^\epsilon}{\partial y^2} + \rho_R b_R(y)\sigma(y) s \tfrac{\partial^2 w^\epsilon}{\partial s \partial y} \\
\mathcal{L}_2 w^\epsilon &= q_2 \mu s \tfrac{\partial w^\epsilon}{\partial x} + \tfrac{1}{2} q_2^2 \sigma^2(y) s^2 \tfrac{\partial^2 w^\epsilon}{\partial x^2} + q_2 \sigma^2(y) s^2 \tfrac{\partial^2 w^\epsilon}{\partial x \partial s} + \rho_R b_R(y) q_2 \sigma(y) s \tfrac{\partial^2 w^\epsilon}{\partial x \partial y} \\
\mathcal{H}^\epsilon &= \sup_{\vartheta \in U} \mathcal{J}^\epsilon(t,s,y,\vartheta) \\
\mathcal{J}^\epsilon(t,s,y,\vartheta) &= \lambda^a(\delta^a)\left(w^\epsilon(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w^\epsilon(t,s,y,q_1,x)\right) \\
&+ \lambda^b(\delta^b)\Big(w^\epsilon(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w^\epsilon(t,s,y,q_1,x)\Big) \\
l &= \Big(C^2_{\mathbb{Q}}(t,s,y) + \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(h^2(S_T)\right) - 2C_{\mathbb{Q}}(t,s,y)C_{\mathbb{P}}(t,s,y)\Big)\left(\lambda^a(\delta^a_L)+\lambda^b(\delta^b_L)\right) \\
k &= \Delta^2\sigma^2(y)s^2
\end{aligned}
\tag{3.232}
$$

Proof. Generalising results obtained in [Ph, Subsection 3.4.1, p43]. Suppose that w^ϵ satisfies

$$
w^\epsilon(t,z) \geq \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w^\epsilon(T, Z^{t,z}_{\vartheta}(T)) + \int_t^T f(s, Z^{t,z}_{\hat{\vartheta}_L}(s), \hat{\vartheta}_L) ds\right) \tag{3.233}
$$

By linearity of expectation (3.233) becomes

$$
w^\epsilon(t,z) \geq \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w^\epsilon(T, Z^{t,z}_{\vartheta}(T))\right) + \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(\int_t^T f(s, Z^{t,z}_{\hat{\vartheta}_L}(s), \hat{\vartheta}_L) ds\right) \tag{3.234}
$$

Note that by the exact same methods as in proof of Proposition 3.16

$$
w^\epsilon(t,z) \geq \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(w^\epsilon(T, Z^{t,z}_{\vartheta}(T))\right) \tag{3.235}
$$

can be shown to result in (3.231) for $\epsilon = 0$. So we will focus on deriving the expression for the integral term in (3.233). Using the assumption from Remark 3.23

$$
f(s, Z^{t,z}_{\hat{\vartheta}_L}(s), \hat{\vartheta}_L) = -\epsilon\left(Q^2_{1,s}K_s + L_s\right) \tag{3.236}
$$

we subsitute for $f(s, Z^{t,z}_{\hat{\vartheta}_L}(s), \hat{\vartheta}_L)$ to obtain

$$
\mathbb{E}\left(\int_t^T f(s, Z^{t,z}_{\hat{\vartheta}_L}(s), \hat{\vartheta}_L) ds\right) = \mathbb{E}_{t,s,y,q_1,x}\left(\epsilon \int_t^T (Q^2_{1,s}K_s + L_s) ds\right) \tag{3.237}
$$

Let $T = t + \tau$, then by Fubini Theorem 5.19 and fundamental theorem of calculus

$$
\mathbb{E}_{t,s,y,q_1,x}\left(\epsilon \int_t^{t+\tau} (Q^2_{1,s}K_s + L_s) ds\right) = -\tau\epsilon\left(\mathbb{E}_{t,s,y,q_1,x}\left(Q^2_{1,t}K_t\right) + \mathbb{E}_{t,s,y,q_1,x}\left(L_t\right)\right) \tag{3.238}
$$

Using definitions from Proposition 3.22 we have

$$
\begin{aligned}
\mathbb{E}_{t,s,y,q_1,x}\left(Q^2_{1,t}K_t\right) &= q_1^2\Delta^2\sigma^2(y)s^2 \\
\mathbb{E}_{t,s,y,q_1,x}\left(L_t\right) &= \Big(C^2_{\mathbb{Q}}(t,s,y) + \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(h^2(S_T)\right) - 2C_{\mathbb{Q}}(t,s,y)C_{\mathbb{P}}(t,s,y)\Big)\left(\lambda^a(\delta^a_L)+\lambda^b(\delta^b_L)\right)
\end{aligned}
\tag{3.239}
$$

Dividing by τ and taking limit as $h \to 0$

$$
-\tau\epsilon\mathbb{E}_{t,s,y,q_1,x}\left(Q^2_{1,t}K_t + L_t\right) = -\epsilon\left(q_1^2 k + l\right) \tag{3.240}
$$

Let

$$\begin{aligned}\mathcal{J}^{\epsilon}(t,s,y,\vartheta) &= \lambda^{a}(\delta^{a})\left(w^{\epsilon}(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^{a})-w^{\epsilon}(t,s,y,q_1,x)\right)\\ &+ \lambda^{b}(\delta^{b})\left(w^{\epsilon}(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^{b})-w^{\epsilon}(t,s,y,q_1,x)\right)\\ &- \epsilon\left(q_1^2k+l\right)\end{aligned} \tag{3.241}$$

Then by results from proof of Proposition 3.16 we get

$$0=\frac{\partial w^{\epsilon}}{\partial t}+\mathcal{L}_1 w^{\epsilon}+\mathcal{L}_2 w^{\epsilon}+\sup_{\vartheta\in U}\mathcal{J}^{\epsilon}(t,s,y,\vartheta) \tag{3.242}$$

Recall from Proposition 3.22 that $\epsilon\left(q_1^2k+l\right)$ is already defined on optimal controls $\hat{\vartheta}_L=(\hat{\delta}^a_L,\hat{\delta}^b_L)$ from Proposition 3.17. Therefore we can take $\epsilon\left(q_1^2k+l\right)$ term outside the supremum and move it to the other side of the equation which will result in

$$\epsilon\left(q_1^2k+l\right)=\frac{\partial w^{\epsilon}}{\partial t}+\mathcal{L}_1 w^{\epsilon}+\mathcal{L}_2 w^{\epsilon}+\sup_{\vartheta\in U}\mathcal{J}^{\epsilon}(t,s,y,\vartheta) \tag{3.243}$$

Regular terminal condition for this equation is

$$w^{\epsilon}(T,s,y,q_1,x)=g(s,y,q_1,x)=x+q_1h(s) \tag{3.244}$$

□

Remark 3.25. Setting $\epsilon=0$ in (3.231), transforms PDE into homogenous type, see Definition 1.88. By Remark 1.106 this is one of the sources of non-uniformity. This leads to use of singular perturbation technique to solve (3.231) as stated by [ElAA3]. We will not be defining this method rigorously, as it is outside the purpuse of this study, and simply follow proofs by [ElAA3].

Proposition 3.26. ***(General optimal controls).*** *Let all assumptions from Proposition 3.24 hold. Given the following ansatz to PDE 3.41*

$$w^{\epsilon}(t,s,y,q_1,x)=x+\sum_{n=0}^{\infty}\epsilon^{n}\phi_n(t,s,y,q_1) \tag{3.245}$$

optimal controls $\hat{\delta}^a=\delta^a,\hat{\delta}^b=\delta^b$ *satisfy*

$$\begin{aligned}\hat{\delta}^{a} &= -\tfrac{\lambda^{a}(\delta^{a})}{\partial\lambda^{a}(\delta^{a})/\partial\delta^{a}}-\left(M^{a}(t,s,y,q_1)+\epsilon^{2}R^{a}(t,s,y,q_1)\right)\\ \hat{\delta}^{b} &= -\tfrac{\lambda^{b}(\delta^{b})}{\partial\lambda^{b}(\delta^{b})/\partial\delta^{b}}+\left(M^{b}(t,s,y,q_1)+\epsilon^{2}R^{b}(t,s,y,q_1)\right)\end{aligned} \tag{3.246}$$

where

$$
\begin{aligned}
M^a(t,s,y,q_1) &= M_0(t,s,y) + \epsilon M_1(t,s,y,q_1) \\
M^b(t,s,y,q_1) &= M_0(t,s,y) + \epsilon M_2(t,s,y,q_1) \\
M_0(t,s,y) &= C_{\mathbb{Q}}(t,s,y) - \theta_1(t,s,y) \\
M_1(t,s,y,q_1) &= -\theta_3(t,s,y) + (1-2q_1)\theta_4(t,s,y) \\
M_2(t,s,y,q_1) &= -\theta_3(t,s,y) - (1+2q_1)\theta_4(t,s,y) \\
R^a(t,s,y,q_1) &= \textstyle\sum_{n=2}^{\infty} \epsilon^{n-2} \left(\phi_n(t,s,y,q_1-1) - \phi_n(t,s,y,q_1)\right) \\
R^b(t,s,y,q_1) &= -\textstyle\sum_{n=2}^{\infty} \epsilon^{n-2} \left(\phi_n(t,s,y,q_1+1) - \phi_n(t,s,y,q_1)\right)
\end{aligned}
\tag{3.247}
$$

Proof. Suppose that w^ϵ is a solution to (3.231), given by perturbation series with respect to ϵ

$$
w^\epsilon(t,s,y,q_1,x) = x + \sum_{n=0}^{\infty} \epsilon^n \phi_n(t,s,y,q_1) \tag{3.248}
$$

If $\epsilon = 0$ then utility is linear, as is the case in Proposition 3.16, hence we assume that ϕ_0 has the following form

$$
\phi_0(t,s,y,q_1) = \theta_0(t,s,y) + q_1\theta_1(t,s,y) \tag{3.249}
$$

As utility contains a constraint on the square of the options inventory q_1, we assume that ϕ_1 has the following form

$$
\phi_1(t,s,y,q_1) = \theta_2(t,s,y) + q_1\theta_3(t,s,y) + q_1^2\theta_4(t,s,y) \tag{3.250}
$$

Then we substitute ϕ_0, ϕ_1 into the expression for M_1 in (3.247) and get

$$
M_1(t,s,y,q_1) = -\theta_3(t,s,y) + (1-2q_1)\theta_4(t,s,y) \tag{3.251}
$$

Using expressions for M_0, R^a in (3.247) we have

$$
w^\epsilon(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w^\epsilon(t,s,y,q_1,x) = \delta^a + M_0(t,s,y) + \epsilon M_1(t,s,y) + \epsilon^2 R^a(t,s,y,q_1) \tag{3.252}
$$

Let

$$
\begin{aligned}
f^a(\delta^a,\epsilon) &= \lambda^a(\delta^a)\left(w^\epsilon(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w^\epsilon(t,s,y,q_1,x)\right) \\
&= \lambda^a(\delta^a)\left(\delta^a + M_0(t,s,y) + \epsilon M_1(t,s,y) + \epsilon^2 R^a(t,s,y,q_1)\right) \\
&= \lambda^a(\delta^a)\left(\delta^a + M^a(t,s,y,q_1) + \epsilon^2 R^a(t,s,y,q_1)\right)
\end{aligned}
\tag{3.253}
$$

where we have used definition of M^a from (3.247). We are looking for δ^asuch that

$$
\nabla\left(f^a(\delta^a,\epsilon)\right) = \frac{\partial f^a(\delta^a,\epsilon)}{\partial\delta^a} = 0 \tag{3.254}
$$

Then

$$
\frac{\partial}{\partial\delta^a}\left[f^a(\delta^a,\epsilon)\right] = \frac{\partial}{\partial\delta^a}\left[\lambda^a(\delta^a)\delta^a\right] + \frac{\partial}{\partial\delta^a}\left[\lambda^a(\delta^a)\right]\left(M^a(t,s,y,q_1) + \epsilon^2 R^a(t,s,y,q_1)\right) \tag{3.255}
$$

By product rule

$$\frac{\partial}{\partial \delta^a}\left[\lambda^a(\delta^a)\delta^a\right] = \delta^a \frac{\partial \lambda^a(\delta^a)}{\partial \delta^a} + \lambda^a(\delta^a)\frac{\partial \delta^a}{\partial \delta^a} = \delta^a \frac{\partial \lambda^a(\delta^a)}{\partial \delta^a} + \lambda^a(\delta^a) \tag{3.256}$$

Hence

$$\frac{\partial}{\partial \delta^a}\left[f^a(\delta^a,\epsilon)\right] = \delta^a \frac{\partial \lambda^a(\delta^a)}{\partial \delta^a} + \lambda^a(\delta^a) + \frac{\partial \lambda^a(\delta^a)}{\partial \delta^a}\left(M^a(t,s,y,q_1) + \epsilon^2 R^a(t,s,y,q_1)\right) = 0 \tag{3.257}$$

We denote the point at which extremum is attained on f^a by $\hat{\delta}^a$

$$\hat{\delta}^a \quad = \quad -\tfrac{\lambda^a(\hat{\delta}^a)}{\partial \lambda^a(\hat{\delta}^a)/\partial \hat{\delta}^a} - \left(M^a(t,s,y,q_1) + \epsilon^2 R^a(t,s,y,q_1)\right) \tag{3.258}$$

By same methods as above, we obtain M^b, R^b and $\hat{\delta}^b$. In which case f^b will be defined by

$$\begin{aligned} f^b(\delta^b,\epsilon) &= \lambda^b(\delta^b)\left(w^\epsilon(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w^\epsilon(t,s,y,q_1,x)\right) \\ &= \lambda^b(\delta^b)\left(\delta^b + M^b(t,s,y,q_1) + \epsilon^2 R^b(t,s,y,q_1)\right) \end{aligned} \tag{3.259}$$

$\square$

Proposition 3.27. ***(Optimal controls [ElAA3, Section 5.1]).*** *Let all assumptions from Proposition 3.26 hold. Suppose $\beta = 0.5$ in (3.37), so order arrival intensity is of the form $\lambda(\delta) = A\left(B+\delta^2\right)^{-\gamma}$. Suppose that optimal controls for risk-neutral market maker from previous subsection satisfy*

$$\hat{\delta}^a_L = \frac{-\gamma M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}}{2\gamma-1},\ \hat{\delta}^b_L = \frac{\gamma M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma-1)\,B}}{2\gamma-1} \tag{3.260}$$

Then optimal controls for risk-averse market maker satisfy

$$\hat{\delta}^a_\epsilon = \frac{-\gamma M^a + \sqrt{\gamma^2 (M^a)^2 + (2\gamma-1)\,B}}{2\gamma-1} + O(\epsilon^2),\ \hat{\delta}^b_\epsilon = \frac{\gamma M^b + \sqrt{\gamma^2 (M^b)^2 + (2\gamma-1)\,B}}{2\gamma-1} + O(\epsilon^2) \tag{3.261}$$

Linear approximation of optimal controls for risk-averse market maker is

$$\begin{aligned} \hat{\delta}^a(\epsilon) &= \hat{\delta}^a_L + \epsilon\left(-\tfrac{\gamma M_1}{2\gamma-1} + \tfrac{\gamma^2 M_1 M_0}{(2\gamma-1)\sqrt{\gamma^2 M_0^2+(2\gamma-1)B}}\right) + O(\epsilon^2) \\ \hat{\delta}^b(\epsilon) &= \hat{\delta}^b_L + \epsilon\left(\tfrac{\gamma M_2}{2\gamma-1} + \tfrac{\gamma^2 M_2 M_0}{(2\gamma-1)\sqrt{\gamma^2 M_0^2+(2\gamma-1)B}}\right) + O(\epsilon^2) \end{aligned} \tag{3.262}$$

where $M^a = M_0+\epsilon M_1$ and $M^b = M_0+\epsilon M_2$. Note that we drop arguments from M functions for simplicity.

Proof. Using expressions for optimal controls in Proposition 3.26

$$\frac{\partial \lambda^a(\delta^a)}{\partial \delta^a} = \frac{\partial}{\partial \delta^a}\left[\frac{A}{(B+(\delta^a)^2)^\gamma}\right] = A\frac{\partial}{\partial \delta^a}\left[\frac{1}{(B+(\delta^a)^2)^\gamma}\right] = -A2\gamma\delta^a(B+(\delta^a)^2)^{-\gamma-1} \tag{3.263}$$

By Proposition 3.14, $\gamma > 1$, so $\gamma - 1 > 0$, then

$$\frac{\partial \lambda^a(\delta^a)}{\partial \delta^a} = -A2\gamma\delta^a(B+(\delta^a)^2)^{-\gamma-1} = -\frac{A2\gamma\delta^a}{(B+(\delta^a)^2)^{\gamma+1}} = -2\gamma\delta^a\lambda^a(\delta^a)\frac{1}{B+(\delta^a)^2} \tag{3.264}$$

then

$$\frac{\lambda^a(\delta^a)}{\partial \lambda^a(\delta^a)/\partial \delta^a} = -\frac{\lambda^a(\delta^a)}{2\gamma\delta^a\lambda^a(\delta^a)\frac{1}{B+(\delta^a)^2}} = -\frac{B+(\delta^a)^2}{2\gamma\delta^a} = -\left(\frac{B}{2\gamma\delta^a} + \frac{(\delta^a)^2}{2\gamma\delta^a}\right) = -\left(\frac{B}{2\gamma\delta^a} + \frac{\delta^a}{2\gamma}\right) \tag{3.265}$$

Substituting $\frac{\lambda^a(\delta^a)}{\partial\lambda^a(\delta^a)/\partial\delta^a}$ into (3.246) for δ^a, we get

$$\begin{aligned}
\delta^a &= -\tfrac{\lambda^a(\delta^a)}{\partial\lambda^a(\delta^a)/\partial\delta^a} - \left(M^a + \epsilon^2 R^a\right) \\
\delta^a &= \tfrac{B}{2\gamma\delta^a} + \tfrac{\delta^a}{2\gamma} - \left(M^a + \epsilon^2 R^a\right) \\
2\gamma(\delta^a)^2 &= B + (\delta^a)^2 - 2\gamma\delta^a \left(M^a + \epsilon^2 R^a\right) \\
2\gamma(\delta^a)^2 - (\delta^a)^2 &= B - 2\gamma\delta^a \left(M^a + \epsilon^2 R^a\right) \\
(\delta^a)^2 (2\gamma - 1) &= B - 2\gamma\delta^a \left(M^a + \epsilon^2 R^a\right)
\end{aligned} \tag{3.266}$$

We have a quadratic equation

$$(2\gamma - 1)(\delta^a)^2 + 2\gamma\left(M^a + \epsilon^2 R^a\right)\delta^a - B = 0 \tag{3.267}$$

Solving for δ^a we have two points at which $\delta^a = 0$

$$\begin{aligned}
\delta^a_+ &= \tfrac{-\gamma\left(M^a+\epsilon^2 R^a\right)+\sqrt{\gamma^2(M^a+\epsilon^2R^a)^2+(2\gamma-1)B}}{2\gamma-1} \\
\delta^a_- &= \tfrac{-\gamma\left(M^a+\epsilon^2 R^a\right)-\sqrt{\gamma^2(M^a+\epsilon^2R^a)^2+(2\gamma-1)B}}{2\gamma-1}
\end{aligned} \tag{3.268}$$

