Lévy Processes, Representations and Models with Applications in Finance

by

WING YAN YIP

A thesis presented for the degree of
Doctor of Philosophy of Imperial College London

Department of Mathematics
Imperial College London
Huxley Building
180 Queens Gate
London SW7 2BZ
United Kingdom

JULY 26, 2008

Abstract

Lévy processes are becoming increasingly important in Mathematical Finance. This thesis aims to contribute to the development of theoretical representations of Lévy processes and their financial applications. The first part of the thesis presents a computational explicit formula of the chaotic representation property (CRP) for the powers of increments of a Lévy process. The formula can be used to obtain the integrands of the CRP in terms of the orthogonalised compensated power jump processes and the CRP in terms of Poisson random measures.

The second part of the thesis presents hedging strategies for European and exotic options in a Lévy market. By applying Taylor's theorem, dynamic hedging portfolios are constructed and in the case of European options, static hedging is also implemented. It is shown that perfect hedging can be achieved by investing in power jump assets, moment swaps or some traded financial derivatives depending on the same underlying asset. Note that variance swaps are special cases of moment swaps and are traded in OTC (Over-The-Counter) markets. We can also hedge by constructing the minimal variance portfolios that invest in the risk-free bank account, the underlying stock and variance swaps. The numerical algorithms and performance of the hedging strategies are presented.

The third part of the thesis contributes to the design of an option trading strategy, where the stock price is driven by a Lévy process. The trading strategy is based on comparing the deviations between the density implied by historical time series and that implied by current market prices of the options. The performance of the trading strategy under different market conditions is reported and optimal parameters are obtained using efficient frontier analysis. The analysis compares the expected returns with the Conditional Value at Risks (CVaRs). Simulation results show that the trading strategy has a high earning potential.

Declaration of Independence

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Wing Yan Yip

Acknowledgements

I would like to start by thanking my supervisors Professor Sofia Olhede and Professor David Stephens. I would like to thanks the Overseas Research Students Awards Scheme and the Royal Bank of Scotland (RBS) for their financial support. I have to thank Mr Andrew Walker, Mr Simon Wenden and Dr Robin Thompson in RBS for their support and encouragement. I have learned a great deal from them about real life financial applications and they provided very useful advice on Part III of this thesis.

Contents

List of Tables

List of Figures

Introduction

This thesis contributes to the development of theoretical representations and financial applications of Lévy processes and comprises of three main parts. The first part is concerned with chaotic representations of a Lévy process. The second part investigates hedging strategies for European and exotic options in a Lévy market with the use of the representation property of Lévy processes. The third part proposes a trading (speculating) strategy investing in European options in a Lévy market.

Part I

To price and hedge derivative securities, it is crucial to have a good model for the evolution of the underlying asset. Despite the popularity of the Black-Scholes model, empirical evidence suggests that it is not sufficiently flexible to describe some of the important statistical properties observed in realised market data. Cont (2001), Schoutens (2003, Chapter 4) and Cont & Tankov (2003, Chapter 1) discussed various stylised empirical facts emerging from statistical analysis of price variations in various types of financial markets. There are two main problems that give rise to the need of more general models. Firstly, abrupt downward jumps have been observed in stock price processes while the Brownian motion, that is used in the Black-Scholes model, is a continuous process. Barndorff-Nielsen & Shephard (2006) performed hypothesis tests on exchange data under the null of no jumps, which was found to be rejected frequently. Secondly, the log return data series has heavy tails and is negatively skewed, which cannot be described by the normal distribution implied by the Black-Scholes formulation. To remedy these problems, market models driven by Lévy processes, that is, processes with independent and stationary increments, were introduced to replace the Black-Scholes model in describing the dynamics of asset price process.

The Lévy-Khintchine formula (see equation (2.1) below) states that a Lévy process can be decomposed into three components: a deterministic drift component, a Brownian component and a pure jump component. A Brownian motion is a special case of a Lévy process in which the pure jump component equals zero. The pure jump component provides more flexibility in describing the shape of the distribution of the log asset price processes since heavy tails and asymmetry are potential characteristics of processes with jumps.

In this thesis, we focus on the different stochastic representations of Lévy processes. The chaotic representation of a square integrable functional of a Lévy process is an expansion via its expectation plus a sum of iterated stochastic integrals, see Solé *et al.* (2006) for a recent review of such representations. The chaotic representations are important in mathematical finance since they provide the decomposition of a random variable adapted to the filtration generated by the underlying Lévy process into orthogonal components. Such representations are useful in the construction of hedging strategies of financial derivatives. We discuss this in more details in Part II of this introduction. There are two different types of chaos expansions: Itô (1956) proved a Chaotic Representation Property (CRP) for any square integrable functional of a general Lévy process. Note that the special cases of the CRP for Brownian motion and Poisson process are in much simpler forms and are commonly treated in the literature. The CRP is written in terms of multiple integrals with respect to a two-parameter random measure associated with the Lévy process. Nualart & Schoutens (2000) proved the existence of a new version of the CRP, which states that every square integrable Lévy functional can be represented as its expectation plus an infinite sum of stochastic integrals with respect to the orthogonalised compensated power jump processes of the underlying Lévy process. Benth *et al.* (2003) and Solé *et al.* (2006) derived the relationships between these two representations. However, these representations are computationally intractable. The first part of the thesis addresses this issue. For the powers of increments of a Lévy process, we derive computationally explicit formulae for the integrands of these two chaotic expansions.

Power jump processes are important in mathematical finance. The jumps can be understood both in terms of a Poisson random measure, or equivalently, by using the Power jump processes. Note that Nualart & Schoutens (2000, Proposition 2) proved that all square integrable random variables, adapted to the filtration generated by the Lévy process denoted by $X = \{X_t, t \geq 0\}$, can be represented as a linear combination of powers of increments of X, see Proposition 3.0.1 below. In fact, for any square integrable random variable, F, with derivatives of all orders, we can apply Taylor's theorem to express F in terms of a polynomial of powers of increments of X. Thus, the chaotic representations of certain financial derivatives can be found using this method, which is discussed further in Section 5.1.

The derivation of an explicit formula for the chaotic representation has been the focus of considerable study, see for example Nualart & Schoutens (2001), Léon *et al.* (2002), Løkka (2004) and Eddahbi *et al.* (2005). All the explicit formulae for general Lévy functionals derived in these papers used Malliavin-type derivatives to derive explicit representations of stochastic processes for applications in finance. Malliavin Calculus was originally developed as a new probabilistic technique to find smooth densities for solutions of stochastic differential equations. Note that the use of Malliavin Calculus in finance is mainly in the computation of the Greeks of options. By expressing the Greeks in terms of some stochastic integrals using Malliavin Calculus, their values can then be approximated quickly and accurately, see Davis & Johansson (2006) and Huehne (2005). Accordingly, the explicit formula derived using Malliavin Calculus is generally not designed to be used to find the explicit representation of a contingent claim in terms of the integrals with respect to the power jump processes. The derivative operator D is, in all of these cases, defined by its action on the chaos expansions. In other words, the explicit chaos expansion must in fact be known before D can be applied to find the explicit form of the chaotic representation, thus yielding a circular specification. We will discuss this in further details in Sections 2.3 and 5.3. As pointed out by Solé *et al.* (2007), 'in order for the Malliavin calculus to be genuinely useful, there is the need for practical rules to compute the derivatives.' In the case of Brownian motion, Nualart (1995) proved a chain rule through the identification of the Malliavin derivative with a weak derivative on the canonical space. For Poisson process, Nualart & Vives (1990) proved that the Malliavin derivative coincides with a difference operator on the canonical space. The derivatives with respect to the compensated power jump processes introduced in Léon *et al.* (2002) were only alternative definitions for the derivative and useful formulae were only developed for a jump-diffusion process with only a finite number of jump sizes. As pointed out by Davis & Johansson (2006), a drawback of this approach used in Léon *et al.* (2002) is that there is no general chain rule. In this thesis, we take a different approach by deriving an explicit formula for the power of increment of a Lévy process directly using Itô formula. For any smooth square integral random variable, we apply Taylor's theorem to express it in terms of those power of increments and hence our explicit formulae can be applied.

Apart from the Malliavin approach, Jamshidian (2005) extended the CRP in Nualart & Schoutens (2000) to a large class of semimartingales and derived the explicit representation of the power of a semimartingale with respect to the corresponding non-compensated power jump processes. In this thesis, we derived an explicit representation of the power of a Lévy process with respect to the corresponding orthogonalised compensated power jump processes. Note that Lévy processes are included in the class of semimartingales, see Kannan & Lakshmikantham (2001, Corollary 2.3.21, p.92). The explicit formula derived in this thesis is designed for those stochastic processes with compensators equal to a constant times t only (which is satisfied by all Lévy processes). This formula can be easily extended to semimartingales when the form of the compensator is known. Our result is therefore complementary to Jamshidian's formula, since our explicit formula gives the CRP with respect to the orthogonalised processes, as defined by Nualart & Schoutens (2000). Note that it is a non-trivial extension from the representation in terms of non-compensated power jump processes to that in terms of orthogonalised compensated power jump processes. In other words, Jamshidian's formula can be deduced from ours (in the Lévy case), but ours cannot be deduced from Jamshidian's by a non-trivial calculation. This is further discussed in Remark 3.2.1.

It is important to be able to express the chaos expansion with respect to orthogonalised compensated power jump processes since it facilitates the applications of the CRP. An immediate result of the CRP is the predictable representation property (PRP), which states that every square integrable functional of a Lévy process can be expressed as an expansion via its expectation plus a stochastic integral with predictable integrand. In practical applications, it is often convenient to truncate the representation given by the PRP. The truncated representation of a stochastic process yields a practically implementable approximation to the stochastic process. This approximation would be used for simulating the process, or for a finite number of traded higher order options, providing hedging formulae as will be discussed in Part II. The advantage of expressing the sum in terms of stochastic integrals with respect to the orthogonalised processes is that the error terms omitted will be uncorrelated with the terms remaining in the approximation.

Part II

The second part of the thesis presents hedging strategies for European and exotic options in a Lévy market. By applying Taylor's theorem, we construct dynamic hedging portfolios under different market assumptions, such as the existence of power jump assets or moment swaps. In the case of European options or baskets of European options, static hedging is also implemented. It is shown that perfect hedging can be achieved.

It is well known, see Schoutens (2000, p.71), that Brownian motion has an elegant version of the CRP: every square integrable random variable adapted to the filtration generated by a Brownian motion can be represented as a sum of its mean and an infinite sum of iterated stochastic integrals with respect to the Brownian motion, with deterministic integrands. This is distinct from the CRP for Lévy processes, which are in terms of power jump processes or Poisson random measures rather than the Lévy process itself. The PRP for Brownian motion states that every square integrable random variable adapted to the filtration generated by a Brownian motion can be represented in the same form, but with a single stochastic integral, where the integrand is a predictable process. The PRP implies the completeness of the Black-Scholes option pricing model. The aforementioned predictable process gives the admissible self-financing strategy of replicating a contingent claim whose price only depends on the time to maturity and the current stock price, which can be hedged by investing in a risk-free bank account and the underlying asset.

Unfortunately, this kind of PRP, where the stochastic integral is with respect to the underlying process only, is an exceptional property, which is only possessed by a few martingales, including the Brownian motion, the compensated Poisson process, and the Azéma martingale (see Schoutens (2003) and Dritschel & Protter (1999)). When the underlying asset is driven by a Lévy process, perfect hedging using only a risk-free bank account and the underlying asset is not in general possible. The market is therefore incomplete. However, even in this case, further developments are possible. As mentioned in Part I, Nualart & Schoutens (2000) proved the existence of a new version of the CRP for Lévy processes which satisfy some exponential moment conditions. This new version states that every square integrable random variable adapted to the filtration generated by a Lévy process can be represented as an infinite sum of iterated stochastic integrals with respect to the orthogonalised compensated power jump processes of the underlying Lévy process. The market can thus be completed by allowing trades in these processes while risks due to jumps and fat tails are considered. In light of the new version of the PRP, Corcuera *et al.* (2005) suggested that the market should be enlarged with power jump assets so that perfect hedging could still be implemented. Corcuera *et al.* (2006) used this completeness to solve the portfolio optimisation problem using the martingale method. Another form of commonly traded financial derivative is the variance swap which depends functionally on the volatility of the underlying asset. Since variance swaps are already traded commonly in the over-the-counter (OTC) markets, Schoutens (2005) suggested trading in moment swaps, which are a generalisation of variance swaps. Based on the CRP derived by Itô (1956), Benth *et al.* (2003) derived a minimal variance portfolio for hedging contingent claims in a Lévy market.

Inspired by these papers, we derive practical and implementable hedging strategies based on the PRP derived from Taylor approximations to the option pricing formulae. We apply Taylor's theorem directly to the option pricing formulae and derive perfect hedging strategies by investing in power jump assets, moment swaps or some traded derivatives depending on the same underlying asset. The hedging of the higher moments terms in the Taylor expansion of a contingent claim using other contingent claims in a Lévy market is a technique introduced by this thesis. When these financial derivatives are not available, we demonstrate how to use the minimal variance portfolios derived by Benth *et al.* (2003) to hedge the higher order terms in the Taylor expansion. While we apply Taylor expansions to decompose the pricing formula into an infinite sum of higher moment terms, Corcuera *et al.* (2005) applied the Itô formula to obtain the PRP of a contingent claim. Note that the Itô formula is derived as a result of an elementary Taylor expansion, see Kijima (2002). In practice, when implementing a hedging strategy numerically, we have to discretise the time variable. Hence, it is more natural to work directly from

Taylor's theorem as this discretisation can be acknowledged explicitly. In fact, the delta and gamma hedges commonly used by traders in the market, given in Section 6.2.4, are derived using a Taylor expansion. We construct static and dynamic hedging strategies for European and exotic options in a Lévy market. Although static hedging is only applied to European options, exotic options can be decomposed into a basket of European options so that static hedging can be achieved, in this case see for example Derman *et al.* (1995). It is practically important to be able to statically hedge since static hedging has several advantages over dynamic hedging. Static hedging is less sensitive to the assumption of zero transaction costs (both commissions and the cost of paying individuals to monitor the positions). Moreover, dynamic hedging tends to fail when liquidity dries up or when the market makes large moves, but especially in these situations effective hedging is needed.

We discuss how hedging can be implemented by applying Taylor's theorem to a pricing formula. We investigate the approximation of the derivatives of the pricing formula and present the numerical procedures used to construct the hedging strategies. The Performance of the hedging is assessed and the difficulties encountered are discussed. Thus, this part of the thesis constitutes a practical development for the hedging of contingent claims, where the underlying asset is driven by a Lévy process.

Part III

In the third part of the thesis, we construct an option trading strategy in a Lévy market, where the price processes of the underlying assets are driven by Lévy processes. We compare the risk-neutral density of the log returns of the underlying at maturity implied by the historical data series of the underlying to that implied by the current option prices in the market. This comparison gives a strategy for speculating options in a Lévy market. This represents an important practical advance in utilizing the Lévy process model. According to the European option pricing formula, the price $P\left(S_t, K, r, T-t\right)$ is given by:

$$P\left(S_t, K, r, T-t\right) = \mathrm{e}^{-r(T-t)} \int_0^{\infty} H\left(x, K\right) f^{*}\left(x, S_t\right) \mathrm{d}x, \tag{1}$$

where H is the payoff function of the option, S_t is the current price of the underlying, K is the strike price of the option, r is the continuously compounded risk-free interest rate, $T-t$ is the time to maturity, x is the price of the underlying at maturity and $f^{*}\left(x, S_t\right)$ is the risk-neutral density of the underlying at maturity, depending on the current option price, S_t. This pricing formula states that the price of an European option today is given by the discounted expected payoff with respect to a risk-neutral measure. Aït-Sahalia *et al.* (2001), Blaskowitz (2001), Blaskowitz & Schmidt (2002) and Blaskowitz *et al.* (2004) considered the profitability of trading on the deviations of the risk-neutral density of the underlying inferred from the historical time series and that implied by the option prices under the Black-Scholes model. There are well-known indications, such as the volatility smile, suggesting that the Black-Scholes model is not sufficiently flexible to capture the statistical behaviour of the underlying. More importantly, it is assumed in the Black-Scholes model that, the market is complete and there is a unique equivalent martingale measure (EMM). Hence, there should not be any deviations of the two density functions if the model can accurately reflect the market prices of the options. In an incomplete market, there are infinitely many EMMs, which give different option prices. This is because an EMM gives an arbitrage-free price of an option but not necessarily the market price of the option. In other words, the market chooses an EMM and the market prices of options are obtained under such a measure. Therefore, if we choose a change of measure method to obtain a risk-neutral density from the historical data of the underlying, deviations between the two densities are expected since they are obtained from two *different* EMMs. Under an incomplete model, choosing an EMM, rather than using the market implied one, is essentially specifying the investors' risk preference. Hence, objective comparison (that is, independent of investors' preference) of the two risk-neutral densities inferred from the historical time series of the underlying and implied from the option prices, as in the papers cited above, is indeed not possible in an incomplete (realistic) market. To allow for realistic comparison of the two risk-neutral densities, we must adopt a model which would lead to the existence of non-unique EMMs. Lévy model is a straight forward extension to the Black-Scholes model since the extra parameters handle the skewness and kurtosis explicitly.

We fit the two sets of data, that is, the historical series of the underlying and the current option prices, to a market model to obtain two sets of parameters. We then simulate the underlying from today to maturity with these two sets of parameters to see which options are overpriced. We sell the overpriced options and also buy far out of money options to prevent infinite loss and hold them until maturity, which is discussed in further detail in Section 8.5. We use the Variance Gamma (VG) model, introduced by Madan *et al.* (1998), to describe the dynamics of the underlying price process. A VG process is a Brownian motion with a stochastic time change determined by a Gamma process. Note that other stochastic models can be used, for example, the stochastic volatility model using a VG process (VGSAM), introduced by Carr *et al.* (2003). The VG model is used because of its simplicity and ability to handle skewness and kurtosis, which correspond to asymmetry and fat tails of the distribution function, respectively. Since in a Lévy market model, the market is incomplete and there are infinitely many EMMs, we have to choose one to obtain the risk-neutral density implied by the historical data of the underlying. For simplicity, we use the mean-correcting martingale measure, see Schoutens (2003, Section 6.2.2). Miyahara (2005) discussed the different properties of a few common kinds of equivalent martingale measures (EMMs) for geometric Lévy processes. The choice of the EMM in this thesis is left to the preference of the investor. The thesis instead will focus on constructing the option trading strategy after an EMM is chosen and a stochastic model is fitted to market data. Therefore, the most appropriate choices of the EMM and the stochastic model for the underlying price process are out of the scope of this discussion.

We choose the mean-correcting martingale measure to obtain the risk-neutral density of the historical time series and compare it to the risk-neutral density implied by the option prices, hence identifying overpriced options today under our subjective belief that the mean-correcting martingale measure gives 'more accurate' prices. The performance of the trading strategy under different market conditions are reported and it suggests that the trading strategy has a high earning potential.

The trading strategy presented is a speculative strategy since we believe that the prices of the underlying should behave according to its historical performance and investors in the market are too risk averse. Although it has a high earning potential, occasionally it would lead to losses. To make the strategy more attractive to risk averse investors, we can combine the trading strategy with risk-free investment to guarantee the capital, which is known as portfolio insurances in finance, see Leland (1979). We discuss this in further details in Section 8.8.

These results presented in this thesis thus comprise both theoretical and practical developments for the usage of Lévy processes in practice.

Chapter 1

Background and Notation

In the introduction, we have discussed the motivation behind the use of market models driven by Lévy processes. In this chapter, we give fundamental results about probability theory, Lévy processes and martingales. Bertoin (1996), Sato (1999) and Applebaum (2004) provide comprehensive details of Lévy processes and stochastic calculus. Schoutens (2003) and Cont & Tankov (2003) provide recent overviews of financial applications of Lévy processes.

1.1 Martingales and random measures

In this section, we give the definitions of martingales and random measures, which are important components of stochastic calculus.

A *stochastic process* is a family $\{X_t : t \geq 0\}$ of random variables on $\mathbb{R}^d$ with parameter $t \in [0, \infty)$ defined on a common probability space. Let $\{X_t\}$ and $\{Y_t\}$ be two stochastic processes. If

$$P\left[X_t = Y_t\right] = 1 \quad \text{for } t \in [0, \infty),$$

then $\{Y_t\}$ is called a *modification* of $\{X_t\}$. If, for every $t \geq 0$ and $\varepsilon > 0$, the stochastic process $\{X_t\}$ on $\mathbb{R}^d$ satisfies

$$\lim_{s \to t} P\left[|X_s - X_t| > \varepsilon\right] = 0,$$

it is said to be *stochastically continuous* or *continuous in probability.* Suppose $\mathcal{F}$ is a σ-algebra of subsets of a given set Ω. A *filtration* is a family $\{\mathcal{F}_t, t \geq 0\}$ of sub σ-algebra of $\mathcal{F}$ such that

$$\mathcal{F}_s \subseteq \mathcal{F}_t \ \text{ for all } s \leq t.$$

A probability space $(\Omega, \mathcal{F}, P)$ is said to be *filtered* if it is equipped with such a family

$(\mathcal{F}_t, t \geq 0)$. Suppose $X = \{X_t, t \geq 0\}$ is a stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}, P)$. It is said to be *adapted* to the filtration (or $\mathcal{F}_t$-adapted) if X_t is $\mathcal{F}_t$-measurable for each $t \geq 0$. Note that any process $\{X_t\}$ is adapted to its own filtration $\mathcal{F}_t^X = \sigma\{X_s : 0 \leq s \leq t\}$, which is known as the *natural filtration*. We have $E[X_s|\mathcal{F}_s] = X_s$ a.s. if $\{X_t\}$ is adapted, that is, $\mathcal{F}_s$ contains all the information required to predict the behaviour of $\{X_t : t \geq 0\}$ up to and including time s. A *stopping time* is a random variable $T : \Omega \to [0, \infty]$ such that the event $(T \leq t) \in \mathcal{F}_t$ for each $t \geq 0$.

Definition 1.1.1 (Martingales) *A martingale is an adapted process X defined on a filtered probability space satisfying $E[|X_t|] < \infty$ for all $t \geq 0$ and $E[X_t|\mathcal{F}_s] = X_s$ a.s. for all $0 \leq s < t < \infty$. The mapping $t \to E[X_t]$ is constant if X is a martingale. Let X be d-dimensional and its i-th element at time t be $X_{i,t}$. A submartingale is an adapted process X satisfying $E[|X_t|] < \infty$ for all $t \geq 0$ and $E[X_{i,t}|\mathcal{F}_s] \geq X_{i,s}$ a.s. for all $0 \leq s < t < \infty$ and $1 \leq i \leq d$. X is a supermartingale if $-X$ is a submartingale. Let $M = \{M_t, t \geq 0\}$ be an adapted process. If there exists a sequence of stopping times $\tau_1 \leq \cdots \leq \tau_n \to \infty$ a.s. such that each of the processes $\left\{M_{\min(t,\tau_n)}, t \geq 0\right\}$ is a martingale, then M is a local martingale.*

Note that a driftless process may not be a martingale, but if $E[X_\tau] = E[X_0]$ for any stopping time τ then X is a martingale. A familiar example of a martingale is the Wiener process. If $\{S_t, 0 \leq t \leq T\}$ is a martingale then for any simple predictable process[1] ϕ, the stochastic integral $\int_0^t \phi_s dS_s$ is also a martingale.

Suppose $\mathcal{I}$ is some index set and $X = \{X_i, i \in \mathcal{I}\}$ is a family of random variables. X is said to be *uniformly integrable* if

$$\lim_{n\to\infty} \sup_{i\in\mathcal{I}} E\left[|X_i| \, 1_{\{|X_i|>n\}}\right] = 0.$$

A process X is said to be in the *Dirichlet class* or *class D* if $\{X_\tau, \tau \in \mathcal{T}\}$ is uniformly integrable, where $\mathcal{T}$ is the family of all finite stopping times on our filtered probability space. A process X is said to be *integrable* if $E(|X_t|) < \infty$ for all $t > 0$. A process X said to be *predictable* if the mapping $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ given by $X(t, \omega) = X_t(\omega)$ is measurable with respect to the smallest σ-algebra generated by all adapted left-continuous mappings from $\mathbb{R}^+ \times \Omega \to \mathbb{R}$.

Definition 1.1.2 (Random measure) *Let $(S, \mathcal{A})$ be a measurable space and $(\Omega, \mathcal{F}, P)$ be a probability space. A collection of random variables $\{M(B), B \in \mathcal{A}\}$ is said to be a*

[1] A stochastic process $(\phi_t)_{t\in[0,T]}$ is called a simple predictable process if it can be represented as $\phi_t = \phi_0 1_{t=0} + \sum_{i=0}^n \phi_i 1_{]T_i,T_{i+1}]}(t)$, where $T_0 = 0 < T_1 < T_2 < \cdots < T_n < T_{n+1} = T$ are nonanticpating random times and each ϕ_i is bounded random variable whose value is revealed at T_i (it is $\mathcal{F}_{T_i}$-measurable).

random measure, denoted by M, on $(S, \mathcal{A})$ if the following are satisfied:
(1) $M(\emptyset) = 0$.
(2) For any sequence $\{A_n, n \in \mathbb{N}\}$ of mutually disjoint sets in $\mathcal{A}$,

$$M\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} M(A_n) \quad a.s.$$

(σ-additivity).
(3) Given any disjoint family $(B_1, ..., B_n)$ in $\mathcal{A}$, the random variables $M(B_1), ..., M(B_n)$ are independent.

Next we give the definition of variation of a mapping. Before doing so, we first recall the concept of a compact space. A topological space S is *compact* if, for every collection $\{U_i\}_{i\in I}$ of open sets in S whose union is S, there exists a finite subcollection $\{U_{i_j}\}_{j=1}^{n}$ whose union is also S. A compact subset of $\mathbb{R}^d$ is a bounded closed subset.

Definition 1.1.3 (Variation) *Suppose $\mathcal{P} = \{a = t_1 < t_2 < \cdots < t_n < t_{n+1} = b\}$ is a partition of the interval $[a, b]$ in $\mathbb{R}$ and let its mesh (the width of the largest sub-interval) be $\delta = \max_{1\le i\le n} |t_{i+1} - t_i|$. The variation $\mathrm{var}_{\mathcal{P}}(g)$ of a cádlág mapping $g_t : [a, b] \to \mathbb{R}^d$ over the partition $\mathcal{P}$ is given by*

$$\mathrm{var}_{\mathcal{P}}(g) = \sum_{i=1}^{n} |g(t_{i+1}) - g(t_i)|.$$

A cádlág mapping g is said to have finite or bounded variation on $[a, b]$ if $V(g) = \sup_{\mathcal{P}} \mathrm{var}_{\mathcal{P}}(g) < \infty$. g is said to have finite variation if it is defined on the whole of $\mathbb{R}$ (or $\mathbb{R}^+$) and has a finite or bounded variation on each compact interval. Every non-decreasing g is of finite variation. Conversely, g can always be written as the difference of two non-decreasing functions if it is of finite variation, since

$$g = \frac{V(g) + g}{2} - \frac{V(g) - g}{2}.$$

A stochastic process $\{X_t, t \ge 0\}$ is of finite variation if for almost all $\omega \in \Omega$, the paths $\{X_t(\omega), t \ge 0\}$ are of finite variation.

In the following, we give a brief introduction of Lévy processes, see Sato (1999) and Applebaum (2004) for a detailed discussion.

Definition 1.1.4 (Lévy process) *A Lévy process,* $\{X_t : t \geq 0\}$, *is a stochastic process on* $\mathbb{R}^d$ *satisfying:*
(1) The random variables $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, ..., X_{t_n} - X_{t_{n-1}}$ *are independent for any choice of* $n \geq 1$ *and* $0 < t_0 < t_1 < \cdots < t_n$.
(2) $X_0 = 0$ *a.s.*
(3) The distribution of $X_{s+t} - X_s$ *is independent of* s.
(4) The process is stochastically continuous.
(5) There exists $\Omega_0 \in \mathcal{F}$ *with* $P[\Omega_0] = 1$ *such that, for every* $\omega \in \Omega_0$, $X_t(\omega)$ *is right-continuous in* $t \geq 0$ *with left limits in* $t > 0$*, that is,* $\{X_t\}$ *is cádlág.*
If only (1)-(4) are satisfied, $\{X_t\}$ *is called a Lévy process in law. Note that every Lévy process in law has a cádlág modification that is a Lévy process.*

The characteristic function of a random variable uniquely determines its distribution. The famous Lévy-Khintchine formula, given in Theorem 2.1.1, gives the decomposition of the characteristic function of a Lévy process. The *characteristic function* of a probability measure μ on $\mathbb{R}^d$ is denoted by $\phi_\mu(z)$ and defined by

$$\phi_\mu(z) = \int_{\mathbb{R}^d} \exp(\mathrm{i}\langle z, x\rangle)\, \mu(\mathrm{d}x), \quad z \in \mathbb{R}^d,$$

where $\langle z, x\rangle = \sum_{j=1}^d z_j x_j$. The characteristic function $\phi_X(z)$ of the distribution P_X of a random variable X on $\mathbb{R}^d$ is given by

$$\phi_X(z) = \int_{\mathbb{R}^d} \exp(\mathrm{i}\langle z, x\rangle)\, P_X(\mathrm{d}x) = E[\exp(\mathrm{i}\langle z, x\rangle)]. \tag{1.1}$$

Let μ be a probability measure on $\mathbb{R}^d$. Let $\mathcal{M}_1(\mathbb{R}^d)$ denote the set of all Borel probability measures on $\mathbb{R}^d$. We define the *convolution* of two probability measures as follows:

$$(\mu_1 * \mu_2)(A) = \int_{\mathbb{R}^d} \mu_1(A - x)\, \mu_2(\mathrm{d}x)$$

for each $\mu_i \in \mathcal{M}_1(\mathbb{R}^d)$, $i = 1, 2$, and each $A \in \mathcal{B}(\mathbb{R}^d)$, where we note that $A - x = \{y - x, y \in A\}$. We define $\mu^n = \mu * \cdots * \mu$ (n times) and say that μ has a *convolution* nth *root*, if there exists a measure $\mu^{1/n} \in \mathcal{M}_1(\mathbb{R}^d)$ for which $(\mu^{1/n})^n = \mu$. If, for any positive integer n, there is a probability measure μ_n on $\mathbb{R}^d$ such that $\mu = (\mu_n)^n$, then μ is said to be *infinitely divisible*. The next theorem shows that infinitely divisibility is closely related to Lévy processes.

Theorem 1.1.5 *Suppose* $\{X_t : t \geq 0\}$ *is a Lévy process in law on* $\mathbb{R}^d$. *Then, for any* $t \geq 0$, P_{X_t} *is infinitely divisible and if* $P_{X_1} = \mu$, *we have* $P_{X_t} = \mu^t$. *Conversely, let*

μ be an infinitely divisible distribution on $\mathbb{R}^d$. Then there exists a Lévy process in law $\{X_t : t \geq 0\}$ such that $P_{X_1} = \mu$.

Note that a measure μ on $\mathbb{R}^d$ is infinitely divisible if and only if for each $n \in \mathbb{N}$, there exists $\mu^{1/n}$ such that $\phi_u(z) = \left[\phi_{\mu^{1/n}}(z)\right]^n$ for each $z \in \mathbb{R}^d$. Some famous examples of Lévy process include Brownian motions (see Definition 1.3.1), Poisson processes, compound Poisson processes, Gamma processes, Inverse Gaussian processes, Generalized Inverse Gaussian processes and Variance Gamma processes (see Section 6.4.1).

In the following, we give the definition of a Poisson random measure of a Lévy process.

Definition 1.1.6 (Poisson Random Measure) *Let $X = \{X_t, 0 \leq t \leq T\}$ be a Lévy process. Its Poisson random measure counts the jumps up to time t that are in a given Borel set A :*

$$N(t, A) = \#\{0 \leq s \leq t; \Delta X_s \in A\}.$$

Note that N is a function of three variables: time t, the Borel set A and the sample point ω. Fixing A, $N(A)$ is a Poisson random variable with intensity $\nu(A)$, where ν is the Lévy measure of X. Therefore, $E[N(A)] = \nu(A)$.

In the following, we give the definition of a semimartingale, which is an important generalisation of Lévy process. In stochastic calculus (see Section 2.2), semimartingales are important in that they are stable under stochastic integration while Lévy processes are not. In other words, a stochastic integral with respect to a semimartingale is also a semimartingale while a stochastic integral with respect to a Lévy process may not be a Lévy process anymore, but will be a semimartingale. A semimartingale is also stable under other operations such as change of measure, change of filtration and 'time change'.

Definition 1.1.7 (Semimartingale) *A process $X = \{X_t, t \geq 0\}$ is called a semimartingale if it is an adapted process such that, for each $t \geq 0$,*

$$X_t = X_0 + M_t + C_t.$$

where $M = \{M_t, t \geq 0\}$ is a local martingale and $C = \{C_t, t \geq 0\}$ is an adapted process of finite variation.

Note that every finite variation process, for example a Poisson process, is a semimartingale. Moreover, every square integrable martingale, for example a Wiener process, is a semimartingale. Any linear combination of a finite number of semimartingales is a semimartingale, for example, all Lévy processes are semimartingales because a Lévy process can be split into a sum of a square integrable martingale and a finite variation process using the Lévy-Itô decomposition, given in Theorem 2.1.4. Every (local) martingale is a semimartingale. A deterministic process is a semimartingale if and only if it is of finite variation so all infinite variation deterministic processes are examples of processes that are not semimartingales.

1.2 Important concepts in mathematical finance

In this section, we recall some of the important concepts in mathematical finance. We recall the definitions of a self-financing portfolio, market completeness, equivalent martingale measure and the fundamental theorems of asset pricing.

We start with the definition of a self-financing portfolio. Suppose there are K assets in the market, $A^{(1)}, A^{(2)}, ..., A^{(K)}$. Let $S_t^A(\omega)$ be the price of asset A at time t under market scenario ω. Assume we hold a portfolio consisting of shares (possibly held short) of each of the traded assets $A^{(j)}$ and we may adjust our portfolio as time progresses. Let $\theta_t^{A^{(j)}}(\omega)$ be the amount of asset $A^{(j)}$ held in a dynamically rebalanced portfolio during the t-th trading period (that is, during the period following completion of trading at time t until the beginning of trading at time $t+1$) under scenario ω, then the sequence $\left\{\theta_t^{A^{(j)}}, 0 \le t \le T\right\}$ must be adapted to the natural filtration. Denote the total value of the portfolio θ after rebalancing at time t in scenario ω by $V_t^\theta(\omega)$ and we have

$$V_t^\theta(\omega) = \sum_{i=1}^{K} \theta_t^{A^{(j)}}(\omega)\, S_t^{A^{(i)}}(\omega).$$

Note that V_t^θ may not equal V_{t+1}^θ, as the share prices of the underlying assets $A^{(j)}$ will generally change between times t and $t+1$. Assuming there is no transaction cost, if we do not invest (or withdraw) additional resources in our portfolio at time $t+1$, the total value of the portfolio just before rebalancing at time $t+1$ must be the same as its value just after, that is,

$$\sum_A \theta_t^A(\omega)\, S_{t+1}^A(\omega) = \sum_A \theta_{t+1}^A(\omega)\, S_{t+1}^A(\omega),$$

which is equal to

$$V_{t+1}^\theta(\omega) - V_t^\theta(\omega) = \sum_A \theta_t^A \left(S_{t+1}^A(\omega) - S_t^A(\omega)\right). \tag{1.2}$$

A dynamically rebalanced portfolio satisfying (1.2) is called *self-financing* since it requires no investments or withdrawals except at the initial time t.

In the introduction, we mentioned that the market under the Black-Scholes model is complete while the Lévy market is incomplete. Here we give a proper definition of market completeness. Define a *contingent claim*, with maturity date T, to be a non-negative $\mathcal{F}_T$-measurable random variable.

Definition 1.2.1 (Market Completeness) *A market model is said to be complete if every contingent claim can be replicated by a dynamic trading strategy: For any contingent claim H, adapted to the natural filtration generated by the price of the underlying process $\{S_t, t \in [0,T]\}$, there exists a self-financing strategy $\left(\phi_t^0, \phi_t\right)$ such that*

$$H = V_0 + \int_0^T \phi_t \mathrm{d}S_t + \int_0^T \phi_t^0 \mathrm{d}B_t, \quad P\text{-a.s.} \tag{1.3}$$

where V_0 is the initial investment and $\{B_t, t \in [0,T]\}$ is a risk-free bank account.

In the market driven by Lévy processes, contingent claims do not in general possess the representation in (1.3) and hence the market is incomplete. Later we will give the relationship between market completeness and the uniqueness of equivalent martingale measure. We first give the definition of an equivalent martingale measure. Suppose P represents the probability of occurrence of scenarios in the market and let r be the continuously compounded risk-free interest rate.

Definition 1.2.2 (Equivalent Martingale Measure) *Let P, Q be two probability measures defined on $(\Omega, \mathcal{F}_T)$. Q is called an equivalent martingale measure of P, denoted $Q \sim P$, if*
(1) Q is equivalent to P, that is, they have the same null sets (events which are impossible under P are also impossible under Q and vice versa).
(2) the discounted stock price process $\tilde{S} = \left\{\tilde{S}_t = \exp\left(-rt\right) S_t, t \geq 0\right\}$ is a martingale under Q.

We then introduce the risk-neutral pricing formula and risk-neutral measure. Suppose Q is an equivalent martingale measure to P and $\Pi_t\left(H\right)$ be the value of a contingent claim with payoff H and maturity T at time t. The *risk-neutral pricing formula* is given by

$$\Pi_t\left(H\right) = \mathrm{e}^{-r(T-t)} E^Q\left[H|\mathcal{F}_t\right], \tag{1.4}$$

that is, the value of a random payoff is given by its discounted expectation under Q, which is called a *risk-neutral* measure.

Lemma 1.2.3 *Let Q be a risk-neutral measure and let $\{X_t\}$ be the value of a portfolio. Under Q, the discounted portfolio value $\mathrm{e}^{-rt}X_t$ is a martingale.*

Next we give the definition of an arbitrage.

Definition 1.2.4 (Arbitrage) *An arbitrage is a portfolio value process* $\{X_t\}$ *satisfying* $X_0 = 0$ *and also satisfying for some time* $T > 0$,

$$P\{X_T \geq 0\} = 1, \quad P\{X_T > 0\} > 0. \tag{1.5}$$

An arbitrage is a way of trading such that one starts with zero capital and at some time later T sure not to lose any money and also has a positive probability of making money. We can summarize in the following theorem in the discrete case:

Theorem 1.2.5 (Fundamental Theorem of Asset Pricing in discrete time) *The market model defined by* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ *and asset prices* $\{S_t, t \in [0,T]\}$ *is arbitrage-free if and only if there exists a probability measure* $Q \sim P$ *such that the discounted assets* $\left\{\mathrm{e}^{-rt} S_t, t \in [0,T]\right\}$ *are martingales with respect to* Q.

The next theorem gives the relationship between market completeness and the uniqueness of equivalent martingale measure in the discrete case. The next theorem gives the relationship between market completeness and the uniqueness of equivalent martingale measure.

Theorem 1.2.6 (Second Fundamental Theorem of Asset Pricing) *A market defined by the assets* $\left(B_t, S_t^1, ..., S_t^d\right)_{t \in [0,T]}$, *described as stochastic processes on* $(\Omega, \mathcal{F}, P)$, *is complete if and only if there is a unique martingale measure* Q *equivalent to* P.

In continuous time the situation is far more complicated and this has been the focus of considerable study, see Bingham & Kiesel (2001). We need the following definitions. Let $S(t) = (S_0(t), S_1(t), ..., S_d(t))$ for a vector of prices of $d+1$ assets at time t. Let φ be a trading strategy, which is a $\mathbb{R}^{d+1}$ vector stochastic process $\varphi = (\varphi(t))_{t=1}^T = (\varphi_0(t,\omega), \varphi_1(t,\omega), ..., \varphi_d(t,\omega))_{t=1}^T$ which is predictable, that is, each $\varphi_i(t)$ is $\mathcal{F}_{t-1}$-measurable.

Definition 1.2.7 *The wealth process of the trading strategy* φ *is defined to be the scalar product*

$$V_\varphi(t) = \varphi(t) \cdot S(t) = \sum_{i=0}^{d} \varphi_i(t) S_i(t) \qquad \text{for } t > 0 \text{ and } V_\varphi(0) = \varphi(1) \cdot S(0).$$

Definition 1.2.8 *A simple predictable trading strategy is* δ*-admissible if the relative wealth process* $V_\varphi(t) \geq -\delta$ *for every* $t \in [0,T]$.

Definition 1.2.9 *A price process S satisfies NFLVR (no free lunch with vanishing risk) if for any sequence (φ_n) of simple trading strategies such that φ_n is δ_n-admissible and the sequence δ_n tends to zero, we have*

$$V_{\varphi_n}(T) \to 0 \qquad \text{in probability as} \qquad n \to \infty.$$

In continuous time, the fundamental theorem of asset pricing is stated as follow.

Theorem 1.2.10 (Fundamtal Theorem of Asset Pricing in continuous time) *In a financial market model with bounded prices, there exists an equivalent martingale measure if and only if the condition NFLVR holds.*

The theorem is proved in Delbaen & Schachermayer (1994). We refer the reader to the academic literature on this topic, see Delbaen & Schachermayer (1998), Cherny & Shiryaev (2002), Harrison & Pliska (1981), Harrison & Kreps (1979) and Schachermayer (2002). A market driven by Lévy processes is incomplete and there are infinitely many equivalent martingale measures.

1.3 Brownian motion and the Black-Scholes model

In this section, we recall the definition of a Brownian motion and the famous Black-Scholes model in finance. Please refer to Schoutens (2003) and Cont & Tankov (2003) for a detailed discussion.

Definition 1.3.1 (Brownian motion) *A standard Brownian motion $X = \{X_t, t \geq 0\}$ is a stochastic process on some probability space $(\Omega, \mathcal{F}, P)$ such that*
(1) $X_0 = 0$ a.s.,
(2) X has independent increments,
(3) X has stationary increments,
(4) $X_{t+s} - X_t$ is normally distributed with mean 0 and variance $s > 0$: $X_{t+s} - X_t \sim N(0, s)$.

We denote a standard Brownian motion by $W = \{W_t, t \geq 0\}$. Brownian motion is also an example of martingale, defined in Definition 1.1.1.

Proposition 1.3.2 (Martingale property) *Let $\mathcal{F} = \mathcal{F}^W = \{\mathcal{F}_t, 0 \leq t \leq T\}$ be the natural filtration of W. For all $0 \leq s \leq t$,*

$$E\left[W_t|\mathcal{F}_s\right] = E\left[W_t|W_s\right] = W_s.$$

Note that from this property, we have $E\left[W_t W_s\right] = \min\{s, t\}$.

The proof is given in Klebaner (2005, Theorem 3.7). The path of a Brownian motion is very special in that it is continuous but nowhere differentiable. The following proposition gives a proper description of its properties.

Proposition 1.3.3 (Path properties) *The paths of Brownian motion are continuous, which means that W_t is a continuous function of t, but has very erratic paths, which are nowhere differentiable and are of infinite variation (see Definition 1.1.3). The paths fluctuate between positive and negative values since we have*

$$P\left(\sup_{t\geq 0} W_t = +\infty \text{ and } \inf_{t\geq 0} W_t = -\infty\right) = 1.$$

Another nice property of a Brownian motion is that it has the scaling property. By multiplying a Brownian motion with a constant and change the time variable accordingly, we get another Brownian motion:

Proposition 1.3.4 (Scaling property) *For every $c \neq 0$, $\tilde{W} = \left\{\tilde{W}_t = cW_{t/c^2}, t \geq 0\right\}$ is also a standard Brownian motion.*

In the Black-Scholes model, the stock price $S = \{S_t, t \geq 0\}$ is modelled by the stochastic differential equation:

$$\mathrm{d}S_t = S_t\left(\mu \mathrm{d}t + \sigma \mathrm{d}W_t\right), \quad S_0 > 0,$$

where W_t is a standard Brownian motion, the parameters μ and $\sigma > 0$ represent the mean rate of return of the stock and the degree of fluctuation of the stock respectively. Applying the Itô formula (Theorem 2.2.2), we have

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right), \tag{1.6}$$

known as the *geometric Brownian motion.* The model assumes that investors can trade continuously up to some fixed finite planning horizon T and the uncertainly is modelled by a filtered probability space $(\Omega, \mathcal{F}, P)$. The market is assumed to be frictionless, that is, there are no transaction costs, no bid/ask spread, no taxes, no margin requirements, no restrictions on short sales, no transaction delays and the markets are perfectly liquid, market participants act as price takers and prefer more to less. There are two assets in the market: a risk-free bank account modelled by $B = \{B_t = \exp(rt), 0 \leq t \leq T\}$, where r is the continuously compounded risk-free interest rate, and a risky stock S which pays a continuous dividend yield $q \geq 0$ and modelled by the geometric Brownian motion (1.6). As noted in the introduction, the Black-Scholes model is complete (see Definition 1.2.1) because of the PRP of Brownian motion.

To derive the dynamic of S under the risk-neutral measure, we need the following results. Consider now a measurable space $(E, \mathcal{E})$ with measures μ_1 and μ_2 defined on it.

Definition 1.3.5 (Absolute continuity) *A measure μ_2 is said to be absolutely continuous with respect to μ_1 if for any measurable set A*

$$\mu_1(A) = 0 \Rightarrow \mu_2(A) = 0.$$

Theorem 1.3.6 (Radon-Nikodym theorem) *If μ_2 is absolutely continuous with respect to μ_1 then there exists a measurable function $Z : E \to [0, \infty[$ such that for any measurable set A*

$$\mu_2(A) = \int_A Z \mathrm{d}\mu_1 = \mu_1(Z 1_A).$$

The function Z is called the density or Radon-Nikodym derivative of μ_2 with respect to μ_1 and denoted as $\frac{\mathrm{d}\mu_2}{\mathrm{d}\mu_1}$. For any μ_2-integrable function f

$$\mu_2(f) = \int_E f \mathrm{d}\mu_2 = \mu_1(fZ) = \int_E \mathrm{d}\mu_1 Z f.$$

Theorem 1.3.7 (Cameron-Martin Theorem) *Let (X, P) and (X, Q) be two Brownian motions on $(\Omega, \mathcal{F}_T)$ with volatilities $\sigma^P > 0$ and $\sigma^Q > 0$ and drifts μ^P and μ^Q. P and Q are equivalent if $\sigma^P = \sigma^Q$ and singular otherwise. When they are equivalent the Radon-Nikodym derivative is*

$$\frac{\mathrm{d}P}{\mathrm{d}Q} = \exp\left\{\frac{\mu^Q - \mu^P}{\sigma^2} X_T - \frac{1}{2}\frac{\left(\mu^Q - \mu^P\right)^2}{\sigma^2} T\right\}.$$

A more general version of this result, valid for diffusion processes with random drift and volatility is known as the Girsanov Theorem, see Jacod & Shiryaev (2002) and Revuz & Yor (1999). Hence, using the Girsanov Theorem, we can transform W in (1.6) to get a new Brownian motion $\tilde{W}$, say. Then the discounted stock price $\tilde{S}_t = e^{-rt} S_t$ is driven by

$$\mathrm{d}\tilde{S}_t = \sigma \tilde{S}_t \mathrm{d}\tilde{W}_t$$

and the martingale property is explicit.

Let K and T be the strike and maturity of a contingent claim respectively. Let V_t be the price of a contingent claim at time t with payoff function $G(S_T)$, depending solely on the value of the stock at maturity. If $G(S_T)$ is a sufficiently integrable function, the price of the contingent claim is given by $V_t = F(t, S_t)$, which solves the following *Black-Scholes partial differential equation*,

$$\begin{aligned} \frac{\partial}{\partial t} F(t,s) + (r-q) s \frac{\partial}{\partial s} F(t,s) + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2}{\partial s^2} F(t,s) - rF(t,s) &= 0, \\ F(T,s) &= G(s). \end{aligned}$$

The explicit formulae for European call and put options are given by

$$\begin{aligned} C(K,T) &= \exp(-qt) S_0 N(d_1) - K \exp(-rT) N(d_2), \\ P(K,T) &= -\exp(-qt) S_0 N(-d_1) + K \exp(-rT) N(-d_2), \end{aligned}$$

where

$$\begin{aligned} d_1 &= \frac{\log(S_0/K) + \left(r - q + \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}, \\ d_2 &= d_1 - \sigma\sqrt{T} \end{aligned}$$

and $N(\cdot)$ is the cumulative probability distribution function for a Normally distributed random variable.

As mentioned in the introduction, Black-Scholes model has been proved to be insufficient in describing the behaviour of the price processes in financial markets. Barndorff-Nielsen & Shephard (2006) performed hypothesis tests on exchange data under the null of no jumps, which were found to be rejected frequently. In fact, at intraday scales, prices move essentially by jumps and even at the scale of months, the discontinuous behaviour cannot be ignored in general. Only after coarse-graining their behaviour over longer time scales do we obtain something similar to Brownian motion. Another problem is that the log return data series have heavy tails and are negatively skewed. Although an appropriate choice of a nonlinear diffusion coefficient in the Black-Scholes model can generate processes with arbitrary heavy tails, we often end up choosing extreme value for the parameters, see Cont & Tankov (2003, Chapter 1). Even so, the diffusion processes are still continuous, that is, no jumps can be truly created by diffusion models. Heavy left tails of the distributions of asset price processes corresponds to large sudden jumps in the price processes. Without the ability to create jumps, diffusion models underestimate the risks incurred from jumps in the market. To remedy these problems, market models driven by Lévy processes (see Definition 1.1.4) were introduced to replace Black-Scholes model in describing the dynamics of asset price processes. A Lévy process has independent and stationary increments generated by a so-called infinitely divisible distribution, which has a one-to-one relationship with the Lévy process, see Theorem 1.1.5. General Lévy processes allow jumps and provide more flexibility in describing log asset price processes since heavy tails and asymmetry can be handled by extra parameters of the infinitely divisible distributions. Since large sudden moves are generic properties of models with jumps, fine-tuning of parameters to extreme values is not required as in diffusion models. Models with jumps capture the unexpected, sudden price movement, which is perceived as risk in the market. As Cont & Tankov (2003) pointed out, 'the question of using continuous or discontinuous models has important consequences for the representation of risk and is not a purely statistical issue.' Apart from the inability to replicate price movements, the Black-Scholes model also fails to reproduce the main features of option prices in the market. The well-known volatility surface is obtained by plotting the implied volatilities of the Black-Scholes model across maturities and across strikes. If the option pricing model is describing the market perfectly, the value of the implied volatilities should be constant throughout. However, this is not the case in practice. In fact, the main driving force behind the generalisation of the Black-Scholes model is to improve the calibration of option prices in the market.

1.4 Orthogonalised processes

In this section, we introduce the orthogonalised compensated power jump processes introduced by Nualart & Schoutens (2000) and give the alternative notation used by Jamshidian (2005). We derive the explicit formula for the CRP in terms of orthogonalised compensated power jump processes in Part I following Nualart and Schoutens notation but since Jamshidian derived an explicit formula for the CRP in terms of non-compensated power jump processes, we include Jamshidian's notation for comparison. Let $X = \{X_t, t \geq 0\}$ be a Lévy process (see Definition 1.1.4). In the rest of the thesis, we assume that all Lévy measures concerned satisfy, for some $\varepsilon > 0$ and $\lambda > 0$,

$$\int_{(-\varepsilon,\varepsilon)^c} \exp(\lambda |x|) \, \nu(\mathrm{d}x) < \infty. \tag{1.7}$$

This condition implies that for $i \geq 2$, $\int_{-\infty}^{+\infty} |x|^i \nu(\mathrm{d}x) < \infty$, and that the characteristic function $E[\exp(\mathrm{i}uX_t)]$ is analytic in a neighborhood of 0.

1.4.1 Nualart and Schoutens notation

Denote the i-th *power jump process* by $X_t^{(i)} = \sum_{0<s\leq t}(\Delta X_s)^i$, $i \geq 2$, and for completeness let $X_t^{(1)} = X_t$. In general, it is not true that $X_t = \sum_{0<s\leq t} \Delta X_s$; this holds only in the bounded variation case (see Definition 1.1.3), with $\sigma^2 = 0$. By definition, the quadratic variation of X_t, $[X,X]_t = \sum_{0<s\leq t}(\Delta X_s)^2 = X_t^{(2)}$ when $\sigma^2 = 0$. These power jump processes are also Lévy processes and jump at the same time as X_t, but with jump sizes equal to the i-th powers of those of X_t, see Nualart & Schoutens (2000).

Clearly $E[X_t] = E[X_t^{(1)}] = m_1 t$, where $m_1 < \infty$ is a constant and by Protter (2004, p.32), we have

$$E[X_t^{(i)}] = E[\sum_{0<s\leq t}(\Delta X_s)^i] = t\int_{-\infty}^{\infty} x^i \nu(\mathrm{d}x) = m_i t < \infty, \quad \text{for } i \geq 2, \tag{1.8}$$

thus defining m_i. Nualart & Schoutens (2000) introduced the *compensated power jump process* (or *Teugels martingale*) of order i, $\left\{Y_t^{(i)}\right\}$, defined by

$$Y_t^{(i)} = X_t^{(i)} - E[X_t^{(i)}] = X_t^{(i)} - m_i t \quad \text{for } i = 1,2,3,.... \tag{1.9}$$

$Y_t^{(i)}$ is constructed to have a zero mean. It was shown by Nualart & Schoutens (2000, Section 2) that there exist constants $a_{i,1}, a_{i,2}, ..., a_{i,i-1}$ such that the processes defined by

$$H_t^{(i)} = Y_t^{(i)} + a_{i,i-1}Y_t^{(i-1)} + \cdots + a_{i,1}Y_t^{(1)}, \tag{1.10}$$

for $i \geq 1$ are a set of pairwise strongly orthogonal martingales, and this implies that for $i \neq j$, the process $H_t^{(i)}H_t^{(j)}$ is a martingale, see Léon *et al.* (2002). For convenience, we define $a_{i,i} = 1$. Nualart & Schoutens (2000) proved that this strong orthogonality is equivalent to the existence of an orthogonal family of polynomials with respect to the measure

$$\mathrm{d}\eta(x) = \sigma^2 \mathrm{d}\delta_0(x) + x^2\nu(\mathrm{d}x),$$

where $\delta_0(x) = 1$ when $x = 0$ and zero otherwise, that is, the polynomials p_n defined by

$$p_n(x) = \sum_{j=1}^{n} a_{n,j}x^{j-1}$$

are orthogonal with respect to the measure η:

$$\int_{\mathbb{R}} p_n(x)\, p_m(x)\, \mathrm{d}\eta(x) = 0, \quad n \neq m.$$

1.4.2 Jamshidian's notation

In Jamshidian (2005), which extends the CRP to semimartingales, the power jump processes and compensators were denoted and defined differently from Nualart & Schoutens (2000). The power jump processes were defined in Jamshidian (2005) by

$$[X]_t^{(2)} = [X^c]_t + \sum_{s\leq t} (\Delta X_s)^2 \text{ and } [X]_t^{(n)} = \sum_{s\leq t} (\Delta X_s)^n \text{ for } n = 3, 4, 5, ..., \tag{1.11}$$

where $[X^c]_t = [X]_t^c$ is the continuous finite-variation (not martingale) part of $[X]_t^{(2)}$. Note that Jamshidian suppressed the time index t, but we add it here for clarification. The *compensator*, $\langle X\rangle_t^{(n)}$, is the predictable right-continuous finite variation process such that $[X]_t^{(n)} - \langle X\rangle_t^{(n)}$ is a uniformly integrable martingale. The compensated power jump process, denoted by $X_t^{(n)}$, is thus defined by

$$X_t^{(n)} = [X]_t^{(n)} - \langle X\rangle_t^{(n)} \text{ for } n = 2, 3, 4, ... \tag{1.12}$$

Part I

Martingale Representations for Lévy processes

An explicit formula for the chaotic representation of the powers of increments of a Lévy process, $(X_{t+t_0} - X_{t_0})^n$, is presented. There are two different chaos expansions of a square integrable functional of a Lévy process: one with respect to the compensated Poisson random measure and the other with respect to the orthogonal compensated powers of the jumps of the Lévy process. Computationally explicit formulae for both of these chaos expansions of $(X_{t+t_0} - X_{t_0})^n$ are given in this part. Simulation results verify that the representation is satisfactory. The CRP of a number of financial derivatives can be found by expressing them in terms of $(X_{t+t_0} - X_{t_0})^n$ using Taylor expansion.

This part is arranged as follow: Chapter 2 gives a quick review of martingale representations in the literature. We give the explicit formulae for the CRP for the powers of increments of a Lévy process X in terms of power jump processes in Chapter 3 and in terms of Poisson random measure in Chapter 4. We show that our formula is an non-trivial extension of Jamshidian's formula in the Lévy case, which is an important subclass of semimartingales. Chapter 5 gives discussion and further applications of the topic. Section 5.1 gives the representation of a common kind of Lévy functionals with the use of Taylor's theorem. Simulation results for the explicit formulae are given in Section 5.2. Section 5.3 discusses the explicit formula derived by Løkka (2004) and Section 5.4 gives the Lévy measures of the orthogonalised compensated power jump processes, $H_t^{(i)}$. Some concluding remarks are provided at the end of this part. Proofs and plots are included in Appendix A.

Chapter 2

Martingale representations in the literature

2.1 Lévy representations

The following theorem, called the Lévy-Khintchine formula, is fundamental to Lévy models and representations, see Sato (1999). Let X be a random variable and let ϕ_X be its characteristic function as defined in (1.1).

Theorem 2.1.1 (The Lévy-Khintchine formula) *For every infinitely divisible random variable* $X \in \mathbb{R}^d$

$$\phi_X(z) = \exp\left[-\frac{1}{2}\langle z, Az\rangle + \mathrm{i}\langle \gamma, z\rangle + \int_{\mathbb{R}^d}\big(\exp(\mathrm{i}\langle z, x\rangle) - 1 - \mathrm{i}\langle z, x\rangle 1_{\{|x|<1\}}(x)\big)\,\nu(\mathrm{d}x)\right], \quad z \in \mathbb{R}^d, \quad (2.1)$$

where A *is a symmetric nonnegative-definite* $d \times d$ *matrix,* ν *is a measure on* $\mathbb{R}^d$ *satisfying*

$$\nu(\{0\}) = 0 \quad \textit{and} \quad \int_{\mathbb{R}^d} \min\left(|x|^2, 1\right)\nu(\mathrm{d}x) < \infty, \quad (2.2)$$

and $\gamma \in \mathbb{R}^d$. *This representation by* A, ν *and* γ *is unique. Conversely, let* A *be a symmetric nonnegative-definite* $d \times d$ *matrix,* ν *be a measure satisfying (2.2), and* $\gamma \in \mathbb{R}^d$. *Then there exists an infinitely divisible distribution* μ *whose characteristic function is given by (2.1).* (γ, A, ν) *is known as the generating triplet of* μ *and* ν *is the Lévy measure of* X.

Proof. See Cont & Tankov (2003, Section 3.4) for an outline of the proof. □

If $d = 1$, we can write

$$\phi_X(z) = \exp\left[\mathrm{i}\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{-\infty}^{+\infty} \left(\exp(\mathrm{i}ux) - 1 - \mathrm{i}ux1_{\{|x|<1\}}\right)\nu(\mathrm{d}x)\right],$$

where $\gamma \in \mathbb{R}$ and ν satisfies (2.2). In this case, the Lévy triplet is given by $\left(\gamma, \sigma^2, \nu\right)$.

Next we give another important decomposition formula for Lévy processes. Before doing so, we have to give the definition of compensated Poisson random measure, following the notation of Cont & Tankov (2003). Note that the definition of Poisson random measure is given in Definition 1.1.6.

Definition 2.1.2 (Compensated Poisson random measure) *Suppose N is a Poisson random measure on $[0,T] \times \mathbb{R}^d$ with intensity $\mu(\mathrm{d}t, \mathrm{d}x)$. The compensated Poisson random measure is defined by $\tilde{N}(A) = N(A) - \mu(A) = N(A) - E[N(A)]$.*

Theorem 2.1.3 *For every measurable set $A \subset \mathbb{R}^d$ with $\mu([0,T] \times A) < \infty$, $N_t(A) = ([0,t] \times A)$ defines a counting process, $\tilde{N}_t(A) = N([0,t] \times A) - \mu([0,t] \times A)$ is a martingale and if $A \cap B = \emptyset$ then $N_t(A)$ and $N_t(B)$ are independent.*

The famous Lévy-Itô decomposition states that a Lévy process can be decomposed into a sum of deterministic component, Brownian motion and integrals with respect to non-compensated and compensated Poisson random measures:

Theorem 2.1.4 (The Lévy-Itô decomposition) *For any d-dimensional Lévy process X, there exists $b \in \mathbb{R}^d$, a Brownian motion $W_t^{(A)}$ with covariance matrix*

$$A = cov\left\{W_t^{(A)}\left(W_t^{(A)}\right)^T\right\}$$

and an independent Poisson random measure N on $\mathbb{R}^+ \times \left(\mathbb{R}^d - \{0\}\right)$ with the corresponding compensated Poisson random measure $\tilde{N}$, such that, for each $t \geq 0$,

$$X_t = bt + W_t^{(A)} + \int_{|x|<1} x\tilde{N}(t, \mathrm{d}x) + \int_{|x|\geq 1} xN(t, \mathrm{d}x).$$

Hence, for any stochastic process built from a Lévy process, the positions and the amplitudes of its jumps are described by a Poisson random measure and various quantities involving the jump times and jump sizes can be expressed as integrals with respect to this measure.

2.2 Stochastic calculus

The martingale representations we study consist of infinite sums of stochastic integrals. This section gives an introduction to stochastic calculus and related concepts.

A stochastic process $\phi = \{\phi_t, 0 \leq t \leq T\}$ is called a *simple predictable process* if it can be represented as

$$\phi_t = \phi_0 1_{\{t=0\}} + \sum_{i=0}^{n} \phi_i 1_{]T_i, T_{i+1}]}(t),$$

where $T_0 = 0 < T_1 < T_2 < \cdots < T_n < T_{n+1} = T$ are nonanticipating random times and each ϕ_i is a bounded $\mathcal{F}_{T_i}$-measurable random variable, that is, its value is revealed at T_i. The *stochastic integral* of the predictable process ϕ with respect to a process $S = \{S_t, 0 \leq t \leq T\}$ is defined by

$$\int_0^t \phi_u \mathrm{d}S_u = \phi_0 S_0 + \sum_{i=0}^{n} \phi_i \left(S_{T_{i+1}\wedge t} - S_{T_i \wedge t}\right).$$

In financial applications, if S represents the price process of a financial asset, then ϕ represents the trading strategy of a dynamic portfolio. The stochastic integral $\int_0^t \phi_u \mathrm{d}S_u$ represents the capital accumulated between 0 and t by the strategy ϕ. The value of the portfolio at time t is given by $V_t(\phi) = \phi_t S_t$. The *cost process* associated to the strategy ϕ is given by

$$C_t(\phi) = V_t(\phi) - \int_0^t \phi_u \mathrm{d}S_u = \phi_t S_t - \int_0^t \phi_u \mathrm{d}S_u.$$

Recall Section 1.2, if the cost is (almost surely) equal to zero, the strategy ϕ is said to be self-financing. In this case, we have

$$V_t(\phi) = \int_0^t \phi_u \mathrm{d}S_u = \phi_0 S_0 + \int_{0+}^t \phi_u \mathrm{d}S_u.$$

Stochastic integrals have the martingale-preserving property. If $S = \{S_t, t \in [0,T]\}$ is a martingale, then for any simple predictable process ϕ, the stochastic integral $\int_0^t \phi_u \mathrm{d}S_u$ is also a martingale. Moreover, if $X = \{X_t, t \in [0,T]\}$ is a real-valued nonanticipating cádlág process, $\sigma = \{\sigma_t, t \geq 0\}$ and $\phi = \{\phi_t, t \geq 0\}$ are real-valued simple predictable processes, then $S_t = \int_0^t \sigma_u \mathrm{d}X_u$ is a nonanticipating cádlág process and

$$\int_0^t \phi_u \mathrm{d}S_u = \int_0^t \phi_u \sigma_u \mathrm{d}X_u.$$

Definition 2.2.1 (Nonanticipating random time) *Given an information flow $\mathcal{F}_t$, a* *positive random variable $T \geq 0$ is a nonanticipating random time ($\mathcal{F}_t$-stopping time) if*

$$\forall t \geq 0, \quad \{T \leq t\} \in \mathcal{F}_t.$$

We now discuss stochastic integrals with respect to Poisson random measure, N, defined in Definition 1.1.6. A function $\phi : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is called a *simple predictable function* if

$$\phi(t,y) = \sum_{i=1}^{n} \sum_{j=1}^{m} \phi_{ij} 1_{]T_i, T_{i+1}]}(t) 1_{A_j}(y),$$

where $T_1 \leq T_2 \leq \cdots \leq T_n$ are nonanticipating random times, $\left\{\phi_{ij}, j = 1, ..., m\right\}$ are bounded $\mathcal{F}_{T_i}$-measurable random variables and $\{A_j, j = 1, ..., m\}$ are disjoint subsets of $\mathbb{R}^d$ with $\mu([0,T] \times A_j) < \infty$. The stochastic integral with respect to N is defined by

$$\int_0^t \int_{\mathbb{R}^d} \phi(s,y) N(\mathrm{d}s, \mathrm{d}y) = \sum_{i,j=1}^{n,m} \phi_{ij} \left[N_{T_{i+1}\wedge t}(A_j) - N_{T_i \wedge t}(A_j)\right]$$

and the stochastic integral with respect to the compensated Poisson random measure, $\tilde{N}$, defined in Definition 2.1.2, is given by

$$\int_0^t \int_{\mathbb{R}^d} \phi(s,y) \tilde{N}(\mathrm{d}s, \mathrm{d}y) = \sum_{i,j=1}^{n,m} \phi_{ij} \left[\tilde{N}_{T_{i+1}\wedge t}(A_j) - \tilde{N}_{T_i \wedge t}(A_j)\right].$$

Our explicit formula for the CRP is derived from the famous Itô formula, see Cont & Tankov (2003, Section 8.3). We firstly give the simplest Itô formula which is with respect to Brownian motion, which implies the market completeness of the Black-Scholes model, see Definition 1.2.1.

Theorem 2.2.2 (Itô formula for Brownian integrals) *If f is a differentiable function and $X_t = \int_0^t \sigma_s \mathrm{d}W_s$, then*

$$f(X_t) = f(0) + \int_0^t f'(X_s) \sigma_s \mathrm{d}W_s + \int_0^t \frac{1}{2}\sigma_s^2 f''(X_s) \mathrm{d}s.$$

Note that σ_s is the integrand and is not to be confounded with the σ of the Lévy triplet.

Recall that the Lévy process is a generalisation of Brownian motion with jumps. Therefore, the Itô formula for scalar Lévy process includes a term to deal with the discontinuity:

Theorem 2.2.3 (Itô formula for one dimensional Lévy process) *If $X = \{X_t, t \geq 0\}$ is a Lévy process with Lévy triplet $\left(\sigma^2, \nu, \gamma\right)$ and $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function,* *then*

$$\begin{aligned} f\left(X_t\right) &= f\left(0\right)+\int_0^t \frac{\sigma^2}{2}\frac{\partial f^2}{\partial^2 x}\left(s,X_{s-}\right)\mathrm{d}s+\int_0^t \frac{\partial f}{\partial x}\left(s,X_{s-}\right)\mathrm{d}X_s \\ &\quad +\sum_{\substack{0\leq s\leq t\\ \Delta X_s\neq 0}}\left[f\left(X_{s-}+\Delta X_s\right)-f\left(X_{s-}\right)-\Delta X_s\frac{\partial f}{\partial x}\left(s,X_{s-}\right)\right]. \end{aligned}$$

The Itô formula for multidimensional Lévy process extends directly from the last theorem and is given by

Theorem 2.2.4 (Itô formula for multidimensional Lévy process) *If the stochastic process* $X_t = \left(X_t^1,...,X_t^d\right)$ *is a multidimensional Lévy process with characteristic triplet* $\left(\Sigma,\nu,\gamma\right)$, *then for any function* $f:[0,T]\times\mathbb{R}^d\rightarrow\mathbb{R}$, *continuous in time and differentiable in* $\mathbb{R}^d$,

$$\begin{aligned} f\left(t,X_t\right)-f\left(0,0\right) &= \int_0^t\sum_{i=1}^d\frac{\partial f}{\partial x_i}\left(s,X_{s-}\right)\mathrm{d}X_s^i+\int_0^t\frac{\partial f}{\partial s}\left(s,X_s\right)\mathrm{d}s \\ &\quad +\frac{1}{2}\int_0^t\sum_{i,j=1}^d\Sigma_{ij}\frac{\partial^2 f}{\partial x_i\partial x_j}\left(s,X_s\right)\mathrm{d}s \\ &\quad +\sum_{0\leq s\leq t}^{\Delta X_s\neq 0}\left[f\left(s,X_{s-}+\Delta X_s\right)-f\left(s,X_{s-}\right)-\sum_{i=1}^d\Delta X_s^i\frac{\partial f}{\partial x_i}\left(s,X_{s-}\right)\right]. \end{aligned}$$

Recall that the Lévy process is a special case of semimartingale, defined in Definition 1.1.7. We therefore give the Itô formula for semimartingale as well:

Theorem 2.2.5 (Itô formula for semimartingale) *If* $X=\{X_t,t\geq 0\}$ *is a semimartingale, then for any function* $f:[0,T]\times\mathbb{R}\rightarrow\mathbb{R}$, *continuous in time and differentiable in* $\mathbb{R}$,

$$\begin{aligned} f\left(t,X_t\right)-f\left(0,X_0\right) &= \int_0^t\frac{\partial f}{\partial s}\left(s,X_s\right)\mathrm{d}s+\int_0^t\frac{\partial f}{\partial x}\left(s,X_{s-}\right)\mathrm{d}X_s \\ &\quad +\frac{1}{2}\int_0^t\frac{\partial^2 f}{\partial x^2}\left(s,X_{s-}\right)\mathrm{d}\left[X,X\right]_s^c \\ &\quad +\sum_{\substack{0\leq s\leq t\\ \Delta X_s\neq 0}}\left[f\left(s,X_s\right)-f\left(s,X_{s-}\right)-\Delta X_s\frac{\partial f}{\partial x}\left(s,X_{s-}\right)\right], \end{aligned}$$

where $\left[X,X\right]^c$ *denotes the continuous part of* $\left[X,X\right]$.

2.3 Chaotic Representation Property in the literature

Itô (1956) proved a CRP for any square integrable functional for a general Lévy process. Nonetheless, only in the Brownian and Poisson cases can the representation of the functional be expressed in terms of multiple integrals with respect to the Brownian motion and Poisson process respectively, see Itô (1951) and Nualart & Vives (1990). The representation of a process in terms of its mean plus a stochastic integral with respect to the underlying process is known as the PRP, which is an immediate result of the CRP. For general Lévy processes, it is necessary to introduce a two-parameter random measure to construct a PRP. The representation is then written using multiple integrals with respect to this two-parameter random measure. In other words, the PRP in terms of a stochastic integral with respect to the underlying process only is lost for the general Lévy case when using Itô's representation. This kind of PRP is important since it provides the market completeness of the Black-Scholes option pricing model. Recall that a market is said to be complete if every contingent claim can be replicated by investing in the underlying stock and a risk-free bond, see Definition 1.2.1. The predictable process gives the self-financing admissible strategy of replicating a contingent claim, see Section 1.2. To obtain a similar property in the general Lévy case, Nualart & Schoutens (2000) proved the existence of a new version of the CRP, which satisfies some exponential moment conditions. This new CRP states that every square integrable random variable adapted to the filtration generated by a Lévy process can be represented as its expectation plus an infinite sum of zero mean stochastic integrals with respect to the orthogonalised compensated power jump processes of the underlying Lévy process. Hence, the market can be completed even in the case of a general Lévy process if trades in these processes are allowed.

Trying to derive an explicit formula for the CRP has been the focus of considerable study. However, it is important to note that previous results for general Lévy functionals available in the literature, namely, the Clark-Ocone-Haussman formulae derived to obtain the integrands of the predictable, or chaotic, representation are **not** truly explicit. The explicit chaos expansion must be known, for these formulae to be applied, making the specification circular. We will discuss this in further detail later.

Nualart & Schoutens (2001) presented a version of the Clark-Ocone formula for functions of a Lévy process using the solution of a Partial Differential Integral Equation (PDIE). The Clark-Ocone formula gives the values of the predictable integrands of the CRP. This version of the formula works for processes derived from certain Backward Stochastic Differential Equations (BSDEs).

Léon *et al.* (2002) developed the basic theory for Malliavin calculus for Lévy processes and derived the Clark-Ocone formula, to give a predictable representation. Simple Lévy processes, which are a sum of a Brownian motion and a finite number of independent Poisson processes with different jump sizes, were studied in their paper. The stochastic integrals in the PRP studied in these cases were with respect to the Brownian motion and compensated Poisson processes rather than the orthogonalised compensated power jump processes, $H_t^{(i)}$'s, introduced originally by Nualart & Schoutens (2000) for the representation. Useful formulae presented in the paper for the calculation of the Clark-Ocone formula were derived in this case. The predictable representation derived using the Clark-Ocone formula is not truly explicit, as again the explicit chaos expansions must be given before the formula can be applied.

Løkka (2004) derived a Clark-Ocone-Haussman formula which provides a representation for Itô's expansion in the case of pure jump Lévy processes. This formula has a different form to the formula of Léon *et al.* (2002) since it is based on a different chaotic representation. Again the Clark-Ocone-Haussman formula derived is not truly explicit. The author derived an explicit formula for a common kind of functionals of Lévy processes in Proposition 8 of the paper, which is discussed in Section 5.3 in this thesis.

Benth *et al.* (2003) and Solé *et al.* (2006) derived the relationship between the chaos expansion in terms of iterated stochastic integrals with respect to power jump processes, and the expansion in terms of iterated integrals with respect to Poisson random measure. Note that Itô (1956) expressed the chaos expansions in terms of multiple integrals but one may convert it to iterated integrals as done by Løkka (2004). Solé *et al.* (2006) gave the relationship between the Nualart & Schoutens (2000) representation and the Itô (1956) representation but this is actually equivalent to the Benth *et al.* (2003) relationship. Thanks to these relations, our explicit formula can be applied to find the explicit representation for Itô's expansion. Benth *et al.* (2003) also gave the explicit representation of the minimal variance portfolio, in markets where the stock prices are modeled by Lévy martingales, using Malliavin calculus.

Eddahbi *et al.* (2005) derived a formula, denoted the Stroock formula, for the kernels of the chaotic decomposition of a smooth random variable as functionals of the underlying Lévy process using a Malliavin type derivative. The formula was used to obtain the chaos expansion of the price of an European call option and its underlying asset. Note that the formulae presented in Nualart & Schoutens (2001), Léon *et al.* (2002) and Løkka (2004) give forms for the integrands in the predictable representation while this Stroock formula gives forms for the integrands in the chaotic representation. As the terms of the chaotic expansion are orthogonal and uncorrelated, the chaotic approach enables the study of the asymptotic behaviour of the variance of the integrals, which is useful in deriving practical hedging strategies. As in Léon *et al.* (2002), the CRP was only applied to simple Lévy processes and the stochastic integrals in the chaos expansion were with respect to the Brownian motion and compensated Poisson processes, rather than the orthogonalised compensated power jump processes. The explicit chaos expansion has to be known before the Stroock formula can be applied.