We denote the point at which maximum is attained on f^a by $\hat{\delta}^a$, since $\delta^a_+ > 0$ that point would be $\hat{\delta}^a = \delta^a_+$, which we will denote by $\hat{\delta}^a_\epsilon$. Then

$$\hat{\delta}^a_\epsilon = -\frac{\gamma\left(M^a + \epsilon^2 R^a\right)}{2\gamma - 1} + \frac{\sqrt{\gamma^2\left(M^a + \epsilon^2 R^a\right)^2 + (2\gamma - 1)B}}{2\gamma - 1} \tag{3.269}$$

Recall from Definition 3.9 that $B > 0$ and $\gamma > 1$. With M^a, R^a fixed these is $C > 0$ such that

$$\left|-\frac{\gamma\left(M^a + \epsilon^2 R^a\right)}{2\gamma - 1}\right| \leq C\left|\epsilon^2\right| \tag{3.270}$$

Which by Definition 1.92 means that

$$-\frac{\gamma\left(M^a + \epsilon^2 R^a\right)}{2\gamma - 1} = O(\epsilon^2) \tag{3.271}$$

With M^a, R^a fixed there are $C_0, C_1, C_2 > 0$ such that

$$\sqrt{\gamma^2\left(M^a + \epsilon^2 R^a\right)^2 + (2\gamma - 1)B} \leq C_0\sqrt{1 + C_1\epsilon^2 + C_2\epsilon^4} \tag{3.272}$$

Next we will use Definition 1.95 to investigate the right hand side of (3.272), where

$$\lim_{\epsilon\to 0}\frac{C_0\sqrt{1 + C_1\epsilon^2 + C_2\epsilon^4}}{C_0\left(1 + \frac{1}{2}C_1\epsilon^2 + C_2\epsilon^4\right)} = \frac{C_0\sqrt{1}}{C_0(1)} = 1,\ \lim_{\epsilon\to 0}\frac{C_0\sqrt{1 + C_1\epsilon^2 + C_2\epsilon^4}}{C_0\left(1 + \frac{1}{2}C_1\epsilon^2\right)} = \frac{C_0\sqrt{1}}{C_0(1)} = 1 \tag{3.273}$$

Then as $\epsilon \to 0$

$$\sqrt{\gamma^2\left(M^a + \epsilon^2 R^a\right)^2 + (2\gamma - 1)B} \leq C_0\left(1 + \frac{1}{2}C_1\epsilon^2\right) \tag{3.274}$$

Which by Definition 1.92 means that

$$\sqrt{\gamma^2\left(M^a + \epsilon^2 R^a\right)^2 + (2\gamma - 1)B} = O(\epsilon^2) \tag{3.275}$$

Hence

$$\hat{\delta}^a = \frac{-\gamma M^a + \sqrt{\gamma^2 (M^a)^2 + (2\gamma - 1)B}}{2\gamma - 1} + O(\epsilon^2) \tag{3.276}$$

Note that when $\epsilon = 0$, optimal spreads for risk-averse market marker $\hat{\delta}^a$ reduce to risk-neutral ones $\hat{\delta}^a_L$, as $M^a = M_0 + \epsilon M_1$ and so

$$\hat{\delta}^a_0 = \frac{-\gamma M_0 + \sqrt{\gamma^2 M_0^2 + (2\gamma - 1)B}}{2\gamma - 1} = \hat{\delta}^a_L \tag{3.277}$$

Applying first order Taylor expansion from Definition 1.101 on function $\hat{\delta}^a_\epsilon$ as $\epsilon \to 0$

$$\hat{\delta}^a_\epsilon = \hat{\delta}^a_0 + \epsilon \left. \frac{\partial \hat{\delta}^a_\epsilon}{\partial \epsilon} \right|_{\epsilon=0} + O(\epsilon^2) \tag{3.278}$$

Differentiation of $\hat{\delta}^a$ with respect to ϵ yields

$$\begin{aligned} \frac{\partial \hat{\delta}^a_\epsilon}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon} \left[\frac{-\gamma\big(M_0 + \epsilon M_1 + \epsilon^2 R^a\big) + \sqrt{\gamma^2 (M_0 + \epsilon M_1 + \epsilon^2 R^a)^2 + (2\gamma - 1)B}}{2\gamma - 1} \right] \\ &= \frac{1}{2\gamma - 1} \left(-\gamma \frac{\partial}{\partial \epsilon} \left[M_0 + \epsilon M_1 + \epsilon^2 R^a \right] + \frac{\partial}{\partial \epsilon} \left[\sqrt{\gamma^2 \left(M_0 + \epsilon M_1 + \epsilon^2 R^a \right)^2 + (2\gamma - 1) B} \right] \right) \\ &= \frac{1}{2\gamma - 1} \left(-\gamma \left(2\epsilon R^a + M_1 \right) + \frac{M_1 \gamma^2 (M_0 + \epsilon M_1)}{\sqrt{\gamma^2 (M_0 + \epsilon M_1)^2 + (2\gamma - 1)B}} \right) \end{aligned} \tag{3.279}$$

Evaluating $\frac{\partial \hat{\delta}^a_\epsilon}{\partial \epsilon}$ at $\epsilon = 0$

$$\begin{aligned} \left. \frac{\partial \hat{\delta}^a_\epsilon}{\partial \epsilon} \right|_{\epsilon=0} &= \frac{1}{2\gamma - 1} \left(-\gamma M_1 + \frac{\gamma^2 M_0 M_1}{\sqrt{\gamma^2 M_0^2 + (2\gamma - 1)B}} \right) \\ &= -\frac{\gamma M_1}{2\gamma - 1} + \frac{\gamma^2 M_1 M_0}{(2\gamma - 1)\sqrt{\gamma^2 M_0^2 + (2\gamma - 1)B}} \end{aligned} \tag{3.280}$$

Substituting evaluated derivative into (3.278) and using (3.277) we get a linear approximation of $\hat{\delta}^a$

$$\hat{\delta}^a_\epsilon = \hat{\delta}^a_L + \epsilon \left(-\frac{\gamma M_1}{2\gamma - 1} + \frac{\gamma^2 M_1 M_0}{(2\gamma - 1)\sqrt{\gamma^2 M_0^2 + (2\gamma - 1)B}} \right) + O(\epsilon^2) \tag{3.281}$$

By same methods as above, we can obtain expression for optimal controls $\hat{\delta}^b$ and their linear approximation.

□

Proposition 3.28. ***(Solution [ElAA3, Section 5.1]).*** *Let all assumptions from Propositions 3.24 and 3.27 hold. Solution to HJB equation (3.231) is*

$$\begin{aligned} w^\epsilon(t, s, y, q_1, x) &= x + \phi_0(t, s, y, q_1) + \epsilon \phi_1(t, s, y, q_1) \\ \phi_0(t, s, y, q_1) &= \theta_0(t, s, y) + q_1 \theta_1(t, s, y) \\ \phi_1(t, s, y, q_1) &= \theta_2(t, s, y) + q_1 \theta_3(t, s, y) + q_1^2 \theta_4(t, s, y) \end{aligned} \tag{3.282}$$

where

$$\begin{aligned}
\phi_0(t,s,y,q_1) &= \theta_0(t,s,y) + q_1\theta_1(t,s,y) \\
\phi_1(t,s,y,q_1) &= \theta_2(t,s,y) + q_1\theta_3(t,s,y) + q_1^2\theta_4(t,s,y) \\
\theta_4(t,s,y) &= -\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T K_u du\right) \\
\theta_3(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_{1,1}(u,S_u,Y_u)du\right) \\
\theta_2(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \left(\mathcal{J}_{1,0}(u,S_u,Y_u) - L_u\right) du\right) \\
\mathcal{J}_{1,1}(t,s,y) &= -2\theta_4(t,s,y)\left(\lambda^a(\hat{\delta}^a_L) - \lambda^b(\hat{\delta}^b_L)\right) \\
\mathcal{J}_{1,0}(t,s,y) &= \lambda^a(\hat{\delta}^a_L)\left(-\theta_3(t,s,y) + \theta_4(t,s,y)\right) - \lambda^b(\hat{\delta}^b_L)\left(-\theta_3(t,s,y) - \theta_4(t,s,y)\right)
\end{aligned} \tag{3.283}$$

Proof. To ease up notation we denote

$$\delta^a_\epsilon = \hat{\delta}^a_\epsilon, \ \delta^a_L = \hat{\delta}^a_L \tag{3.284}$$

Recall functions f^a_0, f^a from proofs of Propositions 3.17 and 3.26, where

$$\begin{aligned}
f^a_0(\delta^a_L) &= \lambda^a(\delta^a_L)\left(\delta^a_L + M_0\right) \\
f^a(\delta^a_\epsilon, \epsilon) &= \lambda^a(\delta^a_\epsilon)\left(\delta^a_\epsilon + M_0 + \epsilon M_1 + \epsilon^2 R^a\right) \\
&= f^a_0(\delta^a_\epsilon) + \epsilon\lambda^a(\delta^a_\epsilon)M_1 + \epsilon^2\lambda^a(\delta^a_\epsilon)R^a \\
f^a(\delta^a_\epsilon, 0) &= f^a_0(\delta^a_\epsilon)
\end{aligned} \tag{3.285}$$

Applying Taylor expansion from Definition 1.101 on function $f^a(\delta^a_\epsilon, \epsilon)$ as $\epsilon \to 0$

$$\begin{aligned}
f^a(\delta^a_\epsilon, \epsilon) &= f^a(\delta^a_\epsilon, 0) + \epsilon \left.\tfrac{\partial f^a(\delta^a_\epsilon, \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} + O(\epsilon^2) \\
&= f^a_0(\delta^a_\epsilon) + \epsilon \left.\tfrac{\partial f^a(\delta^a_\epsilon, \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} + O(\epsilon^2)
\end{aligned} \tag{3.286}$$

where

$$\begin{aligned}
\tfrac{\partial f^a(\delta^a_\epsilon, \epsilon)}{\partial \epsilon} &= \tfrac{\partial}{\partial \epsilon}\left[f^a_0(\delta^a_\epsilon) + \epsilon\lambda^a(\delta^a_\epsilon)M_1 + \epsilon^2\lambda^a(\delta^a_\epsilon)R^a\right] \\
&= \tfrac{\partial}{\partial \epsilon}\left[\epsilon\lambda^a(\delta^a_\epsilon)M_1\right] + \tfrac{\partial}{\partial \epsilon}\left[\epsilon^2\lambda^a(\delta^a_\epsilon)R^a\right] \\
&= \lambda^a(\delta^a_\epsilon)M_1 + 2\epsilon\lambda^a(\delta^a_\epsilon)R^a
\end{aligned} \tag{3.287}$$

Evaluating $\frac{\partial f^a(\delta^a_\epsilon, \epsilon)}{\partial \epsilon}$ at $\epsilon = 0$

$$\left.\frac{\partial f^a(\delta^a_\epsilon, \epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = \lambda^a(\delta^a_\epsilon)M_1 + 2\epsilon\lambda^a(\delta^a_\epsilon)R^a|_{\epsilon=0} = \lambda^a(\delta^a_\epsilon)M_1 \tag{3.288}$$

Substituting evaluated derivative into (3.286) and we get

$$f^a(\delta^a_\epsilon, \epsilon) = f^a_0(\delta^a_\epsilon) + \epsilon\lambda^a(\delta^a_\epsilon)M_1 + O(\epsilon^2) \tag{3.289}$$

Applying Taylor expansion from Definition 1.101 on function $f^a(\delta^a, \epsilon)$ as $\delta^a \to \delta^a_L$

$$f^a(\delta^a_\epsilon, \epsilon) = f^a(\delta^a_L, \epsilon) + (\delta^a_\epsilon - \delta^a_L)\left.\frac{\partial f^a(\delta^a_\epsilon, \epsilon)}{\partial \delta^a_\epsilon}\right|_{\delta^a_\epsilon = \delta^a_L} + O((\delta^a_\epsilon - \delta^a_L)^2) \tag{3.290}$$

Substituting for $f^a(\delta^a_L, \epsilon)$ from (3.289) and for $(\delta^a - \delta^a_L)$ from linear approximation in (3.262) into (3.290), we get

$$\begin{aligned} f^a(\delta^a_\epsilon,\epsilon) &= f^a_0(\delta^a_L)+\epsilon\lambda^a(\delta^a_L)M_1 \\ &+ \epsilon\left(-\tfrac{\gamma M_1}{2\gamma-1}+\tfrac{\gamma^2 M_0 M_1}{(2\gamma-1)\sqrt{\gamma^2M_0^2+(2\gamma-1)B}}\right)\tfrac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}\Big|_{\delta^a_\epsilon=\delta^a_L} \\ &+ O(\epsilon^2+(\delta^a_\epsilon-\delta^a_L)^2) \end{aligned} \tag{3.291}$$

Applying Taylor expansion from Definition 1.101 on $\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}$ as $\epsilon\to 0$

$$\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}=\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}\bigg|_{\epsilon=0}+O(\epsilon) \tag{3.292}$$

where

$$\begin{aligned} \tfrac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon} &= \tfrac{\partial}{\partial\delta^a_\epsilon}\left[f^a_0(\delta^a_\epsilon)+\epsilon\lambda^a(\delta^a_\epsilon)M_1+\epsilon^2\lambda^a(\delta^a_\epsilon)R^a\right] \\ &= \tfrac{\partial f^a_0(\delta^a_\epsilon)}{\partial\delta^a_\epsilon}+\tfrac{\partial\lambda^a(\delta^a_\epsilon)}{\partial\delta^a_\epsilon}\left(\epsilon M_1+\epsilon^2R^a\right) \end{aligned} \tag{3.293}$$

Evaluating $\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}$ at $\epsilon=0$

$$\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}\bigg|_{\epsilon=0}=\frac{\partial f^a_0(\delta^a_\epsilon)}{\partial\delta^a_\epsilon} \tag{3.294}$$

Substituting evaluated derivative into (3.292) and we get

$$\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}=\frac{\partial f^a_0(\delta^a_\epsilon)}{\partial\delta^a_\epsilon}+O(\epsilon) \tag{3.295}$$

Recall that δ^a_L is optimal control from Proposition 3.18, so

$$\nabla\left(f^a_0(\delta^a_L)\right)=\frac{\partial f^a_0(\delta^a)}{\partial\delta^a}\bigg|_{\delta^a=\delta^a_L}=0 \tag{3.296}$$

Hence

$$\frac{\partial f^a(\delta^a_\epsilon,\epsilon)}{\partial\delta^a_\epsilon}\bigg|_{\delta^a_\epsilon=\delta^a_L}=\frac{\partial f^a_0(\delta^a_\epsilon)}{\partial\delta^a_\epsilon}\bigg|_{\delta^a_\epsilon=\delta^a_L}=0 \tag{3.297}$$

Substituting evaluated derivative into (3.291) which becomes

$$f^a(\delta^a_\epsilon,\epsilon)=f^a_0(\delta^a_L)+\epsilon\lambda^a(\delta^a_L)M_1(t,s,y,q_1)+O(\epsilon^2+(\delta^a_\epsilon-\delta^a_L)^2) \tag{3.298}$$

Applying Taylor expansion from Definition 1.101 on $(\delta^a_\epsilon-\delta^a_L)^2$ as $\epsilon\to 0$

$$(\delta^a_\epsilon-\delta^a_L)^2=(\delta^a_\epsilon-\delta^a_L)^2\big|_{\epsilon=0}+\epsilon\,\frac{\partial}{\partial\epsilon}\left[(\delta^a_\epsilon-\delta^a_L)^2\right]\bigg|_{\epsilon=0}+O(\epsilon^2) \tag{3.299}$$

Note that from (3.262) we have that

$$\delta^a_\epsilon-\hat{\delta}^a_L=\epsilon\left(-\frac{\gamma M_1}{2\gamma-1}+\frac{\gamma^2M_0M_1}{(2\gamma-1)\sqrt{\gamma^2M_0^2+(2\gamma-1)B}}\right)+O(\epsilon^2) \tag{3.300}$$

Hence

$$(\delta^a_\epsilon-\delta^a_L)^2=\epsilon^2\left(-\frac{\gamma M_1}{2\gamma-1}+\frac{\gamma^2M_0M_1}{(2\gamma-1)\sqrt{\gamma^2M_0^2+(2\gamma-1)B}}\right)^2+O(\epsilon^4) \tag{3.301}$$

Evaluating $(\delta_\epsilon^a - \delta_L^a)^2$ at $\epsilon = 0$, we get

$$(\delta_\epsilon^a - \delta_L^a)^2\Big|_{\epsilon=0} = 0 \tag{3.302}$$

Taking derivative of $(\delta_\epsilon^a - \delta_L^a)^2$ with respect to ϵ

$$\begin{aligned}\frac{\partial}{\partial\epsilon}\left[(\delta_\epsilon^a - \delta_L^a)^2\right] &= \frac{\partial}{\partial\epsilon}\left[\epsilon^2\left(-\frac{\gamma M_1}{2\gamma-1} + \frac{\gamma^2 M_0 M_1}{(2\gamma-1)\sqrt{\gamma^2 M_0^2+(2\gamma-1)B}}\right)^2\right] \\ &= 2\epsilon\left(-\frac{\gamma M_1}{2\gamma-1} + \frac{\gamma^2 M_0 M_1}{(2\gamma-1)\sqrt{\gamma^2 M_0^2+(2\gamma-1)B}}\right)^2\end{aligned} \tag{3.303}$$

Evaluating $\frac{\partial}{\partial\epsilon}\left[(\delta_\epsilon^a - \delta_L^a)^2\right]$ at $\epsilon = 0$, we get

$$\frac{\partial}{\partial\epsilon}\left[(\delta_\epsilon^a - \delta_L^a)^2\right]\Bigg|_{\epsilon=0} = 0 \tag{3.304}$$

Substituting evaluated derivatives into (3.299), we obtain

$$(\delta_\epsilon^a - \delta_L^a)^2 = O(\epsilon^2) \tag{3.305}$$

Therefore (3.298) is

$$f^a(\delta_\epsilon^a, \epsilon) = f_0^a(\delta_L^a) + \epsilon\lambda^a(\delta_L^a)M_1(t,s,y,q_1) + O(\epsilon^2) \tag{3.306}$$

By same methods as before we can obtain

$$f^b(\delta_\epsilon^a, \epsilon) = f_0^b(\delta_L^b) - \epsilon\lambda(\delta_L^b)M_2(t,s,y,q_1) + O(\epsilon^2) \tag{3.307}$$

Recall from (3.232) that $\mathcal{J}^\epsilon$ is defined as

$$\begin{aligned}\mathcal{J}^\epsilon(t,s,y,\vartheta) &= \lambda^a(\delta^a)\left(w^\epsilon(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w^\epsilon(t,s,y,q_1,x)\right) \\ &+ \lambda^b(\delta^b)\left(w^\epsilon(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w^\epsilon(t,s,y,q_1,x)\right)\end{aligned} \tag{3.308}$$

Recall functions f^a, f^b from proof of Proposition 3.26 are defined as

$$\begin{aligned}f^a(\delta_\epsilon^a, \epsilon) &= \lambda^a(\delta_\epsilon^a)\left(w^\epsilon(t,s,y,q_1-1,x+C_{\mathbb{Q}}(t,s,y)+\delta^a) - w^\epsilon(t,s,y,q_1,x)\right) \\ f^b(\delta_\epsilon^b, \epsilon) &= \lambda^a(\delta_\epsilon^b)\left(w^\epsilon(t,s,y,q_1+1,x-C_{\mathbb{Q}}(t,s,y)+\delta^b) - w^\epsilon(t,s,y,q_1,x)\right)\end{aligned} \tag{3.309}$$

Hence

$$\mathcal{J}^\epsilon(t,s,y,\hat\vartheta) = f^a(\delta_\epsilon^a, \epsilon) + f^b(\delta_\epsilon^b, \epsilon) \tag{3.310}$$

Substituting expressions for f^a, f^a from (3.306) and (3.307), we get

$$\begin{aligned}\mathcal{J}^\epsilon(t,s,y,\hat\vartheta) &= f_0^b(\delta_L^b) + f_0^a(\delta_L^a) + \epsilon\lambda^a(\delta_L^a)M_1(t,s,y,q_1) - \epsilon\lambda^b(\delta_L^b)M_2(t,s,y,q_1) + O(\epsilon^2) \\ &= \mathcal{J}_0(t,s,y) + \epsilon\left(\lambda^a(\delta_L^a)M_1(t,s,y,q_1) - \lambda^b(\delta_L^b)M_2(t,s,y,q_1)\right) + O(\epsilon^2)\end{aligned} \tag{3.311}$$

Then expressing $\mathcal{J}^\epsilon$ in powers of ϵ

$$\mathcal{J}^\epsilon(t,s,y,\hat\vartheta) = \mathcal{J}_0(t,s,y) + \epsilon\mathcal{J}_1(t,s,y,q_1) + O(\epsilon^2) \tag{3.312}$$

where

$$\mathcal{J}_1(t,s,y,q_1) = \lambda^a(\delta_L^a)M_1(t,s,y,q_1) - \lambda^b(\delta_L^b)M_2(t,s,y,q_1) \tag{3.313}$$

Recall from Proposition 3.26 that

$$\begin{aligned} M_1(t,s,y,q_1) &= -\theta_3(t,s,y) + (1-2q_1)\theta_4(t,s,y) \\ M_2(t,s,y,q_1) &= -\theta_3(t,s,y) - (1+2q_1)\theta_4(t,s,y) \end{aligned} \tag{3.314}$$

Let $\theta_3 = \theta_3(t,s,y)$ and $\theta_4 = \theta_4(t,s,y)$, then

$$\begin{aligned} \mathcal{J}_1(t,s,y,q_1) &= \lambda^a(\delta_L^a)M_1(t,s,y,q_1) - \lambda^b(\delta_L^b)M_2(t,s,y,q_1) \\ &= \lambda^a(\delta_L^a)\left(-\theta_3 + (1-2q_1)\theta_4\right) - \lambda^b(\delta_L^b)\left(-\theta_3 - (1+2q_1)\theta_4\right) \\ &= \lambda^a(\delta_L^a)\left(-\theta_3 + \theta_4 - 2q_1\theta_4\right) - \lambda^b(\delta_L^b)\left(-\theta_3 - \theta_4 - 2q_1\theta_4\right) \\ &= \lambda^a(\delta_L^a)\left(-\theta_3 + \theta_4\right) - \lambda^b(\delta_L^b)\left(-\theta_3 - \theta_4\right) - q_1 2\theta_4\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right) \end{aligned} \tag{3.315}$$

Let

$$\begin{aligned} \mathcal{J}_{1,0}(t,s,y) &= \lambda^a(\delta_L^a)\left(-\theta_3 + \theta_4\right) - \lambda^b(\delta_L^b)\left(-\theta_3 - \theta_4\right) \\ \mathcal{J}_{1,1}(t,s,y) &= -2\theta_4\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right) \end{aligned} \tag{3.316}$$

Then

$$\mathcal{J}_1(t,s,y,q_1) = \mathcal{J}_{1,0}(t,s,y) + q_1\mathcal{J}_{1,1}(t,s,y) \tag{3.317}$$

Recall from (3.282), we are searching for a solution of the form

$$\begin{aligned} w^\epsilon(t,s,y,q_1,x) &= x + \phi_0(t,s,y,q_1) + \epsilon\phi_1(t,s,y,q_1) \\ \phi_0(t,s,y,q_1) &= \theta_0(t,s,y) + q_1\theta_1(t,s,y) \\ \phi_1(t,s,y,q_1) &= \theta_2(t,s,y) + q_1\theta_3(t,s,y) + q_1^2\theta_4(t,s,y) \end{aligned} \tag{3.318}$$

Using terminal condition of (3.231)

$$w^\epsilon(T,s,y,q_1,x) = x + \phi_0(T,s,y,q_1) + \epsilon\phi_1(T,s,y,q_1) = x + q_1 h(s) \tag{3.319}$$

we deduce that

$$\phi_0(T,s,y,q_1) = q_1 h(s), \;\; \phi_1(T,s,y,q_1) = 0 \tag{3.320}$$

By perturbation method from Definition 1.104, we split PDE (3.231) into two PDEs on basis of powers of ϵ

$$\begin{aligned} &\epsilon^0 \quad \begin{cases} \frac{\partial\phi_0}{\partial t} + \mathcal{L}_1(x+\phi_0) + \mathcal{L}_2(x+\phi_0) + \mathcal{J}_0(t,s,y) = 0 \\ \phi_0(T,s,y,q_1) = q_1 h(s) \end{cases} \\ &\epsilon^1 \quad \begin{cases} \frac{\partial\phi_1}{\partial t} + \mathcal{L}_1\phi_1 + \mathcal{L}_2\phi_1 + \mathcal{J}_1(t,s,y,q_1) = q_1^2 k + l \\ \phi_1(T,s,y,q_1) = 0 \end{cases} \end{aligned} \tag{3.321}$$

We then further split each PDE in (3.321) on basis of powers of q_1. Recall that we already solved PDE for ϵ^0 in Proposition 3.19 where

$$\begin{aligned}\theta_1(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(h(S_T)\right) - \mu\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \Delta_u S_u du\right)\\ \theta_0(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_0(u,S_u,Y_u)du\right)\end{aligned} \tag{3.322}$$

We need to solve PDE for ϵ^1, where we again split it on basis of powers of q_1

$$\begin{aligned}
&q_1^2 \quad \begin{cases}\frac{\partial\theta_4}{\partial t} + \mathcal{L}_1\theta_4 - k = 0\\ \theta_4(T,s,y) = 0\end{cases}\\
&q_1^1 \quad \begin{cases}\frac{\partial\theta_3}{\partial t} + \mathcal{L}_1\theta_3 + \mathcal{J}_{1,1}(t,s,y) = 0\\ \theta_3(T,s,y) = 0\end{cases}\\
&q_1^0 \quad \begin{cases}\frac{\partial\theta_2}{\partial t} + \mathcal{L}_1\theta_2 + \mathcal{J}_{1,0}(t,s,y) - l = 0\\ \theta_2(T,s,y) = 0\end{cases}
\end{aligned} \tag{3.323}$$

Recall from Proposition 3.24 that

$$\begin{aligned} l &= \left(C^2_{\mathbb{Q}}(t,s,y) + \mathbb{E}^{\mathbb{P}}_{t,s,y,q_1,x}\left(h^2(S_T)\right) - 2C_{\mathbb{Q}}(t,s,y)C_{\mathbb{P}}(t,s,y)\right)\left(\lambda^a(\delta^a_L) + \lambda^b(\delta^b_L)\right)\\ k &= \Delta^2\sigma^2(y)s^2\end{aligned} \tag{3.324}$$

Note that S, Y are both Ito processes, so by Feynman-Kac representation from Theorem 1.91, we get solutions

$$\begin{aligned}\theta_4(t,s,y) &= -\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T K_u du\right)\\ \theta_3(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_{1,1}(u,S_u,Y_u)du\right)\\ \theta_2(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \left(\mathcal{J}_{1,0}(u,S_u,Y_u) - L_u\right)du\right)\end{aligned} \tag{3.325}$$

Therefore solution to HJB equation (3.231) is

$$\begin{aligned} w^\epsilon(t,s,y,q_1,x) &= x + \phi_0(t,s,y,q_1) + \epsilon\phi_1(t,s,y,q_1)\\ \phi_0(t,s,y,q_1) &= \theta_0(t,s,y) + q_1\theta_1(t,s,y)\\ \phi_1(t,s,y,q_1) &= \theta_2(t,s,y) + q_1\theta_3(t,s,y) + q_1^2\theta_4(t,s,y)\end{aligned} \tag{3.326}$$

□

Proposition 3.29. ***(Explicit expressions).*** *Let all assumptions from Propositions 3.24 and 3.27 hold. Let $\tau = (T-t)$ and assume that*

$$\mathcal{J}_{1,1}(t,s,y) = -2\theta_4(t,s,y)\left(\lambda^a(\delta^a_L) - \lambda^b(\delta^b_L)\right),\ K_t = \Delta^2_t\sigma^2(Y_t)S^2_t \tag{3.327}$$

Suppose that functions θ_3, θ_4 from Proposition 3.28 satisfy

$$\theta_3(t,s,y) = \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_{1,1}(u,S_u,Y_u)du\right),\ \theta_4(t,s,y) = -\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T K_u du\right) \tag{3.328}$$

Then explicit expressions for θ_3, θ_4 are

$$\theta_3(t,s,y) = 2\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right)\tau^2\Delta^2\sigma^2(y)s^2, \;\; \theta_4(t,s,y) = -\tau\Delta^2\sigma^2(y)s^2 \tag{3.329}$$

Proof. Let $\tau = T - t$, then by application of mean value Theorem 5.4 and utilisation of initial conditions (3.20) we get

$$\begin{aligned} \theta_4(t,s,y) &= -\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T K_u du\right) \\ &= -\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \Delta_u^2\sigma^2(Y_u)S_u^2 du\right) \\ &= -\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\tau\Delta_t^2\sigma^2(Y_t)S_t^2\right) \\ &= -\tau\Delta^2\sigma^2(y)s^2 \end{aligned} \tag{3.330}$$

Substituting for θ_4 and by the same methods we obtain

$$\begin{aligned} \theta_3(t,s,y) &= \mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \mathcal{J}_{1,1}(u,S_u,Y_u)du\right) \\ &= -2\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right)\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\int_t^T \theta_4(u,S_u,Y_u)du\right) \\ &= -2\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right)\mathbb{E}^{\mathbb{P}}_{t,s,y}\left(\tau\theta_4(t,S_t,Y_t)du\right) \\ &= -2\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right)\tau\theta_4(t,s,y) \\ &= 2\left(\lambda^a(\delta_L^a) - \lambda^b(\delta_L^b)\right)\tau^2\Delta^2\sigma^2(y)s^2 \end{aligned} \tag{3.331}$$

□

Proposition 3.30. ***(Verification [ElAA3, Section 5]).*** *Let $\mathcal{S} = [t,T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$ and $\bar{\mathcal{S}} = [t,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{N} \times \mathbb{R}$. Let w^ϵ be the solution to HJB equation (3.231) defined by (3.282) and let v^ϵ be the value function from Remark 3.23. Then*

$$w^\epsilon(t,s,y,q_1,x) = v^\epsilon(t,s,y,q_1,x) \text{ for } (t,s,y,q_1,x) \in \bar{\mathcal{S}} \tag{3.332}$$

Proof. The first step in proving this proposition is to verify if ϕ_1 is bounded, specifically functions $\theta_2, \theta_3, \theta_4$. As for ϕ_0 have already shown that $w(t,s,y,q_1,x) = \theta_0(t,s,y) + q_1\theta_1(t,s,y)$ is bounded in the proof of Proposition 3.20. Rest of the proof of this proposition will follow the same methods as in proof of Proposition 3.20. □

3.6 Heston model

Definition 3.31. (State process [ElAA3, Chapter 6]). Refer to Subsection 3.1 for notations. We introduce new assumptions. Let $S = (S_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process, which we call spot price process and it is assumed to satisfy

$$\begin{cases} dS_T &= rS_T dT + S_T\sqrt{Y_T}d\widetilde{W}_T^{(1)} \\ S_t &= s \text{ a.s.} \end{cases} \tag{3.333}$$

Let $Y = (Y_s)_{s\in[t,T]}$ be an $\mathbb{R}$-valued process, which we call variance process and it is assumed to satisfy

$$\begin{cases} dY_T &= -\kappa_I(Y_T - \theta_I)dT + \eta_I\sqrt{Y_T}d\widetilde{W}_T^{(2)} \\ Y_t &= y \text{ a.s.} \end{cases} \tag{3.334}$$

We also impose following assumption to ensure that Y is strictly positive

$$2\kappa_I\theta_I > \eta^2 \tag{3.335}$$

Proposition 3.32. ***(Valuation equation [Ga, Chapter 1]).*** *Let S, Y be processes from Definition 3.31. Let V, V_1 be continious and $\mathbb{R}$-valued functions, denoting european options price. Suppose that portfolio value Π is defined by*

$$\Pi_t = V(t, S_t, Y_t) - \Delta S_t - \nu V_1(t, S_t, Y_t) \tag{3.336}$$

where V and V_1 denote european options with expiry at time T. Suppose there exist $\nu, \Delta \in \mathbb{R}$ such that following identities hold

$$\frac{\partial V}{\partial S} - \nu\frac{\partial V_1}{\partial S} - \Delta = 0, \ \frac{\partial V}{\partial Y} - \nu\frac{\partial V_1}{\partial Y} = 0 \tag{3.337}$$

These identities correspond to delta and vega heding of portfolio Π. Specifically, hedging portfolio against changes in spot price S and variance Y. Moreover, assume that portfolio return must be equal to risk-free rate r, i.e. following identity is satisfied

$$\frac{d\Pi_T}{\Pi_t} = \int_t^T r du \tag{3.338}$$

Then V satisfies following PDE, called valuation equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}s^2 y\frac{\partial^2 V}{\partial s^2} + \rho\eta y s\frac{\partial^2 V}{\partial s\partial y} + \frac{1}{2}\eta^2 y\frac{\partial^2 V}{\partial y^2} + rs\frac{\partial V}{\partial s} - rV - \kappa(y-\theta)\frac{\partial V}{\partial y} = 0 \tag{3.339}$$

Proof. We will outline proof used in [Ga, Chapter 1]. Firstly, we apply the Ito Lemma from Theorem 1.61 to function V over time interval t, T. Then use hedging relations from (3.337) to deduce that stochastic integrals are equal to zero, where we end up with

$$\begin{aligned} d\Pi_T &= \int_t^T \left[\tfrac{\partial V}{\partial u} + \tfrac{1}{2}S_u^2 Y_u\tfrac{\partial^2 V}{\partial S^2} + \tfrac{1}{2}\eta^2 Y_u\tfrac{\partial^2 V}{\partial Y^2} + \rho\eta Y_u S_u\tfrac{\partial^2 V}{\partial S\partial Y}\right] du \\ &- \nu\int_t^T \left[\tfrac{\partial V_1}{\partial u} + \tfrac{1}{2}S_u^2 Y_u\tfrac{\partial^2 V_1}{\partial S^2} + \tfrac{1}{2}\eta^2 Y_u\tfrac{\partial^2 V_1}{\partial Y^2} + \rho\eta Y_u S_u\tfrac{\partial^2 V_1}{\partial S\partial Y}\right] du \end{aligned} \tag{3.340}$$

Next, use assumption that portfolio return must be equal to risk-free rate r from (3.338) to get the following

$$\begin{aligned}(V(t,S_t,Y_t)-\Delta S_t-\nu V_1(t,S_t,Y_t))\int_t^T r du &= \int_t^T \left[\frac{\partial V}{\partial u}+\frac{1}{2}S_u^2 Y_u\frac{\partial^2 V}{\partial S^2}+\frac{1}{2}\eta^2 Y_u\frac{\partial^2 V}{\partial Y^2}+\rho\eta Y_u S_u\frac{\partial^2 V}{\partial S\partial Y}\right]du\\ &- \nu\int_t^T \left[\frac{\partial V_1}{\partial u}+\frac{1}{2}S_u^2 Y_u\frac{\partial^2 V_1}{\partial S^2}+\frac{1}{2}\eta^2 Y_u\frac{\partial^2 V_1}{\partial Y^2}+\rho\eta Y_u S_u\frac{\partial^2 V_1}{\partial S\partial Y}\right]du\end{aligned} \tag{3.341}$$

Set $T=t+\tau$ and apply mean value Theorem 5.4 to remove integrals. Use initial conditions from Definition 3.31 for processes S,Y, dive by τ and take limit as $\tau\to 0$ to get a following PDE

$$\begin{aligned}r\left(V(t,s,y)-\Delta s-\nu V_1(t,s,y)\right) &= \frac{\partial V}{\partial t}+\frac{1}{2}s^2 y\frac{\partial^2 V}{\partial s^2}+\frac{1}{2}\eta^2 y\frac{\partial^2 V}{\partial y^2}+\rho\eta y s\frac{\partial^2 V}{\partial s\partial y}\\ &- \nu\left[\frac{\partial V_1}{\partial t}+\frac{1}{2}s^2 y\frac{\partial^2 V_1}{\partial s^2}+\frac{1}{2}\eta^2 y\frac{\partial^2 V_1}{\partial y^2}+\rho\eta y s\frac{\partial^2 V_1}{\partial s\partial y}\right]\end{aligned} \tag{3.342}$$

Substitute for Δ using hedging relations from (3.337), collect V,V_1 terms to opposite sides, substitute for ν using hedging relations from (3.337) and divide both sides by $\frac{\partial V}{\partial y}$, where we end up with the following

$$\begin{aligned}&\frac{1}{\partial V_1/\partial y}\left[\frac{\partial V_1}{\partial t}+\frac{1}{2}s^2 y\frac{\partial^2 V_1}{\partial s^2}+\frac{1}{2}\eta^2 y\frac{\partial^2 V_1}{\partial y^2}+\rho\eta y s\frac{\partial^2 V_1}{\partial s\partial y}+rs\frac{\partial V_1}{\partial s}-rV_1\right] =\\ &\frac{1}{\partial V/\partial y}\left[\frac{\partial C}{\partial t}+\frac{1}{2}s^2 y\frac{\partial^2 V}{\partial s^2}+\frac{1}{2}\eta^2 y\frac{\partial^2 V}{\partial y^2}+\rho\eta y s\frac{\partial^2 V}{\partial s\partial y}+rs\frac{\partial V}{\partial s}-rV\right]\end{aligned} \tag{3.343}$$