All the explicit formulae for general Lévy functionals derived in these papers use the Malliavin type derivatives to derive explicit representations of stochastic processes for applications in finance. The derivative operator D is, in all of these cases, defined by its action on the chaos expansions themselves. In other words, the explicit chaos expansion must in fact be known before D can be applied to find the explicit form of the predictable or chaotic representation, thus yielding a circular specification. For example, Léon *et al.* (2002, Definition 1.7) defined the derivative of F in the l-direction by:

$$D_t^{(l)} F = \sum_{n=1}^{\infty} \sum_{i_1,\ldots,i_n} \sum_{k=1}^{n} 1_{\{i_k=l\}} J_{n-1}^{\left(i_1,\ldots,\widehat{i_k},\ldots,i_n\right)} \left(f_{i_1,\ldots,i_n} \left(\cdots, t, \cdots \right) 1_{\Sigma_n^{(k)}(t)} \left(\cdot \right) \right),$$

and Løkka (2004, Section 3) defined the derivative operator by:

$$D_{t,z} F = \sum_{n=1}^{\infty} n I_{n-1} \left(f_n \left(\cdot, t, z \right) \right),$$

where

$$I_n \left(f_n \right) = \int_{[0,T]^n \times \mathbb{R}_0^n} f_n \left(t_1, ..., t_n, z_1, ..., z_n \right) \mathrm{d} \left(\mu - \pi \right)^{\otimes n}.$$

Please refer to the corresponding papers for notation. Note that both of these definitions require the knowledge of the functions $\{f_{i_1,\ldots,i_n}\}$'s or $f_n \left(t_1, ..., t_n, z_1, ..., z_n \right)$'s, which are the integrands of the chaos expansion of F.

Jamshidian (2005) extended the CRP in Nualart & Schoutens (2000) to a large class of semimartingales and derived the explicit representation of the power of a Lévy process with respect to the corresponding non-compensated power jump processes, which is discussed further in Remark 3.2.1. Note that Lévy processes are included in the class of semimartingales, see Kannan & Lakshmikantham (2001, Corollary 2.3.21, p.92). Our formula for the CRP derived in Chapter 3 gives the explicit representation with respect to the orthogonalised compensated power jump processes as defined in Nualart & Schoutens (2000). Our result is therefore complementary to Jamshidian's formula.

Corcuera *et al.* (2005) suggested enlarging the market by a series of assets related to the power jump processes of the underlying Lévy processes. Using the martingale representation with respect to the compensated power jump processes, the market could be completed. Corcuera *et al.* (2006) used this completeness to solve the portfolio optimisation problem by the martingale method.

Having discussed existing representations of Lévy processes and relationships between the results, for a subset of functionals of Lévy processes, we simplify the CRP to an explicit form.

Chapter 3

The chaotic representation with respect to power jump processes

In this chapter we first derive the explicit formula for the CRP when the random variable, F, in (3.1) is the power of the increment of a pure jump Lévy process and extend it subsequently to a general Lévy process. In the following, we quote the chaotic representation property (CRP) in terms of orthogonalised compensated power jump processes derived by Nualart & Schoutens (2000). The CRP is important in that it implies the predictable representation property (PRP), which provides the hedging portfolio for a contingent claim. Based on the PRP of Lévy processes, Corcuera *et al.* (2005) completed the market by introducing power jump assets. In Part II of this thesis, we further investigate the perfect hedging strategies in a Lévy market. In the following, we firstly quote Proposition 2 in Nualart & Schoutens (2000), which explains the importance of our result for the powers of increments of a Lévy process.

Proposition 3.0.1 (Proposition 2 in Nualart & Schoutens (2000)) *Let*

$$\mathcal{P} = \{X_{t_1}^{k_1}(X_{t_2} - X_{t_1})^{k_2}...(X_{t_n} - X_{t_{n-1}})^{k_n} : n \geq 0,\ 0 \leq t_1 < t_2 < \cdots < t_n, k_1, ..., k_n \geq 1\}$$

be a family of stochastic processes. Then $\mathcal{P}$ is a total family in $L^2(\Omega, \mathcal{F}_T, P)$, that is, the linear subspace spanned by $\mathcal{P}$ is dense in $L^2(\Omega, \mathcal{F}_T, P)$, where $\mathcal{F}_T = \sigma\{X_t, 0 \leq t \leq T\}$ and we write $\mathcal{F} = \mathcal{F}_T$ for simplicity. This means that each element in $L^2(\Omega, \mathcal{F}, P)$ can be represented as a linear combination of elements in $\mathcal{P}$.

Although we only derive the explicit formula for the powers of increments of a Lévy process, this proposition shows that every random variable adapted to the filtration can be represented in terms of these powers of increments. We show in Section 5.1 that we use our explicit formula and Taylor's Theorem to obtain the chaos expansion for a general Lévy functional. The famous CRP by Nualart & Schoutens (2000) is in terms of an infinite sum of orthogonalised compensated power jump processes:

Theorem 3.0.2 (Chaotic Representation Property (CRP)) *Every random variable F in $L^2(\Omega, \mathcal{F}, P)$ has a representation of the form*

$$F = E(F) + \sum_{j=1}^{\infty} \sum_{i_1,...,i_j \geq 1} \int_0^{\infty} \int_0^{t_1-} \cdots \int_0^{t_{j-1}-} f_{(i_1,...,i_j)}(t_1, ..., t_j) \mathrm{d}H_{t_j}^{(i_j)} ... \mathrm{d}H_{t_2}^{(i_2)} \mathrm{d}H_{t_1}^{(i_1)}, \tag{3.1}$$

where the $f_{(i_1,...,i_j)}$'s are functions in $L^2(\mathbb{R}_+^j)$ and H's are defined in equation (1.10). This result means that every random variable in $L^2(\Omega, \mathcal{F}, P)$ can be expressed as its expectation plus an infinite sum of zero mean stochastic integrals with respect to the orthogonalised compensated power jump processes of the underlying Lévy process. Note that this representation does not explicitly allow for calculation of the integrands. The PRP is an immediate result of the CRP:

Theorem 3.0.3 (Predictable Representation Property (PRP)) *The CRP implies that every random variable F in $L^2(\Omega, \mathcal{F}, P)$ has a representation of the form*

$$F = E[F] + \sum_{i=1}^{\infty} \int_0^{\infty} \phi_s^{(i)} \mathrm{d}H_s^{(i)}, \tag{3.2}$$

where H's are defined in equation (1.10) and $\phi_s^{(i)}$'s are predictable, that is, they are $\mathcal{F}_{s-}$-measurable.

3.1 Pure jump case

Let us first outline the form of the representation to introduce the reader to the flavour of the results in this section. Suppose $t_0 \geq 0$ and let $G = \{G_t, t \geq 0\}$ be a pure jump Lévy process with no Brownian part (that is, $\sigma^2 = 0$ in the Lévy triplet), $G^{(i)} = \left\{G_t^{(i)}, t \geq 0\right\}$ be its i-th power jump process and $\widehat{G}^{(i)} = \left\{\widehat{G}_t^{(i)}, t \geq 0\right\}$ be its i-th compensated power jump process. Calculation of $(G_{t+t_0} - G_{t_0})^k$ for $k = 2, 3, 4$ are given in Appendix A.2. From the Itô formula,

$$\begin{aligned}&(X_{t+t_0} - X_{t_0})^k \\ &= \frac{\sigma^2}{2} k(k-1) \left((X_{t+t_0} - X_{t_0})^{k-2} t - \int_0^t s \, \mathrm{d}(X_{s+t_0} - X_{t_0})^{k-2} \right)\end{aligned} \tag{3.3}$$

$$+\sum_{j=1}^{k}\binom{k}{j}\int_{t_0}^{t+t_0}\left(X_{s-}-X_{t_0}\right)^{k-j}\mathrm{d}Y_s^{(j)} \tag{3.4}$$

$$+\sum_{j=1}^{k-1}\binom{k}{j}m_j\left(t+t_0\right)\left(X_{t+t_0}-X_{t_0}\right)^{k-j} \tag{3.5}$$

$$-\sum_{j=1}^{k-1}\binom{k}{j}m_j\int_{t_0}^{t+t_0}s\ \mathrm{d}\left(X_s-X_{t_0}\right)^{k-j}+m_kt. \tag{3.6}$$

The detailed derivation is given in Appendix A.1. Based on the structure of the expressions for $\left(G_{t+t_0}-G_{t_0}\right)^3$ and $\left(G_{t+t_0}-G_{t_0}\right)^4$, where detailed calculation is given in Appendices A.2.2 and A.2.3, we desire to derive a general formula for $\left(G_{t+t_0}-G_{t_0}\right)^k,\ k=1,2,3,...,$ as this forms a starting point for the representation of X. This derivation will be implemented in a series of steps. Firstly, we notice that the numbers of stochastic integrals in $\left(G_{t+t_0}-G_{t_0}\right)^3$ and $\left(G_{t+t_0}-G_{t_0}\right)^4$ are less than the possible full representation specified in the simplified version of the CRP, where the stochastic integrals are with respect to compensated power jump processes Y's, derived by Nualart & Schoutens (2000):

$$\begin{aligned}\left(X_{t+t_0}-X_{t_0}\right)^k &= f^{(k)}\left(t,t_0\right)+\sum_{j=1}^{k}\sum_{\substack{(i_1,...,i_j)\\ \in\{1,...,k\}^j}}\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\cdots\int_{t_0}^{t_{j-1}-}\\ &\qquad f^{(k)}_{(i_1,...,i_j)}\left(t,t_0,t_1,...,t_j\right)\mathrm{d}Y_{t_j}^{(i_j)}\cdots\mathrm{d}Y_{t_2}^{(i_2)}\mathrm{d}Y_{t_1}^{(i_1)},\end{aligned}$$

where the $f^{(k)}_{(i_1,...,i_j)}$'s are deterministic functions in $L^2\left(R_+^j\right)$. For example, in the representation of $\left(G_{t+t_0}-G_{t_0}\right)^2$, we have only three stochastic integrals

$$\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)},\ \int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\ \text{and}\ \int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}$$

in the representation, which we shall represent via the list $\{(1,1),(1),(2)\}$. We can do an equivalent representation of $\left(G_{t+t_0}-G_{t_0}\right)^3$ and $\left(G_{t+t_0}-G_{t_0}\right)^4$ to get the following two lists:

$$\{(1,1,1),(1,1),(1,2),(2,1),(1),(2),(3)\}.$$

$$\begin{aligned}&\{(1,1,1,1),(1,1,1),(1,1,2),(1,2,1),(2,1,1),\\ &(1,1),(1,2),(2,1),(2,2),(1,3),(3,1),(1),(2),(3),(4)\}.\end{aligned}$$

In general, the list of the orders of the compensated power jump processes of the stochastic integrals in $(G_{t+t_0} - G_{t_0})^k$ depends on the collection of numbers

$$\mathcal{I}_k = \left\{ (i_1, i_2, ..., i_j) \mid j \in \{1, 2, ..., k\}, \ i_p \in \{1, 2, ..., k\} \text{ and } \sum_{p=1}^{j} i_p \leq k \right\}. \tag{3.7}$$

This construction is explained in the beginning of the proof of Theorem 3.1.4 (Appendix A.4) using induction. A typical element $(i_1, i_2, ..., i_j)$ in $\mathcal{I}_k$ indexes a multiple stochastic integral j-times repeated with respect to the power jump processes with powers $i_1, i_2, ..., i_j$ and indexed $t_j, t_{j-1}, ..., t_1$. That is, $(i_1, i_2, ..., i_j)$ indexes the integral

$$\int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}} \mathrm{d}\widehat{G}_{t_j}^{(i_1)} \cdots \mathrm{d}\widehat{G}_{t_2}^{(i_{j-1})} \mathrm{d}\widehat{G}_{t_1}^{(i_j)}.$$

Next we consider the terms in the representation not involving any stochastic integrals. That is, in $(G_{t+t_0} - G_{t_0})^2$, $m_1^2 t^2 + m_2 t$ is considered; in $(G_{t+t_0} - G_{t_0})^3$, $m_1^3 t^3 + 3m_1 m_2 t^2 + m_3 t$ is considered, and in $(G_{t+t_0} - G_{t_0})^4$,

$$m_1^4 t^4 + 6 m_1^2 m_2 t^3 + \left(4 m_1 m_3 + 3 m_2^2\right) t^2 + m_4 t$$

is considered. We use (3.3)-(3.6) to derive the representation. This time the representation can be simplified a great deal since we are not considering any stochastic integrals. Denote the terms which do not contain any stochastic integral in $(G_{t+t_0} - G_{t_0})^k$ by $C_{t+t_0-t_0}^{(k)} = C_t^{(k)}$, and we refer this as the deterministic part of the representation.

Proposition 3.1.1 $C_0^{(r)} = 0$ *for all* r, $C_t^{(0)} = 1$, $C_t^{(1)} = m_1 t$, *and for* $k = 2, 3, 4, ...$,

$$C_t^{(k)} = \sum_{j=1}^{k-1} \binom{k}{j} m_j t C_t^{(k-j)} - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_0^t t_1 \ \mathrm{d}C_{t_1}^{(k-j)} + m_k t. \tag{3.8}$$

Proof. The results for $C_0^{(r)}$ and $C_t^{(0)}$ are trivial. For $k = 1$, $(G_{t+t_0} - G_{t_0}) = \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(1)} + m_1 t$ and hence $C_t^{(1)} = m_1 t$. For $k \geq 2$, the terms in (3.3) are equal to zero since G_t has no Brownian part. The term in (3.4) contains a stochastic integral and hence from (3.5) and (3.6), we have

$$C_t^{(k)} = \sum_{j=1}^{k-1} \binom{k}{j} m_j \left(t + t_0\right) C_t^{(k-j)} - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_{t_0}^{t+t_0} t_1 \ \mathrm{d}C_{t_1 - t_0}^{(k-j)} + m_k t.$$

Putting $u = t_1 - t_0$ in the second term, we have

$$\begin{aligned} C_t^{(k)} &= \sum_{j=1}^{k-1} \binom{k}{j} m_j (t + t_0) C_t^{(k-j)} - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_0^t (u + t_0) \, \mathrm{d}C_u^{(k-j)} + m_k t \\ &= \sum_{j=1}^{k-1} \binom{k}{j} m_j t C_t^{(k-j)} - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_0^t t_1 \, \mathrm{d}C_{t_1}^{(k-j)} + m_k t. \end{aligned}$$

Note that $C_t^{(k)}$ is independent of t_0. □

Thus, given Proposition 3.1.1, $C_t^{(k)}$ can be expressed in terms of m_i's for any given k and easily coded. We will show in the followings that in the calculation of $(G_{t+t_0} - G_{t_0})^k$, all the $C_t^{(j)}$'s, $j = 0, 1, ..., k$ are required. In fact the coefficients of the stochastic integrals in the representation depend only on $C_t^{(j)}$'s, $j = 0, 1, ..., k$, as stated in Theorem 3.1.4 below.

The next proposition gives the representation for $C_t^{(k)}$ in a non-recursive form. Let

$$\mathcal{L}_k = \left\{ (i_1, i_2, ..., i_l) \, | l \in \{1, 2, ..., k\}, i_q \in \{1, 2, ..., k\}, i_1 \geq i_2 \geq \cdots \geq i_l, \sum_{q=1}^{l} i_q = k \right\}. \tag{3.9}$$

The number of distinct values in a tuple $\phi_k = \left(i_1^{(k)}, i_2^{(k)}, ..., i_l^{(k)}\right)$ in $\mathcal{L}_k$ is less than or equal to l. When it is less than l, it means some of the value(s) in the tuple are repeated. Let the number of times $r \in \{1, 2, 3, .., k\}$ appears in the tuple $\phi_k = \left(i_1^{(k)}, i_2^{(k)}, ..., i_l^{(k)}\right)$ be $p_r^{\phi_k}$.

Proposition 3.1.2

$$C_t^{(k)} = \sum_{\phi_k = \left(i_1^{(k)}, i_2^{(k)}, ..., i_l^{(k)}\right) \in \mathcal{L}_k} \frac{1}{l!} \left(i_1^{(k)}, i_2^{(k)}, ..., i_l^{(k)}\right)! \left(p_1^{\phi_k}, p_2^{\phi_k}, ..., p_k^{\phi_k}\right)! \left[\prod_{q \in \phi_k} m_q\right] t^l \tag{3.10}$$

where $i_1^{(k)}, ..., i_l^{(k)}$ *are the elements of* ϕ_k*,* $p_j^{\phi_k}$*'s are defined above and* $\left(i_1^{(k)}, i_2^{(k)}, ..., i_l^{(k)}\right)!$ *is the multinomial coefficient:* $\left(i_1^{(k)}, i_2^{(k)}, ..., i_l^{(k)}\right)! = \frac{\left(\sum_{j=1}^{l} i_j^{(k)}\right)!}{i_1^{(k)}! i_2^{(k)}! \cdots i_l^{(k)}!}$.

Proof. The proof is included in Appendix A.3.

Proposition 3.1.3 *Let* $\Pi_{(i_1, i_2, ..., i_j), t}^{(k)}$ *be the coefficient of*

$$\int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}\widehat{G}_{t_j}^{(i_1)} \cdots \mathrm{d}\widehat{G}_{t_2}^{(i_{j-1})} \mathrm{d}\widehat{G}_{t_1}^{(i_j)}$$

in $(G_{t+t_0} - G_{t_0})^k$. *Then*

$$\Pi^{(k)}_{(i_1,i_2,\ldots,i_j),t} = (i_1, i_2, \ldots, i_j, n)! C_t^{(n)} \text{ where } n = k - \sum_{p=1}^{j} i_p. \tag{3.11}$$

Proof. The proof of Proposition 3.1.3 is contained in the proof of Theorem 3.1.4. □

For example, say we want to determine the coefficient of $\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1}$ in the representation of $(G_{t+t_0} - G_{t_0})^4$, that is, we want to find $\Pi^{(4)}_{(1,1),t}$. To derive this coefficient, we first note that $n = 2$ and so $\Pi^{(4)}_{(1,1),t} = \frac{4!}{1!1!2!} C_t^{(2)} = 12\left(m_2 t + m_1^2 t^2\right)$, which is true according to the calculation of $(G_{t+t_0} - G_{t_0})^4$ given in Appendix A.2.3. Now we put the above results together to get a general formula for $(G_{t+t_0} - G_{t_0})^k$.

Theorem 3.1.4 *Let $G = \{G_t, t \geq 0\}$ be a Lévy process with no Brownian part satisfying condition (1.7). Then the power of its increment can be expressed by:*

$$(G_{t+t_0} - G_{t_0})^k = \sum_{\theta_k \in \mathcal{I}_k} \Pi^{(k)}_{\theta_k,t} \mathcal{S}_{\theta_k,t,t_0} + C_t^{(k)}, \tag{3.12}$$

where $\mathcal{I}_k$ is defined in (3.7), $\Pi^{(k)}_{\theta_k,t}$ is defined in Proposition 3.1.3, the $C_t^{(k)}$ are constants defined in Proposition 3.1.2 and $\mathcal{S}_{(i_1,i_2,\ldots,i_j),t,t_0}$ is defined as the integral:

$$\mathcal{S}_{(i_1,i_2,\ldots,i_j),t,t_0} = \int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\cdots\int_{t_0}^{t_{j-1}-} \mathrm{d}\widehat{G}^{(i_1)}_{t_j}\cdots \mathrm{d}\widehat{G}^{(i_{j-1})}_{t_2}\mathrm{d}\widehat{G}^{(i_j)}_{t_1}.$$

Proof. The proof is included in Appendix A.4. □

To derive the explicit formula for the power of increment of a Lévy process with respect to orthogonalised compensated power jump processes, we need the following proposition.

Proposition 3.1.5 *The n-th compensated power jump processes, $Y^{(n)} = \left\{Y_t^{(n)}, t \geq 0\right\}$, of a general Lévy processes satisfying condition (1.7), can be expressed in terms of the orthogonalised compensated power jump processes, $H^{(i)} = \left\{H_t^{(i)}, t \geq 0\right\}$ for $i = 1, 2, \ldots, n$, by*

$$Y_t^{(n)} = H_t^{(n)} + \sum_{k=1}^{n-1} b_{n,k} H_t^{(k)},$$

where $b_{n,k}$ denotes the sum of all the elements of the set $\mathcal{M}^{n,k}$, which is defined by

$$\mathcal{M}^{n,k} = \left\{(-1)^{j-1} a_{i_1,i_2} a_{i_2,i_3} \cdots a_{i_{j-1},i_j} : i_1 = n, i_j = k, i_p > i_q \text{ if } p < q, i_p \in \mathbb{N} \text{ for all } p\right\},$$

and $\mathcal{M}^{n,n} = \{1\}$.

Proof. The proof is included in Appendix A.5. □

Theorem 3.1.6 *Let* $G = \{G_t, t \geq 0\}$ *be a Lévy process with no Brownian part satisfying condition (1.7). Then the power of its increment in terms of stochastic integrals with respect to the orthogonalised compensated power jump processes,* $H^{(j)}$*'s, is given by the following equation:*

$$(G_{t+t_0} - G_{t_0})^k = \sum_{\theta_k \in \mathcal{I}_k} \Pi^{(k)}_{\theta_k,t} \mathcal{S}^{(H)}_{\theta_k,t,t_0} + C^{(k)}_t, \tag{3.13}$$

where $\mathcal{I}_k$ *is defined in (3.7),* $\Pi^{(k)}_{\theta_k,t}$ *is defined in Proposition 3.1.3,* $C^{(k)}_t$ *is defined in Proposition 3.1.2 and* $\mathcal{S}^{(H)}_{(i_1,i_2,\ldots,i_j),t,t_0}$ *is defined as the integral:*

$$\begin{aligned}\mathcal{S}^{(H)}_{(i_1,i_2,\ldots,i_j),t,t_0} &= \sum_{k_1=1}^{i_1} \cdots \sum_{k_{j-1}=1}^{i_{j-1}} \sum_{k_j=1}^{i_j} b_{i_1,k_1} \cdots b_{i_{j-1},k_{j-1}} b_{i_j,k_j} \\ &\quad \times \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}H^{(k_1)}_{t_j} \cdots \mathrm{d}H^{(k_{j-1})}_{t_2} \mathrm{d}H^{(k_j)}_{t_1},\end{aligned}$$

$b_{n,k}$ *is defined in Proposition 3.1.5.*

Proof. From Proposition 3.1.5, we have

$$\begin{aligned}\mathcal{S}_{(i_1,i_2,\ldots,i_j),t,t_0} &= \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}\widehat{G}^{(i_1)}_{t_j} \cdots \mathrm{d}\widehat{G}^{(i_{j-1})}_{t_2} \mathrm{d}\widehat{G}^{(i_j)}_{t_1} \\ &= \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}\left[\sum_{k_1=1}^{i_1} b_{i_1,k_1} H^{(k_1)}_{t_j}\right] \cdots \\ &\qquad \mathrm{d}\left[\sum_{k_{j-1}=1}^{i_{j-1}} b_{i_{j-1},k_{j-1}} H^{(k_{j-1})}_{t_2}\right] \mathrm{d}\left[\sum_{k_j=1}^{i_j} b_{i_j,k_j} H^{(k_j)}_{t_1}\right] \\ &= \sum_{k_1=1}^{i_1} \cdots \sum_{k_{j-1}=1}^{i_{j-1}} \sum_{k_j=1}^{i_j} b_{i_1,k_1} \cdots b_{i_{j-1},k_{j-1}} b_{i_j,k_j} \\ &\quad \times \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}H^{(k_1)}_{t_j} \cdots \mathrm{d}H^{(k_{j-1})}_{t_2} \mathrm{d}H^{(k_j)}_{t_1}.\end{aligned}$$

Hence, by using Theorem 3.1.4, we complete the proof. □

Corollary 3.1.7 *By Theorem 3.1.4,*

$$\begin{aligned}
&\left(G_{t+t_0}-G_{t_0}\right)^m\left(G_{t+t_0}-G_{t_0}\right)^n \\
&= \left(\sum_{\theta_m\in\mathcal{I}_m}\Pi^{(m)}_{\theta_m,t}\mathcal{S}^{(H)}_{\theta_m,t,t_0}+C^{(m)}_t\right)\left(\sum_{\theta_n\in\mathcal{I}_n}\Pi^{(n)}_{\theta_n,t}\mathcal{S}^{(H)}_{\theta_n,t,t_0}+C^{(n)}_t\right) \\
&= \sum_{\theta_{m+n}\in\mathcal{I}_{m+n}}\Pi^{(m+n)}_{\theta_{m+n},t}\mathcal{S}^{(H)}_{\theta_{m+n},t,t_0}+C^{(m+n)}_t.
\end{aligned}$$

Hence, we can convert the product of two iterative stochastic integrals of orders m and n as a weighted sum of iterative stochastic integrals of orders $m+n,\ m+n-1,...,2,1.$

Note in Theorems 3.1.4 and 3.1.6, the integrands of the stochastic integrals do **not** involve t_0 nor any of the integrating variables $t_1,t_2,...,t_j$. They are completely characterised by $C^{(p)}_t$'s, where $p=0,1,...,k$. To find the chaotic representation of $\left(G_{t+t_0}-G_{t_0}\right)^k$, we only need to know the moments of G_t, $m_1=E\left[G_t\right]/t$ and $m_p=\int_{-\infty}^{\infty}x^p\nu\left(\mathrm{d}x\right)$ for $p=2,...,k$. This result is intuitive as $\left(G_{t+t_0}-G_{t_0}\right)$ is a stationary process.

3.2 General case

Next we want to derive the formula for the power of the increments of Lévy processes when $\sigma\neq 0$. Recall $X=\{X_t,t\geq 0\}$ denotes a general Lévy process, $X^{(i)}=\left\{X^{(i)}_t,t\geq 0\right\}$ denotes its i-th power jump process and $Y^{(i)}=\left\{Y^{(i)}_t,t\geq 0\right\}$ denotes its i-th compensated power jump process as defined in (1.9). We define $A_1\left(X_{t+t_0},X_{t_0};k\right)$ and $A_2\left(X_{t+t_0},X_{t_0};k\right)$ such that $\left(X_{t+t_0}-X_{t_0}\right)^k=A_1\left(X_{t+t_0},X_{t_0};k\right)+A_2\left(X_{t+t_0},X_{t_0};k\right)$, where $A_1\left(X_{t+t_0},X_{t_0};k\right)$ comprises all the terms not containing σ in $\left(X_{t+t_0}-X_{t_0}\right)^k$. By expressing $A_2\left(X_{t+t_0},X_{t_0};k\right)$ using (3.3)-(3.6), it may directly be noted:

$$\begin{aligned}
\left(X_{t+t_0}-X_{t_0}\right)^k &= \frac{\sigma^2}{2}k\left(k-1\right)\left(\left(X_{t+t_0}-X_{t_0}\right)^{k-2}t-\int_{t_0}^{t+t_0}\left(s-t_0\right)\mathrm{d}\left(X_s-X_{t_0}\right)^{k-2}\right) \\
&\quad+\sum_{j=1}^{k}\binom{k}{j}\int_{t_0}^{t+t_0}A_2\left(X_{s-},X_{t_0};k-j\right)\mathrm{d}Y^{(j)}_s \\
&\quad+\sum_{j=1}^{k-1}\binom{k}{j}m_j\left(t+t_0\right)A_2\left(X_{t+t_0},X_{t_0};k-j\right) \\
&\quad-\sum_{j=1}^{k-1}\binom{k}{j}m_j\int_{t_0}^{t+t_0}s\ \mathrm{d}\left[A_2\left(X_s,X_{t_0};k-j\right)\right] \\
&\quad+A_1\left(X_{t+t_0},X_{t_0};k\right).
\end{aligned} \tag{3.14}$$

Calculation of $(X_{t+t_0} - X_{t_0})^k$ for $k = 3, 4, 5$ are given in Appendix A.6.

Proposition 3.2.1 *For any Lévy process $X = \{X_t, t \geq 0\}$ satisfying condition (1.7),*

$$(X_{t+t_0} - X_{t_0})^k = A_1 (X_{t+t_0}, X_{t_0}; k) + \sum_{n=1}^{\lfloor k/2 \rfloor} \frac{k!}{(k-2n)!} \frac{1}{n!} \frac{1}{2^n} \sigma^{2n} A_1 (X_{t+t_0}, X_{t_0}; k-2n)\, t^n.$$

Proof. The proof uses the same techniques as the proof of Theorem 3.1.4. Note that $A_1 (X_{t+t_0}, X_{t_0}; p)$, where $p = 1, 2, ..., k$, are given by Theorem 3.1.4. □

Proposition 3.2.1 gives the formula of $(X_{t+t_0} - X_{t_0})^k$ in terms of a summation of A_1, where $\lfloor k/2 \rfloor + 1$ calculations of A_1 are needed. The next theorem gives the formula in an alternative form, which requires A_1 to be computed once only.

Definition 3.2.2 *Let $C_{t,\sigma}^{(k)}$ be the terms obtained by replacing m_2 with $m_2 + \sigma^2$ in $C_t^{(k)}$ (Proposition 3.1.2) and $\Pi_{(i_1,i_2,...,i_j),t,\sigma}^{(k)}$ be the terms obtained by replacing $C_t^{(k)}$ with $C_{t,\sigma}^{(k)}$ in $\Pi_{(i_1,i_2,...,i_j),t}^{(k)}$ (Proposition 3.1.3).*

We then note the following theorem.

Theorem 3.2.3 *For any Lévy process $X = \{X_t, t \geq 0\}$ with $\sigma^2 \neq 0$ and satisfying condition (1.7), the representation of $(X_{t+t_0} - X_{t_0})^n$ is given by Theorem 3.1.4 with m_2 replaced by $(m_2 + \sigma^2)$, that is,*

$$(X_{t+t_0} - X_{t_0})^n = \sum_{\theta_n \in \mathcal{I}_n} \Pi_{\theta_n,t,\sigma}^{(n)} \mathcal{S}'_{\theta_n,t,t_0} + C_{t,\sigma}^{(n)},$$

where $\mathcal{I}_n$ is defined in (3.7), $\Pi_{\theta_n,t,\sigma}^{(n)}$ and $C_{t,\sigma}^{(n)}$ are defined in Definition 3.2.2, and the stochastic integral $\mathcal{S}'_{(i_1,i_2,...,i_j),t,t_0}$ is defined by:

$$\mathcal{S}'_{(i_1,i_2,...,i_j),t,t_0} = \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}Y_{t_j}^{(i_1)} \cdots \mathrm{d}Y_{t_2}^{(i_{j-1})} \mathrm{d}Y_{t_1}^{(i_j)}.$$

Proof. We define a new class of power jump processes by:

$$\begin{aligned} \widetilde{X}_t^{(2)} &= X_t^{(2)} + \sigma^2 t, \\ \widetilde{X}_t^{(j)} &= X_t^{(j)} \quad \text{for } j = 1 \text{ and } j = 3, 4, 5, ... \end{aligned} \tag{3.15}$$

We also define a new class of compensators

$$\begin{aligned} \tilde{m}_2 t &= \left(m_2 + \sigma^2\right) t, \\ \tilde{m}_j t &= m_j t \quad \text{for } j = 1 \text{ and } j = 3, 4, 5, ... \end{aligned}$$

Hence, by definition, the compensated power jump processes, $\tilde{Y}_t^{(i)} = \tilde{X}_t^{(i)} - \tilde{m}_i t = X_t^{(i)} - m_i t = Y_t^{(i)}$ for all $i \geq 1$. Thus the representation of $(X_{t+t_0} - X_{t_0})^k$ in terms of the stochastic integrals with respect to $Y_t^{(i)}$ is the same no matter we start from using $X_t^{(i)}$ or $\widetilde{X}_t^{(i)}$. To calculate the expression using $\widetilde{X}_t^{(i)}$, we use equation (2) in Nualart & Schoutens (2000), namely:

$$
\begin{aligned}
&(X_{t+t_0} - X_{t_0})^k \\
&= \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \, \mathrm{d}X_s^{(j)} \\
&\quad + \frac{\sigma^2}{2} k(k-1) \left((X_{t+t_0} - X_{t_0})^{k-2} t - \int_0^t s \, \mathrm{d}(X_{s+t_0} - X_{t_0})^{k-2} \right) \\
&= \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \, \mathrm{d}X_s^{(j)} + \frac{\sigma^2}{2} k(k-1) \int_0^t \left(X_{(s+t_0)-} - X_{t_0} \right)^{k-2} \mathrm{d}s \\
&= \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \, \mathrm{d}X_s^{(j)} + \frac{\sigma^2}{2} k(k-1) \int_{t_0}^{t+t_0} (X_{u-} - X_{t_0})^{k-2} \, \mathrm{d}u \\
&= \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \, \mathrm{d}X_s^{(j)} + \binom{k}{2} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-2} \, \mathrm{d}\left(\sigma^2 s\right).
\end{aligned}
$$

By (3.15), we have

$$
(X_{t+t_0} - X_{t_0})^k = \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \, \mathrm{d}\tilde{X}_s^{(j)}.
$$

Using exactly the same calculation as the one leading to (3.3)-(3.6), we have

$$
\begin{aligned}
&(X_{t+t_0} - X_{t_0})^k \\
&= \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \, \mathrm{d}Y_s^{(j)} + \sum_{j=1}^{k-1} \binom{k}{j} \tilde{m}_j (t + t_0)(X_{t+t_0} - X_{t_0})^{k-j} \\
&\quad - \sum_{j=1}^{k-1} \binom{k}{j} \tilde{m}_j \int_{t_0}^{t+t_0} s \, \mathrm{d}(X_s - X_{t_0})^{k-j} + \tilde{m}_k t.
\end{aligned}
$$

This is exactly the equation (3.4)-(3.6) we based on in the derivation of Theorem 3.1.4, except that m_j is now replaced by $\tilde{m}_j$. Hence we now have a simple formula for the representation of $(X_{t+t_0} - X_{t_0})^k$ in terms of the stochastic integrals with respect to $Y_t^{(i)}$

by replacing m_j with $\tilde{m}_j$ in the formula given by Theorem 3.1.4. In other words, we have

$$(X_{t+t_0} - X_{t_0})^n = \sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,t,\sigma} \mathcal{S}'_{\theta_n,t,t_0} + C^{(n)}_{t,\sigma},$$

where $\Pi^{(n)}_{\theta_n,t,\sigma}$ and $C^{(n)}_{t,\sigma}$ are defined in Definition 3.2.2. Note that this representation does not depend on the power jump processes directly since it is in terms of the compensated power jump processes, $Y^{(j)}_t$'s. So it does not matter if we change the definition of the power jump processes, as long as we change the compensators accordingly, we will get the same compensated power jump processes. □

Theorem 3.2.4 *For any Lévy process $X = \{X_t, t \geq 0\}$ with $\sigma^2 \neq 0$ and satisfying condition (1.7), the representation of $(X_{t+t_0} - X_{t_0})^n$ is given by Theorem 3.1.6 with m_2 replaced with $\left(m_2 + \sigma^2\right)$, that is,*

$$(X_{t+t_0} - X_{t_0})^n = \sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,t,\sigma} \mathcal{S}'^{(H)}_{\theta_n,t,t_0} + C^{(n)}_{t,\sigma},$$

where $\mathcal{I}_n$ is defined in (3.7), $\Pi^{(n)}_{\theta_n,t,\sigma}$ and $C^{(n)}_{t,\sigma}$ are defined in Definition 3.2.2 and the stochastic integral $\mathcal{S}'_{(i_1,i_2,\ldots,i_j),t,t_0}$ is defined by:

$$\begin{aligned}\mathcal{S}'^{(H)}_{(i_1,i_2,\ldots,i_j),t,t_0} &= \sum_{k_1=1}^{i_1} \cdots \sum_{k_{j-1}=1}^{i_{j-1}} \sum_{k_j=1}^{i_j} b_{i_1,k_1} \cdots b_{i_{j-1},k_{j-1}} b_{i_j,k_j} \\ &\quad \times \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{j-1}-} \mathrm{d}H^{(k_1)}_{t_j} \cdots \mathrm{d}H^{(k_{j-1})}_{t_2} \mathrm{d}H^{(k_j)}_{t_1},\end{aligned}$$

$b_{n,k}$ is defined in Proposition 3.1.5.

Proof. It follows directly from Theorems 3.1.6 and 3.2.3. □

Remark 3.2.1 *As noted in Section 1.4.2, Jamshidian (2005) derived an explicit formula for the chaotic representation of $(X_t)^k$ in terms of the non-compensated power jump processes, $X^{(j)}$'s, when X is a semimartingale. Our explicit formula gives the representation in terms of orthogonalised compensated power jump processes, $H^{(j)}$'s. In the following, we show that our formula is an non-trivial extension of Jamshidian's one in the Lévy case, which is an important subclass of semimartingales. We note the notation used by Jamshidian in Section 1.4.2. If $X = \{X_t, t \geq 0\}$ is a Lévy process, we can see that $[X^c]_t = [X]^c_t = \sigma^2 t$ (where the superscript c stands for continuous part of the process)* *and hence* $[X]_t^{(2)} = \sigma^2 t + \sum_{s\leq t} (\Delta X_s)^2$. *With Jamshidian's notation, the* σ^2 *is implicitly included in the* $[X]_t^{(2)}$.