The left hand side is a function of V_1 only and the right hand side is a function of V only, for this to hold both sides must be equal to some function f depending on t,s,y. Assume that the function is defined by $f(t,s,y)=\kappa(y-\theta)$. In which case we obtain the valuation equation (3.339). □

Proposition 3.33. ***(Forward valuation equation [Ga, Chapter 2]).*** *Let all assumptions from Proposition 3.32 hold. Let $\tau=T-t$ and $x=\ln(F_{t,T}/K)$ where $F_{t,T}$ is the τ-forward value of stock S,* defined by

$$F_{t,T}=S_t\exp\left(\int_t^T r du\right)=S_t e^{r\tau} \tag{3.344}$$

Suppose we only consider the τ-forward value of option V, denoted by function C and defined by

$$C(\tau,x,y)=V(t,s,y)e^{r\tau} \tag{3.345}$$

Then we can obtain PDE

$$0 = -\frac{\partial C}{\partial \tau}-\frac{1}{2}y\frac{\partial C}{\partial x}+\frac{1}{2}y\frac{\partial^2 C}{\partial x^2}+\rho\eta y\frac{\partial^2 C}{\partial x\partial y}+\frac{1}{2}\eta^2 y\frac{\partial^2 C}{\partial y^2}-\kappa(y-\theta)\frac{\partial C}{\partial y} \tag{3.346}$$

Proof. We will outline proof used in [Ga, Chapter 2]. Perform change of variables on partial derivatives of (3.339) using the chain rule and substitute these into (3.339) to obtain

$$\begin{aligned} 0 &= -\frac{\partial V}{\partial \tau} + \frac{1}{2}s^2 y\left(-\frac{1}{s^2}\frac{\partial V}{\partial x} + \frac{1}{s^2}\frac{\partial^2 V}{\partial x^2}\right) + \rho\eta y s\left(\frac{1}{s}\frac{\partial^2 V}{\partial x \partial y}\right) + \frac{1}{2}\eta^2 y\frac{\partial^2 V}{\partial y^2} + rs\left(\frac{1}{s}\frac{\partial V}{\partial x}\right) - rV - \kappa(y-\theta)\frac{\partial V}{\partial y} \\ &= -\frac{\partial V}{\partial \tau} - \frac{1}{2}y\frac{\partial V}{\partial x} + \frac{1}{2}y\frac{\partial^2 V}{\partial x^2} + \rho\eta y\frac{\partial^2 V}{\partial x \partial y} + \frac{1}{2}\eta^2 y\frac{\partial^2 V}{\partial y^2} + r\left(\frac{\partial V}{\partial x} - V\right) - \kappa(y-\theta)\frac{\partial V}{\partial y} \end{aligned} \tag{3.347}$$

Deduce that considering τ-forward value of stock S and option V defined in (3.345) implies that τ-forward returns of portfolio must be zero and so (3.347) becomes (3.346). □

Proposition 3.34. ***(Ansatz [Ro, Chapter 1: p 4-5]).*** *Let all assumptions from Proposition 3.33 hold. Suppose that V is a call option* defined by

$$V(t,s,y) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left((S_T-K)^+\middle|\mathcal{F}_t\right) \tag{3.348}$$

Then its is τ-forward price satisfies

$$C(\tau,x,y) = K\left(e^x P_1 - P_0\right) \tag{3.349}$$

where $K>0$ is the strike price, P_0 and P_1 represent probability that $S_T > K$, given $S_t = s$ and $Y_t = y$.

Proof. We will outline proof used in [Ro, Chapter 1: p 4-5]. First, apply linearity of expectation on (3.348) to obtain the following

$$\begin{aligned} V(t,s,y) &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left(S_T\mathbf{1}_{\{S_T>K\}}\middle|\mathcal{F}_t\right) - Ke^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left(\mathbf{1}_{\{S_T>K\}}\middle|\mathcal{F}_t\right) \\ &= S_tP_1 - Ke^{-r\tau}P_0 \end{aligned} \tag{3.350}$$

Deduce that $P_0 = \mathbb{Q}\left(S_T > K\middle|\mathcal{F}_t\right)$ and for P_1 perform change of measure using Radon-Nikodym derivative from Theorem 5.23 of risk-neutral measure $\mathbb{Q}$ with respect to numeraire S induced measure $\mathbb{S}$. Use the Radon-Nikodym derivative to deduce that $P_1 = \mathbb{S}\left(S_T > K\middle|\mathcal{F}_t\right)$. And finally, use (3.345) and (3.344) to derive C as in (3.349). □

Proposition 3.35. ***(Probabilities PDEs [Ga, Chapter 2]).*** *Let all assumptions from Propositions 3.33 and 3.34 hold. Assume that solution to (3.346) is of the form*

$$C(\tau,x,y) = K\left(e^x P_1 - P_0\right) \tag{3.351}$$

Suppose that $P_0 = P_0(\tau,x,y)$ and *$P_1 = P_1(\tau,x,y)$. Then P_0 and P_1 will satisfy following equations*

$$\begin{cases} -\frac{\partial P_j}{\partial \tau} + \frac{1}{2}y\frac{\partial^2 P_j}{\partial x^2} - \left(\frac{1}{2}-j\right)y\frac{\partial P_j}{\partial x} + \frac{1}{2}\eta^2 y\frac{\partial^2 P_j}{\partial y^2} + \rho\eta y\frac{\partial^2 P_j}{\partial x\partial y} + (a-b_jy)\frac{\partial P_j}{\partial y} \text{ for } j=0,1 \\ P_j(0,x,y) = \mathbf{1}_{\{x>0\}} \text{ for } j=0,1 \end{cases} \tag{3.352}$$

where

$$a = \kappa\theta,\ b_j = \kappa - j\rho\eta \tag{3.353}$$

Proof. We will outline proof used in [Ga, Chapter 2]. First deduce that option C will be exercised given that $x > 0$, using this derive the terminal conditions for (3.352), i.e. $P_j(0, x, y) = \mathbf{1}_{\{x>0\}}$ for $j = 0, 1$. Note that since $x = \ln(S_t e^{r(T-t)}/K)$, then at time T, $x = \ln(S_T e^{r(T-T)}/K) = \ln(S_T/K)$. Then assume that there are two european call options with payoffs $\mathbf{1}_{\{S_T>K\}}$ and $S_T\mathbf{1}_{\{S_T>K\}}$ denoted by C_0 and C_1 . Use (3.348) to deduce that the τ-forward prices of these two options are

$$\begin{aligned} C_0(\tau, x, y) &= \mathbb{Q}\left(S_T > K | \mathcal{F}_t\right) = P_0(\tau, x, y) \\ C_1(\tau, x, y) &= Ke^x \mathbb{S}\left(S_T > K | \mathcal{F}_t\right) = Ke^x P_1(\tau, x, y) \end{aligned} \tag{3.354}$$

Perform change of variables using (3.354) on partial derivatives of (3.346) using the chain rule and substitute these into (3.346) to obtain following two equations for P_0 and P_1

$$\begin{aligned} 0 &= -\tfrac{\partial P_0}{\partial \tau} - \tfrac{1}{2}y\left(\tfrac{\partial P_0}{\partial x} - \tfrac{\partial^2 P_0}{\partial x^2}\right) + \rho\eta y\tfrac{\partial^2 P_0}{\partial x \partial y} + \tfrac{1}{2}\eta^2 y\tfrac{\partial^2 P_0}{\partial y^2} - \kappa(y-\theta)\tfrac{\partial P_0}{\partial y} \\ 0 &= -\tfrac{\partial P_1}{\partial \tau} + \tfrac{1}{2}y\left(\tfrac{\partial P_1}{\partial x} + \tfrac{\partial^2 P_1}{\partial x^2}\right) + \rho\eta y\left(\tfrac{\partial^2 P_1}{\partial x \partial y} + \tfrac{\partial P_1}{\partial y}\right) + \tfrac{1}{2}\eta^2 y\tfrac{\partial^2 P_1}{\partial y^2} - \kappa(y-\theta)\tfrac{\partial P_1}{\partial y} \end{aligned} \tag{3.355}$$

Set $a = \kappa\theta$, $b_j = \kappa - j\rho\eta$ and the result follows. □

Proposition 3.36. ***(Solution [Ga, Chapter 2]).*** *Let all assumptions from Propositions 3.33 and 3.35 hold. Suppose that the solution to (3.352) is of the following form*

$$\tilde{P}_j(\tau, u, y) = \tilde{P}_j(0, u, y)e^{C_j(\tau,u)\theta + D_j(\tau,u)y} \text{ for } j = 0, 1 \tag{3.356}$$

Then functions D and C satisfy the following identities

$$D_j(\tau, u) = r_j^-\left(\frac{1 - e^{-d_j\tau}}{1 - g_j e^{-d_j\tau}}\right),\ C_j(\tau, u) = \kappa\left(r_j^-\tau - \frac{2}{\eta^2}ln\left(\frac{1 - g_j e^{-d_j\tau}}{1 - g_j}\right)\right) \text{ for } j = 0, 1 \tag{3.357}$$

where

$$\begin{aligned} \alpha_j &= -\tfrac{1}{2}u^2 - \tfrac{1}{2}iu + iuj \\ \beta_j &= b_j - \rho\eta iu \\ \gamma &= \tfrac{1}{2}\eta^2 \\ d_j &= \sqrt{\beta_j^2 - 4\gamma\alpha_j} \\ r_j^- &= \tfrac{\beta_j - d_j}{2\gamma} \\ g_j &= \tfrac{r_j^-}{r_j^+} \end{aligned} \tag{3.358}$$

Proof. We will outline proof used in [Ga, Chapter 2] and [Ro, Chapter 1, p13-14, Chapter 3, p67-69]. Using Definition 1.109 we define the Fourier transform of function P by

$$\tilde{P}_j(\tau, u, y) = \int_{-\infty}^{\infty} e^{-iux}P_j(\tau, x, y)dx \text{ for } j = 0, 1 \tag{3.359}$$

and the inverse Fourier transform by

$$P_j(\tau, x, y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{iux}\tilde{P}_j(\tau, u, y)du \text{ for } j = 0, 1 \tag{3.360}$$

Using (3.359) deduce that $\tilde{P}_j(0, u, y) = \frac{1}{iu}$. Next, substitute the (3.360) into (3.352), by taking partial derivatives using Fourier transform and integration by parts, which will result in

$$-\frac{\partial \tilde{P}_j}{\partial \tau} - \frac{1}{2}yu^2\tilde{P}_j - \left(\frac{1}{2}-j\right)yiu\tilde{P}_j + \frac{1}{2}\eta^2 y\frac{\partial^2 \tilde{P}_j}{\partial y^2} + \rho\eta yiu\frac{\partial \tilde{P}_j}{\partial y} + (a-b_j y)\frac{\partial \tilde{P}_j}{\partial y} = 0 \tag{3.361}$$

Set

$$\alpha_j = -\frac{1}{2}u^2 - \frac{1}{2}iu + iuj,\ \beta_j = b_j - \rho\eta iu,\ \gamma = \frac{1}{2}\eta^2 \tag{3.362}$$

Then

$$y\left(\gamma\frac{\partial^2 \tilde{P}_j}{\partial y^2} - \beta\frac{\partial \tilde{P}_j}{\partial y} + \alpha\tilde{P}_j\right) + a\frac{\partial \tilde{P}_j}{\partial y} = \frac{\partial \tilde{P}_j}{\partial \tau} \text{ for } j=0,1 \tag{3.363}$$

Take derivatives of proposed solution (3.356), substitute them into (3.363) and divide by $\tilde{P}$ we get a following equation

$$y\left(\gamma D_j^2 - \beta_j D_j + \alpha_j\right) + aD = \frac{\partial C_j}{\partial \tau}\theta + \frac{\partial D_j}{\partial \tau}y \text{ for } j=0,1 \tag{3.364}$$

Recall that $a = \kappa\theta$ and deduce that (3.364) is satisfied if

$$\frac{\partial C_j}{\partial \tau} = \kappa D_j,\ \frac{\partial D_j}{\partial \tau} = \gamma D_j^2 - \beta_j D_j + \alpha_j \text{ for } j=0,1 \tag{3.365}$$

Deduce from (3.356) that $C(0,u), D(0,u) = 0$ and then use this to solve ODEs in (3.365) to obtain (3.357). □

Remark 3.37. **(Solution under realworld measure [ElAA2, Appendix 4]).** Let all assumptions from Proposition 3.33 hold. We will assume that under real-world measure $\mathbb{P}$ function C satisfies

$$C_j(\tau,u) = \frac{1+iu}{\theta}\mu\tau + \kappa\left(r_j^-\tau - \frac{2}{\eta^2}ln\left(\frac{1-g_je^{-d_j\tau}}{1-g_j}\right)\right) \text{ for } j=0,1, \tag{3.366}$$

Proposition 3.38. ***(Inversion [Ro, Chapter 3, p63-66], [Ga, Chapter 2, p20-21])****. Let all assumptions from Propositions 3.33 and 3.36 hold. Let $z_T = \ln(S_T)$ and $k = \ln(K)$. Suppose that characteristic function of z_T under measures $\mathbb{Q}$ and $\mathbb{S}$ satisfies*

$$\phi_j(u) = e^{C_j(\tau,u)\theta + D_j(\tau,u)y + iuz} \textit{ for } j=0,1 \tag{3.367}$$

Let F be a cumulative distribution function and f be probability density function of random variable z. Then P_0 and P_1 from (3.349) satisfy

$$P_j(\tau,x,y) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left(\frac{\phi_j(u)e^{-iuk}}{iu}\right)du \textit{ for } j=0,1 \tag{3.368}$$

Proof. We will outline proof used in [Ro, Chapter 3: p 63-66]. To simply notation, we set $\phi = \phi_j$ and drop measures $\mathbb{Q}$ and $\mathbb{S}$ from notation and denote probabilities by $\mathbb{P}$. Recall the Definition 5.31 of characteristic function ϕ and Definition 1.109 of Fourier transform of ϕ

$$\phi(u) = \mathbb{E}\left(e^{iuz_T}\right) = \int_{-\infty}^{\infty} e^{iuz}f(z)dz,\ f(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iuz}\phi(u)du \tag{3.369}$$

By Definition 5.29 of cumulative distribution function, the probability of $z_T > k$ is

$$\mathbb{P}(z_T > k) = 1 - F(k) = \int_k^\infty f(z)dz,\ F(k) = \mathbb{P}(z_T \leq k) = \int_{-\infty}^k f(z)dz \tag{3.370}$$

Using Fourier transform from (3.369) we substitute for f, apply Fubini's Theorem 5.19 and integrate e^{-iux} over interval $[0, \infty]$ to obtain the following

$$\mathbb{P}(z_T > k) = \frac{1}{2\pi}\int_{-\infty}^\infty \phi(u)\frac{e^{-iuk}}{iu}du - \frac{1}{2\pi}\lim_{R\to\infty}\int_{-\infty}^\infty \phi(u)\frac{e^{-iuR}}{iu}du \tag{3.371}$$

For the 1st integral on the right hand side of (3.371) apply Euler's formula from Definition 5.6 on the integrand and by Definition 5.8 deduce that integrand in odd in imaginary part and even in its real part. Use the properties of odd and even functions from Proposition 5.10 to deduce that integral is symmetric around $k = 0$ in its real part and is zero in imaginary part to obtain the following

$$\frac{1}{2\pi}\int_{-\infty}^\infty \phi(u)\frac{e^{-iuk}}{iu}du = \frac{1}{\pi}\int_0^\infty \Re\left(\frac{\phi(u)e^{-iuk}}{iu}\right)du \tag{3.372}$$

For the 2nd integral on the right hand side of (3.371) substitute for ϕ from (3.369) and apply Fubini's Theorem 5.19 to get

$$\frac{1}{2\pi}\lim_{R\to\infty}\int_{-\infty}^\infty \phi(u)\frac{e^{-iuR}}{iu}du = \frac{1}{2\pi}\lim_{R\to\infty}\int_{-\infty}^\infty f(z)\left(\int_{-\infty}^\infty \frac{e^{iu(z-R)}}{iu}du\right)dz \tag{3.373}$$

Next, apply Euler's formula from Definition 5.6 on the inner integrand of (3.373) and use the previous methods for odd and even functions to deduce that

$$\int_{-\infty}^\infty \frac{e^{iu(z-R)}}{iu}du = \int_{-\infty}^\infty \frac{\sin(u(z-R))}{u}du \tag{3.374}$$

Use the Proposition 1.110 to evaluate integral in (3.374) and substitute this result into (3.373) to obtain

$$\frac{1}{2\pi}\lim_{R\to\infty}\int_{-\infty}^\infty f(z)\left(\int_{-\infty}^\infty \frac{\sin(u(z-R))}{u}du\right)dz = \frac{1}{2}\lim_{R\to\infty}\int_{-\infty}^\infty f(z)sgn(z-R)dz \tag{3.375}$$

Using Definition 5.29 of cumulative distribution function, split the range of integral in (3.375) into $[R, \infty]$, $[-\infty, R]$, evaluate these integrals and take limit to obtain the following

$$\frac{1}{2}\lim_{R\to\infty}\int_{-\infty}^\infty f(z)sgn(z-R)dz = \frac{1}{2}\lim_{R\to\infty}(1 - 2F(R)) = -\frac{1}{2} \tag{3.376}$$

Substitute (3.372) and (3.376), into (3.371) and the required result follows. □

Proposition 3.39. ***(Call option price and delta).*** *Let all assumptions from Propositions 3.34 and 3.38 hold. Then call option price at time t satisfies*

$$V(t,s,y) = s\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left(\frac{\phi_1(u)e^{-iuk}}{iu}\right)du\right) - Ke^{-r(T-t)}\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left(\frac{\phi_0(u)e^{-iuk}}{iu}\right)du\right) \tag{3.377}$$

where $S_t = s$ a.s from Definition 3.31. Let Δ denote delta of the call option V with strike K and time to *expiry* $\tau = T - t$. *Then* Δ *satisfies*

$$\Delta_t = \frac{\partial V}{\partial s} = P_1 + \frac{s}{\pi}\int_0^{\infty} \Re\left(\frac{\phi_1(u)e^{-iuk}}{s}\right) du + \frac{s}{\pi}\int_0^{\infty} \Re\left(\frac{\phi_0(u)e^{-iuk}}{s}\right) du \tag{3.378}$$

Proof. To prove (3.377), substitute for $P_0 = P_0(\tau, u, y)$ and $P_1 = P_1(\tau, u, y)$ from results in (3.38). To prove (3.378) differentiate V with respect to s to obtain by product rule

$$\frac{\partial V}{\partial s} = P_1 + s\frac{\partial P_1}{\partial s} - K\frac{\partial P_0}{\partial s} \tag{3.379}$$

Note that only ϕ is function of s as from (3.367), $z = \ln(s)$. Then differentiate P_1 and P_0 using chain rule, insert these back into (3.379) and the result follows. $\square$

3.7 Simulation

To simulate state process we are use Heston model as in [ElAA3, Section 6]. As defined in Definition 3.31 parameters of this model under the risk-neutral measure $\mathbb{Q}$ are $(r, \kappa_I, \theta_I, \eta_I, \rho_I)$ and under the real world measure $\mathbb{R}$ are $(\mu, \kappa_R, \theta_R, \eta_R, \rho_R)$. For the purpuses of comparison we introduce a "zero-intelligence" market maker in the same manner as in [ElAA3, Subsection 6.1]. We assume that zero intelligence market maker quotes a symmetric spread $\delta_{ZI} + \delta_{ZI}$ determined by options vega ν_{BS}, these are defined by

$$\delta_{ZI} = 0.005 \times \nu_{BS},\ \nu_{BS} = \frac{\partial C_{BS}(t, s, y)}{\partial y} \tag{3.380}$$

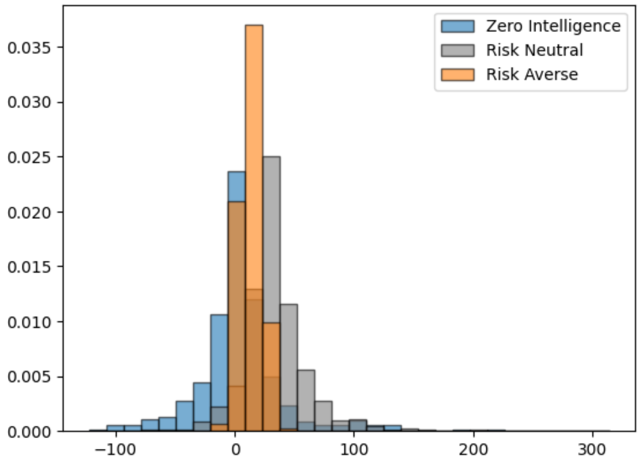
Hence the prices for ask and bid quotes are

$$C^a_{ZI,t} = C_{\mathbb{Q}}(t, S_t, Y_t) + \delta_{ZI},\ C^b_{ZI,t} = C_{\mathbb{Q}}(t, S_t, Y_t) - \delta_{ZI} \tag{3.381}$$

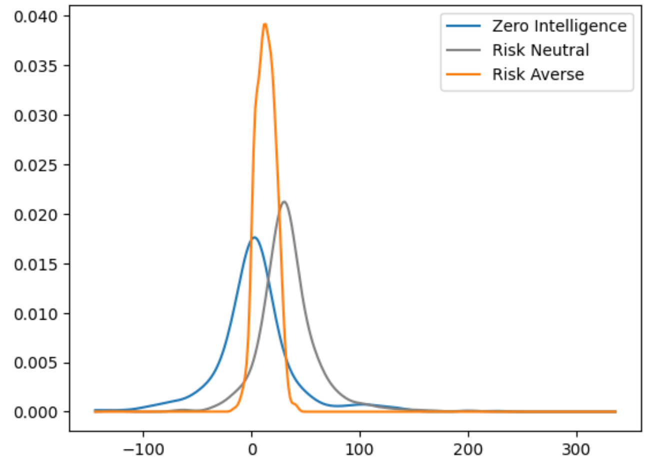
We assume here that C_{BS} is standard Black–Scholes model call option price. In contrast to [ElAA3, Subsection 6.1] we introduce both the risk-neutral market maker who quotes $\hat{\delta}^a_L + \hat{\delta}^b_L$ spread as determined by optimal controls in Proposition 3.18 and risk-averse market maker who quotes $\hat{\delta}^a_\epsilon + \hat{\delta}^b_\epsilon$ spread as determined by optimal controls in Proposition 3.27. We use Heston call price from Proposition 3.39 as the risk-neutral $C_{\mathbb{Q}}$ and risk-averse $C_{\mathbb{P}}$ option prices from assumptions Subsection 3.2. Note on the assumption on differences between $C_{\mathbb{Q}}$ and $C_{\mathbb{P}}$ in Remark 3.37. We assume the initial values for spot price and variance to be $S_0 = 100$ and $Y_0 = 0.04$ respectively. Furthermore, we assume strike price $K = 100$ and a month time to expiry $\tau = \frac{1}{12}$. Simulation is very computationally demanding owing to integration of complex function involved in Heston pricing and as such we will not be simulating as many paths and iterations as in [ElAA3, Section 6]. Instead we simulate paths of the state process in Proposition 3.13 with 500 data points simulated 500 times, aggregating data at each step (for the details of the script used in simulation and instructions see Appendix 5.4). We then use the final values of relevant variables to examine their statistics. As discussed in the introduction to this dissertation, the most relevant variables for market marker are inventory and earnings. Figures 3.2 and 3.3 shows the statistics for simulation with following parameters

$$(r, \kappa_I, \theta_I, \eta_I, \rho_I) = (0.001, 4, 0.04, 0.5, -0.4),\ (\mu, \kappa_R, \theta_R, \eta_R, \rho_R) = (0.001, 4, 0.04, 0.5, -0.4) \tag{3.382}$$

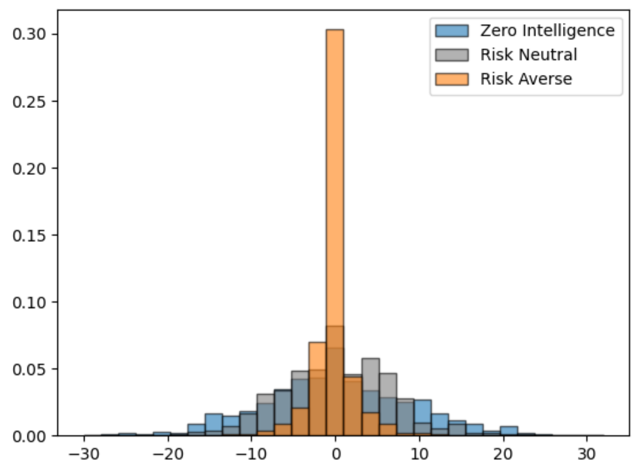
Figure 3.2: Distributions



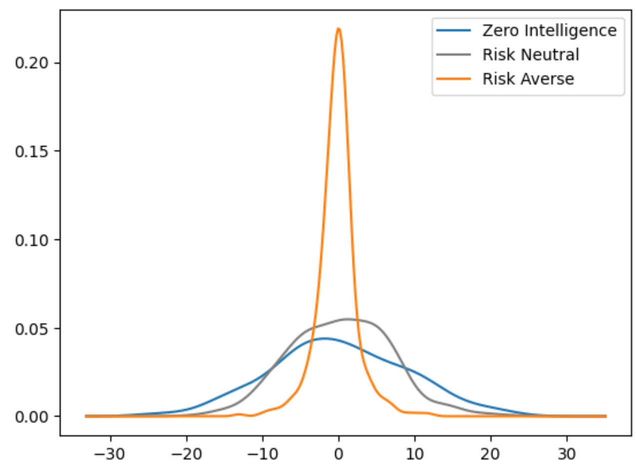
Histogram of final wealth X_T



Density estimate of final wealth X_T



Histogram of final options inventory $Q_{1,T}$



Density estimate of final options inventory $Q_{1,T}$

As expected Figure 3.2 shows that risk-averse market maker is much better at managing inventory risk, which results in higher probability of having terminal inventory around zero. Reduction of risk at expense of profitability is also observed between risk-averse and risk-neutral market makers. Zero intelligence agent returns and risk compare poorly with both risk-averse and risk-neutral market makers. Note that in [ElAA3, Section 6] it is not stated which is "Our agent". From these result we can deduce that it would be the risk-neutral market maker.

Figure 3.3: Descriptive statistics for final wealth X_T

Market maker	Mean	Std	Skewness	Kurtosis
Zero intelligence	2.989894	38.370224	0.757432	5.929867
Risk-neutral	33.516593	28.317222	0.765899	6.910679
Risk-averse	13.318708	8.924925	0.047191	−0.23592

Since our simulated data is for one month, while in [ElAA3, Section 6] it is four, differences in statistics are expected. Disrepancies also come from unknown parameters used in [ElAA3, Section 6] for the market impact function from Definition 3.9. We assumed these to be $\gamma = 1.5$, $A = 500$ and $B = 1$.

Figure 3.4: Descriptive statistics for final options inventory $Q_{1,T}$

Market maker	Mean	Std	Skewness	Kurtosis
Zero intelligence	0.10800	8.95929	0.026925	−0.17365
Risk-neutral	0.284000	6.570059	0.108511	0.088679
Risk-averse	−0.146000	2.612726	0.062784	4.22917

However, comparison of descriptive statistics between different market makers confirm what we have established earlier. In that from the standpoint of managing both risk and return, risk-averse market maker outperforms the other two.

4 Conclusion

We have introduced the relevant background results for the study of stochastic control of processes driven by Lévy noise. We then used these background results to derive and prove the results in [AS] for stock market making and in [ElAA3] for options market making. Findings from the simulation support the results for options market making and go in line with what was found in [ElAA3, Section 6]. This indicates that stochastic control is a viable method of solving the market making problem under simplifying assumptions. A future area of research would concern the relaxation of these assumptions. Specifically, the infinite liquidity assumption for stock in options market making problem and introduction of latency as a state variable influencing order execution. Another area of future research is the resolution of the market-making problem in options across different strikes and expiries using stochastic control.

5 Appendix

5.1 Notations

$\mathbb{R}$	set of real numbers
$\mathbb{R}^+$	set of positive real numbers
$\mathbb{R}^{n\times m}$	the $n \times m$ matrices of real numbers
$\mathbb{N}$	set of natural numbers
$\mathbb{C}$	set of complex numbers
$C^n(A)$	continious and n-times differentiable in set A taking values in $\mathbb{R}$
$\mathcal{B}(A)$	Borel set in A
$f(a-)$	left limit of function $f(x)$ as $x \uparrow a$
$f(a+)$	right limit of function $f(x)$ as $x \downarrow a$
$a \wedge b$	the minimum of a and b
$a \vee b$	the maximum of a and b
$(a-b)^-$	$a - b \wedge 0$
$(a-b)^+$	$a - b \vee 0$
$\propto$	directly proportional
∇	gradient
∂A	boundary of the set A
$\bar{A}$	closure of the set A
$f^{(n)}(a)$	n-th derivative of function f evaluated at a
$\Re(z)$	real part of complex number z
$\Im(z)$	imaginary part of complex number z
$\overline{z}$	conjugate of complex number z
$\stackrel{d}{=}$	equality in distribution
$\mathbb{E}^{\mathbb{Q}}$	expectation with respect to measure $\mathbb{Q}$
i.i.d.	independently and identically distributed
a.s.	almost surely
HJB	Hamilton-Jacobi-Bellman
SDE	Stochastic differential equation
PDE	Partial differential equation
ODE	Ordinary differential equation

5.2 Basic concepts

For the rest of the section we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ and $(M, \mathcal{B}(M))$ is a measurable space.

Definition 5.1. (Concave and convex functions [Ra, Appendix A, p779-782]). Let $f : [a, b] \to \mathbb{R}$, where $a, b \in \mathbb{R}$. We say that f is concave if

$$f((1-c)x_1 + cx_2) \geq (1-c)f(x_1) + cf(x_2) \text{ for } x_1, x_2 \in [a, b], \text{ for any } c \in [0, 1]. \tag{5.1}$$

We say that f is convex if

$$f((1-c)x_1 + cx_2) \leq (1-c)f(x_1) + cf(x_2) \text{ for } x_1, x_2 \in [a, b], \text{ for any } c \in [0, 1]. \tag{5.2}$$

Proposition 5.2. ***(Concave and convex functions and Taylor series [Ra, Theorem A.1]).*** *Let $f : [a, b] \to \mathbb{R}$ and $f \in C^1([a, b])$, where $a, b \in \mathbb{R}$.*
1) If f is concave then

$$f(x_2) \leq f(x_1) + \left.\frac{df(x)}{dx}\right|_{x=x_1} (x_2 - x_1) \text{ for any } x_1, x_2 \in [a, b] \tag{5.3}$$

2) If f is convex then

$$f(x_2) \geq f(x_1) + \left.\frac{df(x)}{dx}\right|_{x=x_1} (x_2 - x_1) \text{ for any } x_1, x_2 \in [a, b] \tag{5.4}$$

Proof. We using same methods as in [Ra, Theorem A.1]. We use Definition 5.1 for convex f where for fixed $x_1, x_2 \in [a, b]$ and for any $c \in [0, 1]$ we have

$$\begin{aligned} f((1-c)x_1 + cx_2) &\leq (1-c)f(x_1) + cf(x_2) \\ f(x_1 - cx_1 + cx_2) &\leq f(x_1) - cf(x_1) + cf(x_2) \\ f(x_1 + c(x_2 - x_1)) - f(x_1) &\leq cf(x_2) - cf(x_1) \\ \tfrac{f(x_1 + c(x_2 - x_1)) - f(x_1)}{c(x_2 - x_1)} &\leq \tfrac{f(x_2) - f(x_1)}{(x_2 - x_1)} \end{aligned} \tag{5.5}$$

Let $\Delta x = c(x_2 - x_1)$, then

$$\begin{aligned} \tfrac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} &\leq \tfrac{f(x_2) - f(x_1)}{(x_2 - x_1)} \\ \lim_{\Delta x \to 0} \tfrac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} &\leq \tfrac{f(x_2) - f(x_1)}{(x_2 - x_1)} \\ \left.\tfrac{df(x)}{dx}\right|_{x=x_1} &\leq \tfrac{f(x_2) - f(x_1)}{(x_2 - x_1)} \\ (x_2 - x_1) \left.\tfrac{df(x)}{dx}\right|_{x=x_1} &\leq f(x_2) - f(x_1) \end{aligned} \tag{5.6}$$

By the same methods we can obtain for concave f

$$(x_2 - x_1) \left.\tfrac{df(x)}{dx}\right|_{x=x_1} \geq f(x_2) - f(x_1) \tag{5.7}$$

Note that these inequalities satisfy Definition 1.101 of first order Taylor series. $\square$

Theorem 5.3. ***(Concave and convex functions and second derivatives [Ra, Appendix A, p779-782]).*** *Let* $f : [a, b] \to \mathbb{R}$*, where* $a, b \in \mathbb{R}$*. If* $f \in C^2([a, b])$*, then* f *is concave if* $\frac{d^2 f(x)}{dx^2} \leq 0$ *and convex on if* $\frac{d^2 f(x)}{dx^2} \geq 0$.

Theorem 5.4. ***(Mean value [Gr, Theorem 13.5.1]).*** *Let* $f : \mathbb{R} \to \mathbb{R}$*, where* $f \in C^1((a, b)) \cup C^0([a, b])$ *and* $a, b \in \mathbb{R}$*. Then there exist* $x \in (a, b)$ *at which*

$$f(b) - f(a) = (b - a)\frac{df(x)}{dx} \tag{5.8}$$

Definition 5.5. (Complex number [Ru, 1.30 Definition]). Let $i = \sqrt{-1}$ and $z \in \mathbb{C}$, then

$$z = \Re(z) + i \times \Im(z), \ \overline{z} = \Re(z) - i \times \Im(z) \tag{5.9}$$

where $\Re : \mathbb{C} \to \mathbb{R}$ is real part of z, $\Im : \mathbb{C} \to \mathbb{R}$ is imaginary part of z, $\mathbb{C}$ is a set of complex numbers and $\overline{z}$ is conjugate of z

Definition 5.6. (Eulers formula [Gr, Section 3.4, p92-93]). Let $i = \sqrt{-1}$ and $x \in \mathbb{R}$, then

$$e^{ix} = \cos(x) + i\sin(x), \ e^{-ix} = \cos(x) - i\sin(x) \tag{5.10}$$

Definition 5.7. (Sign function [Ro, Chapter 3, p65]). Let $x \in \mathbb{R}$ and $sgn : \mathbb{R} \to \{-1, 0, 1\}$ defined by

$$sgn(x) = \begin{cases} 1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -1 & \text{if} \quad x < 0 \end{cases} \tag{5.11}$$

we call sgn a sign function.

Definition 5.8. (Odd and even functions [Gr, Section 17.2, p846]). Let $f : \mathbb{R} \to \mathbb{R}$. We call f

(1) Even function if $f(-x) = f(x)$

(2) Odd function if $f(-x) = -f(x)$

Definition 5.9. (Periodic function [Gr, Section 17.2, p848]). Let $f : \mathbb{R} \to \mathbb{R}$. Suppose that there exists $T > 0$ such that

$$f(x + T) = f(x) \text{ for every } x \in \mathbb{R} \tag{5.12}$$

Then we say that f is a periodic function of x with period T.

Proposition 5.10. ***(Odd and even functions properties [Gr, Section 17.2, p846]).*** *We have* *following relations for odd and even functions*

$$\begin{aligned} odd \times even &= odd \\ odd \times odd &= even \\ even \times even &= even \\ odd + odd &= odd \\ even + even &= even \end{aligned} \tag{5.13}$$

Let $f : \mathbb{R} \to \mathbb{R}$ be even function, let $g : \mathbb{R} \to \mathbb{R}$ be odd function and suppose that f, g are integrable over the interval $[-a, a]$ where $a > 0$. Then

$$\int_{-a}^{a} f(x)dx = 2\int_{0}^{A} f(x)dx, \int_{-A}^{A} g(x)dx = 0 \tag{5.14}$$

Definition 5.11. (Càdlàg functions [Ap, 2.9 Appendix]). Let $[a, b] \in \mathbb{R}^+$, a function $f : [a, b] \to \mathbb{R}^d$ is said to be càdlàg (right-continious with left limits) if

(1) Right continuity: $f(t+) = \lim_{s \downarrow t} f(s)$ exists and $f(t+) = f(t)$ at each point $t \in [a, b]$

(2) Left limits: $f(t-) = \lim_{s \uparrow t} f(s)$ exists at each point $t \in (a, b]$

For $t \in (a, b]$, $f(t-) = f(t)$ if and only if f is continious at t

Definition 5.12. (σ-field [BZ, Definition 1.1]). Let Ω be non-empty set. A σ-field $\mathcal{A}$ on Ω is a collection of subsets of Ω such that

(1) $\emptyset \in \mathcal{A}$

(2) if $A \in \mathcal{A}$, then $\Omega \setminus A$

(3) if $A_1, A_2, ... \in \mathcal{A}$, then $A_1 \cup A_2 \cup ... \in \mathcal{A}$

Definition 5.13. (Measurable space [Ap, Chapter 1: Subsection 1.1.1]). Measurable space is a tuple $(A, \mathcal{A})$, which consists of a set A and a σ-field $\mathcal{A}$ which assigns a measure to the set A.

Definition 5.14. (Borel measurable function [Ap, Chapter 1: Subsection 1.1.2]). Let $(A_1, \mathcal{A}_1)$ and $(A_2, \mathcal{A}_2)$ be measurable spaces. A function $f : A_1 \to A_2$ is said to be $(\mathcal{A}_1, \mathcal{A}_2)$-measurable if $f^{-1}(A) \in \mathcal{A}_1$ for all $A \in \mathcal{A}_2$. If each $A_1 \subseteq \mathbb{R}^d$, $A_2 \subseteq \mathbb{R}^n$ and $\mathcal{A}_1 = \mathcal{B}(A_1)$, $\mathcal{A}_2 = \mathcal{B}(A_2)$, f is said to be Borel measurable.