Jamshidian (2005) defined $\mathcal{C} = \mathcal{C}^* \cap \mathcal{C}_*$, *where* $\mathcal{C}^*$ *is the set of semimartingales of finite moments with continuous compensators adapted to a Brownian filtration, and* $\mathcal{C}_*$ *is the set of processes with exponentially decreasing law. Jamshidian generalised the CRP from Lévy processes to the set* $\mathcal{C}$. *In proposition 8.2 of Jamshidian (2005), an explicit formula for the chaotic representation with respect to the non-compensated power jump processes for the semimartingales in* $\mathcal{C}$ *when* $t_0 = 0$ *was derived. Jamshidian (2005) defined the power jump processes using the power brackets, see (1.11) and (1.12). The multi-indices were denoted by* $I = (i_1, ..., i_p) \in \mathbb{N}^p$, *where* $\mathbb{N}$ *is the set of natural numbers, and for integers* $1 \leq p \leq n$,

$$\mathbb{N}_n^p = \{I = (i_1, ..., i_p) \in \mathbb{N}^p : i_1 + \cdots + i_p = n\}, \quad p, n \in \mathbb{N}. \tag{3.16}$$

Note that from (3.7), $\mathcal{I}_k = \bigcup_{n=1}^{k} \bigcup_{p=1}^{n} \mathbb{N}_n^p$. *Proposition 8.2 of Jamshidian (2005) states that, for a semimartingale* $X = \{X_t, t \geq 0\}$ *with* $X_0 = 0$, *we have, for all* $n \in N$

$$X_t^n = \sum_{p=1}^{n} \sum_{I \in \mathbb{N}_n^p} \frac{n!}{i_1! \cdots i_p!} \int_0^t \int_0^{t_1-} \cdots \int_0^{t_{p-1}-} \mathrm{d}\,[X]_{t_p}^{(i_1)} \cdots \mathrm{d}\,[X]_{t_2}^{(i_p-1)}\, \mathrm{d}\,[X]_{t_1}^{(i_p)}. \tag{3.17}$$

Since Jamshidian (2005) only considered non-compensated processes, we substitute all the m_j *in (3.8) by zeros (since the compensators in the Lévy case are* $m_j t$), *which makes* $C_t^{(k)} = 0$ *for all* $k \neq 0$. *So* $\Pi^{(k)}_{(i_1,i_2,\ldots,i_j),t}$ *is non-zero only when* $\sum_{p=1}^{j} i_p = k$, *as defined in (3.16). Hence in the Lévy case, Theorem 3.2.3 reduces to (3.17). In other words, Jamshidian's formula can be deduced from ours (in the Lévy case), but ours cannot be deduced from Jamshidian's by a non-trivial calculation.*

Corollary 3.2.5 *The expectation of* $(X_{t+t_0} - X_{t_0})^k$ *is given by* $C_{t,\sigma}^{(k)}$, *which can be obtained by replacing* m_2 *with* $m_2 + \sigma^2$ *in* $C_t^{(k)}$, *given by equation (3.10).*

Proof. As the expectations of all the stochastic integrals are zero, this follows directly from Theorem 3.2.3. □

Corollary 3.2.6 *The expectation of* $\left(H_t^{(1)}\right)^k = \left(\int_0^t \mathrm{d}H_{t_1}^{(1)}\right)^k$ *can be obtained by replacing* m_2 *with* $m_2 + \sigma^2$ *and* m_1 *with* 0 *in* $C_t^{(k)}$, *given by Proposition 3.1.2.*

Proof. From Corollary 3.2.5, $E\left[X_t^k\right]$ can be obtained by replacing m_2 with $m_2 + \sigma^2$ in $C_t^{(k)}$. Since $H_t^{(1)} = X_t - m_1 t$ and

$$(X_t)^k = \left(\int_0^t \mathrm{d}H_{t_1}^{(1)} + m_1 t\right)^k, \tag{3.18}$$

by putting $m_1 = 0$ in (3.18), we can conclude that the expectation of $\left(\int_0^t \mathrm{d}H_{t_1}^{(1)}\right)^k$ can be obtained by replacing m_2 with $m_2 + \sigma^2$ and m_1 with 0 in $C_t^{(k)}$.
□

In the next chapter, we extend our results to chaos expansions in terms of the Poisson random measure, with the use of the relationship between the two chaos expansions derived by Benth *et al.* (2003).

Chapter 4

Chaos expansion with respect to Poisson random measures

Itô (1956) proved a chaos expansion for general Lévy processes in terms of multiple integrals with respect to the compensated Poisson random measure. One may convert the representation to one involving iterated integrals by defining the symmetrisation of a real function. Following Løkka (2004), let f be a real function on $([0,T]\times\mathbb{R})^n$. We define its *symmetrisation* $\tilde{f}$, with respect to the variables $(t_1,x_1),...,(t_n,x_n)$, to be

$$\tilde{f}(t_1,x_1,...,t_n,x_n)=\frac{1}{n!}\sum_{\boldsymbol{\pi}} f(t_{\pi_1},x_{\pi_1},...,t_{\pi_n},x_{\pi_n}), \tag{4.1}$$

where the sum is taken over all permutations $\boldsymbol{\pi}$ of $\{1,...,n\}$. f is said to be *symmetric* if $f=\tilde{f}$.

4.1 Pure jump case

We first consider the representation of pure jump Lévy processes as in Løkka (2004). Let $\tilde{L}_2((\lambda\times\nu)^n)$ be the space of all square integrable symmetric functions on $([0,T]\times\mathbb{R})^n$. In an iterative integral such as (3.1), the time variables $t_1,...,t_n$ are monotonic. For ease of notation so that we do not have to explicitly note the time points and the process values, we let:

$$G_n=\{(t_1,x_1,...,t_n,x_n):0\le t_1\le\cdots\le t_n\le T;x_i\in\mathbb{R},i=1,...,n\}, \tag{4.2}$$

and let $L_2(G_n)$ be the space of functions g such that

$$\|g\|^2_{L_2(G_n)} = \int_{G_n} g^2(t_1, x_1, ..., t_n, x_n)\, \mathrm{d}t_1 \nu(\mathrm{d}x_1) \cdots \mathrm{d}t_n \nu(\mathrm{d}x_n) < \infty,$$

where $\nu(\mathrm{d}x)$ is the Lévy measure of the underlying Lévy process. For $f \in L_2(G_n)$, let

$$J_n(f) = \int_0^T \int_{\mathbb{R}} \cdots \int_0^{t_2} \int_{\mathbb{R}} f(t_1, x_1, ..., t_n, x_n)\, \tilde{N}(\mathrm{d}t_1, \mathrm{d}x_1) \cdots \tilde{N}(\mathrm{d}t_n, \mathrm{d}x_n),$$

an iterative stochastic integral with respect to individual measures, where $\tilde{N}$ is the compensated Poisson random measure defined in Definition 1.1.6. For $f \in \tilde{L}_2((\lambda \times \nu)^n)$, let

$$I_n(f) = \int_{([0,T]\times\mathbb{R})^n} f(t_1, x_1, ..., t_n, x_n)\, \tilde{N}^{\otimes n}(\mathrm{d}\mathbf{t}, \mathrm{d}\mathbf{x}) = n! J_n(f),$$

an stochastic integral with respect to the product measure.

Theorem 4.1.1 (Chaos expansion for Lévy process by Itô (1956)) *Let F be a square integrable random variable adapted to the underlying pure jump Lévy process, X. There exists a unique sequence $\{f_n\}_{n=0}^{\infty}$ where $f_n \in \tilde{L}_2([0,T] \times \mathbb{R})^n$ such that*

$$F = E(F) + \sum_{n=1}^{\infty} I_n(f_n). \tag{4.3}$$

Benth *et al.* (2003) derived relations between the expansion in terms of compensated power jump processes and the expansion in terms of the Poisson random measure. Benth *et al.* (2003) showed that when the underlying Lévy process is a pure jump process, the compensated power jump process defined in (1.9) satisfies the equation

$$Y_t^{(i)} = \int_0^t \int_{\mathbb{R}} x^i \tilde{N}(\mathrm{d}s, \mathrm{d}x), \quad 0 \le t \le T,\ i = 1, 2, ... \tag{4.4}$$

This relationship is very important in the development of the chaotic representation of Lévy processes. Since the introduction of the chaos expansion by Itô (1956), the development of representations in the literature has been focused on expansions with respect to the Poisson random measure. Unfortunately, we cannot trade in the Poisson random measure. Note that trading in a finite set of power jump assets is theoretically possible because the i-th power jump asset contains information of the i-th moment of the Lévy process, given that i is finite. Therefore, it is possible to construct a financial product which contains information of the i-th moment of the underlying process. For example, if we want to hedge the risk introduced by the variance of the underlying process, we can trade in the variance swaps or the second power jump asset[1]. However, the Poisson random measure contains **all** the information of the moments up to infinity and hence it is not clear how to construct such a financial product unless information of all the higher moments are obtained. This limits the application of the CRP in terms of Poisson random measures and also the application of Lévy processes in finance.

In the Black-Scholes world, due to the existence of PRP of Brownian motions, the market is complete and every contingent claim can be replicated by a portfolio investing only in a risk-free bank account and the underlying asset. Nualart & Schoutens (2000) introduced a new version of the CRP in terms of orthogonalised compensated power jump processes. Corcuera *et al.* (2005) suggested trading in some related power jump assets, making perfect hedging possible. The equation (4.4) therefore links the two important expansions together and hence the results derived for expansions in terms of Poisson random measures can be applied to expansions in terms of power jump processes. In this thesis, we first derive the explicit formula for the latter expansion and then apply equation (4.4) to obtain the explicit formula for the former expansion. The CRP in terms of compensated power jump processes can be converted into the CRP in terms of the Poisson random measure as follows:

$$\begin{aligned} F &= E(F)+\sum_{j=1}^{\infty}\sum_{i_1,\dots,i_j\geq 1}\int_0^T\int_0^{t_1-}\cdots\int_0^{t_{j-1}-} f_{(i_1,\dots,i_j)}(t_1,\dots,t_j) \qquad (4.5)\\ &\quad \mathrm{d}Y_{t_j}^{(i_j)}\dots\mathrm{d}Y_{t_2}^{(i_2)}\mathrm{d}Y_{t_1}^{(i_1)}\\ &= E(F)+\sum_{j=1}^{\infty}\sum_{i_1,\dots,i_j\geq 1}\int_0^T\int_{\mathbb{R}}\int_0^{t_1-}\int_{\mathbb{R}}\cdots\int_0^{t_{j-1}-}\int_{\mathbb{R}} x_j^{i_j}\cdots x_2^{i_2}x_1^{i_1}\\ &\quad \times f_{(i_1,\dots,i_j)}(t_1,\dots,t_j)\tilde{N}(\mathrm{d}t_j,\mathrm{d}x_j)\cdots\tilde{N}(\mathrm{d}t_2,\mathrm{d}x_2)\tilde{N}(\mathrm{d}t_1,\mathrm{d}x_1) \qquad (4.6)\\ &= E(F)+\sum_{j=1}^{\infty}\int_0^T\int_{\mathbb{R}}\int_0^{t_1-}\int_{\mathbb{R}}\cdots\int_0^{t_{j-1}-}\int_{\mathbb{R}} g_j(t_1,x_1,\dots,t_j,x_j)\\ &\quad \tilde{N}(\mathrm{d}t_j,\mathrm{d}x_j)\cdots\tilde{N}(\mathrm{d}t_2,\mathrm{d}x_2)\tilde{N}(\mathrm{d}t_1,\mathrm{d}x_1)\\ &= E(F)+\sum_{j=1}^{\infty}J_j(g_j)=E(F)+\sum_{j=1}^{\infty}j!J_j(\tilde{g}_j)=E(F)+\sum_{j=1}^{\infty}I_j(\tilde{g}_j), \end{aligned}$$

[1] In Part II, we discuss the use of power jump assets and moment swaps in perfect hedging of options and pointed out that power jump assets could not actually be traded in reality because they cannot be observed. Nonetheless, moment swaps, which are the generalisations of variance swaps, have high potential to be traded in the market.

where $\tilde{g}_j$ is the symmetrisation (defined in (4.1)) of the function g_j given by

$$g_j(t_1, x_1, ..., t_j, x_j) = \begin{cases} \sum_{i_1,...,i_j \geq 1} x_1^{i_1} \cdots x_j^{i_j} f_{(i_1,...,i_j)}(t_1, ..., t_j), & \text{on } G_j \\ 0 & \text{on } ([0,T] \times \mathbb{R})^j - G_j. \end{cases} \tag{4.7}$$

Therefore, by uniqueness, $\{f_n\}_{n=0}^{\infty}$ in Theorem 4.1.1 is given by $f_n = \tilde{g}_n$, where $n = 1, 2, ...$ This equation provides a simple relationship between the two expansions. From Theorem 3.2.3, we have

$$(X_{t+t_0} - X_{t_0})^n = \sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,t,\sigma} \mathcal{S}'_{\theta_n,t,t_0} + C^{(n)}_{t,\sigma}. \tag{4.8}$$

We can now use this relationship to derive a form for $\tilde{g}_n$ in terms of $\mathcal{I}_n$, $\Pi^{(n)}_{\theta_n,t,\sigma}$ and $C^{(n)}_{t,\sigma}$. Let $\mathcal{K}_{l,s} = \left\{(i_1, ..., i_l) \,|i_j \in \{1, 2, ..., s\} \text{ and} \sum_{j=1}^{l} i_j = s\right\}$. Since the length of a tuple must not be greater than the sum of all the elements in the tuple (because an element must be at least 1), $l \leq s$. By definition, we have $\mathcal{I}_n = \bigcup_{s=1}^{n} \bigcup_{l=1}^{s} \mathcal{K}_{l,s}$. So we can write

$$(X_{t+t_0} - X_{t_0})^n = \sum_{l=1}^{n} \sum_{s=l}^{n} \sum_{\theta_n \in \mathcal{K}_{l,s}} \Pi^{(n)}_{\theta_n,t,\sigma} \mathcal{S}'_{\theta_n,t,t_0} + C^{(n)}_{t,\sigma},$$

where θ_n is the tuple $\left(i_1^{\theta_n}, ..., i_l^{\theta_n}\right)$ with l elements which sum up to s. Therefore, we deduce that for $F = (X_{t+t_0} - X_{t_0})^n$ in (4.5), $f_{(i_1,...,i_l)}(t_1, ..., t_l)$ is given by

$$f_{(i_1,...,i_l)}(t_1, ..., t_l) = \Pi^{(n)}_{\theta_n,t,\sigma}. \tag{4.9}$$

By (4.7), we have then proved the following proposition.

Proposition 4.1.2 *For any pure jump Lévy process $X = \{X_t, t \geq 0\}$ satisfying condition (1.7),*

$$(X_{t+t_0} - X_{t_0})^n = \sum_{l=1}^{n} I_l\left(\tilde{g}_l^{(n)}\right) + C^{(n)}_{t,\sigma},$$

where $\tilde{g}_l^{(n)}$ is the symmetrisation of the function $g_l^{(n)}$ defined by

$$g_l^{(n)}(t_1, x_1, ..., t_l, x_l) = \begin{cases} \sum_{s=l}^{n} \sum_{\theta_n \in \mathcal{K}_{l,s}} x_1^{i_1^{\theta_n}} \cdots x_l^{i_l^{\theta_n}} \Pi^{(n)}_{\theta_n,t,\sigma}, & \text{on } G_l \\ 0 & \text{on } ([0,T] \times \mathbb{R})^l - G_l, \end{cases}$$

where $C_{t,\sigma}^{(n)}$ *and* $\Pi_{\theta_n,t,\sigma}^{(n)}$ *are defined in Definition 3.2.2.*

The following proposition gives a more straightforward representation.

Proposition 4.1.3 *For any pure jump Lévy process* $X = \{X_t, t \geq 0\}$ *satisfying condition (1.7),*

$$\begin{aligned}(X_{t+t_0} - X_{t_0})^n &= \sum_{\theta_n \in \mathcal{I}_n} \int_{t_0}^{t+t_0} \int_{\mathbb{R}} \int_{t_0}^{t_1-} \int_{\mathbb{R}} \cdots \int_{t_0}^{t_{l-1}-} \int_{\mathbb{R}} x_l^{i_l^{\theta_n}} \cdots x_2^{i_2^{\theta_n}} x_1^{i_1^{\theta_n}} \\ &\quad \times \Pi_{\theta_n,t,\sigma}^{(n)} \tilde{N}(\mathrm{d}t_l, \mathrm{d}x_l) \cdots \tilde{N}(\mathrm{d}t_2, \mathrm{d}x_2) \tilde{N}(\mathrm{d}t_1, \mathrm{d}x_1) \\ &\quad + C_{t,\sigma}^{(n)}, \end{aligned} \tag{4.10}$$

where $C_{t,\sigma}^{(n)}$ *and* $\Pi_{\theta_n,t,\sigma}^{(n)}$ *are defined in Definition 3.2.2.*

Proof. This follows directly by replacing $f_{(i_1,\ldots,i_l)}(t_1, \ldots, t_l)$ in (4.6) by (4.9). □

Note that both chaos expansions, that is, the expansion in terms of compensated power jump processes and the expansion in terms of random measure, depend on $\mathcal{I}_n$, $\Pi_{\theta_n,t,\sigma}^{(n)}$ and $C_{t,\sigma}^{(n)}$. From (4.4), we note the relationship between $Y^{(i)}(t)$ and $\tilde{N}(\mathrm{d}s, \mathrm{d}x)$. Because of the simple form of this relationship, we can use Theorem 3.1.4 to derive the explicit representation of (4.10).

4.2 General case

We shall now discuss the general relationship between the two representations. Itô (1956) proved the chaos expansion for general Lévy functionals. In this general case, the stochastic integrals are in terms of both Brownian motion, W, and the compensated Poisson measure, $\tilde{N}(\cdot, \cdot)$. Hence, to unify notation, Benth *et al.* (2003) defined the following notation:

$$\begin{aligned} U_1 = [0, T] \qquad & \text{and} \quad U_2 = [0, T] \times \mathbb{R} \\ \mathrm{d}Q_1(\cdot) = \mathrm{d}W(\cdot) \qquad & \text{and} \quad Q_2(\cdot) = \tilde{N}(\cdot, \cdot) \\ \int_{U_1} g\left(u^{(1)}\right) Q_1\left(\mathrm{d}u^{(1)}\right) &= \int_0^t g(s) W(\mathrm{d}s) \quad \text{and} \\ \int_{U_2} g\left(u^{(2)}\right) Q_2\left(\mathrm{d}u^{(2)}\right) &= \int_0^t \int_{\mathbb{R}} g(s, x) \tilde{N}(\mathrm{d}s, \mathrm{d}x). \end{aligned}$$

The CRP in terms of Brownian motion and Poisson random measures is given by:

Theorem 4.2.1 (Chaos expansion for general Lévy process by Itô (1956)) *Let F be a square integrable random variable adapted to the underlying Lévy process, $X = \{X_t, t \geq 0\}$. We have*

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{j_1,...,j_n=1,2} \tilde{J}_n\left(g_n^{(j_1,\dots,j_n)}\right), \tag{4.11}$$

for a unique sequence $g_n^{(j_1,\dots,j_n)}$ ($j_1,...,j_n = 1,2;\ n = 1,2,...$) of deterministic functions in the corresponding L_2-space, $L_2(G_n)$, where

$$G_n = \left\{\left(u_1^{(j_1)},...,u_n^{(j_n)}\right) \in \Pi_{i=1}^n U_{j_i} : 0 \leq t_1 \leq \cdots \leq t_n \leq T\right\}$$

with $u^{(j_i)} = t$ if $j_i = 1$, and $u^{(j_i)} = (t,x)$ if $j_i = 2$, and

$$\begin{aligned} &\tilde{J}_n\left(g_n^{(j_1,\dots,j_n)}\right) \\ &= \int_{\Pi_{i=1}^n U_{j_i}} g_n^{(j_1,\dots,j_n)}\left(u_1^{(j_1)},...,u_n^{(j_n)}\right) 1_{G_n}\left(u_1^{(j_1)},...,u_n^{(j_n)}\right) Q_{j_1}\left(\mathrm{d}u_1^{(j_1)}\right)\cdots Q_{j_n}\left(\mathrm{d}u_n^{(j_n)}\right). \end{aligned}$$

Similar to the pure jump case, we can derive the explicit formula for the chaos expansion with respect to the Poisson random measure of a general Lévy process, that is, $\sigma \neq 0$. In this case, we have

$$\begin{aligned} Y_t^{(1)} &= \sigma \int_0^t \mathrm{d}W(ds) + \int_0^t \int_{\mathbb{R}} x\tilde{N}(\mathrm{d}s,\mathrm{d}x) \\ Y_t^{(i)} &= \int_0^t \int_{\mathbb{R}} x^i \tilde{N}(\mathrm{d}s,\mathrm{d}x), \quad 0 \leq t \leq T,\ i = 2,3,.... \end{aligned}$$

To derive the relation between the two chaos expansions, we introduce the following notation. Let

$$\begin{aligned} R^{(1)}(\mathrm{d}s,\mathrm{d}x) &= \sigma \mathrm{d}W(\mathrm{d}s) + \int_{\mathbb{R}} x\tilde{N}(\mathrm{d}s,\mathrm{d}x) \\ R^{(i)}(\mathrm{d}s,\mathrm{d}x) &= \int_{\mathbb{R}} x^i \tilde{N}(\mathrm{d}s,\mathrm{d}x), \quad i = 2,3,.... \end{aligned}$$

Hence, similar to (4.5), the CRP with respect to the power jump processes can be written as

$$F = E(F) + \sum_{j=1}^{\infty} \sum_{i_1,...,i_j \geq 1} \int_0^T \int_0^{t_1-} \cdots \int_0^{t_{j-1}-} f_{(i_1,...,i_j)}(t_1,...,t_j) \mathrm{d}Y_{t_j}^{(i_j)} ... \mathrm{d}Y_{t_2}^{(i_2)} \mathrm{d}Y_{t_1}^{(i_1)}$$

$$= E(F) + \sum_{j=1}^{\infty} \sum_{i_1,...,i_j \geq 1} \int_0^T \int_0^{t_1-} \cdots \int_0^{t_{j-1}-} f_{(i_1,...,i_j)}(t_1,...,t_j) \\ R^{(i_j)}(\mathrm{d}t_j, \mathrm{d}x_j) ... R^{(i_2)}(\mathrm{d}t_2, \mathrm{d}x) R^{(i_1)}(\mathrm{d}t_1, \mathrm{d}x).$$

From Theorem 3.2.3,

$$\begin{aligned}(X_{t+t_0} - X_{t_0})^n &= \sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,t,\sigma} \mathcal{S}'_{\theta_n,t,t_0} + C^{(n)}_{t,\sigma} \\ &= \sum_{\theta_n \in \mathcal{I}_n} \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{l-1}-} \Pi^{(n)}_{\theta_n,t,\sigma} \\ &\qquad R^{(i_l^{\theta_n})}(\mathrm{d}t_l, \mathrm{d}x_l) ... R^{(i_2^{\theta_n})}(\mathrm{d}t_2, \mathrm{d}x) R^{(i_1^{\theta_n})}(\mathrm{d}t_1, \mathrm{d}x) + C^{(n)}_{t,\sigma}.\end{aligned}$$

We have then proved the following proposition.

Proposition 4.2.2 *For any Lévy process $X = \{X_t, t \geq 0\}$ satisfying condition (1.7),*

$$\begin{aligned}(X_{t+t_0} - X_{t_0})^n &= \sum_{\theta_n \in \mathcal{I}_n} \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{l-1}-} \Pi^{(n)}_{\theta_n,t,\sigma} \\ &\qquad R^{(i_l^{\theta_n})}(\mathrm{d}t_l, \mathrm{d}x_l) ... R^{(i_2^{\theta_n})}(\mathrm{d}t_2, \mathrm{d}x) R^{(i_1^{\theta_n})}(\mathrm{d}t_1, \mathrm{d}x) + C^{(n)}_{t,\sigma},\end{aligned}$$

where $C^{(n)}_{t,\sigma}$ and $\Pi^{(n)}_{\theta_n,t,\sigma}$ are defined in Definition 3.2.2.

Chapter 5

Discussion and further applications

5.1 The explicit chaos expansions for a common kind of Lévy functionals

Note that we have only found the explicit representations for powers of increments of Lévy processes. In this section, we explain how the explicit formulae for a common kind of Lévy functionals might be obtained using multivariate Taylor expansions.

Assume that a real function g, possessing derivatives of all orders, is such that

$$F = g\left(X_{t_1}, X_{t_2} - X_{t_1}, ..., X_{t_n} - X_{t_{n-1}}\right), \tag{5.1}$$

where the indices $0 \leq t_1 < t_2 < \cdots < t_n$ are known and n is finite. By expressing F in terms of power of increments of X, we can use our explicit formula to obtain the CRP of F. This might seems like a very strong assumption but actually this requirement is frequently met. For example, in financial applications, g might correspond to all pricing functions of contingent claims which depend on the underlying asset at a finite number of time points. Suppose $\{X_t, 0 \leq t \leq T\}$ is the background driving Lévy process and time is now $t = t_n$. Suppose the underlying asset, $\{S_t, 0 \leq t \leq T\}$, is given by the exponential-Lévy model, see Cont & Tankov (2003, Chapter 8.4), $S_t = S_0 \exp(X_t)$, where S_0 is the initial value of the underlying asset at time $t = 0$. Then, for example, we can represent F as the pricing functions of a number of contingent claims listed in Table 5.1.1 (Appendix B.1 gives a detailed description of some of the contingent claims).

Name	**Formula**
Forward and future contracts on a security providing no income	$F_t = S_t \exp\left(r\left(T-t\right)\right) = S_0 \exp\left(X_t + r\left(T-t\right)\right)$, where r is the risk-free interest rate and T is the maturity of the contract.
Forward and future contracts on a security providing a known cash income	$F_t = \left(S_t - I\right)\exp\left(r\left(T-t\right)\right)$ $= \left(S_0 \exp\left(X_t\right) - I\right)\exp\left(r\left(T-t\right)\right)$, where I is the present value of the perfectly predictable income on $S = \{S_t, t \geq 0\}$.
Forward and future contracts on a foreign currency	$F_t = S_t \exp\left(\left(r - r_f\right)\left(T-t\right)\right)$ $= S_0 \exp\left(X_t + \left(r - r_f\right)\left(T-t\right)\right)$, where r_f is the risk-free interest rate of the foreign currency.
Forward and future contracts on commodity	$F_t = \left(S_t + U\right)\exp\left(r\left(T-t\right)\right)$ $= \left(S_0 \exp\left(X_t\right) + U\right)\exp\left(r\left(T-t\right)\right)$, where U is the present value of all storage costs.
European call options	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(S_T - K\right)^+ \lvert \mathcal{F}_t\right]$, where K is the strike, T is the maturity, Q is the risk neutral measure and F_t is the filtration of $S = \{S_t, t \geq 0\}$.
'up-and-out' barrier call options	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(S_T - K\right)^+ 1_{\left\{M_T^S < H\right\}}\right]$, where H is the barrier and $M_t^S = \sup\left\{S_u, 0 \leq u \leq t\right\},\ 0 \leq t \leq T$.
'up-and-in' barrier call options	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(S_T - K\right)^+ 1_{\left\{M_T^S \geq H\right\}}\right]$.
'down-and-out' barrier call options	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(S_T - K\right)^+ 1_{\left\{m_T^S > H\right\}}\right]$, where $m_t^S = \inf\left\{S_u, 0 \leq u \leq t\right\},\ 0 \leq t \leq T$.
'down-and-in' barrier call options	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(S_T - K\right)^+ 1_{\left\{m_T^S \leq H\right\}}\right]$.
Lookback options with a floating strike	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[M_T^S - S_T\right]$.
Lookback options with a fixed strike	$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(M_T^S - K\right)^+\right]$.
Asian call options	$F\left(t, S_t\right) = \frac{\exp(-r(T-t))}{n} E_Q\left[\left(\sum_{k=1}^n S_{t_k} - nK\right)^+ \middle\vert \mathcal{F}_t\right]$.

Table 5.1.1: The contingent claims and their pricing formulae to which Taylor expansions can be applied at some values of S_t.

For an European call option, the option price function before maturity with strike K, maturity T is then given at time t by:

$$F\left(t, S_t\right) = \exp\left(-r\left(T-t\right)\right) E_Q\left[\left(S_T - K\right)^+ \lvert \mathcal{F}_t\right],$$

where Q is the risk-neutral measure and $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ is the natural filtration of $S = \{S_t, t \geq 0\}$.

In (5.1), let $x_1 = X_{t_1}, x_2 = X_{t_2} - X_{t_1}, ..., x_n = X_{t_n} - X_{t_{n-1}}$. If g is not a linear combination of powers of x_i, we need to use the multivariate Taylor's series, see Jeffreys & Jeffreys (1988), about the points $x_i = 0, i = 1, ..., n$ to obtain such a representation:

$$g(x_1, ..., x_n) = \sum_{j=0}^{\infty} \left\{ \frac{1}{j!} \left[\sum_{k=1}^{n} x_k \frac{\partial}{\partial x'_k} \right]^j g(x'_1, ..., x'_n) \right\}_{x'_1=0,...,x'_n=0}. \tag{5.2}$$

Note that this representation exists when g is an analytic function. To show typical elements in this representation, we note the special case of $n = 2$:

$$\begin{aligned} g(x_1, x_2) &= \sum_{j=0}^{\infty} \left\{ \frac{1}{j!} \left[x_1 \frac{\partial}{\partial x'_1} + x_2 \frac{\partial}{\partial x'_2} \right]^j g(x'_1, x'_2) \right\}_{x'_1=0, x'_2=0} \\ &= g(0,0) + \left[x_1 \left. \frac{\partial g}{\partial x'_1} \right|_{x'_1=0, x'_2=0} + x_2 \left. \frac{\partial g}{\partial x'_2} \right|_{x'_1=0, x'_2=0} \right] \\ &\quad + \frac{1}{2!} \left[x_1^2 \left. \frac{\partial^2 g}{\partial x'^2_1} \right|_{x'_1=0, x'_2=0} + 2 x_1 x_2 \left. \frac{\partial^2 g}{\partial x'_1 \partial x'_2} \right|_{x'_1=0, x'_2=0} + x_2^2 \left. \frac{\partial^2 g}{\partial x'^2_2} \right|_{x'_1=0, x'_2=0} \right] \\ &\quad + \cdots . \end{aligned}$$

Let $g^{(l)}_{j_1, j_2, ..., j_l}(\mathbf{0}) = \frac{1}{l!} \left. \frac{\partial^l g}{\partial x'_{j_1} \partial x'_{j_2} \cdots \partial x'_{j_l}} \right|_{x'_1=0,...,x'_n=0}$. As in Corcuera *et al.* (2005, Lemma 2), we assume that

$$\sum_{l=2}^{\infty} \sum_{j_1,...,j_l \in \{1,...,n\}} \left| g^{(l)}_{j_1, j_2, ..., j_l}(\mathbf{0}) \right| R^l < \infty, \tag{5.3}$$

for all $R > 0$. The multivariate Taylor's series in equation (5.4) below expresses F in terms of sum of products of powers of increments of $X = \{X_t, t \geq 0\}$. From Theorem 3.2.4, we can substitute $x_i,\ i = 1, 2, ...$with the sum of iterated integrals with respect to the orthogonalised compensated power jump processes.

For all $F \in L^2(\Omega, \mathcal{F})$ having the form (5.1), let $F = g(x_1, ..., x_n)$ and then we have

$$\begin{aligned} F &= \sum_{j=0}^{\infty} \left\{ \frac{1}{j!} \left[\sum_{k=1}^{n} \left(X_{t_k} - X_{t_{k-1}} \right) \frac{\partial}{\partial x'_k} \right]^j g(x'_1, ..., x'_n) \right\}_{x'_1=0,...,x'_n=0} \\ &= g(0, 0, ..., 0) + \sum_{j=1}^{n} \left(X_{t_j} - X_{t_{j-1}} \right) g^{(1)}_j(\mathbf{0}) + \sum_{j=1}^{n} \left(X_{t_j} - X_{t_{j-1}} \right)^2 g^{(2)}_{j,j}(\mathbf{0}) \end{aligned}$$

$$
\begin{aligned}
&+2\sum_{j_1=1}^{n}\sum_{j_2=1}^{n} 1_{\{j_1\neq j_2\}}\left(X_{t_{j_1}}-X_{t_{j_1-1}}\right)\left(X_{t_{j_2}}-X_{t_{j_2-1}}\right) g^{(2)}_{j_1,j_2}(\mathbf{0})\\
&+\sum_{j=1}^{n}\left(X_{t_j}-X_{t_{j-1}}\right)^3 g^{(3)}_{j,j,j}(\mathbf{0})\\
&+3\sum_{j_1=1}^{n}\sum_{j_2=1}^{n} 1_{\{j_1\neq j_2\}}\left(X_{t_{j_1}}-X_{t_{j_1-1}}\right)^2\left(X_{t_{j_2}}-X_{t_{j_2-1}}\right) g^{(3)}_{j_1,j_1,j_2}(\mathbf{0})\\
&+\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n} 1_{\{j_1\neq j_2\neq j_3\}}\left(X_{t_{j_1}}-X_{t_{j_1-1}}\right)\left(X_{t_{j_2}}-X_{t_{j_2-1}}\right)\\
&\times\left(X_{t_{j_3}}-X_{t_{j_3-1}}\right) g^{(3)}_{j_1,j_2,j_3}(\mathbf{0})+\cdots, \qquad (5.4)
\end{aligned}
$$

where $\left(X_{t_i}-X_{t_{i-1}}\right)^n$'s are given by Theorem 3.2.4 and we assume $X_{t_0}=0$. The sums converge for every $\omega\in\Omega$ because of (5.3).

Since $0\le t_1<t_2<\cdots<t_n$, the product of two iterated integrals with non-overlapping limits results in an iterated integral: if $i\le j-1$, $u,v\in\{1,2,3,...\}$ and ϕ_i,ϕ_j are the predictable integrands,

$$
\begin{aligned}
\int_{t_{i-1}}^{t_i}\phi_i\,\mathrm{d}H^{(u)}_{s_1}\times\int_{t_{j-1}}^{t_j}\phi_j\,\mathrm{d}H^{(v)}_{r_1} &= \int_{t_{j-1}}^{t_j}\int_{t_{i-1}}^{t_i}\phi_i\phi_j\,\mathrm{d}H^{(u)}_{s_1}\mathrm{d}H^{(v)}_{r_1}\\
&= \int_0^{t_j}\int_0^{t_i} 1_{\{s_1>t_{i-1}\}}1_{\{r_1>t_{j-1}\}}\phi_i\phi_j\,\mathrm{d}H^{(u)}_{s_1}\mathrm{d}H^{(v)}_{r_1}\\
&= \int_0^{t_j}\int_0^{r_1} 1_{\{t_i>s_1>t_{i-1}\}}1_{\{r_1>t_{j-1}\}}\phi_i\phi_j\,\mathrm{d}H^{(u)}_{s_1}\mathrm{d}H^{(v)}_{r_1},
\end{aligned}
$$

since $r_1>t_{j-1}\ge t_i$, giving an iterated integral. Hence, we get a chaos expansion of F in terms of iterated integrals with respect to orthogonalised compensated power jump processes.

Note that in some applications, it is only necessary to apply Taylor's theorem directly to F to obtain a PRP representation. Part II of this thesis applied Taylor's theorem directly to obtain the PRP of European and exotic option prices for hedging and the use of the explicit formulae is further discussed.

5.2 Simulations using the explicit formula

To verify the theoretical results given in this part, we simulate the underlying Lévy processes and compare the values of $(X_{t+t_0}-X_{t_0})^n$ with the value given by its chaos expansion. In simulations we apply the stochastic Euler scheme for the stochastic differential equations (SDEs) of general Lévy processes, which is given in Appendix A.7. The rate of convergence of this scheme for Lévy processes was discussed by Protter & Talay (1997). For an introduction to numerical solutions of SDEs, see for example Higham & Kloeden (2002), Higham (2001), Kloeden (2002) and Kloeden & Platen (1999).

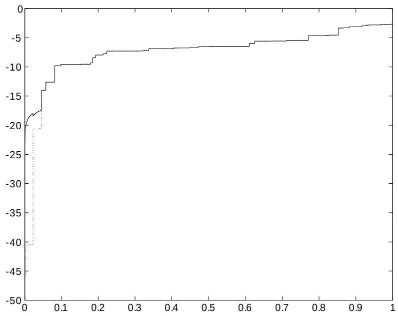

Figure 5.2.1: G_t^4 generated using CRP and directly from the Gamma process in log scale.

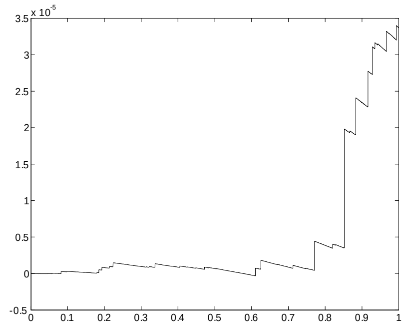

Figure 5.2.2: The difference of the two series in Figure 5.2.1.

For simplicity, we consider Gamma processes as well as a combination of Wiener and Gamma processes. For illustration, we ran simulations for $k = 4$ and $k = 9$ in the pure jump case and $k = 5$ and $k = 8$ for the combined case. The plots produced are shown in Figures A.8.1, A.8.3, A.8.5, A.8.7 in Appendix A.8 respectively. The log scale version of Figure A.8.1 is reproduced as Figure 5.2.1 for illustration. In the second and fourth simulations, we set $t_0 = 0.0099$ and $t_0 = 0.0019$ respectively. These simulations substantiate our explicit formula of the CRP for $t_0 \geq 0$. We see that processes generated using the CRP and those generated directly from the Gamma process jump at the same time points. To see more the two lines more clearly, Figure 5.2.1 is in log scale. Again the two lines are still very close together except in the beginning, where the values are very close to 0 and hence the log of the numbers are very negative. The differences between the two lines are rather due to the numerical rounding errors. The differences between the two lines are plotted in Figures A.8.2, A.8.4, A.8.6, A.8.8 accordingly. Figure A.8.2 is reproduced as Figure 5.2.2 for illustration. Note that the axis of Figures A.8.2, A.8.4, A.8.6, A.8.8 are in much smaller scales than those in Figures A.8.1, A.8.3, A.8.5, A.8.7. In fact, the difference between the two series is so small that we can only see one line in Figure 5.2.1. The difference is due to approximation errors of the stochastic Euler scheme. The errors decrease with the step size Δ. In each of the Figures A.8.1, A.8.3, A.8.5, A.8.7, independent realisations of the Gamma and Wiener processes are used. We note that the line representing the error between the two jumps at the same time points as the Lévy process. Moreover, the jump sizes of the error are proportional to those of the Lévy process. After each jump, the error tends to decrease gradually and then increases again by jumping. It shows that the Euler scheme is more sensitive to jumps in the original process and perform better for Brownian motion which is smooth.