Proposition 5.15. ***([Bj, Proposition A.44]).*** *Every continious function $f : \mathbb{R}^d \to \mathbb{R}$ is Borel measurable*

Definition 5.16. (Measure [Ap, Chapter 1: Subsection 1.1.1]). Let $(A, \mathcal{A})$ be measurable space. A measure on $(A, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [-\infty, \infty]$ that satisfies

(1) Non-negativity: $\mu(B) \geq 0$ for all $B \in \mathcal{A}$

(2) Null empty set: $\mu(\emptyset) = 0$

(3) Countable additivity: given any sequence $(B_n)_{n\in\mathbb{N}}$of pairwise disjoint sets in $\mathcal{A}$

$$\mu\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} \mu(B_n) \tag{5.15}$$

Definition 5.17. (Measure space [Ap, Chapter 1: Subsection 1.1.1]). Measure space is a triple $(A, \mathcal{A}, \mu)$, which consists of a set A, a σ-field $\mathcal{A}$ which assigns a measure to the set A and a measure μ on $(A, \mathcal{A})$.

Proposition 5.18. ***(Lebesgue measure [Bj, Proposition A.49]).*** *Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be measurable space. There exists a unique measure m such that*

$$m([a, b]) = b - a \textit{ for } a, b \in \mathbb{R} \tag{5.16}$$

which is called Lebesgue measure. Suppose we have a continious function f and a finite inteval A, then we can form

(1) Lebesgue integral: $\int_A f(x)dm(x)$

(2) Riemann integral: $\int_A f(x)dx$

Theorem 5.19. ***(Fubini [Ap, Theorem 1.1.7]).*** *Let $(A_1, \mathcal{A}_1, \mu_1)$ and $(A_2, \mathcal{A}_2, \mu_2)$ be measure spaces. If $f : A_1 \times A_2$ is $(\mathcal{A}_1 \times \mathcal{A}_2)$-measurable with*

$$\int\int |f(x, y)|\mu_1(dx)\mu_2(dy) < \infty \tag{5.17}$$

where $\mu_1 : y \to \int |f(x, y)|\mu_1(dx)$ and $\mu_2 : y \to \int |f(x, y)|\mu_2(dy)$. Then

$$\begin{aligned} \int_{A_1\times A_2} f(x, y)(\mu_1 \times \mu_2)(dx, dy) &= \int_{A_2} \left[\int_{A_1} f(x, y)\mu_1(dx)\right] \mu_2(dy) \\ &= \int_{A_1} \left[\int_{A_2} f(x, y)\mu_2(dy)\right] \mu_1(dx) \end{aligned} \tag{5.18}$$

Definition 5.20. (Finite measure [Ap, Subsection 1.1.1, p2]). Let μ be a measure on $(A, \mathcal{A})$. It is a finite measure if $\mu(A) < \infty$.

Definition 5.21. (σ-finite measure [Ap, Subsection 1.1.1, p2]). Let μ be a measure on $(A, \mathcal{A})$. It is a σ-finite measure if there exists a sequence $(B_n)_{n\in\mathbb{N}}$of sets in $\mathcal{A}$ such that $A = \bigcup_{n=1}^{\infty} B_n$ and each $\mu(B_n) < \infty$

Definition 5.22. (Separate measures [Bj, Definition A.50]). Let μ, ν be two separate measures on $(A, \mathcal{A})$. We call μ, ν

(1) Absolutely continious and write $\mu \ll \nu$, if

$$\mu(B) = 0 \Rightarrow \nu(B) = 0 \text{ for all } B \in \mathcal{A} \tag{5.19}$$

(2) Equivalent and write $\mu \sim \nu$, if

$$\mu \ll \nu,\ \nu \ll \mu \tag{5.20}$$

Theorem 5.23. ***(Radon-Nikodym [Bj, Theorem A.52]).*** *Let $(A, \mathcal{A}, \mu)$ be measure space with σ-finite measure μ. Suppose there exists a measure ν on $(A, \mathcal{A})$ such that $\nu \ll \mu$. Then there exists a function $f : A \to \mathbb{R}^+$ such that*

(1) f is $\mathcal{A}$-measurable

(2) $\int_A f(x)d\mu(x) < \infty$

(3) $\nu(x) = \int_B f(x)d\mu(x)$ for all $B \in \mathcal{A}$

It is called the Radon-Nikodym derivative of ν with respect to μ and is denoted by

$$f(x) = \frac{d\nu(x)}{d\mu(x)} \tag{5.21}$$

Definition 5.24. (Probability measure [BZ, Definition 1.2]). Let $\mathcal{A}$ be σ-field on Ω. A probability measure is a function $\mathbb{P} : \mathcal{A} \to [0, 1]$ such that

(1) $\mathbb{P}(\Omega) = 1$

(2) if $A_1, A_2, ... \in \mathcal{A}$ are pairwise disjoint, then $\mathbb{P}(A_1 \cup A_2 \cup ...) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + ...$

The triple $(\Omega, \mathcal{A}, \mathbb{P})$ is called a probability space. Sets $A \in \mathcal{A}$ are called events. An event A is said to occur almost surely (a.s.) whenever $\mathbb{P}(A) = 1$

Definition 5.25. (Measurability and random variables [BZ, Definition 1.3]). If $\mathcal{A}$ is a σ-field on Ω, then a function $X : \Omega \to \mathbb{R}$ is said to be $\mathcal{A}$-measurable if

$$\{X \in B\} \in \mathcal{A} \text{ for every } B \in \mathcal{B}(\mathbb{R}) \tag{5.22}$$

where $\{X \in B\}$ is short-hand notation for $\{\omega \in \Omega : X(\omega) \in B\}$. If $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space, then function X is called a random variable.

Definition 5.26. (Generated σ-field [BZ, Definition 1.4]). The σ-field $\sigma(X)$ generated by a random variable $X : \Omega \to \mathbb{R}$ consists of all sets of the form $\{X \in B\}$ for $B \in \mathcal{B}(\mathbb{R})$

Definition 5.27. (Smallest σ-field [BZ, Definition 1.5]). The σ-field $\sigma\{X_i : i \in I\}$ generated by a collection $\{X_i : i \in I\}$ of random variables is defined to be the smallest σ-field containing all events of the form $\{X_i \in B\}$ for $B \in \mathcal{B}(\mathbb{R}), i \in I$

Definition 5.28. (Borel σ-field [Ap, Subsection 1.1.1, p2]). The Borel σ-field $\mathcal{B}(\mathbb{R}^d)$ is the smallest σ-field of subsets of $\mathbb{R}^d$ that contains all the open sets. Sets $B \in \mathcal{B}(\mathbb{R}^d)$ are called Borel sets and are defined by

$$\mathcal{B}(B) = \{A \cap B : A \in \mathcal{B}(\mathbb{R}^d)\} \tag{5.23}$$

Definition 5.29. (Distribution function [BZ, Definition 1.6]). Every random variable $X : \Omega \to \mathbb{R}$ gives rise to a probability measure

$$\mathbb{P}_X(B) = \mathbb{P}(\{X \in B\}) \text{ for } B \in \mathcal{B}(\mathbb{R}) \tag{5.24}$$

on $\mathbb{R}$ defined on the Borel σ-field $\mathcal{B}(\mathbb{R})$. We call $\mathbb{P}_X$ the distribution (or law) of X. The function $F_X : \mathbb{R} \to [0, 1]$ defined b

$$F_X(x) = P(\{X \leq x\}) \text{ for } x \in \mathbb{R} \tag{5.25}$$

is called the distribution function of X.

Definition 5.30. (Density function [BZ, Definition 1.7]). If there is a Borel function $f_X : \mathbb{R} \to \mathbb{R}$ such that for any Borel set $B \subset \mathbb{R}$

$$\mathbb{P}(\{X \in B\}) = \int_B f_X(x)dx \tag{5.26}$$

then X is said to be a random variable with absolutely continious distribution and f_X is called the density of X. If there is a (finite or infinite) sequence of pairwise distinct real numbers $x_1, x_2, ...$ such that for any Borel set $B \subset \mathbb{R}$

$$\mathbb{P}(\{X \in B\}) = \sum_{x_i \in B} \mathbb{P}(\{X = x_i\}) \tag{5.27}$$

then X is said to have discrete distribution with values $x_1, x_2, ...$ and mass $\mathbb{P}(\{X = x_i\})$ at x_i.

Definition 5.31. (Characteristic function [Ap, Subsection 1.1.6, p16]). Let X be an $\mathbb{R}^d$-valued random variable and f_X its density. Let $\phi_X : \mathbb{R}^d \to \mathbb{C}$ be defined by

$$\phi_X(u) = \mathbb{E}\left(e^{iuX}\right) = \int_{\mathbb{R}^d} e^{iux} f_X(x)dx \text{ for each } u \in \mathbb{R}^d \tag{5.28}$$

where ϕ_X is called characteristic function of X. If $M_X(u) = \phi_X(-iu)$ exists for u close to or at $u = 0$, then M_X is called moment generating function of X.

Definition 5.32. (Expectation and integrability [BZ, Definition 1.9]). A random variable $X : \Omega \to \mathbb{R}$ is said to be integrable if

$$\int_\Omega |X(\omega)| d\mathbb{P}(\omega) < \infty \tag{5.29}$$

Then

$$\mathbb{E}\left(X(\omega)\right) = \int_\Omega X(\omega) d\mathbb{P}(\omega) \tag{5.30}$$

exists and is called the expectation of $X = X(\omega)$.

Definition 5.33. (Independent random variables [BZ, Definition 1.13]). We say that random variables $X_1, ..., X_n$ are independent if for any $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$, the events

$$\{X_1 \in B_1\}, ..., \{X_n \in B_n\} \tag{5.31}$$

are independent. In general, a (finite or infinite) collection of random variables is said to be independent if any finite number of random variables from this collection are independent.

Proposition 5.34. ***(Uncorrelated random variables [BZ, Proposition 1.1]).*** *If integrable random variables $X_1, ..., X_n : \Omega \to \mathbb{R}$ are independent, then they are uncorrelated, i.e.*

$$\mathbb{E}(X_1 X_2 ... X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)...\mathbb{E}(X_n) \tag{5.32}$$

given that $X_1 X_2 ... X_n$ is also integrable.

Definition 5.35. (Independent σ-fields [BZ, Definition 1.14]). Any finite number of σ-fields $\mathcal{G}_1, ..., \mathcal{G}_n \in \mathcal{A}$ are independent if any n, the events

$$A_1 \in \mathcal{G}_1, ..., A_n \in \mathcal{G}_n \tag{5.33}$$

are independent. In general, a (finite or infinite) collection of σ-fields is said to be independent if any finite number of them are independent.

Definition 5.36. (Independence of random variables from σ-fields [BZ, Definition 1.15]). A collection with any finite number of random variables $X_1, ..., X_n$ and σ-fields $\mathcal{G}_1, ..., \mathcal{G}_m$, is called independent if the σ-fields

$$\sigma(X_1), ..., \sigma(X_n), \mathcal{G}_1, ..., \mathcal{G}_m \tag{5.34}$$

are independent.

Definition 5.37. (Conditional expectation [BZ, Definition 2.1]). Let $\mathcal{A}$ be a σ-field. For any integrable random variable X and any event $B \in \mathcal{A}$ such that $\mathbb{P}(B) \neq 0$ the conditional expectation of X given B is defined by

$$\mathbb{E}(X|B) = \frac{1}{\mathbb{P}(B)} \int_B X d\mathbb{P} \tag{5.35}$$

Definition 5.38. (Conditional expectation given a discrete random variable [BZ, Definition 2.2]). Let X be an integrable random variable and let Y be a random variable taking possible values $y_1, y_2, ...$ such that

$$P(\{Y = y_n\}) \neq 0 \text{ for each } n \tag{5.36}$$

then Y is a discrete random variable and the conditional expectation of X given Y is defined to be a random variable $\mathbb{E}(X|Y)$ such that

$$\mathbb{E}(X(\omega)|Y(\omega)) = \mathbb{E}(X|\{Y = y_n\}) \text{ if } Y(\omega) = y_n \tag{5.37}$$

for any $n = 1, 2, ...$

Proposition 5.39. ***([BZ, Proposition 2.1]).*** *If X is an integrable random variable and Y is a discrete random variable, then*

(1) $\mathbb{E}(X|Y)$ is $\sigma(Y)$-measurable

(2) For any $A \in \sigma(Y)$

$$\int_A \mathbb{E}(X|Y) d\mathbb{P} = \int_A X d\mathbb{P} \tag{5.38}$$

Definition 5.40. (Conditional expectation given an arbitrary random variable [BZ, Definition 2.3]). Let X be an integrable random variable and let Y be an arbitrary random variable. Then the conditional expectation of X given Y is defined to be a random variable $\mathbb{E}(X|Y)$ such that

(1) $\mathbb{E}(X|Y)$ is $\sigma(Y)$-measurable

(2) For any $A \in \sigma(Y)$

$$\int_A \mathbb{E}(X|Y) d\mathbb{P} = \int_A X d\mathbb{P} \tag{5.39}$$

Remark 5.41. **(Conditional probability given a random variable [BZ, Remark 2.1]).** We define the conditional probability of an event $A \in \mathcal{A}$ given Y by

$$\mathbb{P}(A|Y) = \mathbb{E}(\mathbf{1}_A|Y) \tag{5.40}$$

Definition 5.42. (Conditional expectation given a σ-field [BZ, Definition 2.4]). Let X be an integrable random variable on $(\Omega, \mathcal{A}, \mathbb{P})$, and let $\mathcal{G} \subset \mathcal{A}$. Then the conditional expectation of X given $\mathcal{G}$ is defined to be a random variable $\mathbb{E}(X|\mathcal{G})$ such that

(1) $\mathbb{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable

(2) For any $A \in \mathcal{G}$

$$\int_A \mathbb{E}(X|\mathcal{G}) d\mathbb{P} = \int_A X d\mathbb{P} \tag{5.41}$$

Remark 5.43. **(Conditional probability given a σ-field [BZ, Remark 2.2]).** The conditional probability of an event $A \in \mathcal{A}$ given a σ-field $\mathcal{G}$ can be defined by

$$\mathbb{P}(A|\mathcal{G}) = \mathbb{E}(\mathbf{1}_A|\mathcal{G}) \tag{5.42}$$

This notion extends to conditioning on a random variable Y in the sense that

$$\mathbb{E}(X|\sigma(Y)) = \mathbb{E}(X|Y) \tag{5.43}$$

where $\sigma(Y)$ is σ-field generated by Y.

Proposition 5.44. ***(Conditional expectation properties [BZ, Proposition 2.4]).*** *Conditional expectation has the following properties*

(1) Linearity: $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$

(2) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$

(3) Taking out what is known: $\mathbb{E}(XY|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$ *if* X *is* $\mathcal{G}$*-measurable and* XY *is integrable*

(4) $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ *if* X is independent of $\mathcal{G}$

(5) Tower property: $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(X|\mathcal{H})$ *if* $\mathcal{H} \subset \mathcal{G}$

(6) Positivity: if $\mathbb{P}(X \geq 0) = 1$*, then* $\mathbb{P}(\mathbb{E}(X|\mathcal{G}) \geq 0) = 1$

(7) Monotinicity: if $\mathbb{P}(X \leq Y) = 1$*, then* $\mathbb{E}(X) \leq \mathbb{E}(Y)$

where $a, b \in \mathbb{R}$*,* X, Y *are integrable random variables on* $(\Omega, \mathcal{A}, \mathbb{P})$ *and* $\mathcal{G}, \mathcal{H} \in \mathcal{A}$ *are* σ*-fields on* Ω*.*

Definition 5.45. (Variance [BZ, Definition 1.10]). Suppose we have a random variable $X : \Omega \to \mathbb{R}$, which is square integrable

$$\int_\Omega |X(\omega)|^2 d\mathbb{P}(\omega) < \infty \tag{5.44}$$

Then the variance of X is defined by

$$Var(X) = \int_\Omega (X(\omega) - \mathbb{E}(X))^2 \, d\mathbb{P}(\omega) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \tag{5.45}$$

Proposition 5.46. ***(Linearity of variance).*** *Suppose we have two square integrable* $\mathbb{R}$*-valued random variables* X, Y*, which are not independent and let* $a, b \in \mathbb{R}$*. Then*

$$Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y) \pm 2abCov(X, Y) \tag{5.46}$$

where

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2, \; Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \tag{5.47}$$

If X, Y are independent, then

$$Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y) \tag{5.48}$$

where

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2, \; Cov(X, Y) = 0 \tag{5.49}$$

Proof. By Definition 5.45 and linearity of expectation

$$\begin{aligned} Var(aX) &= \mathbb{E}\left((aX - a\mathbb{E}(X))^2\right) \\ &= \mathbb{E}\left(a^2X^2 + a^2(\mathbb{E}(X))^2 - a^2 2X\mathbb{E}(X)\right) \\ &= a^2\mathbb{E}\left(X^2\right) + a^2\left(\mathbb{E}(X)\right)^2 - a^2 2\mathbb{E}(X)\mathbb{E}(X) \\ &= a^2\mathbb{E}\left(X^2\right) - a^2\left(\mathbb{E}(X)\right)^2 \end{aligned} \tag{5.50}$$

If are X, Y are not independent

$$\begin{aligned} Cov(aX, bY) &= \mathbb{E}\left((aX - a\mathbb{E}(X))(bY - b\mathbb{E}(Y))\right) \\ &= \mathbb{E}\left(abXY - abX\mathbb{E}(Y) - abY\mathbb{E}(X) + ab\mathbb{E}(X)\mathbb{E}(Y)\right) \\ &= ab\mathbb{E}\left(XY\right) - ab\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right) - ab\mathbb{E}\left(Y\right)\mathbb{E}\left(X\right) + ab\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right) \\ &= ab\mathbb{E}\left(XY\right) - ab\mathbb{E}\left(Y\right)\mathbb{E}\left(X\right) \\ &= abCov(X, Y) \end{aligned} \tag{5.51}$$

Then

$$\begin{aligned} Var(aX \pm bY) &= \mathbb{E}\left[\left((aX \pm bY) - \mathbb{E}(aX \pm bY)\right)^2\right] \\ &= \mathbb{E}\left[(aX \pm bY)^2 + \left(\mathbb{E}(aX \pm bY)\right)^2 - 2(aX \pm bY)\mathbb{E}(aX \pm bY)\right] \\ &= \mathbb{E}\left((aX \pm bY)^2\right) + \left(\mathbb{E}(aX \pm bY)\right)^2 - 2\mathbb{E}(aX \pm bY)\mathbb{E}(aX \pm bY) \\ &= \mathbb{E}\left((aX \pm bY)^2\right) - \left(\mathbb{E}(aX \pm bY)\right)^2 \\ &= \mathbb{E}\left(a^2X^2 + b^2Y^2 \pm 2abXY\right) - \left(a\mathbb{E}(X) \pm b\mathbb{E}(Y)\right)^2 \\ &= a^2\mathbb{E}\left(X^2\right) + b^2\mathbb{E}\left(Y^2\right) \pm 2ab\mathbb{E}\left(XY\right) - a^2\left(\mathbb{E}(X)\right)^2 - b^2\left(\mathbb{E}(Y)\right)^2 \mp 2ab\mathbb{E}(X)\mathbb{E}(Y) \\ &= a^2\mathbb{E}\left(X^2\right) - a^2\left(\mathbb{E}(X)\right)^2 + b^2\mathbb{E}\left(Y^2\right) - b^2\left(\mathbb{E}(Y)\right)^2 \pm 2ab\left(\mathbb{E}\left(XY\right) - \mathbb{E}(X)\mathbb{E}(Y)\right) \\ &= a^2Var(X) + b^2Var(Y) \pm 2abCov(X, Y) \end{aligned} \tag{5.52}$$