5.3 Discussion on Proposition 8 of Løkka (2004)

Løkka (2004, Proposition 8) derived an explicit expression for the chaos expansion of an infinitely differentiable and square integrable functional of a pure jump Lévy process. For every $n \in \mathbb{N}$ and $m \leq n$, define the sets $\mathcal{A}_m^n$ by

$$\mathcal{A}_m^n = \left\{ (a_1, ..., a_m) \in \{1, ..., n\}^m : a_i < a_{i+1} \ \forall i = 1, ..., m-1 \right\}.$$

Assume the underlying Lévy process has no Brownian part. Let $g \in C_0^\infty\left(\mathbb{R}^k\right)$ be such that $g\left(X_{s_1}, ..., X_{s_k}\right) \in L^2\left(\mathcal{F}_T, P\right)$. Then Løkka (2004, Proposition 8) claimed that

$$g\left(X_{s_1}, ..., X_{s_k}\right) = G\left(0, ...0\right) + \sum_{n=1}^{\infty} I_n\left(f_n\right),$$

where $G\left(x_1, ..., x_k\right) = E\left[g\left(x_1 + X_{s_1}, ..., x_k + X_{s_k}\right)\right]$ is in $C_0^\infty\left(\mathbb{R}^k\right)$, and

$$\begin{aligned} & f_n\left(t_1, ..., t_n, z_1, ..., z_n\right) \\ & = \frac{1}{n!}\left\{ \sum_{m=1}^{n} \sum_{\sigma \in \mathcal{A}_m^n} (-1)^{n-m} G\left(z_{\sigma_1} 1_{[0,s_1]}\left(t_{\sigma_1}\right) + \cdots + z_{\sigma_m} 1_{[0,s_1]}\left(t_{\sigma_m}\right), ..., \right.\right. \\ & \qquad \left.\left. z_{\sigma_1} 1_{[0,s_k]}\left(t_{\sigma_1}\right) + \cdots + z_{\sigma_m} 1_{[0,s_k]}\left(t_{\sigma_m}\right)\right) + (-1)^n G\left(0, ..., 0\right)\right\}. \end{aligned}$$

Note that this approach requires the ability to evaluate

$$\begin{aligned} G\left(x_1, ..., x_k\right) &= E\left[g\left(x_1 + X_{s_1}, ..., x_k + X_{s_k}\right)\right] \\ &= \int_{\mathbb{R}^k} g\left(x_1 + y_1, ..., x_k + y_k\right) \mathrm{d}F_{X_{s_1}, ..., X_{s_k}} \mathrm{d}\mathbf{y}, \end{aligned} \tag{5.5}$$

where $F_{X_{s_1}, ..., X_{s_k}}$ is the distribution function of $X_{s_1}, ..., X_{s_k}$. We cannot use Monte Carlo since we need to express

$$G\left(z_{\sigma_1} 1_{[0,s_1]}\left(t_{\sigma_1}\right) + \cdots + z_{\sigma_m} 1_{[0,s_1]}\left(t_{\sigma_m}\right), ..., z_{\sigma_1} 1_{[0,s_k]}\left(t_{\sigma_1}\right) + \cdots + z_{\sigma_m} 1_{[0,s_k]}\left(t_{\sigma_m}\right)\right) \tag{5.6}$$

in terms of z_{σ_1}, $z_{\sigma_2}, ..., z_{\sigma_m}$, which are the integrating variables in $I_n\left(\cdot\right)$. To use Monte Carlo, the values of $z_{\sigma_1}, ..., z_{\sigma_m}$ have to be known constants. Hence, it is not possible to calculate (5.6) using Monte Carlo. Analytic calculation of (5.5) is therefore required.

Apart from these computational issues, we also want to clarify a result in the paper. In the proof, Løkka stated that 'By Theorem 4, the random variable

$$\exp\left(\int_0^T \int_{\mathbb{R}_0} \mathrm{i}z\xi(y,t)(\mu-\pi)(\mathrm{d}z,\mathrm{d}t) - \int_0^T \int_{\mathbb{R}_0} \left[\mathrm{e}^{\mathrm{i}z\xi(y,t)} - 1 - \mathrm{i}z\xi(y,t)\right]\nu(\mathrm{d}z)\,\mathrm{d}t\right) \quad (5.7)$$

has a chaos expansion given by $1+\sum_{n=1}^{\infty} I_n\left((1/n!)\left(\mathrm{e}^{\mathrm{i}z\xi(y,t)}-1\right)^{\otimes n}\right)$.', where

$$\xi(y,t) = y_1 1_{[0,s_1]}(t) + \cdots + y_k 1_{[0,s_k]}(t).$$

Obviously, $\xi(y,t)$ is not continuous in t since it comprises of indicator functions in t. We can find the derivation of the above chaos expansion in the proof of Theorem 4 of the paper. We notice that the result derived is for random variable defined in equation 6 (page 872) of the paper:

$$\begin{aligned} Y_T &= \exp\left\{\int_0^T \int_{\mathbb{R}_0} h(t)\gamma(z)(\mu-\pi)(\mathrm{d}z,\mathrm{d}t)\right. \\ &\qquad \left. - \int_0^T \int_{\mathbb{R}_0} \left(\mathrm{e}^{h(t)\gamma(z)} - 1 - h(t)\gamma(z)\right)\pi(\mathrm{d}z,\mathrm{d}t)\right\}, \end{aligned} \quad (5.8)$$

where $h \in C([0,T])$. That is, the function $h(t)$ must be continuous in t. However, in (5.7), the corresponding function of t, $\xi(y,t)$, is **not** continuous in t. Therefore, the results derived for Y_T cannot be applied to (5.7). Nonetheless, (5.7) does have the chaotic representation given by $1+\sum_{n=1}^{\infty} I_n\left((1/n!)\left(\mathrm{e}^{\mathrm{i}z\xi(y,t)}-1\right)^{\otimes n}\right)$. It is because in the derivation of the chaotic representation for (5.8), the condition $h \in C([0,T])$ is not needed. Løkka stated on page 874 that (5.8) solves

$$\mathrm{d}Y_t = \int_{\mathbb{R}_0} Y_{t-}\left(\mathrm{e}^{h(t)\gamma(z)} - 1\right)(\mu-\pi)(\mathrm{d}z,\mathrm{d}t)$$

by the Itô formula. Here we discuss the derivation of this result in details. Let Z be a process with stochastic integral

$$\mathrm{d}Z_t = \int_{\mathbb{R}_0} H(t,x)\,\tilde{N}(\mathrm{d}t,\mathrm{d}x),$$

where $H(t,x) \in L^2(\Omega,\mathcal{F},P)$. By the Itô formula for function of integrals with respect to the compensated Poisson measure, see (Applebaum (2004, Theorem 4.4.7)), we have for each $f \in C^2(\mathbb{R})$ with probability 1 that,

$$\begin{aligned} f(Z_t) - f(Z_0) &= \int_0^t \int_{\mathbb{R}_0} \left[f(Z_{s-} + H(s,x)) - f\left(Z_{s_-}\right) \right] \tilde{N}(\mathrm{d}t, \mathrm{d}x) \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \int_{\mathbb{R}_0} \left[f(Z_{s-} + H(s,x)) - f\left(Z_{s_-}\right) \right. \\ &\quad \left. - H(s,x) \frac{\mathrm{d}}{\mathrm{d}Z_{s-}} f(Z_{s-}) \right] \nu(\mathrm{d}x). \end{aligned}$$

Note that the notation used by Løkka are equal to

$$(\mu - \pi)(\mathrm{d}z, \mathrm{d}t) \equiv \tilde{N}(\mathrm{d}t, \mathrm{d}x) \text{ and } \pi(\mathrm{d}z, \mathrm{d}t) \equiv \nu(\mathrm{d}x).$$

Therefore, if we put $Z = Y$ and $f(Y) = \log(Y)$, we have

$$\begin{aligned} f(Y_{s-} + H(s,x)) - f\left(Y_{s_-}\right) &= \log\left(Y_{s-} + Y_{t-}\left(\mathrm{e}^{h(t)\gamma(z)} - 1\right)\right) - \log(Y_{s-}) \\ &= \log(Y_{s-}) + \log\left(1 + \mathrm{e}^{h(t)\gamma(z)} - 1\right) - \log(Y_{s-}) \\ &= h(t)\gamma(z) \end{aligned}$$

and

$$H(s,x) \frac{\mathrm{d}}{\mathrm{d}Y_{s-}} f(Y_{s-}) = \left(\mathrm{e}^{h(t)\gamma(z)} - 1\right).$$

Therefore the result follows. In the derivation, we do not need any condition on $h(t)$ apart from $Y_{t-}\left(\mathrm{e}^{h(t)\gamma(z)} - 1\right) \in L^2(\Omega, \mathcal{F}, P)$. Hence the condition $h \in C([0,T])$ for the chaotic representation of (5.8) is not necessary.

5.4 Lévy measures of the orthogonalised processes $H^{(i)}$'s

In this section, we calculate the Lévy measure of the i-th orthogonalised compensated power jump process of a Lévy process, $H^{(i)} = \left\{H_t^{(i)}, t \geq 0\right\}$, defined in (1.10). To obtain the results for general Lévy processes, we first establish some results for pure jump processes. Let $G = \{G_t, t \geq 0\}$ be a pure jump Lévy process, $G^{(i)} = \left\{G_t^{(i)}, t \geq 0\right\}$ be its i-th power jump process and $\widehat{G}^{(i)} = \left\{\widehat{G}_t^{(i)}, t \geq 0\right\}$ be its i-th compensated power jump process.

Nualart & Schoutens (2000) proved that the orthogonalisation in (1.10) is related with classical orthogonal polynomials with respect to the underlying Lévy process X_t by identifying the polynomials $P(.), Q(.)$ such that $\int_0^\infty P(x)\, Q(x)\, x^2 \nu(\mathrm{d}x) = 0$, where $\nu(\mathrm{d}x)$ is the Lévy measure of X. In the standard Gamma case, Nualart & Schoutens

(2000) considered $\int_0^\infty P(x)Q(x)xe^{-x}\mathrm{d}x$ since $\nu(\mathrm{d}x)$ of $G(1,1)$ is $1_{(x>0)}\frac{e^{-x}}{x}\mathrm{d}x$. For general $Gamma(a,b)$, the Lévy measure is $\nu_G(\mathrm{d}x) = \frac{ae^{-bx}}{x}1_{(x>0)}\mathrm{d}x$. To generalise the orthogonalising procedure, we consider $\int_0^\infty P(x)Q(x)x^2\frac{ae^{-bx}}{x}\mathrm{d}x$, which is equal to zero if $P(.),Q(.)$ are orthogonal with respect to $Gamma(a,b)$. Put $u = bx$, and consider the stochastic integral $\int_0^\infty P\left(\frac{u}{b}\right)Q\left(\frac{u}{b}\right)ue^{-u}\mathrm{d}u = 0$. By Koekeok & Swarttouw (1998), an orthogonalisation of $\left\{1, \frac{x}{b}, \frac{x^2}{b^2}, ...\right\}$ gives the Laguerre polynomials $L_n^{(1)}(x)$. Hence, we can see that orthogonality of the compensated power jump processes in the $Gamma(a,b)$ case is given by $L_n^{(1)}(bx)$, which is independent of the first parameter of the distribution. The coefficients used in the orthogonalisation of $\widehat{G}^{(i)} = \left\{\widehat{G}_t^{(i)}, t \geq 0\right\}$ are independent of time t. The Laguerre polynomial $L_n^{(\alpha)}(x)$ can be expressed as

$$L_n^{(\alpha)}(x) = \frac{1}{n!}\sum_{k=0}^{n}\frac{(-n)_k}{k!}(\alpha+k+1)_{n-k}x^k.$$

The pochhammer symbols $(a)_k$ are defined by

$$(a)_0 = 1, \quad (a)_1 = a \quad \text{and} \quad (a)_k = a(a+1)(a+2)...(a+k-1) \quad \text{for } k = 2,3,.... \tag{5.9}$$

and by Gradshteyn & Ryzhik (1965, 8.971(6)), we have the following recursive relation:

$$(n+1)L_{n+1}^{\alpha}(x) - (2n+\alpha+1-x)L_n^{\alpha}(x) + (n+\alpha)L_{n-1}^{\alpha}(x) = 0 \quad \text{for } n = 1,2,....$$

Apart from using the Laguerre polynomials, we can use the following formula to find the coefficients, $a_{i,j}$, in equation (1.10):

Lemma 5.4.1 *For $i \geq j$, $i,j = 1,2,3,...$,*

$$a_{i,j} = (-1)^{i-j}\binom{i-1}{j-1}\frac{m_{i+1}}{m_{j+1}}$$

when the underlying Lévy process is $Gamma(a,b)$ and we have $m_i = \frac{a(i-1)!}{b^i}$.

Proof. The proof is given in Appendix A.9. The proof does not rely on any properties of the Laguerre polynomials, but it is instead derived from the properties of orthogonalised compensated power jump processes and the Gamma Lévy measure. By using the property of the Laguerre polynomials, see Weisstein (1999b),

$$L_n^{(k)}(x) = \sum_{m=0}^{n}(-1)^m\frac{(n+k)!}{(n-m)!(k+m)!m!}x^m,$$

the result follows immediately. The proof in Appendix A.9 gives an insight to how the $a_{i,j}$ of other Lévy processes can be calculated if the corresponding orthogonal polynomials cannot be recognised as known sets of polynomials. □

The Lévy measure of $G^{(j)}$ with $Gamma(1,1)$ is given in Nualart & Schoutens (2000, p.119). Using the same method we generalise it and also find the Lévy measure of the compensated power jump process $\widehat{G}^{(j)} = \left\{\widehat{G}_t^{(j)}, t \geq 0\right\}$. The method makes use of the exponential formula in Bertoin (1996): Let f be a complex-valued Borel function and $\int_{-\infty}^{\infty} \left|1 - e^{f(x)}\right| \nu(\mathrm{d}x) < \infty$, we have for every $t \geq 0$,

$$E\left(\exp\left\{\sum_{0<s\leq t} f(X_s)\right\}\right) = \exp\left\{t\int_{-\infty}^{\infty}\left(e^{f(x)}-1\right)\nu(\mathrm{d}x)\right\}. \tag{5.10}$$

Proposition 5.4.2 *If the condition in equation (1.7) is satisfied, the Lévy triplet of* $G^{(j)}$ *is given by*

$$\left(\frac{a}{b^j}\int_0^b \exp(-z)\, z^{j-1}\mathrm{d}z, 0, \frac{a\exp\left(-bz^{\frac{1}{j}}\right)}{jz} 1_{(z>0)}\mathrm{d}z\right).$$

Proof. The proof is given in Appendix A.10. □

$\widehat{G}^{(j)} = \left\{\widehat{G}_t^{(j)}, t \geq 0\right\}$ is obtained by subtracting a positive drift from the pure jump process $G^{(j)} = \left\{G_t^{(j)}, t \geq 0\right\}$. Since the drift is deterministic, it is clear that the compensated power jump process $\widehat{G}^{(j)}$ is also a Lévy process using the Lévy-Khintchine formula given in (2.1). The Lévy measure of $\widehat{G}^{(j)}$ is the same as that of $G^{(j)}$ and the additional drift is given by $-m_j t$, where $m_j = \int_{-\infty}^{+\infty} x^j \nu_G(\mathrm{d}x)$ for $i \geq 2$, $m_1 t = E\left[G_t^{(1)}\right]$ and $\nu_G(\mathrm{d}x)$ is the Lévy measure of the original Gamma process. Using the Lévy-Khintchine formula, we can easily show that the Lévy triplet of $\widehat{G}^{(j)}$ is given by

$$\left(-m_j + \frac{a}{b^j}\int_0^b \exp(-z)\, z^{j-1}\mathrm{d}z, 0, \frac{a\exp\left(-bz^{\frac{1}{j}}\right)}{jz} 1_{(z>0)}\mathrm{d}z\right).$$

Recall the i-th orthogonalised compensated power jump process, $H^{(j)} = \left\{H_t^{(j)}, t \geq 0\right\}$, has the form

$$H_t^{(j)} = \widehat{G}_t^{(j)} + a_{j,j-1}\widehat{G}_t^{(j-1)} + a_{j,j-2}\widehat{G}_t^{(j-2)} + ... + a_{j,1}\widehat{G}_t^{(1)}. \tag{5.11}$$

It is obvious that $H^{(j)}$ is also a Lévy process since $\widehat{G}^{(j)}, \widehat{G}^{(j-1)}, ..., \widehat{G}^{(1)}$ are Lévy processes. The equation (5.11) can be represented alternatively as,

$$\begin{aligned} H_t^{(j)} &= -\left[m_j + a_{j,j-1}m_{j-1} + a_{j,j-2}m_{j-2} + ... + a_{j,1}m_1\right] t \\ &\quad + \left[G_t^{(j)} + a_{j,j-1}G_t^{(j-1)} + a_{j,j-2}G_t^{(j-2)} + ... + a_{j,1}G_t^{(1)}\right]. \end{aligned}$$

Hence $H^{(j)}$ can be considered as the weighted sum of the pure jump processes

$$G_t^{(j)} + a_{j,j-1}G_t^{(j-1)} + a_{j,j-2}G_t^{(j-2)} + ... + a_{j,1}G_t^{(1)} \tag{5.12}$$

plus a drift $-\left[m_j + a_{j,j-1}m_{j-1} + a_{j,j-2}m_{j-2} + ... + a_{j,1}m_1\right] t$. Note that the jumps in $H^{(j)}$ can be negative. Therefore,

$$\begin{aligned} E\left[\exp\left(\mathrm{i}\theta H_t^{(j)}\right)\right] &= \exp\left(-\mathrm{i}\theta\left[m_j + a_{j,j-1}m_{j-1} + a_{j,j-2}m_{j-2} + ... + a_{j,1}m_1\right] t\right) \\ &\quad \times E\left\{\exp\left(\mathrm{i}\theta\left[G_t^{(j)} + a_{j,j-1}G_t^{(j-1)} + a_{j,j-2}G_t^{(j-2)} + ... + a_{j,1}G_t^{(1)}\right]\right)\right\}. \end{aligned}$$

We cannot substitute the Lévy measure of $\widehat{G}^{(i)}, i = 1, .., j$ directly into the above formula since we do not know the joint Lévy measure for the correlated processes. Instead, we can try to calculate the characteristic function of $H^{(j)}$ using (5.10) directly. Following Nualart & Schoutens (2000, p. 119), we put

$$f^{(j)}(x) = \mathrm{i}\theta\left\{x^j + a_{j,j-1}x^{j-1} + a_{j,j-2}x^{j-2} + ... + a_{j,1}x\right\},$$

and hence

$$\begin{aligned} & E\left\{\exp\left(\mathrm{i}\theta\left[G_t^{(j)} + a_{j,j-1}G_t^{(j-1)} + a_{j,j-2}G_t^{(j-2)} + ... + a_{j,1}G_t^{(1)}\right]\right)\right\} \\ & = \exp\left\{t\int_0^\infty \left(e^{f^{(j)}(x)} - 1\right)\frac{ae^{-bx}}{x}\mathrm{d}x\right\}. \end{aligned} \tag{5.13}$$

Let $h^{(j)}(x) = f^{(j)}(x)/\mathrm{i}\theta$ and put $z = h^{(j)}(x)$, assuming that $h^{(j)}(x) = z$ has $k \leq j$ number of distinct real roots. There are k possible values of x in terms of z, that is, $x = h_1^{(j)}(z)$ or $x = h_2^{(j)}(z)$ etc. Let $p_1 < p_2 < ... < p_{k-1}$ be the turning points of the function $h^{(j)}(x) - z = 0$ such that $h_1^{(j)}(z) \in (0, p_1)$, $h_k^{(j)}(z) \in (p_{k-1}, \infty)$ and $h_l^{(j)}(z) \in (p_{l-1}, p_l)$ for $l = 2, 3, ..., k-1$. Note that the number of turning points of the function $h^{(j)}(x) - z = 0$ can be greater than $k-1$ but we just consider $k-1$ of them. For convenience, put $p_0 = 0$ and $p_k = \infty$. Let $\alpha, \beta \in \mathbb{R}$ and $\alpha \neq \beta$.

Proposition 5.4.3 *The Lévy measure of $H_t^{(j)}$ for $j \geq 2$ is given by:*

$$\nu_H^{(j)}(\mathrm{d}z) = a\left[\sum_{i=1}^{k} g\left(h^{(j)}(p_{i-1}), h^{(j)}(p_i); z\right) \frac{e^{-bh_i^{(j)}(z)}}{h_i^{(j)}(z)} \mathrm{d}h_i^{(j)}(z)\right] \mathrm{d}z,$$

where

$$g(\alpha, \beta; z) = \begin{cases} 1_{\{\alpha < z < \beta\}} & \text{if } \alpha < \beta \\ -1_{\{\beta < z < \alpha\}} & \text{if } \alpha > \beta \end{cases}.$$

The Lévy triplet of $H_t^{(j)}$ is given by

$$\left(-\left[m_j + a_{j,j-1}m_{j-1} + a_{j,j-2}m_{j-2} + ... + a_{j,1}m_1\right] + \int_{-1}^{1} z\nu_H^{(j)}(\mathrm{d}z), 0, \nu_H^{(j)}(\mathrm{d}z)\right).$$

Proof. Starting from (5.13) and using the above argument, we can rearrange to arrive at the form of the Lévy-Khintchine formula and get the results. The proof is given in Appendix A.11 □

Example We verify this result for $H^{(2)} = \left\{H_t^{(2)}, t \geq 0\right\}$. Using Proposition 5.4.3, we can show that the Lévy measure of $H^{(2)}$ is given by:

$$\begin{aligned} \nu_H^{(2)}(\mathrm{d}z) = a\Bigg[& \frac{e^{-\left[1-\sqrt{1+zb^2}\right]}}{\left[1-\sqrt{1+zb^2}\right]} \frac{b^2}{2} \frac{1}{\sqrt{1+zb^2}} 1_{\left(-\frac{1}{b^2}<z<0\right)} \\ & + \frac{e^{-\left[1+\sqrt{1+zb^2}\right]}}{\left[1+\sqrt{1+zb^2}\right]} \frac{b^2}{2} \frac{1}{\sqrt{1+zb^2}} 1_{\left(-\frac{1}{b^2}<z<\infty\right)} \Bigg] \mathrm{d}z. \end{aligned} \tag{5.14}$$

We show that $\nu_H^{(2)}(\mathrm{d}z)$ is a valid Lévy measure, that is, $\int_{-\infty}^{+\infty} (1 \wedge z^2)\, \nu_H^{(2)}(\mathrm{d}z) < \infty$, in Appendix A.12.

Summary of Part I

Lévy processes were introduced in mathematical finance to improve the performance of some of the financial models which are based on using Brownian motion as the underlying process and to model stylised features observed in financial processes. The derivation of an explicit formula for the CRP has been the focus of considerable study, for previous work, see Léon *et al.* (2002), Benth *et al.* (2003), Løkka (2004) and Eddahbi *et al.* (2005). The immediate result of the CRP is the predictable representation property (PRP), which gives the hedging formulae for contingent claims in the financial market. The CRP expresses the functional of a Lévy process in terms of an infinite sum of stochastic integrals with respect to orthogonalised compensated power jump processes. This provides a clear representation of the structure of the Lévy functional. The chaos expansion explains how the Lévy functional depends on the underlying Lévy process in terms of the power jump processes, which are related to the moment structure of the underlying process. In this part, we derived a computational explicit formula for the construction of the CRP of the powers of increments of Lévy processes in terms of orthogonalised compensated power jump processes and its CRP in terms of Poisson random measures. Jamshidian (2005) extended the CRP in terms of power jump processes to a large class of semimartingales and we showed that our formula is an non-trivial extension of the one given by Jamshidian (2005) in the Lévy case, which is an important subclass of semimartingales. Our explicit formula shows that the integrands of the stochastic integrals in the CRP of the powers of increments of Lévy processes do not depend on the integrating variables nor the starting time. This makes the construction and simulation of the CRP much easier to implement. The coefficients of the CRP depend on the m_i's which represent the moments of the process with respect to its Lévy measure. In this part, we considered only Lévy processes and their compensators are always of the form $m_i t$. Using the same calculation, it is trivial to extend the representation to semimartingales whose stochastic compensators have known representations. The CRP of the pricing functions for some common financial derivatives can be found by expressing the pricing functions in terms of powers of increments of the underlying Lévy process using a Taylor expansion.

Part II

Hedging strategies and minimal variance portfolios for European and exotic options in a Lévy market

In Part I of this thesis, we gave the two versions of the chaotic representation property (CRP) in terms of orthogonalised compensated power jump processes and also in terms of the Poisson random measure. The power jump processes are closely related to the power jump assets, see Corcuera *et al.* (2005), which will be used for perfect hedging in this part. The CRP in terms of the Poisson random measure is used in the derivation of the minimal variance portfolio. The CRP is important as it implies the predictable representation property (PRP), which provides the hedging portfolio for a contingent claim. After Nualart & Schoutens (2000) proved the existence of the CRP and PRP for Lévy processes in terms of orthogonalised compensated power jump processes, Corcuera *et al.* (2005) suggested completing the market by trading in the related power jump assets with the use of the PRP, which is derived from the Itô formula. The trading strategy was expressed in terms of a sum of stochastic integrals with respect to some tradable assets. However, the use of stochastic integrals implies that the hedging period, Δt, and the changes of values of the tradable assets have to be very small in order for the stochastic integrals to be implemented by discrete approximation. In reality this would not be practical and especially if the assets are driven by Lévy processes, we expect the changes in values of the assets to be non-trivial. We get around this problem by deriving hedging strategies for European and exotic options in a Lévy market in terms of Taylor's Theorem such that the change of time and changes of values of the tradable assets can be acknowledged explicitly. Moreover, by expressing the change of value of the contingent claim to be hedged in terms of an expansion with respect to the powers of increments of the underlying stock, we can explicitly consider the terms relating to different moments of the underlying stock and consider hedging these terms separately. In this part, dynamic hedging portfolios are constructed under different market assumptions, such as the existence of power jump assets or moment swaps. Static hedging is implemented in the case of European options or baskets of European options. It is shown that perfect hedging can be achieved. Delta and gamma hedging strategies are extended to higher moment hedging by investing in other traded derivatives depending on the same underlying asset. This development is of practical importance as such other derivatives might be readily available. Moment swaps or power jump assets are not liquidly traded. It is shown how minimal variance portfolios can be used to hedge the higher order terms in a Taylor expansion of the pricing function, investing only in a risk-free bank account, the underlying asset and potentially variance swaps. The numerical algorithms and performance of the hedging strategies are presented, showing the practical utility of the derived results. We derive the hedging portfolio directly from the Taylor expansion and investigate the performance of the hedging strategies. In our simulation study, we use the Variance Gamma (VG) model, which is convenient to use as it is analytically tractable and easy to simulate.

This part is arranged as follows: Chapter 6.1 introduces the hedging instruments used in this part, namely the variance swaps, moment swaps and power jump assets. Chapter 6.2 gives hedging strategies using the approximation formulae obtained from applying Taylor's theorem to the pricing formulae and investing in variance swaps, moment swaps or power jump assets. We extend the delta and gamma hedging strategies to higher moment hedging by investing in some traded derivatives depending on the same underlying asset. Chapter 6.3 demonstrates how to use the minimal variance portfolios derived by Benth *et al.* (2003) to hedge the higher order terms in the Taylor expansion, investing only in a risk-free bank account, the underlying asset and potentially variance swaps. Chapter 6.4 gives the approximation procedure of the hedging strategies and the performance of the hedging strategies implemented on a set of different types of options as illustration of the performance of the proposed method. Some concluding remarks are provided at the end of this part. Proofs and tables are included in Appendix B.

Chapter 6

Perfect hedging strategies

An investment made to specifically reduce or cancel out risk in another investment is called a *hedge*. The strategy designed to minimise the exposure to an unwanted risk in finance is called a *hedging strategy*. Under the Black-Scholes model, the PRP of Brownian motions allows perfect hedging of European options. Unfortunately, the derivation of hedging strategies of options in an incomplete market is not as simple and has been the focus of considerable study in the literature, see for example Carr *et al.* (2001), He *et al.* (2005) and Cont *et al.* (2005). In this thesis, by extending the ideas of Corcuera *et al.* (2005), Schoutens (2005) and Benth *et al.* (2003), we derive and implement some hedging strategies for European and exotic options. Numerical procedures are provided and performance of the hedging strategies is discussed.

The predictable representation property, given in (3.2), is useful in option hedging. For option pricing functions which are infinitely differentiable in the stock price, we can simply apply the Itô formula to obtain such a predictable representation. After Nualart & Schoutens (2000) proved the existence of the CRP and PRP for Lévy processes in terms of orthogonalised compensated power jump processes, Corcuera *et al.* (2005) suggested completing the market by trading in the related power jump assets with the use of the PRP, which is derived from the Itô formula. Assuming power jump assets are traded in the market, Corcuera *et al.* (2005) derived a self-financing replicating portfolio for a contingent claim whose payoff function only depends on the stock price at maturity. Their hedging formula is derived from the Itô formula and given in terms of an infinite sum of stochastic integrals. In this thesis, we use a different approach to determine a self-financing replicating portfolio, which, in some cases, can be used in both static and dynamic hedging with a flexible Δt, where Δt denotes the time change during the hedging period. We will apply Taylor's theorem directly to the option pricing formulae to obtain hedging portfolios. Note that delta and gamma hedging commonly used by traders in the market, discussed in Section 6.2.4, are based on Taylor's theorem, see Hull (2003). In the literature, the results on option hedging using CRP, given in (3.1), has previously focused on the theoretical aspects of the problem, see, for example, Corcuera *et al.* (2005) and Løkka (2004). We aim to investigate the problem from a practical point of view by providing methods to obtain the hedging portfolios explicitly using numerical methods and shall discuss the difficulties encountered. When implementing stochastic processes computationally, it is necessary to discretise the time variable. Hence, it is natural to work directly from Taylor's theorem, which can be considered as a discrete version of Itô formula. As a matter of fact, Taylor's theorem was used to derived the delta and gamma hedges commonly used by traders in the market, given in Section 6.2.4. Our approach can also be applied to barrier options, whose pricing functions are given in Appendix B.1, in the case of dynamic hedging.

In the followings, we shall derive hedging strategies using Taylor's theorem. Firstly, we specify the model of the underlying asset, $S = \{S_t, t \geq 0\}$. Following Corcuera *et al.* (2005, Theorem 3), we assume

$$\frac{\mathrm{d}S_t}{S_{t-}} = b\mathrm{d}t + \mathrm{d}X_t, \tag{6.1}$$

where $X = \{X_t, t \geq 0\}$ is a general Lévy process. For example, in our simulation in Section 8.7, we assume X is a Variance Gamma (VG) process, which will be discussed in more details in Section 6.4.1. Let the risk-free bank account be

$$B_t = \exp(rt), \tag{6.2}$$

where r is the continuously compounded risk-free interest rate. Let $F(t,x)$ be the option pricing function at time $t < T$ and stock price equal to x, where T is the maturity of the option. Let $D_1^i F(t,x)$ be the i-th derivative of $F(t,x)$ with respect to the first variable (time), and $D_2^i F(t,x)$ be the i-th derivative of $F(t,x)$ with respect to the second variable (stock price). Suppose $F(t,x)$ is continuous and infinitely differentiable in the second variable and satisfies $\sup_{x<K, t\leq t_0} \sum_{n=2}^{\infty} |D_2^n F(t,x)| R^n < \infty$ for all $K, R > 0, t_0 > 0$.

Let Δt be the time change during the hedging period and $\Delta S_t = S_{t+\Delta t} - S_t$. Applying Taylor's theorem twice to the option pricing formula, $F(t, S_t)$, we obtain

$$\begin{aligned} & F(t+\Delta t, S_t + \Delta S_t) - F(t, S_t) \\ &= [F(t+\Delta t, S_t+\Delta S_t) - F(t+\Delta t, S_t)] + [F(t+\Delta t, S_t) - F(t, S_t)] \\ &= \sum_{i=1}^{\infty} \frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i + \sum_{i=1}^{\infty} \frac{D_1^i F(t, S_t)}{i!} (\Delta t)^i, \end{aligned} \tag{6.3}$$

which is true as long as the derivatives $D_2^i F(t+\Delta t, S_t)$ and $D_1^i F(t, S_t)$ exist for $i =$

$1, 2, 3, \ldots$. Hence, the change of value of F during time t to $t + \Delta t$ can be hedged by investing in $\frac{D_2^i F(t+\Delta t, S_t)}{i!}$ units of $(\Delta S_t)^i$ and $\frac{D_1^i F(t,S_t)}{i!}$ units of $(\Delta t)^i$ for $i = 1, 2, \ldots$. Note that it is not necessary to apply the multivariate Taylor's theorem since the value of Δt is known at time t. Let $M^{(q)}(t, x)$ be the price of a financial derivative such that $M^{(q)}(0, S_0) = F(0, S_0)$ and

$$M^{(q)}(t + \Delta t, S_t + \Delta S_t) - M^{(q)}(t, S_t) = \sum_{i=1}^{q} \frac{D_2^i F(t + \Delta t, S_t)}{i!} (\Delta S_t)^i + \sum_{i=1}^{\infty} \frac{D_1^i F(t, S_t)}{i!} (\Delta t)^i, \tag{6.4}$$

where q is a positive integer. Therefore, we have

$$\lim_{q \to \infty} M^{(q)}(T, S_T) = F(T, S_T),$$

that is, the value of the financial derivative $M^{(q)}$ is asymptotic to F as q goes to infinity. Our aim is to construct a self-financing hedging portfolio for $M^{(q)}$. Note that the hedging error at time Δt,

$$\begin{aligned}&[F(t + \Delta t, S_t + \Delta S_t) - F(t, S_t)] - \left[M^{(q)}(t + \Delta t, S_t + \Delta S_t) - M^{(q)}(t, S_t)\right] \\ &= \sum_{i=q+1}^{\infty} \frac{D_2^i F(t + \Delta t, S_t)}{i!} (\Delta S_t)^i,\end{aligned}$$

can be approximated using standard techniques in calculating the remainder terms in a Taylor expansion. Let $\mathcal{P}_t^{(i)}$ be the value of a basket of financial derivatives such as the risk-free bank account, the underlying stock, variance swaps, moment swaps, power jump assets or other financial derivatives depending on the same underlying stock such that

$$(\Delta S_t)^i = \Delta \mathcal{P}_t^{(i)} = \mathcal{P}_{t+\Delta t}^{(i)} - \mathcal{P}_t^{(i)} \text{ for } i = 2, 3, \ldots.$$

Note that $\mathcal{P}_t^{(i)}$ is a basket of assets that would not lead to arbitrage opportunities. We will show later how to construct such a basket of tradable assets. Therefore, we have

$$\begin{aligned}M^{(q)}(t + \Delta t, S_t + \Delta S_t) - M^{(q)}(t, S_t) &= \sum_{i=1}^{\infty} \frac{D_1^i F(t, S_t)}{i!} (\Delta t)^i + D_2^1 F(t + \Delta t, S_t) \Delta S_t \\ &\quad + \sum_{i=2}^{q} \frac{D_2^i F(t + \Delta t, S_t)}{i!} \Delta \mathcal{P}_t^{(i)}.\end{aligned} \tag{6.5}$$

The self-financing portfolio to hedge $M^{(q)}(t + \Delta t, S_t + \Delta S_t) - M^{(q)}(t, S_t)$ is then
(i) Invest $\sum_{i=1}^{\infty} D_1^i F(t, S_t) (\Delta t)^i / i! (\exp(r\Delta t) - 1)$ in a risk-less bank account such that

at time $t+\Delta t$, the deposit is worth $\sum_{i=1}^{\infty} D_1^i F\left(t, S_t\right)(\Delta t)^i \exp\left(r\Delta t\right)/i!\left(\exp\left(r\Delta t\right)-1\right)$ and the change of value of the investment is $\sum_{i=1}^{\infty} \frac{D_1^i F(t,S_t)}{i!}\left(\Delta t\right)^i$;
(ii) Invest $D_2^1 F\left(t+\Delta t, S_t\right)$ in the underlying stock;
(iii) Invest $\frac{D_2^i F(t+\Delta t,S_t)}{i!}$ in $\mathcal{P}_t^{(i)}$ for $i=2,3,...,q$.

In real life application, we have to find a reasonable value for q and we discuss methods of choosing q in Section 8.7. Note that the approximation in (6.4) only requires the existence of $D_1^i F\left(t, S_t\right)$ for $i=1,2,3,...$ and $D_2^i F\left(t+\Delta t, S_t\right)$ for $i=1,2,3,...,q$. The value of q determines how many financial derivatives we need to invest in, in order to hedge the option up to a pre-specified level of accuracy. If $q=1$, it is only necessary to hedge the deterministic term $\sum_{i=1}^{\infty} \frac{D_1^i F(t,S_t)}{i!}\left(\Delta t\right)^i$ by investing in a risk-free bank account and the term $D_2^1 F\left(t+\Delta t, S_t\right)\Delta S_t$ by investing in the underlying stock, which is a simple extension to the delta hedging discussed in Section 6.2.4. If $q=2$, we can hedge by investing in a risk-free bank account, the underlying stock and the variance swaps currently traded in the market, which is discussed in Section 6.2.1. If $q\geq 3$, we can consider perfect hedging in three cases: (a) trading in moment swaps, discussed in Section 6.2.2, (b) trading in power jump assets, discussed in Section 6.2.3 and (c) trading in some financial derivatives depending on the same underlying assets, discussed in Section 6.2.5. Note that (a) and (b) are not liquidly traded in the market while (c) might be more readily available. If all of these financial derivatives are not available for trading, we can employ the minimal variance portfolios derived in Section 6.3.

The approximation in (6.4) can be used in both static and dynamic hedging for European options by just changing Δt. The reason why static hedging may not be applicable to exotic options is because if during the hedging period, Δt, the value of the $S_{t+\Delta s}$, where $\Delta s<\Delta t$ is explicitly occurring in the formulae, then this must be used in the calculation of the option price. In this case, we have to apply Taylor's theorem with respect to both $\Delta S_t=\left(S_{t+\Delta t}-S_t\right)$ and $\left(S_{t+\Delta s}-S_t\right)$. In the case of dynamic hedging, we can assume that the minimum time period for a change of value of S to take place is equal to Δt, the hedging period. Although static hedging can only be applied to European options, some exotic options can be decomposed into a basket of European options such that static hedging can still be achieved, see for example Derman *et al.* (1995). In Section 6.4.4, we show the approximation results for both static hedging (Δt equals to 3 months) and dynamic hedging (Δt equals to 5 minutes) for European options and dynamic hedging for barrier options. The advantage of static hedging over dynamic hedging is that in real life, transaction costs and bid-ask spreads of option prices are not negligible. The replicating portfolio is not truly self-financing since extra investment must be made to pay for these additional costs. Hence, it is preferable to hedge statically rather than dynamically as the costs involved will be less and constant rebalancing is not required. In the literature and in practice, it is common to assume that ΔS_t is very small such that the approximation in (6.4) can be truncated without loss of accuracy; this is the main assumption behind the delta and gamma hedges commonly used by traders in the market. However, in real life, the price of every traded asset in the market moves by a tick size, such as 0.5 or 1. After a very short period of time, the price of the traded asset either stays unchanged or moves by a multiple of the tick size. Hence, the assumption of ΔS_t being very small in hedging is not sufficiently accurate. It would not in general be reasonable to assume that ΔS_t is small when modelling S as a process with jumps. Thus, we consider $\Delta S_t \geq 1$ for both static and dynamic hedging in our simulation analysis in Section 6.4.4.

6.1 Hedging instruments

In this section, we consider the use of moment swaps (including variance swaps) and power jump assets in our hedging strategies. Recall in the Black-Scholes world, the PRP is in terms of a stochastic integral with respect to a Brownian motion. Therefore, a contingent claim can be hedged by investing merely in a risk-free bank account and the underlying asset. However, the PRP for Lévy processes involves stochastic integrals with respect to power jump processes, which are related to the higher moments of the underlying Lévy process. In equation (6.4), they are represented through $\frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i$. To hedge these terms, we need to invest in some financial derivatives related to these higher moments. We show how moment swaps introduced by Schoutens (2005) and power jump assets by Corcuera *et al.* (2005) can be used to construct $\mathcal{P}_t^{(i)}$ used in the hedging portfolio given in (6.5).

6.1.1 Variance swaps and moment swaps

Variance swaps, introduced by Demeterfi *et al.* (1999), are commonly traded over-the-counter (OTC) derivatives. Schoutens (2005) generalised variance swaps to moment swaps, which are not liquidly traded in the market. Windcliff *et al.* (2006) gave a detailed discussion on volatility swaps.

There are two common contractual definitions of returns of stock price. Let the sampling points of the contract be $\{s_1, s_2, ..., s_n\}$, where the s's are equally spaced with length Δs. The *actual return* is defined to be

$$R_{\text{actual},i} = \frac{S_{s_{i+1}} - S_{s_i}}{S_{s_i}} \tag{6.6}$$

and the *log return* is defined to be

$$R_{\log,i} = \log\left(\frac{S_{s_{i+1}}}{S_{s_i}}\right). \tag{6.7}$$

The annualised realised variance, $\sigma^2_{\text{realised}}$, is defined by

$$\sigma^2_{\text{realised}} = \frac{1}{\Delta s\,(n-2)}\sum_{i=1}^{n-1} R_i^2,$$

where R_i is either the actual return or log return of the stock price. We can now give the definition of a variance swap, introduced by Demeterfi *et al.* (1999).

Definition 6.1.1 *A variance swap is a forward contract on annualised realised variance, $\sigma^2_{realised}$. Its payoff at expiration is equal to*

$$\left(\sigma^2_{realised} - \sigma^2_{strike}\right) N,$$

where $\sigma^2_{realised}$ is the realised stock variance (quoted in annual terms) over the life of the contract, σ^2_{strike} is the pre-defined delivery price for variance, and N is the notional amount of the swap. The holder of a variance swap at expiration receives N dollars for every point by which the stock's realised variance has exceeded the variance delivery price σ^2_{strike}. The annualised realised variance is calculated based on the pre-specified set of sampling points over the period, $\{s_1, s_2, ..., s_n\}$.

In the case of log return, $R_i = R_{\log,i}$, Schoutens (2005) generalised variance swaps to moment swaps. The annualised realised k-th moment, $M^{(k)}_{\text{realised}}$, is defined by

$$M^{(k)}_{\text{realised}} = \frac{1}{\Delta s\,(n-2)}\sum_{i=1}^{n-1} R_i^k.$$

This definition can be easily extended to the case where $R_i = R_{\text{actual},i}$. We can now give the definition of the k-th moment swap.

Definition 6.1.2 *A k-th moment swap is a forward contract on annualised realised k-th moment, $M^{(k)}_{realised}$. Its payoff at expiration is equal to*

$$\left(M^{(k)}_{realised} - M^{(k)}_{strike}\right) N,$$

where $M^{(k)}_{realised}$ is the realised k-th moment (quoted in annual terms) over the life of the contract, $M^{(k)}_{strike}$ is the pre-defined delivery price for the k-th moment, and N is the notional *amount of the swap. The holder of a k-th moment swap at expiration receives N dollars for every point by which the stock's realised k-th moment has exceeded the k-th moment delivery price* $M^{(k)}_{strike}$. *The annualised realised k-th moment is calculated based on the pre-specified set of sampling points over the period,* $\{s_1, s_2, ..., s_n\}$.

6.1.2 Power jump assets

Corcuera *et al.* (2005) suggested enlarging the Lévy market with power jump assets, where the i-th power jump asset is defined by

$$T_t^{(i)} = \exp(rt)\, Y_t^{(i)}, \quad i \geq 2, \tag{6.8}$$

and $Y_t^{(i)}$ is the compensated power jump process defined in (1.9). The authors derived the dynamic hedging portfolio trading in these assets using the Itô formula. Corcuera *et al.* (2005) noted that the 2nd power jump process is related to the realised variance, see Barndorff-Nielsen & Shephard (2002). However, the 2nd power jump asset is not the same as a variance swap and we consider their usages separately in Section 6.2.

6.2 Hedging strategies

In the last section, we introduce two different kinds of financial derivatives involving higher moments, namely, the moment swaps and the power jump assets. In this section, we explain how to use them to construct the basket of financial derivatives, $\mathcal{P}_t^{(i)}$, in order to hedge the terms in equation (6.4). We also discuss the delta and gamma hedges in the literature and we extend them in order to obtain perfect hedging by trading in certain financial derivatives depending on the same underlying asset, which may be available in the market.

In constructing the hedging portfolio in (6.5), we already showed how to hedge the deterministic term $\sum_{i=1}^{\infty} \frac{D_1^i F(t,S_t)}{i!} (\Delta t)^i$. Here we give a more detailed discussion. Let x be the deterministic change in value of the portfolio over a period of time, where x is some known real number. To hedge x, we invest an amount P in a risk-free bank account with continuous compound interest rate r such that the gain from this investment over time Δt is equal to x:

$$P(\exp(r\Delta t) - 1) = x \quad \Rightarrow \quad P = \frac{x}{(\exp(r\Delta t) - 1)}. \tag{6.9}$$

In other words, to hedge x, we invest $\frac{x}{(\exp(r\Delta t)-1)}$ amount of cash into a risk-free bank account paying a compound interest rate of r. For example, to hedge the term

$$\sum_{i=1}^{\infty} \frac{D_1^i F(t, S_t)}{i!} (\Delta t)^i$$

in equation (6.4), we need to invest

$$\frac{\sum_{i=1}^{\infty} \frac{D_1^i F(t,S_t)}{i!} (\Delta t)^i}{(\exp(r\Delta t) - 1)}$$

in a risk-free bank account. Note that the risk free interest rate, r, is almost always non-zero in real life. If it were zero, $\mathrm{d}S_t = S_{t-}\mathrm{d}X$ under the risk-neutral measure and there would be no drift term. Since an option is a function of S_t, there would be no drift term in the option pricing formula and hence there would not be any deterministic term to hedge.

6.2.1 Hedging with variance swaps

To hedge the term $(\Delta S_t)^2$ in equation (6.4), we construct $\mathcal{P}_t^{(2)}$ which invest in a risk-free bank account and variance swaps. If Δt is negligible compared to ΔS_t, from (6.1), we have

$$(\Delta S_t)^2 = S_t^2 (\Delta X_t)^2 . \tag{6.10}$$

Note that we cannot use the variance swaps using log return, $R_{\log,i}$ defined in (6.7) to hedge. It is because $\left[\log\left(\frac{S_{t+\Delta t}}{S_t}\right)\right]^2 = [\log(1 + \Delta X_t)]^2$ since we assume Δt to be negligible. From (6.10), we need $(\Delta X_t)^2$ rather than $[\log(1 + \Delta X_t)]^2$ to hedge, therefore the variance swaps using log returns are not useful in this case. Even if we use the model $S_{t+\Delta t} = S_t \exp(\Delta X_t)$ such that $\log(S_{t+\Delta t}/S_t) = \Delta X_t$, we then have $(\Delta S_t)^2 = (S_{t+\Delta t} - S_t)^2 = S_t^2 [\exp(\Delta X_t) - 1]^2$, which still can not be hedged by the variance swaps using log returns. Therefore, in our case where we apply Taylor's theorem with respect to ΔS_t, we should invest in the variance swaps using absolute returns, $R_{\text{actual},i}$, as defined in (6.6).

Recall in Section 6.1.1 that there is a set of sampling points, $\{s_1, s_2, ..., s_n\}$, for each contract. We invest in the variance swap at time t where the last two sampling points are equal to t and $t + \Delta t$: $s_{n-1} = t$ and $s_n = t + \Delta t$ and maturity equal to $t + \Delta t$. Note that Δt does not have to be negligible here. At maturity, we receive the payoff $\sigma^2_{\text{realised}} - \sigma^2_{\text{strike}}$, where

$$\sigma^2_{\text{realised}} = \frac{1}{\Delta s (n-2)} \sum_{i=1}^{n-1} \left(\frac{S_{t_{i+1}} - S_{t_i}}{S_{t_i}}\right)^2 = \frac{1}{\Delta s (n-2)} \left[\left(\frac{\Delta S_t}{S_t}\right)^2 + \sum_{i=1}^{n-2} \left(\frac{S_{t_{i+1}} - S_{t_i}}{S_{t_i}}\right)^2\right]$$

and the value of $\sum_{i=1}^{n-2} (S_{t_{i+1}} - S_{t_i})^2 / S_{t_i}^2$ is known as time t. In the following, we give the hedging strategy to hedge the term

$$Q_2 = \frac{D_2^2 F\left(t + \Delta t, S_t\right)}{2} \left(\Delta S_t\right)^2 = C_2 \left(\Delta S_t\right)^2 \tag{6.11}$$

in equation (6.4) by constructing $\mathcal{P}_t^{(2)}$.

Proposition 6.2.1 *To hedge the term Q_2 in equation (6.11) we invest in C_2 units of $\mathcal{P}_t^{(2)}$ at time t, consisting of $\Delta s\,(n-2)\,S_t^2$ units of the variance swap with sampling points $\{..., s_{n-1} = t, s_n = t + \Delta t\}$, maturity $t + \Delta t$, strike σ^2_{strike} and*

$$\frac{S_t^2 \Delta s\,(n-2)}{\left[\exp\left(r\Delta t\right) - 1\right]} \left[\sigma^2_{strike} - \frac{1}{\Delta s\,(n-2)} \sum_{i=1}^{n-2} \left(\frac{S_{t_{i+1}} - S_{t_i}}{S_{t_i}}\right)^2\right] + \frac{P_V \Delta s\,(n-2)\,S_t^2}{\left[\exp\left(r\Delta t\right) - 1\right]}$$

units of cash in a risk-free bank account, where P_V is the price of one unit of the variance swap.

Proof. Let

$$\overline{S}_{n,2} = \frac{1}{\Delta s\,(n-2)} \sum_{i=1}^{n-2} \left(\frac{S_{t_{i+1}} - S_{t_i}}{S_{t_i}}\right)^2 = \frac{1}{\Delta s\,(n-2)} \widetilde{S}_{n,2}. \tag{6.12}$$

The initial investment at time t equals the price of the variance swap plus the deposit into the risk-free bank account, which is equal to

$$C_2 \Delta s\,(n-2)\,S_t^2 P_V \left[1 + \frac{1}{\exp\left(r\Delta t\right) - 1}\right] + \frac{C_2 S_t^2 \Delta s\,(n-2)}{\left[\exp\left(r\Delta t\right) - 1\right]} \left[\sigma^2_{\text{strike}} - \overline{S}_{n,2}\right].$$

At maturity, the portfolio is worth

$$\begin{aligned}
&C_2 S_t^2 \Delta s\,(n-2) \left\{ \frac{\left[\sigma^2_{\text{strike}} - \overline{S}_{n,2}\right] e^{r\Delta t}}{e^{r\Delta t} - 1} + \left[\frac{1}{\Delta s\,(n-2)} \left[\left(\frac{\Delta S_t}{S_t}\right)^2 + \widetilde{S}_{n,2}\right] - \sigma^2_{\text{strike}}\right]\right\} \\
&\quad + C_2 P_V \frac{\Delta s\,(n-2)\,S_t^2 e^{r\Delta t}}{e^{r\Delta t} - 1} \\
&= C_2 \left(\Delta S_t\right)^2 + \frac{C_2 S_t^2 \Delta s\,(n-2)}{\left[e^{r\Delta t} - 1\right]} \left[e^{r\Delta t} - e^{r\Delta t} + 1\right] \left[\sigma^2_{\text{strike}} - \overline{S}_{n,2}\right] \\
&\quad + C_2 P_V \frac{\Delta s\,(n-2)\,S_t^2 e^{r\Delta t}}{e^{r\Delta t} - 1} \\
&= C_2 \left(\Delta S_t\right)^2 + C_2 S_t^2 \Delta s\,(n-2) \left[\sigma^2_{\text{strike}} - \overline{S}_{n,2}\right] / \left[\left(e^{r\Delta t} - 1\right)\right] + C_2 P_V \frac{\Delta s\,(n-2)\,S_t^2 e^{r\Delta t}}{e^{r\Delta t} - 1}.
\end{aligned}$$

Hence, the change of value of the hedging portfolio is equal to

$$C_2\left(\Delta S_t\right)^2+C_2\Delta s\left(n-2\right)S_t^2P_V\left[\frac{e^{r\Delta t}}{e^{r\Delta t}-1}-1-\frac{1}{e^{r\Delta t}-1}\right]=C_2\left(\Delta S_t\right)^2,$$

as desired. □

6.2.2 Hedging with moment swaps

In the last section, we explained how to hedge the term Q_2 in equation (6.11) using variance swaps. The idea can be extended easily to moment swaps to hedge the term

$$Q_i=\frac{D_2^iF\left(t+\Delta t,S_t\right)}{i!}\left(\Delta S_t\right)^i=C_i\left(\Delta S_t\right)^i \tag{6.13}$$

for $i=3,4,5,...$, which can be done by investing in the i-th moment swap at time t with sampling points $s_{n-1}=t$ and $s_n=t+\Delta t$ and maturity equal to $t+\Delta t$. At maturity, we receive the payoff $M_{\text{realised}}^{(i)}-M_{\text{strike}}^{(i)}$, where

$$M_{\text{realised}}^{(i)}=\frac{1}{\Delta s\left(n-2\right)}\left[\left(\frac{\Delta S_t}{S_t}\right)^i+\sum_{i=1}^{n-2}\left(\frac{S_{t_{i+1}}-S_{t_i}}{S_{t_i}}\right)^i\right]=\frac{1}{\Delta s\left(n-2\right)}\left[\left(\frac{\Delta S_t}{S_t}\right)^i+\widetilde{S}_{n,i}\right],$$

and the value of $\widetilde{S}_{n,i}$ is known at time t. In the following, we give the hedging strategy to hedge the term Q_i by constructing $\mathcal{P}_t^{(i)}$.

Proposition 6.2.2 *To hedge the terms Q_i defined in (6.13), we invest in C_i units of $\mathcal{P}_t^{(i)}$ at time t, consisting of $\Delta s\left(n-2\right)S_t^i$ units of the i-th moment swap with sampling points $\{...,s_{n-1}=t,s_n=t+\Delta t\}$, maturity $t+\Delta t$ and strike $M_{strike}^{(i)}$, and*

$$\frac{S_t^i\Delta s\left(n-2\right)}{\left[\exp\left(r\Delta t\right)-1\right]}\left[M_{strike}^{(i)}-\frac{1}{\Delta s\left(n-2\right)}\widetilde{S}_{n,i}\right]+\frac{\Delta s\left(n-2\right)S_t^iP_M}{\left[\exp\left(r\Delta t\right)-1\right]}$$

units of cash in a risk-free bank account where P_M is the price of one unit of the moment swap.

Proof. The initial investment at time t equals the price of the moment swap and the deposit into the risk-free bank account:

$$C_i\Delta s\left(n-2\right)S_t^iP_M\left[1+\frac{1}{e^{r\Delta t}-1}\right]+\frac{C_iS_t^i\Delta s\left(n-2\right)}{e^{r\Delta t}-1}\left[M_{\text{strike}}^{(i)}-\frac{1}{\Delta s\left(n-2\right)}\widetilde{S}_{n,i}\right].$$

At maturity, the portfolio is worth

$$
\begin{aligned}
& C_i S_t^i \Delta s\,(n-2) \left\{ \frac{\left[M_{\text{strike}}^{(i)} - \frac{1}{\Delta s(n-2)} \widetilde{S}_{n,i}\right] e^{r\Delta t}}{e^{r\Delta t}-1} \right. \\
& \quad \left. + \left[\frac{1}{\Delta s\,(n-2)} \left[\left(\frac{\Delta S_t}{S_t}\right)^i + \widetilde{S}_{n,i}\right] - M_{\text{strike}}^{(i)}\right]\right\} + C_i P_M \left[\frac{\Delta s\,(n-2)\, S_t^i e^{r\Delta t}}{e^{r\Delta t}-1}\right] \\
& = C_i \left(\Delta S_t\right)^i + \frac{C_i S_t^i \Delta s\,(n-2)}{e^{r\Delta t}-1} \left[e^{r\Delta t} - e^{r\Delta t} + 1\right] \left[M_{\text{strike}}^{(i)} - \frac{1}{\Delta s\,(n-2)} \widetilde{S}_{n,i}\right] \\
& \quad + C_i P_M \left[\frac{\Delta s\,(n-2)\, S_t^i e^{r\Delta t}}{e^{r\Delta t}-1}\right] \\
& = C_i \left(\Delta S_t\right)^i + C_i S_t^i \Delta s\,(n-2) / \left(e^{r\Delta t}-1\right) \\
& \quad \times \left[M_{\text{strike}}^{(i)} - \frac{1}{\Delta s\,(n-2)} \widetilde{S}_{n,i}\right] + C_i P_M \left[\frac{\Delta s\,(n-2)\, S_t^i e^{r\Delta t}}{e^{r\Delta t}-1}\right].
\end{aligned}
$$

Hence, the change of value of the hedging portfolio is equal to

$$
C_i \left(\Delta S_t\right)^i + C_i \Delta s\,(n-2)\, S_t^i P_M \left[\frac{e^{r\Delta t}}{e^{r\Delta t}-1} - 1 - \frac{1}{e^{r\Delta t}-1}\right] = C_i \left(\Delta S_t\right)^i,
$$

as desired. □

6.2.3 Hedging with power jump processes of higher orders

In the last two sections, we discuss how to hedge $\sum_{i=1}^{q} Q_i$ for $q \geq 2$ using variance swaps and moment swaps. If we allow trading in the power jump assets, discussed in Section 6.1.2, we can hedge using power jump assets instead. Since we assume the underlying is driven by the formula (6.1), the famous Doléans-Dade exponential, see Cont & Tankov (2003, Proposition 8.21), has the solution

$$
S_t = S_0 \exp\left(X_t + \left(b - \frac{\sigma^2}{2}\right) t\right) \prod_{0<s\leq t} \left(1 + \Delta X_s\right) \exp\left(-\Delta X_s\right), \tag{6.14}
$$

where b is defined in (6.1) and σ^2 is the Brownian variance parameter. In the following, we consider the simplified case where there is at most one jump of X between t and $t+\Delta t$, and the general case where there can be infinite number of jumps. Note that the latter case might not be realistic because in reality, we only observe a discrete series of the underlying stock S, while the power jump processes of the Lévy process with infinite activity are not observable. Therefore, it appears to be more practical to consider trading in moment swaps rather than power jump processes. We consider both assets for completeness and theoretical interest.

The simplified case

If Δt is negligible compared to ΔS_t, from (6.1) and assuming there is at most one jump of X between t and $t+\Delta t$. From (6.1) and (6.8), we have

$$\begin{aligned}(\Delta S_t)^i &= S_t^i (\Delta X_t)^i = S_t^i \left[\sum_{0<s\leq t+\Delta t} (\Delta X_s)^i - \sum_{0<s\leq t} (\Delta X_s)^i\right] \\ &= S_t^i \left[X_{t+\Delta t}^{(i)} - X_t^{(i)}\right] = S_t^i \left[Y_{t+\Delta t}^{(i)} - Y_t^{(i)} + m_i \Delta t\right] \\ &= S_t^i \left[\exp(-r(t+\Delta t)) T_{t+\Delta t}^{(i)} - \exp(-rt) T_t^{(i)} + m_i \Delta t\right]. \end{aligned} \tag{6.15}$$

Therefore, we can derive the hedging strategy to hedge the term Q_i by constructing $\mathcal{P}_t^{(i)}$:

Proposition 6.2.3 *If Δt is negligible compared to ΔS_t, to hedge Q_i, we invest in C_i units of $\mathcal{P}_t^{(i)}$, consisting of $S_t^i \exp(-r(t+\Delta t))$ units of $T_t^{(i)}$ and*

$$\left\{\frac{S_t^i \exp(-r(t+\Delta t)) T_t^{(i)}}{\exp(r\Delta t) - 1} + \frac{S_t^i \left[-\exp(-rt) T_t^{(i)} + m_i \Delta t\right]}{\exp(r\Delta t) - 1}\right\}$$

units of cash in a risk-free bank account.

Proof. The proof is included in Appendix B.2. □

If Δt is not negligible compared to ΔS_t, assuming $\sigma = 0$ and there is only one jump of X between times t and $t+\Delta t$ as before, we have from (6.14)

$$\begin{aligned}\Delta S_t &= S_{t+\Delta t} - S_t \\ &= S_t \exp(X_{t+\Delta t} - X_t + b\Delta t)(1+\Delta X_t)\exp(-\Delta X_t) - S_t \\ &= S_t [\exp(b\Delta t)(1+\Delta X_t) - 1]. \end{aligned} \tag{6.16}$$

Note that if $\Delta t \to 0$, $\exp(b\Delta t) \to 1$, we have $\Delta S_t = S_t(\Delta X_t)$, as in the case above. Squaring both sides, we have

$$\begin{aligned}(\Delta S_t)^2 &= S_t^2 [\exp(b\Delta t)(1+\Delta X_t) - 1]^2 \\ &= S_t^2 \left\{\exp(2b\Delta t)(\Delta X_t)^2 + 2\exp(b\Delta t)[\exp(b\Delta t) - 1]\Delta X_t + [\exp(b\Delta t) - 1]^2\right\}. \end{aligned}$$

Substituting ΔX_t by $\left[\frac{\Delta S_t}{S_t}+1\right]\exp(-b\Delta t)-1$ using (6.16), we have

$$\begin{aligned}(\Delta S_t)^2 &= S_t^2\left\{\exp(2b\Delta t)(\Delta X_t)^2+[\exp(b\Delta t)-1]^2\right\}\\&\quad +2S_t^2\exp(b\Delta t)[\exp(b\Delta t)-1]\left\{\left[\frac{\Delta S_t}{S_t}+1\right]\exp(-b\Delta t)-1\right\}\\&= 2S_t[\exp(b\Delta t)-1]\Delta S_t+S_t^2\exp(2b\Delta t)(\Delta X_t)^2-S_t^2[\exp(b\Delta t)-1]^2.\end{aligned}$$

Similarly to (6.15) above,

$$\begin{aligned}(\Delta S_t)^2 &= -S_t^2[\exp(b\Delta t)-1]^2+2S_t[\exp(b\Delta t)-1]\Delta S_t\\&\quad +S_t^2\exp(2b\Delta t)\Big[\exp(-r(t+\Delta t))T_{t+\Delta t}^{(2)}\\&\quad -\exp(-rt)T_t^{(2)}+m_2\Delta t\Big].\end{aligned}\tag{6.17}$$

We can then derive the hedging strategy to hedge the term Q_2 in equation (6.4) by constructing $\mathcal{P}_t^{(2)}$ when Δt is not negligible compared to ΔS_t.

Proposition 6.2.4 *If Δt is not negligible compared to ΔS_t, to hedge the term Q_2, we invest in C_2 units of $\mathcal{P}_t^{(2)}$, consisting of $S_t^2\exp(2b\Delta t)\exp(-r(t+\Delta t))$ units of $T_t^{(2)}$ and*

$$\begin{aligned}&\frac{1}{[\exp(r\Delta t)-1]}\left\{S_t^2\exp(2b\Delta t)\exp(-r(t+\Delta t))T_t^{(2)}-S_t^2[\exp(b\Delta t)-1]^2\right.\\&\quad\left.+2S_t[\exp(b\Delta t)-1]\Delta S_t+S_t^2\exp(2b\Delta t)\left[-\exp(-rt)T_t^{(2)}+m_2\Delta t\right]\right\}\end{aligned}$$

units of cash in a risk-free bank account.

Proof. The proof is included in Appendix B.3. □

To hedge Q_i for $i>2$ if Δt is not negligible compared to ΔS_t, we start from (6.16),

$$\begin{aligned}(\Delta S_t)^i &= S_t^i[\exp(b\Delta t)(1+\Delta X_t)-1]^i\\&= S_t^i\left\{\sum_{j=0}^{i}\binom{i}{j}(-1)^{i-j}\exp(jb\Delta t)\left[1+j\Delta X_t+\sum_{k=2}^{j}\binom{j}{k}(\Delta X_t)^k\right]\right\}.\end{aligned}$$

Substituting ΔX_t by $\left[\frac{\Delta S_t}{S_t}+1\right]\exp(-b\Delta t)-1$ using (6.16), we have

$$(\Delta S_t)^i=S_t^i\left\{\sum_{j=0}^{i}\binom{i}{j}(-1)^{i-j}\exp(jb\Delta t)\left[1+j\left(\left[\frac{\Delta S_t}{S_t}+1\right]\exp(-b\Delta t)-1\right)\right.\right.$$

$$
\begin{aligned}
& \quad + \sum_{k=2}^{j} \binom{j}{k} (\Delta X_t)^k \Bigg] \Bigg\} \\
= & \; S_t^i \sum_{j=0}^{i} \binom{i}{j} (-1)^{i-j} \exp(jb\Delta t) \Bigg\{ 1 + j (\exp(-b\Delta t) - 1) + j \exp(-b\Delta t) \frac{\Delta S_t}{S_t} \\
& \quad + \sum_{k=2}^{j} \binom{j}{k} (\Delta X_t)^k \Bigg\}.
\end{aligned}
$$

Let

$$
c_0^{(i,j)} = S_t^i \binom{i}{j} (-1)^{i-j} \exp(jb\Delta t) \{1 + j(\exp(-b\Delta t) - 1)\} \tag{6.18}
$$

$$
c_1^{(i,j)} = S_t^{i-1} \binom{i}{j} (-1)^{i-j} j \exp((j-1) b\Delta t) \tag{6.19}
$$

$$
c_k^{(i,j)} = S_t^i \binom{i}{j} (-1)^{i-j} \exp(jb\Delta t) \binom{j}{k} \quad \text{for } k = 2, 3, ..., j \ , \tag{6.20}
$$

we have

$$
(\Delta S_t)^i = \sum_{j=0}^{i} \left[c_1^{(i,j)} \Delta S_t + \sum_{k=2}^{j} c_k^{(i,j)} (\Delta X_t)^k + c_0^{(i,j)} \right].
$$

Similar to (6.15) above,

$$
\begin{aligned}
(\Delta S_t)^i = & \sum_{j=0}^{i} \Bigg[c_1^{(i,j)} \Delta S_t + c_0^{(i,j)} \\
& + \sum_{k=2}^{j} c_k^{(i,j)} \left[\exp(-r(t+\Delta t)) T_{t+\Delta t}^{(k)} - \exp(-rt) T_t^{(k)} + m_k \Delta t \right] \Bigg].
\end{aligned}
$$

Therefore, we can derive the hedging strategy to hedge the term Q_i by constructing $\mathcal{P}_t^{(i)}$ when Δt is not negligible compared to ΔS_t.

Proposition 6.2.5 *To hedge Q_i for $i > 2$ if Δt is not negligible compared to ΔS_t, we invest in C_i units of $\mathcal{P}_t^{(i)}$, consisting of $\sum_{j=k}^{i} c_k^{(i,j)} \exp(-r(t+\Delta t))$ units of $T_t^{(k)}$ for $k = 2, 3, ...i$, and*

$$
\begin{aligned}
& \frac{1}{[\exp(r\Delta t) - 1]} \sum_{j=0}^{i} \Bigg\{ \sum_{k=2}^{j} c_k^{(i,j)} \exp(-r(t+\Delta t)) T_t^{(k)} \\
& \quad + c_1^{(i,j)} \Delta S_t + \sum_{k=2}^{j} c_k^{(i,j)} \left[-\exp(-rt) T_t^{(k)} + m_k \Delta t \right] + c_0^{(i,j)} \Bigg\}
\end{aligned}
$$

units of cash in a risk-free bank account, where $c_0^{(i,j)}$, $c_1^{(i,j)}$ *and* $c_k^{(i,j)}$ *are defined in (6.18)-(6.20).*

Proof. The initial investment at time t is

$$C_i\left\{\sum_{k=2}^{i}\sum_{j=k}^{i}c_k^{(i,j)}\exp\left(-r\left(t+\Delta t\right)\right)T_t^{(k)}+\sum_{j=0}^{i}\frac{\sum_{k=2}^{j}c_k^{(i,j)}\exp\left(-r\left(t+\Delta t\right)\right)T_t^{(k)}}{\exp\left(r\Delta t\right)-1}\right.$$
$$\left.+\sum_{j=0}^{i}\frac{\left[c_1^{(i,j)}\Delta S_t+\sum_{k=2}^{j}c_k^{(i,j)}\left[-\exp\left(-rt\right)T_t^{(k)}+m_k\Delta t\right]+c_0^{(i,j)}\right]}{\exp\left(r\Delta t\right)-1}\right\}$$
$$=C_i\sum_{j=0}^{i}\left\{\sum_{k=2}^{j}c_k^{(i,j)}\exp\left(-r\left(t+\Delta t\right)\right)T_t^{(k)}+\frac{\sum_{k=2}^{j}c_k^{(i,j)}\exp\left(-r\left(t+\Delta t\right)\right)T_t^{(k)}}{\exp\left(r\Delta t\right)-1}\right.$$
$$\left.+\frac{\left[c_1^{(i,j)}\Delta S_t+\sum_{k=2}^{j}c_k^{(i,j)}\left[-\exp\left(-rt\right)T_t^{(k)}+m_k\Delta t\right]+c_0^{(i,j)}\right]}{\exp\left(r\Delta t\right)-1}\right\}.$$

At maturity, the portfolio is worth

$$C_i\sum_{j=0}^{i}\left\{\sum_{k=2}^{j}c_k^{(i,j)}\exp\left(-r\left(t+\Delta t\right)\right)T_{t+\Delta t}^{(k)}+\frac{\sum_{k=2}^{j}c_k^{(i,j)}\exp\left(-r\left(t+\Delta t\right)\right)T_t^{(k)}}{\exp\left(r\Delta t\right)-1}\exp\left(r\Delta t\right)\right.$$
$$\left.+\frac{c_1^{(i,j)}\Delta S_t+\sum_{k=2}^{j}c_k^{(i,j)}\left[-\exp\left(-rt\right)T_t^{(k)}+m_k\Delta t\right]+c_0^{(i,j)}}{\exp\left(r\Delta t\right)-1}\exp\left(r\Delta t\right)\right\}.$$

The change of value of the portfolio is

$$C_i\sum_{j=0}^{i}\left\{\sum_{k=2}^{j}c_k^{(i,j)}\left[\exp\left(-r\left(t+\Delta t\right)\right)T_{t+\Delta t}^{(k)}-\exp\left(-rt\right)T_t^{(k)}+m_k\Delta t\right]+c_1^{(i,j)}\Delta S_t+c_0^{(i,j)}\right\},$$

as desired. □

The general case

In the case where there are infinite number of jumps from t to $t+\Delta t$, we need the following results on explicit formulae of CRP proved in Part I.

If Δt is negligible compared to ΔS_t, from (6.1), (6.8) and Theorem 3.2.3, we have

$$\left(\Delta S_t\right)^n=S_t^n\left(\Delta X_t\right)^n=S_t^n\left(X_{t+\Delta t}-X_t\right)^n=S_t^n\left[\sum_{\theta_n\in\mathcal{I}_n}\Pi_{\theta_n,\Delta t,\sigma}^{(n)}\mathcal{S}'_{\theta_n,\Delta t,t}+C_{\Delta t,\sigma}^{(n)}\right].\quad(6.21)$$

In order to hedge $(\Delta S_t)^n$, we can invest in the *power jump integral process*:

$$\mathcal{U}_{(i_1,i_2,\ldots,i_j),\Delta t,t} = \exp(r\Delta t)\, \mathcal{S}'_{(i_1,i_2,\ldots,i_j),\Delta t,t}.$$

Note that since $Y^{(i)}$'s are martingales, $\left\{\mathcal{S}'_{(i_1,i_2,\ldots,i_j),\Delta t,t}, t \geq 0\right\}$'s are also martingales. Therefore, the discounted versions of the $\mathcal{U}_{(i_1,i_2,\ldots,i_j),\Delta t,t}$ are Q-martingales:

$$\begin{aligned} E_Q\left[\exp(-r\Delta t)\, \mathcal{U}_{(i_1,i_2,\ldots,i_j),\Delta t,t} | \mathcal{F}_s\right] &= E_Q\left[\mathcal{S}'_{(i_1,i_2,\ldots,i_j),\Delta t,t} | \mathcal{F}_s\right] \\ &= \mathcal{S}'_{(i_1,i_2,\ldots,i_j),s-t,t}, \quad \text{for } t \leq s \leq t+\Delta t. \end{aligned}$$

Hence the market allowing trade in the bond, the stock and the power jump integral assets remains arbitrage-free. From (6.21), we have

$$(\Delta S_t)^n = S_t^n \left[\sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,\Delta t,\sigma} \exp(-r\Delta t)\, \mathcal{U}_{\theta_n,\Delta t,t} + C^{(n)}_{\Delta t,\sigma}\right]$$

Proposition 6.2.6 *If Δt is negligible compared to ΔS_t, to hedge Q_i, we invest in C_i units of $\mathcal{P}_t^{(i)}$, consisting of $S_t^i \Pi^{(i)}_{\theta_i,\Delta t,\sigma} \exp(-r\Delta t)$ units of $\mathcal{U}_{\theta_i,\Delta t,t}$ for $\theta_i \in \mathcal{I}_i$ and $\frac{S_t^i C^{(i)}_{\Delta t,\sigma}}{(\exp(r\Delta t)-1)}$ units of cash in a risk-free bank account.*

Remark 6.2.1 *In this general case, we can only derive simple hedging strategy when Δt is negligible. Note that both power jump assets introduced by Corcuera* et al. *(2005) and power jump integral assets introduced here are imaginary assets. In reality, we only observe a discrete series of stock price, S, while there are an infinite number of jumps between any finite time interval if the underlying Lévy process has infinite activity. In other words, the values of these assets cannot be observed in the market and hence cannot be traded. The moment swaps introduced by Schoutens (2005) depend on the increment of the underlying stock, ΔS, and can hence be observed and traded in reality. We include the discussion on power jump assets for theoretical interest.*

Alternatively, note that in $\mathcal{S}'_{(i_1,i_2,\ldots,i_j),\Delta t,t}$, the integrand $\int_t^{t_1-} \cdots \int_t^{t_{j-1}-} \mathrm{d}Y^{(i_1)}_{t_j} \cdots \mathrm{d}Y^{(i_{j-1})}_{t_2}$ is a predictable function. Since we assume Δt to be very small, we can hedge $(\Delta S_t)^n$ by investing in the power jump assets. Let $\phi^{(q)}_{j,s}$ be the predictable function such that

$$(\Delta S_t)^n = S_t^n \left[\sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,\Delta t,\sigma} \mathcal{S}'_{\theta_n,\Delta t,t} + C^{(n)}_{\Delta t,\sigma}\right] = \sum_{j=1}^{n} \int_t^{t+\Delta t} \phi^{(n)}_{j,s} \mathrm{d}Y^{(j)}_s + S_t^n C^{(i)}_{\Delta t,\sigma}, \quad (6.22)$$

where $\phi^{(n)}_{j,s}$'s can be calculated by rearranging the terms in $\sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,\Delta t,\sigma} \mathcal{S}'_{\theta_n,\Delta t,t}$'s. We then have

$$\begin{aligned}(\Delta S_t)^n &= \sum_{j=1}^{n}\int_t^{t+\Delta t}\phi_{j,s}^{(n)}\mathrm{d}\left[e^{-rs}T_s^{(j)}\right]+S_t^nC_{\Delta t,\sigma}^{(i)}\\ &= \sum_{j=1}^{n}\int_t^{t+\Delta t}\phi_{j,s}^{(n)}\left[-re^{-rs}T_s^{(j)}\mathrm{d}s+e^{-rs}\mathrm{d}T_s^{(j)}\right]+S_t^nC_{\Delta t,\sigma}^{(i)}\\ &= \int_t^{t+\Delta t}\sum_{j=1}^{n}-e^{-2rs}T_s^{(j)}\phi_{j,s}^{(n)}\mathrm{d}e^{rs}+S_t^nC_{\Delta t,\sigma}^{(i)}+\sum_{j=1}^{n}\int_t^{t+\Delta t}\phi_{j,s}^{(n)}e^{-rs}\mathrm{d}T_s^{(j)}.\end{aligned}$$

Hence, to hedge $(\Delta S_t)^n$, we invest $\sum_{j=1}^{n}-e^{-2r\Delta t}T_{t-}^{(j)}\phi_{j,t-}^{(n)}+\frac{S_t^nC_{\Delta t,\sigma}^{(i)}}{\exp(r\Delta t)-1}$ in a risk-less bank account and invest $\phi_{j,t-}^{(n)}e^{-r\Delta t}$ units of $T_t^{(i)}$ for $j=1,2,...,n$.

6.2.4 Delta and gamma hedges in the literature

So far we have discussed the hedging strategies using moment swaps and power jump assets. In this section, we give a brief introduction to delta and gamma hedging strategies and extend it to obtain perfect hedging in a Lévy market in the next section. Let Π be the value of the portfolio under consideration. The delta and gamma dynamic hedging strategies are constructed using a Taylor expansion:

$$\delta\Pi=\frac{\partial\Pi}{\partial S}\delta S+\frac{\partial\Pi}{\partial t}\delta t+\frac{1}{2}\frac{\partial^2\Pi}{\partial S^2}\delta S^2+\frac{1}{2}\frac{\partial^2\Pi}{\partial t^2}\delta t^2+\frac{\partial^2\Pi}{\partial S\partial t}\delta S\delta t+...,\tag{6.23}$$

where $\delta\Pi$ and δS are the changes in Π and S in a small time interval δt. Hull (2003, Chapter 14) gave detailed descriptions of the strategies in finance. The *delta* of a portfolio is defined as the rate of change of the portfolio with respect to the price of the underlying asset, that is, $\frac{\partial\Pi}{\partial S}$. Delta hedging eliminates the first term on the right-hand side of (6.23). The second term is deterministic. Suppose we write a option with price function Π. In delta hedging, we assume $\frac{1}{2}\frac{\partial^2\Pi}{\partial S^2}\delta S^2+\frac{1}{2}\frac{\partial^2\Pi}{\partial t^2}\delta t^2+\frac{\partial^2\Pi}{\partial S\partial t}\delta S\delta t+...=o(1)$, that is,

$$\delta\Pi=\frac{\partial\Pi}{\partial S}\delta S+\frac{\partial\Pi}{\partial t}\delta t+o(1).$$

Hence, if we sell one unit of Π, we should buy $\frac{\partial\Pi}{\partial S}$ unit of the underlying, so that the change of value of the portfolio is

$$\frac{\partial\Pi}{\partial S}\delta S-\delta\Pi=\frac{\partial\Pi}{\partial t}\delta t+o(1),$$

which is deterministic plus a negligible term. A portfolio with zero delta is said to be *delta-neutral*.