If X, Y are independent, then by Proposition 5.34

$$Cov(X, Y) = \mathbb{E}\left(XY\right) - \mathbb{E}\left(Y\right)\mathbb{E}\left(X\right) = \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right) - \mathbb{E}\left(Y\right)\mathbb{E}\left(X\right) = 0 \tag{5.53}$$

then

$$Var(aX \pm bY) = a^2Var(X) + b^2Var(Y) \pm 2abCov(X, Y) = a^2Var(X) + b^2Var(Y) \tag{5.54}$$

□

Theorem 5.47. ***(Jensen inequality [Ap, Section 1.1, p7]).*** *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}$-valued integrable random variable and let $f : \mathbb{R} \to \mathbb{R}$ be integrable and convex. Then we have the inequality*

$$f\left(\mathbb{E}\left(X\right)\right) \leq \mathbb{E}\left(f\left(X\right)\right) \tag{5.55}$$

Definition 5.48. (Stochastic process [BZ, Definition 6.1]). A stochastic process is a collection of

$\mathbb{R}^d$-valued random variables $X(t)$ parametrised by $t \in \mathbb{T}$, where $\mathbb{T} \subset \mathbb{R}$. When $\mathbb{T} = \{1, 2, ...\}$, we say that $X(t)$ is a stochastic process in discrete time (i.e. a sequence of random variables). When $\mathbb{T}$ is an interval in $\mathbb{R}$, like $\mathbb{T} = [0, \infty)$, we say that $X(t)$ is a stochastic process in continious time. We define
(1) a stochastic process $X = (X(t)), t \in \mathbb{T}$ by the map:

$$\mathbb{T} \times \Omega \ni (t, \omega) \to X(t, \omega) \in \mathbb{R}^d \tag{5.56}$$

(2) a sample path of X by the map:

$$\mathbb{T} \ni t \to X(t, \omega) \in \mathbb{R}^d \text{ for each } \omega \in \Omega \tag{5.57}$$

(3) a random variable $X(t)$ by the map:

$$\Omega \ni \omega \to X(t, \omega) \in \mathbb{R}^d \text{ for each } t \in \mathbb{T} \tag{5.58}$$

Definition 5.49. (Filtration [BZ, Definition 6.2]). A collection $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ of σ-fields on Ω parametrised by $t \in \mathbb{T}$, where $\mathbb{T} \subset \mathbb{R}$, is called a filtration if

$$\mathcal{F}_s \subset \mathcal{F}_t \text{ for any } s \leq t \in \mathbb{T} \tag{5.59}$$

Definition 5.50. (Martingale [BZ, Definition 6.3]). A process $X(t)$ parametrised by $t \in \mathbb{T}$, where $\mathbb{T} \subset \mathbb{R}$, is called a martingale (submartingale or supermartingale) with respect to a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ if

(1) $X(t)$ is integrable for each $t \in \mathbb{T}$

(2) $X(t)$ is $(\mathcal{F}_t)$-measurable for each $t \in \mathbb{T}$ (in which case we say that process $X(t)$ is $(\mathcal{F}_t)$-adapted)

(3) $X(s) = \mathbb{E}(X(t)|\mathcal{F}_s)$ (respectively, $\leq$ or $\geq$) for every $s, t \in \mathbb{T}$ such that $s \leq t$

Definition 5.51. (Adapted processes [IW, Definition 5.1]). Let $X = (X(t, \omega)), t \geq 0$ be a $\mathbb{R}^d$-valued process and fix $T \in (0, \infty)$. The process X is called $(\mathcal{F}_t)$-adapted if X_t is $(\mathcal{F}_t)$-measurable for every t. Process X is $(\mathcal{F}_t)$-measurable if the map

$$[0, T] \times \Omega \ni (t, \omega) \to X(t, \omega) \in \mathbb{R}^d \tag{5.60}$$

is $\mathcal{B}([0, T]) \times \mathcal{F} \setminus \mathcal{B}(\mathbb{R}^d)$-measurable.

Definition 5.52. (Predictable processes [IW, Definition 3.3]). Let $X = (X(t, x, \omega)), t \geq 0$ be a $\mathbb{R}^d$-valued process, fix $T \in (0, \infty)$ and $A \in \mathcal{B}(\mathbb{R}^d)$. Process X is $(\mathcal{F}_t)$-predictable if the map

$$[0, T] \times A \times \Omega \ni (t, x, \omega) \to X(t, x, \omega) \in \mathbb{R}^d \tag{5.61}$$

is $\mathcal{A} \setminus \mathcal{B}(\mathbb{R}^d)$-measurable. Here $\mathcal{A}$ is the smallest σ-field on $[0, T] \times A \times \Omega$ with respect to which all X having below properties are measurable:

(1) for each $t \in [0,T]$, mapping $(x,\omega) \to X(t,x,\omega)$ is $\mathcal{B}(A) \times (\mathcal{F}_t)$-measurable

(2) for each $x \in A, \omega \in \Omega$, mapping $t \to X(t,x,\omega)$ is left continious

Definition 5.53. (Progressively measurable process [Pr, Progressive process definition]). Let $X = (X(t,\omega)), t \geq 0$ be a $\mathbb{R}^d$-valued process and let $\mathcal{A}$ be a σ-field defined on $[0,T] \times \Omega$. We call X a $\mathbb{F}$-progressively measurable if the map

$$[0,t] \times \Omega \ni (s,\omega) \to X(s,\omega) \in \mathbb{R}^d \tag{5.62}$$

is $\mathcal{B}([0,t]) \times (\mathcal{F}_t)$-measurable for each $t \in [0,T]$. We call $\mathcal{A}$ a $(\mathcal{F}_t)$-progressively measurable σ-field, if $\mathcal{A}$ is the smallest σ-field such that all $(\mathcal{F}_t)$-progressively measurable processes are measurable.

Definition 5.54. (Stopping time [Ph, Definition 1.1.4]). Let τ we be an $[0,\infty]$-valued random variable. We call τ a $(\mathcal{F}_t)$-stopping time (a random time) if

$$\{\tau \leq t\} = \{\tau(\omega) \leq t : \omega \in \Omega\} \in \mathcal{F}_t \text{ for all } t \geq 0 \tag{5.63}$$

Definition 5.55. (First hitting time [Pr, Chapter 1: Section 1], [Ap, Section 2.2, p91-92]). Let $X = (X(t,\omega)), t \geq 0$ be a $\mathbb{R}^d$-valued and let $A \in \mathcal{B}(\mathbb{R}^d)$, then first hitting time of a process to a set is defined by

$$\tau_A(\omega) = \inf\{t > 0 : X(t,\omega) \in A\} \tag{5.64}$$

with the convention that $\inf\{\emptyset\} = \infty$.

Theorem 5.56. ***(First hitting time is stopping time [Pr, Chapter 1: Theorems 3,4]).*** *Let X be a process, τ_A be first hitting time and $A \in \mathcal{B}(\mathbb{R}^d)$ from Definition 5.55. If X is $(\mathcal{F}_t)$-adapted and càdlàg and set A is either open or closed, then τ_A is stopping time.*

Definition 5.57. (Stopped process [Ap, 2.2 Stopping times]). Let $X = (X(t,\omega)), t \geq 0$ be a $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-adapted process and let τ be a $(\mathcal{F}_t)$-stopping time. We call a random variable

$$X(\tau,\omega) = X(\tau(\omega),\omega) \text{ for each } \omega \in \Omega \tag{5.65}$$

a stopped random variable and we call X a stopped process at τ, which is defined by

$$X = (X(\tau \wedge t,\omega)), t \geq 0 \tag{5.66}$$

(Local martingales [Ok, 7.12. Local martingales]). An $(\mathcal{F}_t)$-adapted process $M = (M_t)_{t\geq 0}$ taking values in $\mathbb{R}^n$ is called a local martingale with respect to the given filtration $\mathbb{F}$ if there exists a sequence $(\tau_k)_{k\in\mathbb{N}}$ of $(\mathcal{F}_t)$-stopping times such that

(1) $\tau_k < \tau_{k+1}$ a.s.

(2) $\tau_k \to \infty$ as $k \to \infty$ a.s.

(3) $(M_{t\wedge\tau_k})_{t\geq 0}$ is an $(\mathcal{F}_t)$-martingale for every k

Definition 5.58. (Total variation [BZ, Definition 6.11]). Total variation of a function $f : [a, b] \to \mathbb{R}^d$ is defined as

$$V_b^a(f) = \sup_{\mathcal{P}} \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)| \tag{5.67}$$

where $\mathcal{P} = (x_0, x_1, ..., x_n)$ is a partition of $[a, b]$, i.e. $a = x_0 < x_1 < ... < x_n = b$.

Definition 5.59. (Process of finite variation [Ap, Subsection 2.3.3, p110]). Let $X = (X_t)_{t\geq 0}$ be $\mathbb{R}^d$-valued process defined on $[0, T] \times \Omega$. We call X a process of finite variation on if

$$V_0^T(X) < \infty \text{ for every } T > 0, \omega \in \Omega \tag{5.68}$$

5.3 Convergence of random variables

For this subsection let $(X_n)_{n=1}^{\infty}$ be a sequence of $\mathbb{R}^d$-valued random variables, let X be an $\mathbb{R}^d$-valued random variable and let $\mathbb{F} = (\mathcal{F}_n)_{n=1}^{\infty}$be a filtration.

Definition 5.60. (Pointwise convergence [Ru, 7.1 Definition]). Let A be a set, suppose we have a sequence of functions $(f_n(x))_{n\in\mathbb{N}}$ defined on A and suppose that the sequence $(f_n(x))_{n\in\mathbb{N}}$ converges for every $x \in A$. We then define a function f by

$$f(x) = \lim_{n\to\infty} f_n(x) \text{ for } x \in A \tag{5.69}$$

where we say that $(f_n(x))_{n\in\mathbb{N}}$ converges pointwise to f on A.

Definition 5.61. (Uniform convergence [Ru, 7.7 Definition]). Let A be a set and suppose we have a sequence of functions $(f_n(x))_{n\in\mathbb{N}}$ defined on A. We say that $(f_n(x))_{n\in\mathbb{N}}$ converges uniformly to a function f on A if for every $\epsilon > 0$ there exists integer N such that

$$|f_n(x) - f(x)| \le \epsilon \text{ for all } n \ge N,\ x \in A \tag{5.70}$$

Theorem 5.62. ***(Lebesgue's dominated convergence [Ap, Theorem 1.1.4]).*** *Let $A \subset \mathbb{R}^d$ and $(f_n(x))_{n\in\mathbb{N}}$be a sequence of Borel measurable functions converging pointwise to f, where $f_n : A \to \mathbb{R}^d$ and $f : A \to \mathbb{R}^d$. Let g be some integrable function such that*

$$|f_n(x)| \le g(x) \textit{ for all } n \in \mathbb{N} \tag{5.71}$$

in which case we say that $f_n(x)$ is dominated by $g(x)$. Then

$$\lim_{n\to\infty} \int_A f_n(x)\mu(dx) = \int_A f(x)\mu(dx) \tag{5.72}$$

Definition 5.63. (Convergence in distribution [Ap, Subsection 1.1.5, p14]). We say that $(X_n)_{n=1}^{\infty}$ converges in distribution if X exists such that

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f(x)\mathbb{P}_{X_n}(dx) = \int_{\mathbb{R}^d} f(x)\mathbb{P}_X(dx) \text{ for all } f \in C^0(\mathbb{R}^d) \tag{5.73}$$

Definition 5.64. (Convergence in probability [Ap, Subsection 1.1.5, p14]). We say that $(X_n)_{n=1}^{\infty}$ converges in probability if X exists such that

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - X| > a\right) = 0 \text{ for all } a > 0 \tag{5.74}$$

Definition 5.65. (Almost sure convergence [Ap, Subsection 1.1.5, p14]). We say that $(X_n)_{n=1}^{\infty}$ converges almost surely (a.s.) if X exists such that

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1 \tag{5.75}$$

Definition 5.66. (Convergence in L^p [Ap, Subsection 1.1.5, p14]). We say that $(X_n)_{n=1}^{\infty}$ converges in L^p if X exists such that

$$\lim_{n\to\infty} \mathbb{E}\left(|X_n - X|^p\right) = 0 \tag{5.76}$$

Definition 5.67. (Convergence relations [Ap, Subsection 1.1.5, p14]). Almost sure convergence implies convergence in probability and in distribution. L^p convergence also implies convergence in probability and in distribution. For convergence in probability we can always find a subsequence that converges almost surely.

Theorem 5.68. ***(Convergence to Lévy process [Ap, Theorem 1.3.7]).*** *If $X = (X(t)), t \geq 0$ is an $\mathbb{R}^d$-valued stochastic process and there exists a sequence of Lévy processes $(X_n)_{n=1}^{\infty}$ with each $X_n = (X_n(t)), t \geq 0$ such that*

(1) Convergence in probability: $\lim_{n\to\infty} \mathbb{P}\left(|X_n(t) - X(t)| > c\right) = 0$ for each $t \geq 0$

(2) $\lim_{n\to\infty} \limsup_{t\to 0} \mathbb{P}\left(|X_n(t) - X(t)| > c\right) = 0$ *for all* $c > 0$

then X is a Lévy process.

Definition 5.69. (Bounded in L^p [BZ, Exercise 4.6]). We say that $(X_n)_{n=1}^{\infty}$ is bounded in L^p if

$$\sup_{n<\infty} \mathbb{E}\left(|X_n|^p\right) < \infty \text{ for some } p > 0 \tag{5.77}$$

Theorem 5.70. ***(Doob's martingale convergence [BZ, Theorem 4.2, Remark 4.1]).*** *Suppose that sequence $(X_n)_{n=1}^{\infty}$ is an $(\mathcal{F}_n)_{n=1}^{\infty}$-martingale. If $(X_n)_{n=1}^{\infty}$ is bounded in L^p and there exists X such that*

$$\lim_{n\to\infty} X_n = X \;\; \textit{a.s. and in } L^p \tag{5.78}$$

then we say that X is an L^p-martingale.

Theorem 5.71. ***(Doob's martingale inequality [Ap, Theorem 2.1.5]).*** *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}$-valued stochastic process that is a positive submartingale. Then for any $p > 1$*

$$\mathbb{E}\left(\sup_{0\leq s\leq t} X_s^p\right) \leq q^p \mathbb{E}\left(X_t^p\right) \;\; \textit{for all } t > 0 \tag{5.79}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. If $X = (X_t)_{t\geq 0}$ is an $\mathbb{R}^d$-valued stochastic process that is a martingale, then

$$\mathbb{E}\left(\sup_{0\leq s\leq t} |X_s|^2\right) \leq 4\mathbb{E}\left(|X_t|^2\right) \;\; \textit{for all } t > 0 \tag{5.80}$$

Definition 5.72. (Uniform integrability [Ap, Subsection 2.2.1, p93]). Let $X = (X_n)_{n\in\mathbb{N}}$ be a collection of $\mathbb{R}$-valued random variables. We say that X is uniformly integrable if

$$\lim_{a\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}\left(|X_n| \mathbf{1}_{\{|X_n|>a\}}\right) = 0 \tag{5.81}$$

Theorem 5.73. ***(Doob-Meyer decomposition [Ap, Subsection 2.2.1, p94]).*** *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}$-valued stochastic process that is a L^2-martingale. Then there exists a unique $(\mathcal{F}_t)$-predictable, integrable and increasing process $A = (A_t)_{t\geq 0}$ with $A_0 = 0$ such that*

$$X_t = X_t^2 - X_0^2 - A_t \text{ for each } t \geq 0 \tag{5.82}$$

is a uniformly integrable martingale. The process $A = \langle X \rangle_t$ is called a quadratic variation, where $\langle X \rangle_t$ short-hand notation for $\langle X, X \rangle_t$.

5.4 Script

The simulation script is written in Python. Instructions for running on Windows:

1. Download and install Anaconda from https://www.anaconda.com/products/individual
2. Launch Anaconda
3. Locate and open Spyder
4. At the top left corner, select “File” and then “Open”
5. Locate the “Simulation.py” file and open it
6. Press “F5” to run

Please note that with 500 simulations script will take some time to run.