The *gamma* of a portfolio is defined as the rate of change of the portfolio's delta with respect to the price of the underlying. It is the second partial derivative of the portfolio with respect to asset price, that is, $\frac{\partial^2 \Pi}{\partial S^2}$. Since a position in the underlying asset itself or a forward contract on the underlying asset both have zero gamma, they cannot be used to change the gamma of a portfolio. To hedge the gamma risk of an option, we need to trade in an instrument, such as another option, which is not linearly dependent on the underlying asset. Let Γ_1 be the gamma of a delta-neutral portfolio and Γ_2 be the gamma of a traded option. If we add w number of traded options to the portfolio, the gamma of the portfolio becomes

$$w\Gamma_2 + \Gamma_1.$$

Therefore, to make the portfolio gamma neutral, we need $w = -\Gamma_1/\Gamma_2$. Note that including the traded options may change the delta of the portfolio. Hence, the position in the underlying asset has to be changed to maintain delta neutrality.

6.2.5 Extension of delta and gamma hedges

In this section, we extend the gamma hedge in order to obtain a perfect hedging strategy in a Lévy market. Note that equation (6.23) is a multivariate Taylor expansion and it is assumed that all the cross derivative terms are negligible. In equation (6.4), we applied Taylor expansions twice to avoid the cross derivative terms, since the value of Δt is deterministic and known at time t. Hence, for fixed n, the approximation by:

$$F(t + \Delta t, S_t + \Delta S_t) - F(t, S_t) = \sum_{i=1}^{\infty} \frac{D_1^i F(t, S_t)}{i!} (\Delta t)^i + \sum_{i=1}^{n} \frac{D_2^i F(t + \Delta t, S_t)}{i!} (\Delta S_t)^i \tag{6.24}$$

is more accurate than

$$F(t + \Delta t, S_t + \Delta S_t) - F(t, S_t) = \sum_{i=1}^{\infty} \frac{D_1^i F(t, S_t)}{i!} (\Delta t)^i + \sum_{i=1}^{n} \frac{D_2^i F(t, S_t)}{i!} (\Delta S_t)^i .$$

Moreover, in the literature, Δt and ΔS_t are assumed to be very small (such that the cross terms and higher terms are negligible). We provide the flexibility of specifying the values of Δt and ΔS_t such that static hedging is possible in some cases.

It is natural to extend the delta and gamma hedging strategies in the last section to the n-th derivative of the portfolio with respect to the underlying asset using the approximation of equation (6.24). Let F be the value of our portfolio to be hedged and there are $n-1$ traded options, F_i, $i = 2, ..., n$, which are linearly independent of each other. Suppose we add w_i number of F_i into our portfolio, $i = 2, ..., n$ and add w_1 number of the underlying asset, which is denoted by F_1. We assume that $D_2^j F_i(t+\Delta t, S_t)$ are nonzero for $j = i$ and can be zero, or not, for $j = 1, 2, ..., i-1, i+1, ..., n$.

To make the portfolio n-th moment neutral, that is, to make the n-th moment of the portfolio zero, we require

$$D_2^n F(t+\Delta t, S_t) + \sum_{i=1}^{n} w_i D_2^n F_i(t+\Delta t, S_t) = 0.$$

To make the portfolio $(n-1)$-th moment neutral, we require

$$D_2^{n-1} F(t+\Delta t, S_t) + \sum_{i=1}^{n} w_i D_2^{n-1} F_i(t+\Delta t, S_t) = 0.$$

In general, to make the portfolio k-th moment neutral for $k = 1, ..., n$, we need

$$D_2^k F(t+\Delta t, S_t) + \sum_{i=1}^{n} w_i D_2^k F_i(t+\Delta t, S_t) = 0 \quad \text{for } k = 1, 2, ..., n.$$

Therefore, we have n equations for n unknown, w_i's. Note that whether the system of equations is solvable depends on the values of $D_2^k F_i(t+\Delta t, S_t)$, $i, k = 1, 2, ..., n$. Therefore, the traded options have to be chosen such that the system of equations are solvable.

6.3 Minimal variance portfolios in a Lévy market

So far we gave the perfect hedging portfolios, given that the moment swaps, power jump assets and certain financial derivatives that depend on the same underlying asset, are available in the market. In this section, we demonstrate how to use the minimal variance portfolios derived by Benth *et al.* (2003) to hedge the higher order terms in the Taylor expansion, investing only in a risk-free bank account, the underlying asset and, if possible, variance swaps.

6.3.1 Minimal variance portfolio

Benth *et al.* (2003) derived the minimal variance hedging portfolio of a contingent claim in a market such that the stock prices are independent Lévy martingales in terms of Malliavin derivatives. In this section, we gave a modified version of their results and will demonstrate how to use them to hedge the terms Q_i in the next section. Following Benth *et al.* (2003), to derive the minimal variance portfolio, we need to confine ourselves to the case of Lévy processes, $\eta = \{\eta(t), 0 \le t \le T\}$, which are martingales on the filtered probability space under consideration. That is, $E[\eta(t)] = 0$ and $E[\eta^2(t)] = \left(\sigma^2 + \int_{\mathbb{R}} x^2 \nu(\mathrm{d}x)\right) t$. Benth *et al.* (2003) called such processes *Lévy martingales of the second order.* From Benth *et al.* (2003, equation (2.1)), $\eta(t)$ has the following representation formula:

$$\eta(t) = \sigma W(t) + \int_0^t \int_{\mathbb{R}} x\tilde{N}(\mathrm{d}s, \mathrm{d}x), \quad \text{for } 0 \leq t \leq T, \tag{6.25}$$

where $\sigma \in \mathbb{R}^+$, $W(t)$ is the standard Brownian motion and $\tilde{N}(\mathrm{d}t, \mathrm{d}x)$ is the compensated Poisson random measure of the Lévy process η, defined in Definition 2.1.2.

Based on the methodology developed by Benth *et al.* (2003), we modify their results to express the minimal variance portfolio for independent securities without referring to Malliavin calculus. Benth *et al.* (2003) assumed the underlying asset is directly represented by the Lévy martingale, that is, $S_t = \eta(t)$. We find it more natural to employ an exponential model and allow a drift term in the model of the underlying asset since the mean of $\eta(t)$ is zero. By extending (6.1), we suppose there are k independent securities prices $S_1, ..., S_k$, modeled as follows:

$$\mathrm{d}S_j(t) = b_j S_j(t_-)\,\mathrm{d}t + S_j(t_-)\,\mathrm{d}\eta_j(t), \quad j = 1, ..., k, \tag{6.26}$$

where $b_j \in \mathbb{R}$.

A *replicable* or *hedgable claim* $\xi \in L_2(\Omega, \mathcal{F}_T, P)$ is a random variable such that there exists a (predictable) adapted process $\varphi(t) = (\varphi_1(t), ..., \varphi_k(t))$, for $0 \leq t \leq T$ (a portfolio) which replicates (or hedges) ξ, that is,

$$\xi = E[\xi] + \sum_{j=1}^{k} \int_0^T \varphi_j(s)\,\mathrm{d}S_j(s), \tag{6.27}$$

where φ is *admissible, that is:*

$$\sum_{j=1}^{k} E\left[\int_0^T \varphi_j^2(s)\,\mathrm{d}s\right] < \infty.$$

$\varphi(t), 0 \leq t \leq T$, is called a *hedging portfolio* and corresponds to the strategies to buy and sell assets $S_j(t)$ for $j = 1, 2, ..., k$. If every $\xi \in L_2(\Omega, \mathcal{F}_T, P)$ is replicable, the market is said to be *complete* (recall Definition 1.2.1). Let $\mathcal{A}$ be the set of all admissible portfolios. A market driven by Lévy processes is incomplete due to the random jumps of the prices, as discussed in Section 1.2. Recall that in an incomplete market, perfect hedging by investing only in a risk-free bank account and the underlying asset is not in general possible. We would like to find an admissible portfolio φ which can replicate a claim which is closest to the claim ξ. If closeness is measured in terms of variance, we want to find $\varphi \in \mathcal{A}$ such that

$$E\left[\left(\xi - E[\xi] - \sum_{j=1}^{k}\int_0^T \varphi_j(s)\,\mathrm{d}S_j(s)\right)^2\right]$$
$$= \inf_{\psi\in\mathcal{A}} E\left[\left(\xi - E[\xi] - \sum_{j=1}^{k}\int_0^T \psi_j(s)\,\mathrm{d}S_j(s)\right)^2\right]. \tag{6.28}$$

This is known as the *minimal variance hedging* for incomplete markets. The portfolio φ satisfying (6.28) is called the *minimal variance portfolio*.

Let $L_2(\Omega) = L_2(\Omega, \mathcal{F}, P)$ and define a measure of the length of ξ by:

$$\|\xi\| = \left(\int_\Omega |\xi(\omega)|^2 P(\mathrm{d}\omega)\right)^{1/2} = \left(E\left[|\xi|^2\right]\right)^{1/2}.$$

Following Benth *et al.* (2003, Definition 3.10 (a)), let $\mathbb{D}_{1,2}$ be the set of all $\xi \in L_2(\Omega)$ such that the chaos expansion defined in (4.11) satisfies the condition

$$\|\xi\|^2_{\mathbb{D}_{1,2}} = E\left[\xi^2\right] + \sum_{n=1}^{\infty}\sum_{j_1,\ldots,j_n=1,2}\int_{U_{j_n}}\left\|g_n^{(j_1,\ldots,j_n)}\left(\cdot, u_n^{(j_n)}\right)\right\|^2_{L_2(G_{n-1})}\mathrm{d}\left\langle Q_{j_n}\right\rangle\left(u_n^{(j_n)}\right) < \infty,$$

where G_n is defined in (4.2).

The chaotic representation given in (4.11) implies that every ξ satisfying some moment conditions can be expressed in the form

$$\xi = E[\xi] + \sum_{j=1}^{k}\int_0^T f_1(\xi; s, j)\,\mathrm{d}W_j(s) + \sum_{j=1}^{k}\int_0^T\int_{\mathbb{R}} f_2(\xi; s, x, j)\,\tilde{N}_j(\mathrm{d}s, \mathrm{d}x), \tag{6.29}$$

where $f_1(\xi; s, j)$ and $f_2(\xi; s, x, j)$ are predictable functions. Recall that in Part I we derived the computationally explicit representation formula for $f_1(\xi; s, j)$ and $f_2(\xi; s, x, j)$ when ξ is the power of increments of a Lévy process and in Section 5.1 gave the method to obtain (6.29) when ξ is a smooth function with respect to the underlying asset.

The minimal variance portfolio consisting of independent securities driven by (6.26), can be obtained by modifying Theorem 4.1 in Benth *et al.* (2003):

Proposition 6.3.1 *For any $\xi \in \mathbb{D}_{1,2}$, the minimal variance portfolio $\varphi = (\varphi_1, ..., \varphi_k)$ in (6.28),*

$$\widehat{\xi} = E[\xi] + \sum_{j=1}^{k}\int_0^T \varphi_j(s)\,\mathrm{d}S_j(s),$$

admits the following representation:

$$\varphi_j(s) = \frac{f_1(\xi; s, j)\,\sigma_j + \int_{\mathbb{R}} x f_2(\xi; s, x, j)\,\nu_j(\mathrm{d}x)}{\left\{\sigma_j^2 + \int_{\mathbb{R}} x^2 \nu_j(\mathrm{d}x)\right\} S_j(s)},$$

where $f_1(\xi; s, j)$ *and* $f_2(\xi; s, x, j)$ *are predictable functions defined in (6.29).*

Proof. The proof is included in Appendix B.4. □

6.3.2 Hedging with minimal variance portfolios

In Section 6.2, we discuss how to hedge the terms Q_i for $i = 2, 3, 4, ...$ perfectly. In the current market, moment swaps (other than variance swaps) and power jump assets are not liquidly traded. In this section, we discuss how to hedge the terms Q_i using a minimal variance portfolio using only a risk-free bank account and the underlying stock if the variance swaps needed are not available; and the case where the variance swaps can be traded.

Although variance swaps are traded in OTC markets, there might be times that the appropriate variance swaps needed are not available. Hence, we firstly discuss how to use a minimal variance portfolio to hedge $\sum_{i=2}^{q} Q_i$ using only a risk-free bank account and the underlying stock. As in the Section 6.2.3, we consider the simplified case where there is at most one jump of X between t and $t + \Delta t$, and the general case where there can be infinite number of jumps.

The simplified case

If Δt is negligible compared to ΔS_t, from (6.15),

$$\begin{aligned} \sum_{i=2}^{q} Q_i &= \sum_{i=2}^{q} C_i S_t^i \left[Y_{t+\Delta t}^{(i)} - Y_t^{(i)} + m_i \Delta t \right] \\ &= \sum_{i=2}^{q} C_i S_t^i \left[\int_t^{t+\Delta t} \mathrm{d}Y_s^{(i)} + m_i \Delta t \right]. \end{aligned} \tag{6.30}$$

Proposition 6.3.2 *If* Δt *is negligible compared to* ΔS_t*, the minimal variance portfolio to hedge* $\sum_{i=2}^{q} Q_i$ *using only a risk-free bank account and the underlying asset is to*
1) invest

$$\sum_{i=2}^{q} \frac{C_i}{\exp(r\Delta t) - 1} S_t^i m_i \Delta t$$

in a risk-free bank account, and
2) buy

$$\frac{1}{[\sigma^2 + m_2]} \sum_{i=2}^{q} C_i S_t^{i-1} m_{i+1}$$

units of the underlying asset, S_t, where m is defined in (1.8).

Proof. The proof is included in Appendix B.5. □

In the followings, we discuss how to hedge the terms $\sum_{i=3}^{q} Q_i$ using a risk-free bank account, the underlying stock and variance swaps. If Δt is negligible compared to ΔS_t, similar to (6.30),

$$\sum_{i=3}^{q} Q_i = \sum_{i=3}^{q} C_i S_t^i \left[\int_t^{t+\Delta t} \mathrm{d}Y_s^{(i)} + m_i \Delta t \right]. \tag{6.31}$$

Therefore, we have the following hedging portfolio.

Proposition 6.3.3 *If Δt is negligible compared to ΔS_t, the minimal variance portfolio to hedge $\sum_{i=3}^{q} Q_i$ by investing in a risk-free bank account, the underlying asset and variance swaps is given by:*
1) buy

$$\phi \Delta s \,(n-2)\, S_t^2$$

units of the variance swap at time t with sampling points $\{..., s_{n-1} = t, s_n = t + \Delta t\}$, maturity $t + \Delta t$ and strike σ^2_{strike}, where

$$\phi = \frac{\sum_{i=3}^{q} C_i S_t^{i-2} \int_{\mathbb{R}} x^i \nu\,(\mathrm{d}x)}{\int_{\mathbb{R}} x^2 \nu\,(\mathrm{d}x)} = \frac{\sum_{i=3}^{q} C_i S_t^{i-2} m_i}{m_2},$$

P_V is the price of one unit of the variance swap, m_i are defined in (1.8) and for VG process, it is given by Lemma 6.4.1.
2) invest nothing in the underlying asset, S_t,
3) invest

$$\frac{1}{e^{r\Delta t} - 1} \left\{ \sum_{i=3}^{q} C_i S_t^i m_i \Delta t + \phi S_t^2 \left\{ \Delta s\,(n-2) \left[\sigma^2_{strike} - \overline{S}_{n,2} \right] + P_V \Delta s\,(n-2) - m_2 \Delta t \right\} \right\}$$

in a risk-free bank account, where $\overline{S}_{n,2}$ is defined in (6.12).

Proof. The proof is included in Appendix B.6. □

The general case

If Δt is negligible compared to ΔS_t, from (6.21),

$$(\Delta S_t)^n = S_t^n \sum_{\theta_n \in \mathcal{I}_n} \Pi^{(n)}_{\theta_n,\Delta t,\sigma} \mathcal{S}'_{\theta_n,\Delta t,t} + S_t^n C^{(n)}_{\Delta t,\sigma},$$

where the expression can be calculated explicitly using Theorem 3.2.3. Let

$$\begin{aligned}\sum_{i=2}^{q} Q_i &= \sum_{i=2}^{q} C_i S_t^i \left[\sum_{\theta_i \in \mathcal{I}_i} \Pi^{(i)}_{\theta_i,\Delta t,\sigma} \mathcal{S}'_{\theta_i,\Delta t,t} + C^{(i)}_{\Delta t,\sigma} \right] \\ &= \sum_{j=1}^{q} C_i \int_t^{t+\Delta t} \phi^{(q)}_{j,s} \mathrm{d}Y_s^{(j)} + \sum_{i=2}^{q} C_i S_t^i C^{(i)}_{\Delta t,\sigma},\end{aligned}$$

where $\phi^{(q)}_{j,s}$ is defined in (6.22).

Proposition 6.3.4 *If Δt is negligible compared to ΔS_t, the minimal variance portfolio to hedge $\sum_{i=2}^{q} Q_i$ using only a risk-free bank account and the underlying asset is to*
1) invest $\sum_{i=2}^{q} \frac{C_i}{\exp(r\Delta t)-1} S_t^i C^{(i)}_{\Delta t,\sigma}$ in a risk-free bank account, and
2) buy $\frac{1}{[\sigma^2+m_2]} \sum_{j=1}^{q} C_i \phi^{(q)}_{j,s} S_t^{-1} m_{i+1}$ units of the underlying stock, S_t, where m_i is defined in (1.8).

Proof. The proof is similar to the proof of Proposition 6.3.2. □

In the following, we discuss how to hedge the terms $\sum_{i=3}^{q} Q_i$ using a risk-free bank account, the underlying stock and variance swaps.

Proposition 6.3.5 *If Δt is negligible compared to ΔS_t, the minimal variance portfolio to hedge $\sum_{i=3}^{q} Q_i$ by investing in a risk-free bank account, the underlying asset and variance swaps is given by:*
1) buy

$$\phi \Delta s \,(n-2)\, S_t^2$$

units of the variance swap at time t with sampling points $\{..., s_{n-1} = t, s_n = t + \Delta t\}$, maturity $t + \Delta t$ and strike σ^2_{strike}, where

$$\phi = \frac{\sum_{i=1}^{q} C_i \phi^{(q)}_{j,s} S_t^{-2} \int_{\mathbb{R}} x^i \nu\,(\mathrm{d}x)}{\int_{\mathbb{R}} x^2 \nu\,(\mathrm{d}x)} = \frac{\sum_{i=1}^{q} C_i \phi^{(q)}_{j,s} S_t^{-2} m_i}{m_2},$$

P_V is the price of one unit of the variance swap, m_i are defined in (1.8) and for VG process, it is given by Lemma 6.4.1.

2) invest nothing in the underlying asset, S_t,
3) invest

$$\frac{1}{e^{r\Delta t}-1}\left\{\sum_{i=2}^{q} C_i S_t^i C_{\Delta t,\sigma}^{(i)} + \phi S_t^2 \left\{\Delta s\left(n-2\right)\left[\sigma_{strike}^2 - \overline{S}_{n,2}\right] + P_V \Delta s\left(n-2\right) - m_2 \Delta t\right\}\right\}$$

in a risk-free bank account, where $\overline{S}_{n,2}$ is defined in (6.12).

Proof. The proof is similar to the proof of Proposition 6.3.3. □

6.4 Simulation

In this chapter, we discuss the Variance Gamma model, the approximation of the derivatives, $D_2^i F\left(t+\Delta t, S_t\right)$, computational implementation and performance of the hedging strategies.

6.4.1 Variance Gamma model

In the literature, many different kinds of models using Lévy processes have been introduced. Schoutens (2003) for example provided a good review on Lévy market models. We work with the Variance Gamma (VG) process, introduced by Madan *et al.* (1998), in Parts II and III of this thesis because of its simplicity and ability to handle skewness and kurtosis. It is analytically tractable and easy to simulate realisations from. In this section, we give a brief introduction to the VG model.

The VG process is a three-parameter stochastic process which generalises Brownian motion. This process is obtained by evaluating Brownian motion (with constant drift and volatility) at a random time change given by a Gamma process. The extra parameters control the skewness and kurtosis of the return distribution. Let

$$B\left(t;\theta,\sigma\right) = \theta t + \sigma W_t,$$

where $W = \{W_t, t \geq 0\}$ is a standard Brownian motion. The process

$$B = \left\{B\left(t;\theta,\sigma\right), t \geq 0\right\}$$

is a Brownian motion with drift θ and volatility σ. The VG process is defined as

$$X\left(t;\sigma,\nu,\theta\right) = B\left(G\left(t;1,\nu\right);\theta,\sigma\right),$$

where $G = \{G(t;\mu,\nu), t \geq 0\}$ is a Gamma process with mean rate μ and variance rate ν, independent of W. The VG process can be expressed as the difference of two independent increasing Gamma processes as

$$X(t;\sigma,\nu,\theta) = G_p(t;\mu_p,\nu_p) - G_n(t;\mu_n,\nu_n), \tag{6.32}$$

where

$$\mu_p = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2}, \qquad \mu_n = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2},$$
$$\nu_p = \left(\frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2}\right)^2 \nu = \mu_p^2\nu, \qquad \nu_n = \left(\frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2}\right)^2 \nu = \mu_n^2\nu.$$

Equation (6.32) facilitates the simulation of the VG process since Gamma processes are easy to simulate realisations from. The Lévy measure of the VG process is given by

$$\nu(x)\,\mathrm{d}x = \begin{cases} \frac{\mu_n^2}{\nu_n}\frac{\exp\left(-\frac{\mu_n}{\nu_n}|x|\right)}{|x|}\mathrm{d}x & \text{for } x < 0 \\ \frac{\mu_p^2}{\nu_p}\frac{\exp\left(-\frac{\mu_p}{\nu_p}x\right)}{x}\mathrm{d}x & \text{for } x > 0. \end{cases} \tag{6.33}$$

The characteristic function of the VG process, $\phi_X(u,t;\sigma,\nu,\theta) = E[\exp(\mathrm{i}uX(t;\sigma,\nu,\theta))]$, is given by

$$\phi_X(u,t;\sigma,\nu,\theta) = \left(1 - \mathrm{i}\theta\nu u + \frac{\sigma^2\nu}{2}u^2\right)^{-\frac{t}{\nu}}. \tag{6.34}$$

This characteristic function is useful in the derivation of the mean-correcting martingale measure discussed in Section 7.3. Let $S = \{S_t, t \geq 0\}$ be the price of the underlying stock price process at time t. Under the real world measure, we assume that the price is driven by

$$S_t = S_0\exp(mt + X(t;\sigma_S,\nu_S,\theta_S) + \omega_S t), \tag{6.35}$$

where m is the mean rate of return on the underlying under the statistical probability measure, $X(t;\sigma_S,\nu_S,\theta_S)$ is a VG process and $\omega_S = \frac{1}{\nu_S}\ln\left(1 - \theta_S\nu_S - \sigma_S^2\nu_S/2\right)$. Although the density function of the log returns of the underlying was derived in Madan *et al.* (1998), it is quite involving and computationally time demanding. We therefore employ the methods of moments to calibrate the models for the historical time series. Note that under the risk-neutral measure,

$$S_t = S_0\exp(rt + X(t;\sigma_{RN},\nu_{RN},\theta_{RN}) + \omega_{RN}t),$$

where r is the risk-free interest rate and $\omega_{RN} = \frac{1}{\nu_{RN}}\ln\left(1 - \theta_{RN}\nu_{RN} - \sigma_{RN}^2\nu_{RN}/2\right)$. The change of measure is explained in Section 7.3 of Part III of this thesis when we compare the risk-neutral densities implied by historical time series and current option prices.

We have the following lemma to calculate the value of the compensators of the power jump process, $m = \int_{-\infty}^{\infty} x^i \nu(\mathrm{d}x)$, defined in (1.8).

Lemma 6.4.1 *For a VG process,*

$$m_n = \int_{-\infty}^{\infty} x^n \nu(\mathrm{d}x) = (n-1)!\nu^{n-1}\left[(-1)^n \mu_n^n + \mu_p^n\right].$$

Proof. The proof is given in Appendix B.7. □

6.4.2 Central difference approximation of arbitrary degree

In this section, we quote the result by Khan & Ohba (2003, Section 1) on central difference approximation of arbitrary p-th degree derivative of a function. This method is employed in the computational implementation of our hedging strategy later in Section 6.4.3. Khan & Ohba (2003, Section 1) showed that Taylor's series based central difference approximation of arbitrary p-th degree derivative of a function $f(t)$ at $t = t_0$ can be written for an order $2N$ as

$$f_0^{(p)} = \frac{1}{T^p} \sum_{k=-N}^{N} d_k^{(p)} f_k, \tag{6.36}$$

where T is the sampling period, f_k denotes the value of function $f(t)$ at $t = t_0 + kT$, $2N$ is an integer bigger than p and the coefficients $d_k^{(p)}$ are given by

$$d_0^{(p)} = 0 \text{ if } p \text{ is odd, otherwise } d_0^{(p)} = -2\sum_{k=1}^{N} d_k^{(p)}, \tag{6.37}$$

and

$$d_k^{(p)} = (-1)^{k+c_1} \frac{p!}{k^{1+c_2}} C_{N,k} \sum_i \frac{1}{X(i)^2}, \quad \text{for } k = -N, -N+1, ..., -1, 1, ..., N-1, N, \tag{6.38}$$

where

$$\begin{aligned} C_{N,k} &= \frac{N!^2}{(N-k)!\,(N+k)!}, \\ c &= \text{largest integer less than or equal to } (p-1)/2, && (6.39) \\ c_1 &= 1 \text{ if } c \text{ is even, otherwise } c_1 = 0, && (6.40) \\ c_2 &= 1 \text{ if } p \text{ is even, otherwise } c_2 = 0, && (6.41) \end{aligned}$$

and the vector X is generated in the following way:

1. Take a vector Y containing all integers from 1 to N except $|k|$ (in Khan & Ohba (2003, p. 121), it was except k, but from the derivation of the formula, it should be $|k|$).

2. The vector X contains the product of all the possible combinations of length c in Y.

Remark 6.4.1 *Khan & Ohba (2003, Section 2) derived a new finite difference approximation method but we find that the values of $\widehat{d}^{(p)}_{(2k-1)/2}$ (please refer to the paper) are too large when $p \geq 3$ and $k = 0$, which may affect the accuracy of the approximation.*

Assuming that the terms $\sum_{i=2}^{\infty} \frac{D_1^i F(t,S_t)}{i!} (\Delta t)^i$ do not contribute to the approximation significantly and can be ignored (which is found to be true in our simulation study), we have

$$F(t + \Delta t, S_t + \Delta S_t) - F(t, S_t) = D_1^1 F(t, S_t) \Delta t + \sum_{i=1}^{q} \frac{D_2^i F(t + \Delta t, S_t)}{i!} (\Delta S_t)^i ,$$

which is true as long as the derivatives $D_2^i F(t + \Delta t, S_t)$ exist for $i = 1, 2, 3, ...$ and $D_1^1 F(t, S_t)$ exists. Note that the assumption $\sum_{i=2}^{\infty} \frac{D_1^i F(t,S_t)}{i!} (\Delta t)^i \approx 0$ is only for simplicity here since we are more interested in finding ways to hedge the first term of equation (6.3). The deterministic terms $\sum_{i=2}^{\infty} \frac{D_1^i F(t,S_t)}{i!} (\Delta t)^i$ can be hedged by investing in a risk-free bank account, as discussed in the beginning of Section 6.2. Since the pricing formulae for options with underlying driven by Lévy processes are in general not analytic, we need to approximate the derivatives of the pricing formulae, $D_2^i F(t + \Delta t, S_t)$, for $i = 1, 2, 3, ...$ In this section, we discuss the numerical procedures used for this purpose.

We test the approximation in equation (6.36) on the exponential function: $f(x) = \exp(x)$. We obtain $f_0^{(p)}$ and use them in the Taylor expansions

$$\exp(x + \Delta x) - \exp(x) = \sum_{i=0}^{\infty} \frac{f_x^{(p)}}{i!} (\Delta x)^i .$$

At $x = 0$ and $\Delta x = 0.001$, $\exp(0.001) = 1.001$. We take the order of approximation, $2N$, to be 22 and calculate the approximation for $p = 1, 2, 3, ..., 20$. The values of $f_0^{(p)}$ and $\exp(\Delta x)$ by Taylor expansion are given in Appendix B.8. Note that $f_0^{(p)}$should be equal to 1 for all $p = 1, 2, 3, ...$ The errors are due to computational rounding since the term T^p decreases very quickly as p increases. As the values of each $d_k^{(p)} f_k$ are very large while their sum, $\sum_{k=-N}^{N} d_k^{(p)} f_k$, must be equal to T^p. The accrued error from rounding the digits of $d_k^{(p)} f_k$ results in $\sum_{k=-N}^{N} d_k^{(p)} f_k$ not equating to T^p. Although the errors of $f_0^{(p)}$ become large as p increases, Taylor expansions still give the correct approximation because $(\Delta x)^i$ decreases very quickly as i increases.

6.4.3 Computational implementation

In this section, we discuss how to calculate the derivatives of the option prices. We note that the most time consuming step in the approximation procedures is the calculation of $\sum_i \frac{1}{X(i)^2}$ in finding $d_k^{(p)}$ in equation (6.38) in the central difference approximation of derivatives. It is because the vector X contains the product of all the possible combinations of length c in Y, where Y contains all integers from 1 to N except $|k|$. For example, if we want to approximate the 31st derivative and set $N = 33$ (the accuracy of the approximation increases with the value of N), $c = 15$ and $k = 1$, the number of values in Y is 32 and the number of possible combinations of length c in Y is

$$C_{15}^{32} = \frac{32!}{15!\,(32-15)!} = 565,722,720,$$

which takes quite a while to calculate. Nevertheless, this calculation is the same for all functions $f(t)$. Therefore, we can build up a look-up table to store values of $C_{N,k}\sum_i \frac{1}{X(i)^2}$ for different N, c and k and use it for all options. Although the calculation for large N can take a very long time, we only need to do this once.

Step 1 For a fixed N, construct the look-up table of $C_{N,k}\sum_i \frac{1}{X(i)^2}$, where $k =$ 0,1,2,...,N and $c = 3, 4, ..., c_{\max}$, where $c_{\max} = N - 1$ (since $2N > p$ and c is the largest integer less than or equal to $\frac{p-1}{2}$). Therefore, the maximum derivative obtainable is $(2N - 1)$-th.

Algorithm	
1.	Construct the look-up table of $C_{N,k}\sum_i \frac{1}{X(i)^2}$ defined in equation (6.38).
2.	Create a vector of length M of VG random variables, where M is a large positive number.
3.	Produce a $k \times i$ matrix of $Z_\Delta^{(i,k)}$, which is defined in (6.43).
4.	Calculate the sample paths of S with different values of the current stock price, S_t.
5.	Use Monte Carlo simulation to calculate the option prices with respect to different values of the current stock price.
6.	Calculate the derivatives with respect to the underlying, $D_2^i F(t + \Delta t, S_t)$, using the look-up table produced in Step 1.
7.	Calculate the first derivative with respect to time, $D_1^1 F(t, S_t)$.

Table 6.4.3.1: The simulation algorithm to calculate the derivatives in Taylor expansions.

Note that we should loop through c and then k. For each value of c, we use a vector to save the intermediate values of $\sum_i \frac{1}{X(i)^2}$ for each k. Therefore, we only need to calculate the combination of choosing c from Y once for each c.

In our simulation, we assume that the stock price process is driven by equation (6.1), where $X = \{X_t, t \geq 0\}$ is a VG process, using different parameters depending on the type of options used in order to demonstrate the hedging performance. For example, in our simulation if we only use $\theta > 0$, the VG distribution is positive skewed, that is, there are higher chances to have positive jumps than negative jumps. Therefore, the stock price is generally increasing and down-and-in and down-and-out options would perform very similarly to European options. It would then be difficult to distinguish their performance in our study. Therefore, to illustrate a range of behaviour, we choose $\theta < 0$ for these options such that the changes in their prices would differ from those of European options. For similar reasons, $\theta = 0$ is used for European options and $\theta > 0$ is used for up-and-out and up-and-in options. Moreover, the parameters have been chosen such that the simulated price process has similar statistical properties to the FTSE data series. Firstly, we need a function to produce a matrix containing a large number of VG variables. A VG process can be expressed as the difference of two Gamma processes. Note that the probability density function (pdf) of the Gamma variables is in the form

$$f_t\left(g;\mu,\nu\right) = \left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 t}{\nu}} \frac{g^{\frac{\mu^2 t}{\nu}-1}\exp\left(-\frac{\mu}{\nu}g\right)}{\Gamma\left(\frac{\mu^2 t}{\nu}\right)},\ g > 0,$$

with mean μt and νt, while the in-built Gamma pdf of the some computing languages is

$$f\left(g;a,b\right) = \frac{x^{a-1}\exp\left(-\frac{x}{b}\right)}{b^a\Gamma\left(a\right)},$$

with mean ab and variance ab^2. If it is the case, we need to input $a = \frac{\mu^2 t}{\nu}$ and $b = \frac{\nu}{\mu}$.

Step 2 Create a vector of length M of VG random variables, where M is a large positive number. M has to be sufficiently large such that the vector gives a sample of random variables reflecting the VG distribution.

In our simulation, we use $M = 1,000,000$. Recall from (6.14) that we can express $S_{t+\Delta t}$ as:

$$S_{t+\Delta t} = S_t \exp\left(\Delta X_t + b\Delta t\right) \prod_{t<s\leq t+\Delta t} \left(1+\Delta X_s\right)\exp\left(-\Delta X_s\right),$$

where $\Delta X_t = X_{t+\Delta t} - X_t$. In the simulations, we assume X has only one jump between t and $t + \Delta t$. Since all Lévy processes have stationary increments, $X_{t+\Delta t} - X_t$ has the same distribution as $X_{\Delta t}$. Therefore we may write:

$$S_{t+\Delta t} = S_t \exp\left(X_{\Delta t} + b\Delta t\right)\left(1 + X_{\Delta t}\right)\exp\left(-X_{\Delta t}\right).$$

Since we use Monte Carlo simulations to calculate the expectation in the option pricing formulae, we need to create a large number of sample paths of S. Also, to calculate the derivatives of the option prices with respect to the current stock price using the finite difference method given in Section 6.4.2, we need to calculate option prices using different values of the current stock price. Note that the value of $S_{t+i\Delta t}$ in the future is equal to the current value of S times some random variable, that is,

$$S^{(k)}_{t+i\Delta t} = S_t Z^{(i,k)}_{\Delta t}, \tag{6.42}$$

where

$$Z^{(i,k)}_{\Delta t} = \prod_{j=1}^{i} \exp\left(X^{(j,k)}_{\Delta t} + b\Delta t\right)\left(1 + X^{(j,k)}_{\Delta t}\right)\exp\left(-X^{(j,k)}_{\Delta t}\right), \tag{6.43}$$

$X^{(j,k)}_{\Delta t}$ are VG random variables and the superscript k represents it is the k-th sample path. Hence, given that the number of sample paths is big enough to reflect the distribution of the values of $Z^{(i,k)}_{\Delta}$'s, we can use the same set of $Z^{(i,k)}_{\Delta}$'s to calculate the option prices for different current stock prices. Therefore, we need to produce a $k \times i$ matrix of $Z^{(i,k)}_{\Delta}$'s.

Step 3 Produce a $k \times i$ matrix of $Z^{(i,k)}_{\Delta}$, which is given by (6.43), using the VG random variables created in Step 2.

Step 4 Using the matrix of $Z^{(i,k)}_{\Delta}$ obtained in Step 3, calculate the sample paths of S by (6.42) with different values of the current stock price, S_t.

Step 5 Use Monte Carlo simulation to calculate the option prices with respect to different values of the current stock price, using the sample paths of S generated in Step 4.

Step 6 Using the finite different method given in Section 6.4.2, calculate the derivatives with respect to the underlying, $D^i_2 F\left(t + \Delta t, S_t\right)$. This makes use of the look-up table produced in Step 1.

Step 7 Similar to Step 6, calculate the first derivative with respect to time, $D^1_1 F\left(t, S_t\right).$

After calculating the derivatives, we show the performance of the proposed hedging strategies in the next section.

6.4.4 Performance of the hedging strategies

In this section, we investigate the performance of the hedging strategies given in Section 6.2 on European options and barrier options. To improve the accuracy of the approximation given in (6.3), we need to re-scale the values of the underlying stock such that the ΔS_t used in the Taylor expansion is small enough. We discuss this in detail below and compare the approximation results with and without re-scaling. We truncate the infinite sum in (6.3) and calculate $\sum_{i=1}^{p} \frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i + D_1^1 F(t, S_t) \Delta t$ for some fixed p. By comparing the values on the L.H.S. and R.H.S. of (6.3), it may be noted that for some $q \in \mathbb{N}$, the terms $\frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i \simeq 0$ for $i > q$. This approximation is very useful, since in practice it is ideal to hedge by investing in as few kinds of products as possible, due to cost of transaction and administration. By fixing a tolerance level, α_{tol}, we can find the smallest value of p such that

$$\left| [F(t+\Delta t, S_t + \Delta S_t) - F(t, S_t)] - \left[D_1^1 F(t, S_t) \Delta t + \sum_{i=1}^{p} \frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i \right] \right| \leq \alpha_{\text{tol}} \tag{6.44}$$

and we call it q. In real applications, this would be implemented on simulated data from a model calibrated on historical data or option data. Effects from nonstationarity are not considered. In order words, for a given tolerance level, α_{tol}, the following approximation is then assumed satisfactory:

$$F(t+\Delta t, S_t + \Delta S_t) - F(t, S_t) = D_1^1 F(t, S_t) \Delta t + \sum_{i=1}^{q} \frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i . \tag{6.45}$$

Thus the magnitude of α_{tol} determines the number of terms required for a Taylor expansion to obtain a satisfactory approximation. In option hedging, we want the number of terms to be as small as possible since we have to invest in an additional financial derivative to hedge each term. In practice as we noted before, transaction costs, bid-ask spreads and the cost of administration make the trades of a large number of different financial derivatives not preferable. Therefore, there is a trade-off between the accuracy of the hedging and the additional costs involved.