Listing 1: Simulation

```
import matplotlib.pyplot as plt
from scipy.stats import norm
import scipy.stats
import pandas as pd
import numpy as np
i=complex(0,1)

#Get last row of struct
def unpack(st):
    dct = dict()
    for x,y in zip(st[:][-1],st.dtype.names):
        dct[y]=x.tolist()
    return pd.DataFrame(dct)

#Integral approximation via trapezoid method
def trapz(f,a,b,n,*args):
    x = np.linspace(a,b,n)
    y = np.nan_to_num(f(x,*args))
    h = (b-a)/(n-1)
    return (h/2)*(y[1:]+y[:-1]).sum()

```

```
#Draw random samples from a multivariate normal distribution
def dWdW(dt,rho):
    mu = np.array([0,0])
    cov = np.array([[1,rho],[rho,1]])
    return np.random.multivariate_normal(mu,dt*cov)

#Draw random samples from Poisson distribution
def dN(dt,lam):
    return np.random.poisson(lam*dt)

class BlackScholes:
    def __init__(self,S_t,sig,tau,r):
        self.S_t = S_t
        self.sig = sig
        self.tau = tau
        self.r = r

    def d1(self,K):
        return (np.log(self.S_t/K)+(self.r+(self.sig**2)/2)*self.tau)/ \
                (self.sig*np.sqrt(self.tau))

    def d2(self,K):
        return (np.log(self.S_t/K)+(self.r-(self.sig**2)/2)*self.tau)/ \
                (self.sig*np.sqrt(self.tau))

    def C(self,K):
        return (self.S_t*norm.cdf(self.d1(K)))- \
                (K*np.exp(-self.r*self.tau)*norm.cdf(self.d2(K)))

    def P(self,K):
        return (K*np.exp(-self.r*self.tau)*norm.cdf(-self.d2(K)))- \
                (self.S_t*norm.cdf(-self.d1(K)))

    def vega(self,K):
        return (self.S_t*norm.pdf(self.d1(K))*np.sqrt(self.tau))

```

```
    def delta_C(self,K):
        return norm.cdf(self.d1(K))

    def delta_P(self,K):
        return norm.cdf(-self.d1(K))

class Heston:
    def __init__(self,kap,the,eta,rho,measure="Q"):
        self.kap = kap
        self.the = the
        self.eta = eta
        self.rho = rho
        self.measure = measure

    #Set state variables
    def set(self,S_t,Y_t,tau,r):
        self.S_t = S_t
        self.Y_t = Y_t
        self.tau = tau
        self.r = r
        self.F = S_t*np.exp(r*tau)
        return self

    #Characteristic function
    def CF(self,u,j):
        m = ((1+i*u)*self.r*self.tau)/self.the if self.measure == "P" else 0
        b_j = self.kap-j*self.rho*self.eta
        alp_j = -.5*u**2+i*u*(j-.5)
        bet_j = b_j-self.rho*self.eta*i*u
        gam = .5*self.eta**2
        d_j = np.sqrt(bet_j**2-4*alp_j*gam)
        rm_j = (bet_j-d_j)/(2*gam)
        rp_j = (bet_j+d_j)/(2*gam)
        g_j = rm_j/rp_j
        D_j = rm_j*((1-np.exp(-d_j*self.tau))/(1-g_j*np.exp(-d_j*self.tau)))
        C_j = m+self.kap*(rm_j*self.tau-2/self.eta**2* \
```

```
                        np.log((1-g_j*np.exp(-d_j*self.tau))/(1-g_j)))
        return np.exp(C_j*self.the+D_j*self.Y_t+i*u*np.log(self.S_t))

    #Probability density function
    def PDF(self,K,j):
        def _I(u,k,j): #Integrand
            return np.real(self.CF(u,j)*np.exp(-i*u*k))
        k = np.log(K)
        I = trapz(_I,1e-8,2e2,1000,k,j)
        return 1/np.pi * I

    #Cumulative distribution function
    def CDF(self,K,j):
        def _I(u,k,j): #Integrand
            return np.real((self.CF(u,j)*np.exp(-i*u*k))/(i*u))
        k = np.log(K)
        I = trapz(_I,1e-8,2e2,1000,k,j)
        return .5 + 1/np.pi * I

    #Call px
    def C(self,K):
        P0 = self.CDF(K,0) #Probability under risk neutral/real world measure
        P1 = self.CDF(K,1) #Probability under numeraire induced measure
        return self.S_t*P1 - K*np.exp(-self.r*self.tau)*P0

    #Put px via put-call parity
    def P(self,K):
        return self.C(K) + K*np.exp(-self.r*self.tau) - self.S_t

    #Call delta
    def delta_C(self,K):
        def _I(u,k,j,S): #Integrand
            return np.real((self.CF(u,j)*np.exp(-i*u*k))/S)
        P_1 = self.CDF(K,1)
        k = np.log(K)
        I_1 = trapz(_I,1e-8,2e2,1000,k,1,self.S_t)
```

```
        I_0 = trapz(_I,1e-8,2e2,1000,k,0,self.S_t)
        return P_1 + self.S_t/np.pi * I_1 - K/np.pi * I_0

    #Put delta via put-call parity
    def delta_P(self,K):
        return self.delta_C(K)-1

class StochasticControl:
    def __init__(self,A,B,gam,mu,r):
        self.A = A
        self.B = B
        self.gam = gam
        self.mu = mu
        self.r = r

    #Set state variables
    def set(self,S_t,Y_t,CQ_t,CP_t,del_t,tau):
        self.S_t = S_t
        self.Y_t = Y_t
        self.CQ_t = CQ_t
        self.CP_t = CP_t
        self.del_t = del_t
        self.tau = tau
        return self

    #Optimal controls risk neutral market maker
    def RN(self):
        M_0 = self.CQ_t - self.CP_t + self.mu*self.tau*self.S_t*self.del_t
        _x = 2*self.gam-1
        _d = np.sqrt(self.gam**2 * M_0**2 + _x*self.B)
        d_a_L = (_d - self.gam*M_0) / _x
        d_b_L = (_d + self.gam*M_0) / _x
        return d_a_L,d_b_L

    #Optimal controls risk averse market maker
    def RA(self,Q1_t,eps):
```

```
        M_0 = self.CQ_t - self.CP_t + self.mu*self.tau*self.S_t*self.del_t
        _x = 2*self.gam-1
        _d = np.sqrt(self.gam**2 * M_0**2 + _x*self.B)
        d_a_L = (_d - self.gam*M_0) / _x
        d_b_L = (_d + self.gam*M_0) / _x

        lam_a_L = self.lam(d_a_L)
        lam_b_L = self.lam(d_b_L)

        the_4 = -self.tau*self.del_t**2*self.Y_t*self.S_t**2
        the_3 = 2*(lam_a_L-lam_b_L)*self.tau**2*self.del_t**2*self.Y_t*self.S_t**2

        M_1 = -the_3+(1-2*Q1_t)*the_4
        M_2 = -the_3-(1+2*Q1_t)*the_4

        M_a = M_0 + eps*M_1
        M_b = M_0 + eps*M_2

        d_a_e = (np.sqrt(self.gam**2 * M_a**2 + _x*self.B) - self.gam*M_a) / _x
        d_b_e = (np.sqrt(self.gam**2 * M_b**2 + _x*self.B) + self.gam*M_b) / _x
        return d_a_e,d_b_e

    #Order arrival intensity under square root market impact
    def lam(self,d):
        return self.A/(self.B+d**2)**self.gam

class Simulation:
    def __init__(self,params:dict):
        #Other params
        self.T = params["T"]
        self.dt = params["dt"]
        self.K = params["K"]
        self.M = params["M"] #Number of sims
        self.N = int(self.T/self.dt)
        self.eps = params["eps"]

```

```python
        #Heston params - measure P
        self.kap_R = params["kap_R"]
        self.the_R = params["the_R"]
        self.eta_R = params["eta_R"]
        self.rho_R = params["rho_R"]

        #Heston params - measure Q
        self.kap_I = params["kap_I"]
        self.the_I = params["the_I"]
        self.eta_I = params["eta_I"]
        self.rho_I = params["rho_I"]

        #Stochastic control params
        self.r = params["r"]
        self.mu = params["mu"]
        self.A = params["A"]
        self.B = params["B"]
        self.gam = params["gam"]

        #State process initial values
        self.S_0 = params["S_0"]
        self.Y_0 = params["Y_0"]
        self.X_0 = params["X_0"]
        self.Q1_0 = params["Q1_0"]
        self.Q2_0 = params["Q2_0"]

        self.SC = StochasticControl(self.A,self.B,self.gam,self.mu,self.r)
        self.HQ = Heston(self.kap_I,self.the_I,self.eta_I,self.rho_I,"Q")
        self.HP = Heston(self.kap_R,self.the_R,self.eta_R,self.rho_R,"P")

        assert 0 < self.eps < 1, "Small␣parameter␣is␣has␣to␣be␣between␣0␣and␣1"
        assert (self.A > 0 and self.B > 0 and self.gam > 1), \
               "A,B␣has␣to␣be␣larger␣then␣0␣and␣gamma␣larger␣then␣1"
        assert (2*self.kap_I*self.the_I > self.eta_I**2 and
               2*self.kap_R*self.the_R > self.eta_R**2), \
               "2*kappa*theta␣has␣to␣be␣larger␣then␣eta^2"
```

```python

    def run(self):
        sp_head = ([(x,np.float64,(self.M,)) for x in ["t","S_t","Y_t","CH_t","CBS_t"]])
        mm_head = ([(x,np.float64,(self.M,)) for x in ["t","X_t","Q1_t","Q2_t","d_a","d_b"]])
        sp = np.zeros(self.N,dtype=sp_head) #State process struct
        rn = np.zeros(self.N,dtype=mm_head) #Risk neutral market maker struct
        ra = np.zeros(self.N,dtype=mm_head) #Risk averse market maker struct
        zi = np.zeros(self.N,dtype=mm_head) #Zero intelligence market maker struct

        for m in range(self.M):
            print(f"Path:{m}")
            S_t,Y_t = self.S_0,self.Y_0
            X_RN_t,X_RA_t,X_ZI_t = self.X_0,self.X_0,self.X_0
            Q1_RN_t,Q1_RA_t,Q1_ZI_t = self.Q1_0,self.Q1_0,self.Q1_0
            Q2_RN_t,Q2_RA_t,Q2_ZI_t = self.Q2_0,self.Q2_0,self.Q2_0
            for n in range(self.N):
                t = n*self.dt
                tau = self.T-t

                #State process
                dW = dWdW(self.dt,self.rho_I)
                dW_1,dW_2 = dW[1],dW[0]
                dS = S_t*self.mu*self.dt + S_t*np.sqrt(Y_t)*dW_2
                dY = self.kap_I*(self.the_I-Y_t)*self.dt + self.eta_I*np.sqrt(Y_t)*dW_1
                Y_t += dY
                S_t += dS
                Y_t=max(Y_t,1e-308) #Bound by float point limit

                #Option prices
                self.HQ.set(S_t,Y_t,tau,self.r)
                self.HP.set(S_t,Y_t,tau,self.mu)
                BS = BlackScholes(S_t,np.sqrt(Y_t),tau,self.r)
                CHQ_t = self.HQ.C(self.K)
                CHP_t = self.HP.C(self.K)
                CBS_t = BS.C(self.K)
                self.SC.set(S_t,Y_t,CHQ_t,CHP_t,self.HQ.delta_C(self.K),tau)
```

```

            #Risk-neutral market maker
            d_RN_a,d_RN_b = self.SC.RN()
            lam_RN_a = self.SC.lam(d_RN_a)
            lam_RN_b = self.SC.lam(d_RN_b)
            dN_RN_a,dN_RN_b = dN(self.dt,lam_RN_a),dN(self.dt,lam_RN_b)
            Q1_RN_t += dN_RN_b - dN_RN_a
            Q2_RN_t = int(-Q1_RN_t*self.HQ.delta_C(self.K)) #Delta hedging
            X_RN_t += (CHQ_t+d_RN_a)*dN_RN_a-(CHQ_t-d_RN_b)*dN_RN_b+Q2_RN_t*dS

            #Risk-averse market maker
            d_RA_a,d_RA_b = self.SC.RA(Q1_RA_t,self.eps)
            lam_RA_a = self.SC.lam(d_RA_a)
            lam_RA_b = self.SC.lam(d_RA_b)
            dN_RA_a,dN_RA_b = dN(self.dt,lam_RA_a),dN(self.dt,lam_RA_b)
            Q1_RA_t += dN_RA_b - dN_RA_a
            Q2_RA_t = int(-Q1_RA_t*self.HQ.delta_C(self.K)) #Delta hedging
            X_RA_t += (CHQ_t+d_RA_a)*dN_RA_a-(CHQ_t-d_RA_b)*dN_RA_b+Q2_RA_t*dS

            #Zero-intelligence market maker
            d_ZI = .005 * BS.vega(self.K)
            lam_ZI = self.SC.lam(d_ZI)
            dN_ZI_a,dN_ZI_b = dN(self.dt,lam_ZI),dN(self.dt,lam_ZI)
            Q1_ZI_t += dN_ZI_b - dN_ZI_a
            Q2_ZI_t = -int(Q1_ZI_t*BS.delta_C(self.K)) #Delta hedging
            X_ZI_t += (CHQ_t+d_ZI)*dN_ZI_a - (CHQ_t-d_ZI)*dN_ZI_b + Q2_ZI_t*dS

            #Data
            sp["t"][n,m]=rn["t"][n,m]=ra["t"][n,m]=zi["t"][n,m]=t
            sp["S_t"][n,m],sp["Y_t"][n,m]=S_t,Y_t
            sp["CH_t"][n,m],sp["CBS_t"][n,m]=CHQ_t,CBS_t

            rn["X_t"][n,m],rn["d_a"][n,m],rn["d_b"][n,m]=X_RN_t,d_RN_a,d_RN_b
            rn["Q1_t"][n,m],rn["Q2_t"][n,m]=Q1_RN_t,Q2_RN_t

            ra["X_t"][n,m],ra["d_a"][n,m],ra["d_b"][n,m]=X_RA_t,d_RA_a,d_RA_b
```

```
                ra["Q1_t"][n,m],ra["Q2_t"][n,m]=Q1_RA_t,Q2_RA_t

                zi["X_t"][n,m],zi["d_a"][n,m],zi["d_b"][n,m]=X_ZI_t,d_ZI,d_ZI
                zi["Q1_t"][n,m],zi["Q2_t"][n,m]=Q1_ZI_t,Q2_ZI_t

        self.sp = unpack(sp)
        self.rn = unpack(rn)
        self.ra = unpack(ra)
        self.zi = unpack(zi)
        return self

    #Histograms for terminal values of selected variable across 3 market makers
    def figures(self,var):
        assert self.M > 5, "Not enough paths simulated to get figures"

        #Plot Histograms
        x = self.zi[var]+self.rn[var]+self.ra[var]
        xmx,xmn = max(x),min(x)
        plt.hist(self.zi[var],density=True,bins=30,label="Zero Intelligence",
                range=(xmn,xmx),alpha=0.6,edgecolor='black')
        plt.hist(self.rn[var],density=True,bins=30,label="Risk Neutral",
                range=(xmn,xmx),alpha=0.6,edgecolor='black',color="gray")
        plt.hist(self.ra[var],density=True,bins=30,label="Risk Averse",
                range=(xmn,xmx),alpha=0.6,edgecolor='black')
        mn,mx = plt.xlim()
        plt.xlim(mn,mx)
        plt.legend()
        plt.title(var)
        plt.show()

        #Plot density estimators
        x = np.linspace(mn,mx,300)
        kde = scipy.stats.gaussian_kde(self.zi[var])
        plt.plot(x,kde.pdf(x),label="Zero Intelligence")
        kde = scipy.stats.gaussian_kde(self.rn[var])
        plt.plot(x,kde.pdf(x),label="Risk Neutral",color="gray")
```

```python
        kde = scipy.stats.gaussian_kde(self.ra[var])
        plt.plot(x,kde.pdf(x),label="Risk Averse")
        plt.legend()
        plt.title(var)
        plt.show()

    #Descriptive statistics for terminal values of selected variable across 3 market makers
    def stats(self,var):
        assert self.M > 5, "Not enough paths simulated to get stats"
        print(f"\n{var}:")

        df = pd.DataFrame(columns=["Market maker","Mean","Variance","Skewness","Kurtosis"])
        mm = ["Zero Intelligence","Risk Neutral","Risk Averse"]
        for x,n in zip([self.zi,self.rn,self.ra],range(3)):
            df.loc[n] = [mm[n],x[var].mean(),x[var].var(),x[var].skew(),x[var].kurtosis()]
        print(df)

#--------------------------------------------------------------------------------
params = {
         "T" : 1/12 #Expiry date
        ,"dt" : (1/12)/500 #Time increment
        ,"K" : 100 #Strike price
        ,"M" : 500 #Number of paths to simulate
        ,"eps" : 0.5 #Small parameter value

        ,"kap_R" : 4 #Kappa in Heston model under measure P
        ,"the_R" : 0.04 #Theta in Heston model under measure P
        ,"eta_R" : 0.5 #Eta in Heston model under measure P
        ,"rho_R" : -0.4 #Rho in Heston model under measure P

        ,"kap_I" : 4 #Kappa in Heston model under measure Q
        ,"the_I" : 0.04 #Theta in Heston model under measure Q
        ,"eta_I" : 0.5 #Eta in Heston model under measure Q
        ,"rho_I" : -0.4 #Rho in Heston model under measure Q

        ,"r" : 0.001 #Spot price process drift under measure Q
```

```
        ,"mu" : 0.001 #Spot price process drift under measure P
        ,"A" : 500 #Order arrival intensity parameter A
        ,"B" : 1 #Order arrival intensity parameter B
        ,"gam" : 1.5 #Order arrival intensity parameter gamma

        ,"S_0" : 100 #Spot price process initial value
        ,"Y_0" : 0.04 #Variance process initial value
        ,"X_0" : 0 #Cash process initial value
        ,"Q1_0" : 0 #Options inventory process initial value
        ,"Q2_0" : 0 #Stock inventory process initial value
        }

S = Simulation(params)
S.run()
S.figures("X_t")
S.stats("X_t")
S.figures("Q1_t")
S.stats("Q1_t")
```

References

[Ap] D. Applebaum, Levy processes and stochastic calculus, 2nd edition, Cambridge University Press, (2009)

[AS] M. Avellaneda and S. Stoikov, High-frequency trading in a limit order book, Quant. Finance 8, no. 3, 217–224 (2008)

[Bj] T. Björk, Arbitrage theory in continious time, 2nd edition, Oxford University Press, (2004)

[BZ] Z. Brzezniak and T. Zastawniak, Basic stochastic processes: A course through exercises, SpringerVerlag London, (1999)

[ElAA3] S. El Aoud and F. Abergel, A stochastic control approach for options market making, Market Microstructure and Liquidity, Vol. 01, No. 01, 1550006 (2015), https://hal.archives-ouvertes.fr/hal-01061852v3

[ElAA2] S. El Aoud and F. Abergel, A stochastic control approach for options market making, Market Microstructure and Liquidity, (2015), https://hal.archivesouvertes.fr/hal-01061852v2

[Fr] A. Friedman, Partial differential equations of parabolic type, Prentice-Hall, (1983)

[FS] W.H. Fleming and H.M. Soner, Controlled markov processes and viscosity solutions, 2nd edition, Springer, (2006)

[Ga] J. Gatheral, The volatility surface, John Wiley & Sons, (2006)

[Gr] M.D. Greenberg, Advanced engineering mathematics, 2nd edition, Prentice-Hall, (1998)

[HS] T. Ho and H. R. Stoll, Optimal dealer pricing under transactions and return uncertainty, J. Finance 9, no. 1, 47–73 (1981)

[IW] N. Ikeda, S. Watanabe, Stochastic differential equations and Diffusion processes, 2nd edition, (1989)

[Mu] J.D. Murray, Asymptotic analysis, SpringerVerlag New York Berlin Heidelberg Tokyo, (1984)

[Na] A.H. Nayfeh, Perturbation methods, John Wiley & Sons, (2000)

[Ok] B. Øksendal, Stochastic differential equations : an introduction with applications, 5th edition, Springer-Verlag Heidelberg New York, (1998)

[Ol] P.J. Olver, Introduction to partial differential equations, Springer, (2014)

[OS] B. Øksendal and A. Sulem, Applied stochastic control of jump diffusions, Springer, (2005)

[Ph] H. Pham, Continuous-time stochastic control and optimization with financial applications, Springer, (2009)

[Pr] E.P. Protter, Stochastic integration and differential equations, Springer Berlin / Heidelberg, (2013)

[Ra] S.S. Rao, Engineering optimization theory and practice, 4th edition, John Wiley & Sons, (2009)

[Ro] F.D. Rouah, The Heston model and its extensions in Matlab and C#, John Wiley & Sons, (2013)

[Ru] W. Rudin, Principles of mathematical analysis, 3rd edition, McGraw-Hill, (1976)

[Ve] F. Verhulst, Methods and applications of singular perturbations, Springer, (2000)