Re-scaling of the stock price

We use the finite difference method given in Section 6.4.2 to find the i-th derivatives with respect to the stock price, that is, $D_2^i F(t+\Delta t, S_t)$, in the approximation provided by equation (6.45). The accuracy of the finite difference method increases as the step size of

S decreases and the closer the step size is to ΔS_t, the smaller the q, which is defined in equation (6.44) as the number of terms required in the Taylor approximation to achieve pre-specified accuracy, will be. Let ϵ be a small positive number such that if the step size of S_t is smaller than ϵ, the accuracy of the finite different method is sufficient. From our simulation analysis, we found $\epsilon = 0.001$ to be sufficient. This assessment was based on investigating the performance of the hedging strategies with ϵ smaller than 0.001, and determining that no noticeable improvement to the calculations was observed. We need to re-scale the stock price process such that

$$\Delta S'_t = \frac{\Delta S_t}{M} = \epsilon,$$

where M is some positive number. We show in Tables 6.4.4.1 and 6.4.4.2 the approximation results for the European options (using the algorithms given in Section 6.4.3, $\Delta S_t = 2$, $\Delta t = 0.25$ and the step size used in the finite difference approximation equals 0.001) with and without re-scaling. The second column gives the p-th derivative, $D_2^p F(t + \Delta t, S_t)$, and the third column gives the approximation by Taylor expansion up to $i = p : D_1^1 F(t, S_t + \Delta S_t)(\Delta t) + \sum_{i=1}^{p} \frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i$. By direct calculation,

$$F(t + \Delta t, S_t + \Delta S_t) - F(t, S_t) = -158.696807.$$

By Taylor expansion,

$$F(t + \Delta t, S_t + \Delta S_t) - F(t, S_t) = D_1^1 F(t, S_t + \Delta S_t)(\Delta t) + \sum_{i=1}^{p} \frac{D_2^i F(t + \Delta t, S_t)}{i!} (\Delta S_t)^i .$$

$$D_1^1 F(t, S_t + \Delta S_t) = -642.787229 \text{ and } D_1^1 F(t, S_t + \Delta S_t)(\Delta t) = -160.696807.$$

p	pth derivative	Up to i=p	p	pth derivative	Up to i=p	p	pth derivative	Up to i=p
1	0.5	-159.6968	8	-4.93E+18	-1.26E+21	15	6.12E+25	1.00E+37
2	681.912643	2567.954	9	-1.06E+12	-1.26E+21	16	-1.56E+37	-1.02E+42
3	0.000207	2567.955	10	3.59E+23	3.67E+26	17	-1.23E+30	-1.02E+42
4	-2.20E+08	-3.53E+09	11	5.91E+16	3.67E+26	18	2.97E+41	7.79E+46
5	-78.004903	-3.53E+09	12	-1.76E+28	-7.22E+31	19	1.86E+34	7.79E+46
6	4.31E+13	2.76E+15	13	-2.25E+21	-7.22E+31	20	-4.32E+45	-4.53E+51
7	1.22E+07	2.76E+15	14	6.13E+32	1.00E+37			

Table 6.4.4.1: The approximation results of European options without re-scaling the stock price.

p	pth derivative	Up to i=p	p	pth derivative	Up to i=p	p	pth derivative	Up to i=p
1	0.5	-159.696807	8	-3.86E-05	-158.697491	15	6.00E-24	-158.696807
2	0.340956	-158.332982	9	-2.10E-18	-158.697491	16	-4.76E-13	-158.696807
3	4.91E-15	-158.332982	10	7.01E-07	-158.696773	17	-1.21E-26	-158.696807
4	-2.76E-02	-158.773808	11	4.83E-20	-158.696773	18	2.27E-15	-158.696807
5	-9.66E-16	-158.773808	12	-8.61E-09	-158.696809	19	6.94E-29	-158.696807
6	1.35E-03	-158.687621	13	-4.11E-22	-158.696809	20	-8.24E-18	-158.696807
7	7.01E-17	-158.687621	14	7.49E-11	-158.696807			

Table 6.4.4.2: The approximation results of European options with re-scaling the stock

price and keeping M varied.

Clearly, the approximation of some higher order derivatives are incorrect without re-scaling. With $\alpha_{\text{tol}} = 0.01$ (defined in (6.44)), $q = 6$ with re-scaling. Without re-scaling, the approximation error is always bigger than α_{tol} when $q \leq 20$.

In real applications, we do not know the value of ΔS_t in advance. Let $\Delta S_{t\,\text{min}}$ be the smallest possible change of stock price from times t to $t + \Delta t$, that is, the tick size of the stock. We should fix M such that

$$\frac{\Delta S_{t\,\text{min}}}{M} = \epsilon.$$

Assume $\Delta S_{t\,\text{min}} = 1$ and $\epsilon = 0.001$, then we should fix $M = 1000$. Hence, for $\Delta S_t = 2$, we have $\Delta S'_t = \frac{2}{1000} = 0.002$. We show the approximation results in Table 6.4.4.3.

p	pth derivative	Up to i=p	p	pth derivative	Up to i=p	p	pth derivative	Up to i=p
1	0.5	-159.696807	8	-4.93E-03	-159.001101	15	-4.28E-21	-158.695859
2	0.681913	-156.969157	9	9.75E-16	-159.001101	16	-1.56E-08	-158.696881
3	-3.68E-13	-156.969157	10	3.59E-04	-158.633662	17	1.38E-21	-158.696881
4	-2.20E-01	-160.495764	11	-3.92E-17	-158.633662	18	2.97E-10	-158.696803
5	9.58E-14	-160.495764	12	-1.76E-05	-158.705907	19	-1.64E-23	-158.696803
6	4.31E-02	-157.737793	13	1.88E-18	-158.705907	20	-4.32E-12	-158.696808
7	-1.23E-14	-157.737793	14	6.13E-07	-158.695859			

Table 6.4.4.3: The approximation results of European options with re-scaling the stock price and keeping M fixed.

With M fixed, $q = 14$. The value of q is much bigger than before, when $\Delta S'_t$ took the value 0.001. This is due to the fact that when the step size of the finite difference method is closer to $\Delta S'_t$, the approximation is more efficient.

Results on performance test

In this section, we give the performance of the static and dynamic hedging strategies on European, up-and-out, up-and-in, down-and-out and down-and-in options. We investigate how many terms in the Taylor expansions are needed to obtain a satisfactory approximation, that is, we determine the value of q for a given α_{tol}, which are defined in the beginning Section 8.7. In our simulations, we set $\alpha_{\text{tol}} = 0.01$. It is because in practice, we are hedging the prices of the options, the lowest price change is 0.01. We assume the current stock price, S_0, is 5000 and the strike price of the options, K, are 5000. Note that our strategies work for all values of K. We consider the cases where the change in the price of the stock price ΔS_t is equal to 1, 2, ..., 10. For static hedging, we assume $\Delta t = 0.25$, approximately 3 months, and the options are expiring in 3 months, that is, $T = 0.25$. For dynamic hedging, we set $\Delta t = 9.5129 \times 10^{-6}$, approximately 5 minutes, and $T = 1.1416 \times 10^{-4}$, approximately 1 hour. The reason why we do not consider very small ΔS_t is that, in practice, most of the assets traded in the market have a tick size of 0.5 or 1. After a very short period of time, ΔS_t can either be zero or a multiple of the tick size. It is not possible for $|\Delta S_t|$ to be smaller than the tick size if it is nonzero. The assumption of ΔS_t being very small when Δt is small would lead to fast convergence of the Taylor expansion, but it is not a realistic assumption, especially in a Lévy market with jumps.

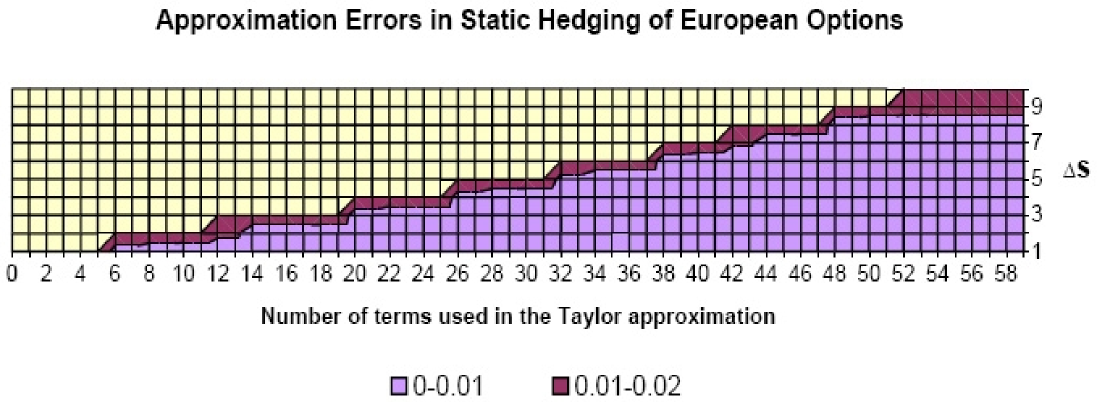

Figure 6.4.4.1: The approximation error in static hedging of European options. The x-axis gives the value of q and the y-axis gives ΔS. The area of the graph is coloured in blue when the approximation error ≤ 0.01 and in purple when the approximation error is between 0.01 and 0.02.

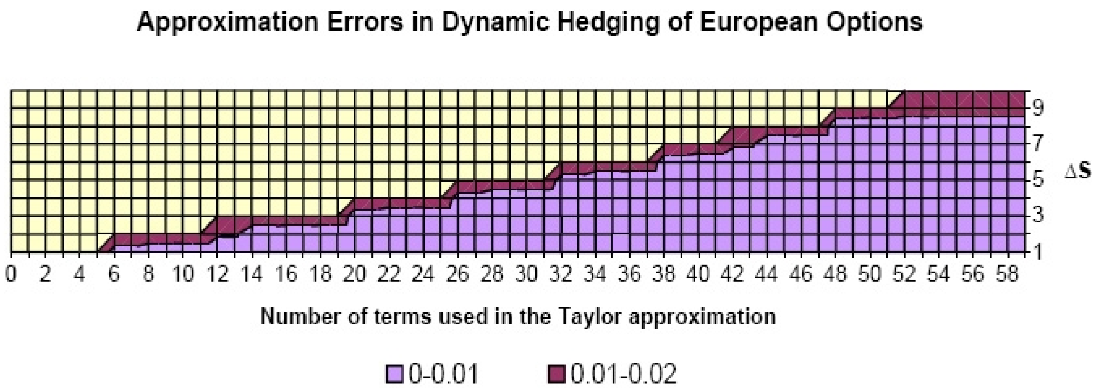

Figure 6.4.4.2: The approximation error in dynamic hedging of European options. The x-axis gives the value of q and the y-axis gives ΔS. The area of the graph is coloured in blue when the approximation error ≤ 0.01 and in purple when the approximation error is between 0.01 and 0.02.

The performance of static and dynamic hedging of European options is given in Figure 6.4.4.1 and Figure 6.4.4.2. We can see that the values of q required are the same in the cases of static and dynamic hedging. The value of q, that is, the number of terms required in the Taylor approximation, such that the error $\leq \alpha_{\text{tol}}$ increases gradually as the value of ΔS_t increases. This verifies the discussion given in the beginning of this section, that is, for a given tolerance level, the number of terms required in the Taylor expansions is finite. In reality, for static hedging where the hedging period lasts for 3 months, we expect ΔS_t to be much bigger than 10. We note that such scenarios are much more computationally expensive. The utility of static hedging is validated by the simulations with smaller jump sizes. The values of q for different values of ΔS_t is also given in Table 6.4.4.4.

The performance of dynamically hedging of up-and-out options is given in Figure 6.4.4.3. We assume the barrier is given by $H = 5010.5$. Contrary to the approximation results of the European options, the increase in the approximation errors as ΔS increases is non-monotonic for this realisation. This is due to that fact that the option becomes worthless once the price of the underlying stock becomes greater than the barrier, H, that is, the option is 'knocked-out'. Therefore, if the option in some of the simulated scenarios becomes worthless, the calculated option price would decrease accordingly. The non-monotonic increase in the approximation error is due to these random knock-outs in our simulations. This result suggests that in determining the value of q, extra care need to be taken. For example, from Figure 6.4.4.1 we determine that $q = 24$ is sufficient to hedge the option if $\Delta S \leq 4$. However, when $q = 25$ at $\Delta S = 4$, the approximation error becomes bigger than the tolerance level, α_{tol}. Hence, we should look at subsequent values of q to ensure the the approximation errors are below the tolerance level. In this case, we should choose $q = 26$. The values of q for different values of ΔS_t is also given in Table 6.4.4.4.

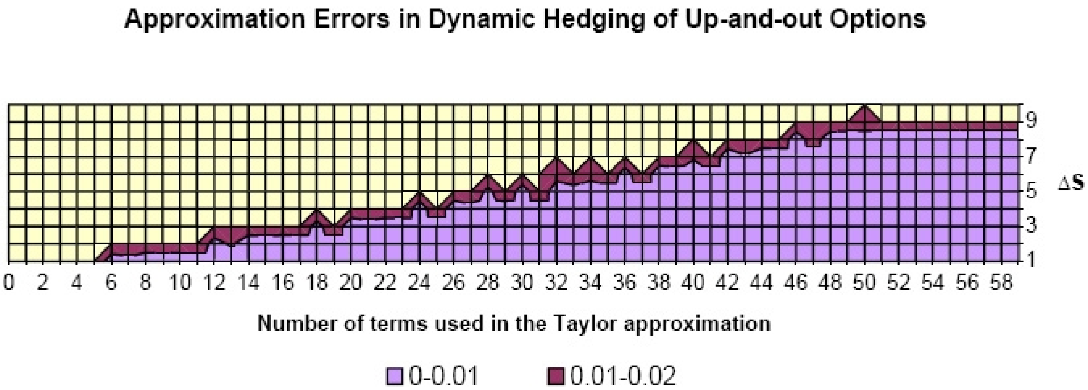

Figure 6.4.4.3: The approximation error in dynamic hedging of up-and-out options. The x-axis gives the value of q and the y-axis gives ΔS. The area of the graph is coloured in blue when the approximation error ≤ 0.01 and in purple when the approximation error is between 0.01 and 0.02.

The performance of dynamically hedging of up-and-in options is given in Figure 6.4.4.2.

We assume the barrier is given by $H = 5010$. We observe that the error is less then α_{tol} when $\Delta S = 1$ at $q = 0$. It is because the change in values of the option is less than α_{tol}. Therefore, no hedging is needed. However, in a real life situation, we would not know for sure that the option price would not change and hence we would have to use Taylor expansions to hedge. If $\Delta S = 0$, we need five terms in the Taylor expansion in order to hedge, that is, $q = 5$. For $\Delta S = 2$, we would need $q = 13$. Note that although the error is less than α_{tol} at $q = 11$, the error is bigger than α_{tol} at $q = 12$ and hence we choose $q = 13$ for the reason outlined in the case of up-and-out options. The values of q for different values of ΔS_t is given in Table 6.4.4.4.

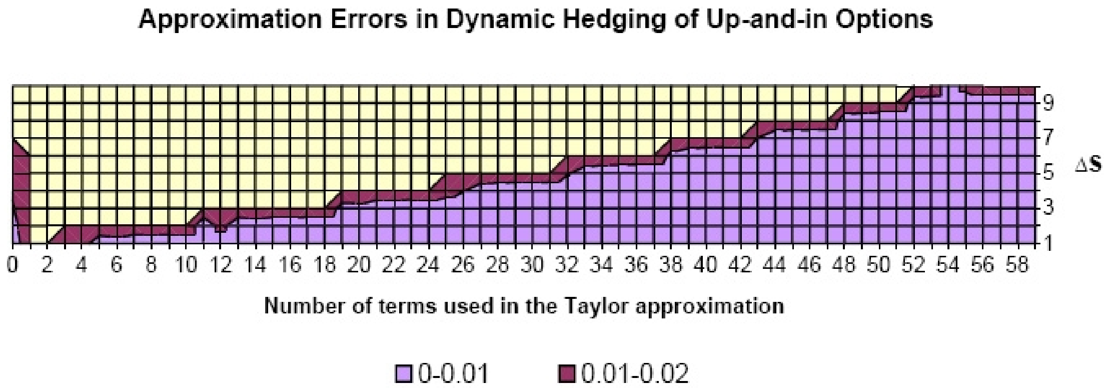

Figure 6.4.4.4: The approximation error in dynamic hedging of up-and-in options. The x-axis gives the value of q and the y-axis gives ΔS. The area of the graph is coloured in blue when the approximation error ≤ 0.01 and in purple when the approximation error is between 0.01 and 0.02.

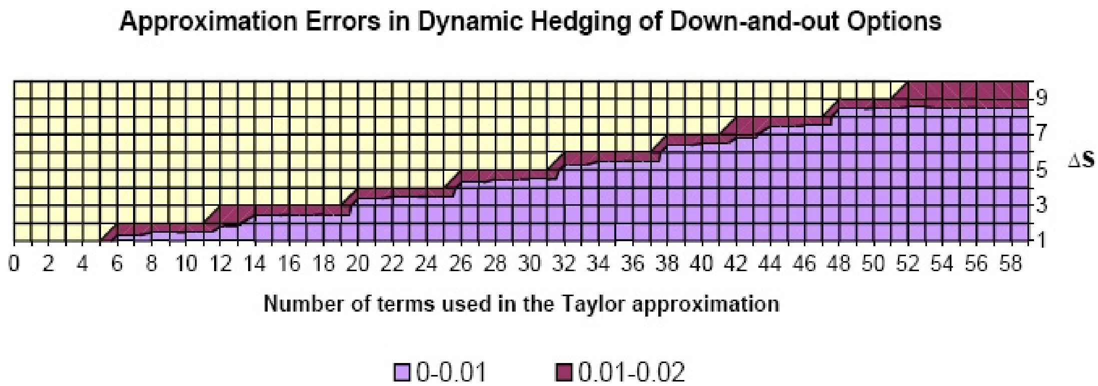

Figure 6.4.4.5: The approximation error in dynamic hedging of down-and-out options. The x-axis gives the value of q and the y-axis gives ΔS. The area of the graph is coloured in blue when the approximation error ≤ 0.01 and in purple when the approximation error is between 0.01 and 0.02.

The performance of dynamically hedging of down-and-out and down-and-in options is given in Figure 6.4.4.5 and Figure 6.4.4.5. We assume their barriers are given by $H = 4900$ and $H = 5000.5$, respectively. The performance of dynamically hedging down-and-out options is very similar to that of European options because we set the barrier to be 100 below the current stock price. During the five minutes hedging period, the stock price would hardly goes down by 100 and therefore the down-and-out options behave like an European options. Comparatively, the performance of dynamically hedging down-and-in options is different from that of European options. The values of q for different values of ΔS_t is also given in Table 6.4.4.4.

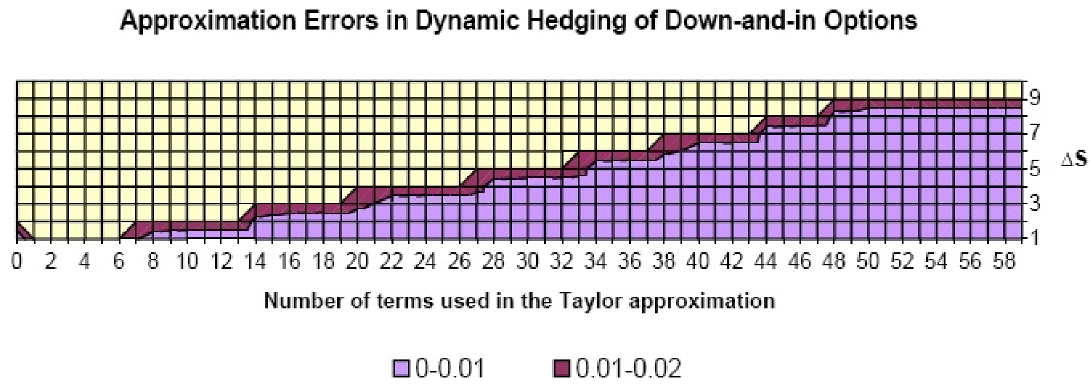

Figure 6.4.4.6: The approximation error in dynamic hedging of down-and-in options. The x-axis gives the value of q and the y-axis gives ΔS. The area of the graph is coloured in blue when the approximation error ≤ 0.01 and in purple when the approximation error is between 0.01 and 0.02.

The performance of hedging some other exotic options, such as lookback options and Asian options, can be obtained similarly since we employ Monte Carlo simulation in calculating the option prices. Recall in Section 6.4.3, as the positive integer N increases, the number of derivatives that can be calculated increases. The results show that the number of derivatives needed to achieve a satisfactory approximation, q, increases rapidly with increasing ΔS_t. Note that the bigger the value of ΔS_t, the slower the convergence rate of Taylor expansion and this is the reason why dynamic hedging is more popular in the literature. From our simulation results, we note that the derivatives, $\frac{D_2^i F(t+\Delta t, S_t)}{i}$, become very small as i increases, but the value of $(\Delta S_t)^i$ increases very rapidly. Therefore, we cannot ignore the terms $\frac{D_2^i F(t+\Delta t, S_t)}{i!} (\Delta S_t)^i$. To enable perfect hedging using moment swaps, power jump assets or some other traded derivatives depending on the same underlying asset, the market has to allow trading in these financial derivatives in a unit as small as $\frac{D_2^i F(t+\Delta t, S_t)}{i}$.

In summary, as long as we can find the q such that the Taylor approximations are accurate for all possible values of ΔS_t under consideration, the perfect hedging using moment swaps, power jump assets or other traded derivatives depending on the same underlying asset works very well.

In static hedging of European options in Figure 6.4.4.1,										
ΔS_t	1	2	3	4	5	6	7	8	≥ 9	
q	6	14	20	26	32	38	44	48	> 59	
In dynamic hedging of European options in Figure 6.4.4.2,										
ΔS_t	1	2	3	4	5	6	7	8	≥ 9	
q	6	14	20	26	32	38	44	48	> 59	
In dynamic hedging of up-and-out options in Figure 6.4.4.3,										
ΔS_t	1	2	3	4	5	6	7	8	≥ 9	
q	6	14	20	26	32	38	42	48	> 59	
In dynamic hedging of up-and-in options in Figure 6.4.4.4,										
ΔS_t	1	2	3	4	5	6	7	8	9	≥ 10
q	5	13	19	26	33	38	43	48	52	>59
In dynamic hedging of down-and-out options in Figure 6.4.4.5,										
ΔS_t	1	2	3	4	5	6	7	8	≥ 9	
q	6	14	20	26	32	38	44	48	> 59	
In dynamic hedging of down-and-in options in Figure 6.4.4.6,										
ΔS_t	1	2	3	4	5	6	7	8	≥ 9	
q	8	14	21	28	34	40	44	48	> 59	

Table 6.4.4.4: The values of q for given ΔS_t in static hedging of European options, dynamic hedging of European, up-and-out, up-and-in, down-and-out and down-and-in options.

Summary of Part II

In this part, we provided some perfect hedging strategies and minimal variance portfolios in a Lévy market. Many financial institutions hold derivative securities in their portfolios, and frequently these securities need to be hedged for extended periods of time. Failure to hedge properly can expose an institution to sudden swings in the values of derivatives, such as options, resulting from large, unanticipated changes in the levels or volatilities of the underlying asset. Research in the techniques employed for hedging derivative securities is therefore of crucial importance. Under the assumption of the famous Black-Scholes model, the market is complete and an European option can be hedged perfectly by investing in a risk-free bank account and the underlying stock. However, there is statistical evidence, such as the volatility smile, that the Black-Scholes model is not sufficiently flexible to model the price process. As a result, the study of Lévy process, which is a generalisation of Brownian motion with jumps, has become increasingly important in mathematical finance. It is well known that if the underlying asset is driven by a Lévy process, the market is not complete, that is, a contingent claim cannot be hedged using only a risk-free bank account and the underlying asset. By applying a Taylor expansion to the pricing formulae, we derived dynamic perfect hedging strategies of European and some exotic options by trading in moment swaps, power jump assets or certain traded derivatives depending on the same underlying asset. In the case of European options, static hedging can also be achieved. We extended the delta and gamma hedging strategies to higher moment hedging by investing in other traded derivatives depending on the same underlying asset. We demonstrated how to use the minimal variance portfolios derived by Benth *et al.* (2003) to hedge the higher order terms in the Taylor expansion, investing only in a risk-free bank account, the underlying asset and, potentially, variance swaps. We explicitly addressed numerical issues in the procedures, such as the approximation of the derivatives in the Taylor expansion, as well as investigated the performance of the hedging strategies. If as many derivatives as the Taylor expansion needed for accuracy can be determined and the financial derivatives required to hedge are available in the specified amounts, perfect hedging is possible.

Part III

Trading on the deviations of history implied and option implied distributions

In this part, we present an option trading strategy which involves comparing the probability distributions of the underlying at maturity implied by the historical data series of the underlying, and the current market prices of the options. The former distribution contains information about the underlying from observing historical data series, thus forming the basis for prediction. The current option prices contain information about future values of the underlying predicted by the investors of currently sold options. Since investors of options are risk averse, the distributions implied by option prices generally have fatter tails than the distribution implied by historical data. Due to these differences in the shape of the two distributions, we may construct an option trading strategy, using the extra information available. The two densities are assumed to follow the Variance Gamma (VG) model for ease of implementation. We simulate the underlying from today to maturity to see which options are overpriced and construct a trading portfolio at the day which both densities are estimated. The performance of the trading strategy under different market conditions are investigated and reported. Simulation results show that the trading strategy has a high earning potential.

This part is arranged as follow: Section 7.1 gives a brief review on skewness and kurtosis trades in the literature, Section 7.2 gives the overview of our trading strategy, Section 7.3 introduces the mean-correcting martingale measure used to obtain the risk neutral density of the history implied distribution and Section 7.4 gives an introduction of an efficient frontier which is used in determining the optimal variables for our strategy. Sections 8.1-8.5 give the algorithm for estimating the densities implied by history and current option prices, investigate the prediction of the direction of movement of the underlying and discuss the use of dynamic trading volume to control maximum possible loss. Section 8.6 introduces the risk-return analysis used to determine the optimal set of parameters of our strategy. Section 8.7 gives the performance of the strategy and Chapter 8.8 introduces portfolio insurances which can be used in combination with our trading strategy. Some concluding remarks are provided at the end of this part. Proofs, figures and tables are included in Appendix C.

Chapter 7

Background on skewness and kurtosis trades

7.1 Skewness and kurtosis trades in the literature

The study of the profitability of trading the deviations of the risk-neutral density of the underlying inferred from the historical time series and the risk-neutral density implied by the option prices has been the interest of the literature, see, for example, Aït-Sahalia *et al.* (2001), Blaskowitz (2001), Blaskowitz & Schmidt (2002) and Blaskowitz *et al.* (2004). Such studies were based on the assumption that the underlying process is driven by the Black-Scholes model. These authors considered so-called 'skewness' and 'kurtosis' trades, as the trades are initiated only if the skewness/kurtosis of the option implied density is higher than that of the history implied density. Corcuera *et al.* (2005) actually considered trading in power jump assets, among which the third and forth power jump processes are related to the skewness and kurtosis of the underlying driving process respectively. We derived some hedging strategies using power jump processes in Part II of this thesis. Although plots of density functions (reproduced as Figures C.1.1 and C.1.2 in Appendix C.1) were included in their papers, no quantitative analysis was carried out to obtain the optimal trading portfolio. The plots of density functions were merely used to explain the concept of making profits from the deviations of the two distributions. In applications, only the values of the skewness and kurtosis of the two distributions were compared. Aït-Sahalia *et al.* (2001) introduced S1, S2, K1 and K2 trades (detailed in Table C.1.1), which were initiated only when the corresponding skewness or kurtosis condition was satisfied and the strikes and volumes of the options bought/sold were pre-specified. Figures C.1.3 and C.1.4 showed the payoffs of S1 and K1 trades respectively. It is clear that the strategies are not practical since the possible loss is unbounded. More importantly, if the option implied risk-neutral density is negative skewed or has a higher kurtosis than the normal distribution, this **does not** imply that it must have the shapes as given in Figures C.1.1 or C.1.2. Hence, the trades described in Table C.1.1 do not take full advantage of the comparison: a full distributional comparison is needed. Each trade should be tailor-made using a dynamic algorithm which produces the best portfolio according to the investors' preference. Blaskowitz (2001) further investigated the profitability of buying and selling a predefined range of options to change the shapes of the payoff functions. Again, the portfolios suggested are static because the strikes and volumes of the options traded are pre-specified. In this part of the thesis, we extract more information from the density implied from the historical time series of the underlying and that implied from current option prices. The densities approximated from the data are compared and the ranges of the underlying where the options are overpriced are calculated and trading volumes are selected to maximise profits.

7.2 Strategy overview

Inspired by the research papers cited in Section 7.1, we develop a model which generates trading strategies for options in a Lévy market. The development of the model is motivated by the observation that the densities of the values of the underlying asset at maturity implied by the historical data of the underlying and the current option prices are different. These deviations suggest that the current market prices of options with certain strikes are mispriced. From the option pricing formulae given in equations (8.3) and (8.4) in Section 8.1, we will see that the option prices are a function of the density of the value of the underlying asset at maturity. Therefore, a mismatch of the densities would indicate the range of strikes at which the options might be mispriced. Since there would be a range of such strikes, we need to define two parameters, one for the left tail, R_L, and one for the right tail, R_R, which indicate exactly which strike we use for the strategy. Totally seven parameters will be introduced to describe the strategy proposed in this paper. The values of these parameters will be optimised later using efficient frontier analysis, as described in Section 7.4. We define the parameter, p, which helps us to predict whether a rise or a fall in the underlying is likely based on observing the density implied from the current option prices. Since these three parameters, R_L, R_R and p, are defined from the density functions, we call them the 'density parameters'. From a practical point of view, we would like to develop a strategy that could not lead to infinite loss and at the same time maximise profit. Therefore, we define two parameters which indicate the maximum and minimum amounts, $L_{\max}$ and $L_{\min}$, of money we would invest in the strategy, given the current level of capital at hand. We also define a parameter indicating the initial capital,

c, assigned for the trading strategy. Finally, we need a parameter to determine the default volume of trading, v, that is, the amounts of options we originally specify to trade in a day. Note that our strategy may change the trading volume according to the maximum and minimum amounts of money we are allowed to invest. The parameters c, v, $L_{\max}$ and $L_{\min}$ are then optimised using efficient frontier analysis. Since these four parameters are related to how much should be invested in the strategy, they are called the 'monetary parameters'. In the following, we describe the definitions of these seven parameters in detail and then perform efficient frontier analysis to obtain the set of optimised parameters in terms of expected annual growth rate and Conditional Value at Risk (CVaR).

7.3 The mean-correcting martingale measure

In this part we calibrate the VG model (see Section 6.4.1) with historical data of the underlying asset of an option and apply a change of measure method to obtain a risk-neutral distribution of the price of the underlying at maturity. A way to obtain an equivalent martingale measure is by mean-correcting the exponential of a Lévy process, see Schoutens (2003, p. 79). For the VG model given in (6.35), this can be done by the parameter m. We estimate all the parameters involved in the process, given the historical data series, with the formula $S_t = S_0 \exp\left(X\left(t; \sigma_S, \nu_S, \theta_S\right)\right)$. Then, together with (6.34) and (6.35), the m parameter for the mean-correcting equivalent martingale measure in the VG model is, according to Schoutens (2003, p. 79), given by

$$\begin{aligned} m + \omega &= r - \ln \phi_X\left(-i, 1; \sigma, \theta, \nu\right) = r - \ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right)^{-\frac{1}{\nu}} \\ &= r + \frac{1}{\nu}\ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right) = r + \omega, \end{aligned}$$

where $\omega = \frac{1}{\nu}\ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right)$ defined in (6.35). Hence, we should set $m = r$. We can check easily that our new discounted stock price is a martingale: for $s < t$,

$$\begin{aligned} &E\left[\exp\left(-rt\right) S_t | \mathcal{F}_s\right] \\ &= E\left[S_0 \exp\left(X\left(t; \sigma, \nu, \theta\right) + \omega t\right) | \mathcal{F}_s\right] \\ &= S_0 E\left[\exp\left\{X\left(t; \sigma, \nu, \theta\right) - X\left(s; \sigma, \nu, \theta\right) + X\left(s; \sigma, \nu, \theta\right) + \omega t\right\} | \mathcal{F}_s\right] \\ &= S_0\left\{E\left[\exp\left\{X\left(t; \sigma, \nu, \theta\right) - X\left(s; \sigma, \nu, \theta\right)\right\}\right] \exp\left(X\left(s; \sigma, \nu, \theta\right) + \omega t\right)\right\} \\ &= S_0\left\{\exp\left\{-\omega\left(t - s\right)\right\} \exp\left(X\left(s; \sigma, \nu, \theta\right) + \omega t\right)\right\} \\ &= S_0\left\{\exp\left(X\left(s; \sigma, \nu, \theta\right) + \omega s\right)\right\} \\ &= \exp\left(-rs\right) S_s. \end{aligned}$$

7.4 Efficient frontier

The efficient frontier was first defined by Markowitz (1952) in his groundbreaking paper that launched portfolio theory. The theory considers a universe of risky investments and explores what might be an optimal portfolio based upon trading in a combination of those possible investments. Using the same idea, we explore the optimal set of parameters based on comparing the expected risk and return level obtained from simulation.

According to Markowitz (1952), every possible asset combination can be plotted in a risk-return space, and the collection of all such possible portfolios defines a region in this space. The line along the upper edge of this region is known as the *efficient frontier* (or 'the Markowitz frontier'). Combinations along this line represent portfolios for which there is lowest risk for a given level of return. Conversely, for a given amount of risk, the portfolio lying on the efficient frontier represents the combination offering the best possible return. Mathematically the efficient frontier is the intersection of the set of portfolios with minimum risk and the set of portfolios with maximum return.

The efficient frontier will be convex since the risk-return characteristics of a portfolio change in a non-linear fashion as its component weightings are changed. The efficient frontier is a parabola (hyperbola) when expected return is plotted against risk. Points below the frontier are suboptimal. A rational investor will hold a portfolio only on the frontier.

In this part of the thesis, we employ efficient frontier analysis to obtain the optimal set of parameters. We define the range of each parameter to be used. For each set of parameters' values, a large number of realisations are simulated to approximate the expected return measured by the Compounded Annual Growth Rates (CAGR)

$$\text{CAGR} = \left(\frac{\text{ending value of total capital}}{\text{starting value of total capital}}\right)^{\frac{1}{\text{number of years}}} - 1. \tag{7.1}$$

Instead of using the variance to represent the risk, we measure risks by the Conditional Value at Risk (CVaR) at level α, where we set $\alpha = 95\%$. CVaR is used because it can measure the downside risk, that is, the likelihood that a security or other investment will decline in price, more effectively. It is because CVaR is subadditive and risk-sensitive, opposed to the more well-known risk measure Value at Risk (VaR). For more details of the comparison between different risk measures, we refer the reader to Martin (2004, Chapter 9). To find the CVaR, we need to find the VaR first, which is the value Q such that $P(X < Q) = \alpha$, for some given tail probability P, where X is the CAGR. The VaR is then $-Q$. The CVaR is the expectation conditional on the loss exceeding Q, again with a minus sign: $\text{CVaR} = -E[X|X < Q]$. It is clear that $|\text{CVaR}| > |\text{VaR}|$ because all the events that give worse outcomes than a loss of Q are being averaged when the CVaR is computed.

Chapter 8

The option trading strategy

In this chapter, we describe the option trading strategy step-by-step and give detailed explanation of the methodology. We first give the algorithm of our strategy in Table 8.1.1. We will discuss it in more detail in the following sections.

8.1 Estimation of the densities implied by history and current option prices (Steps 1-7)

We compare the risk-neutral densities implied by the options traded in the markets and by the historical data series of the underlying. The former distribution contains information about the underlying from observing historical data series, thus forming the basis for prediction. The current option prices contain information about future values of the underlying predicted by the investors of currently sold options. Since investors of options are risk averse, the distributions implied by option prices generally have fatter tails and are more negatively skewed than the distribution implied by historical data. Due to these differences in the shape of the two distributions, an option trading strategy may be constructed, using the extra information available. The two densities are assumed to follow the Variance Gamma (VG) model for ease of implementation. The underlying is simulated forward to see which options are overpriced and construct a trading portfolio at the day at which both densities are estimated. FTSE daily return data from 29th March 2000 to 23rd November 2007 are used as the historical price series and the options traded on each day, with 80 different strike prices and maturity of 3 months are used in estimating the option implied densities. On each trading day, we perform the density comparison and construct an option trading strategy. The length of the historical data series used on each trading day to calibrate the model has to be chosen with care. If the time series is too long, it cannot provide information about current market conditions. If

The Algorithm	
1.	Determine the time window to be used for prediction using historical data series.
2.	Estimate the parameters of the VG model using the historical data series.
3.	Simulate N replicates of the underlying at maturity using the fitted VG model from Step 2.
4.	Estimate the density function by applying a smoothing kernel method.
5.	Calibrate the VG model with the option prices in the market.
6.	Simulate N replicates of the underlying at maturity using the fitted VG model from Step 5.
7.	Estimate the density function by applying a smoothing kernel method as done in Step 4.
8.	Estimate the optimal value of the prediction parameter, p.
9.	Fix the values of the grid for the density parameters: R_R and R_L.
10.	Fix the values of the grid for monetary parameters: c, $L_{\max}$, $L_{\min}$ and v.
11.	The maximum and minimum of loss allowed are given by $cL_{\max}$ and $cL_{\min}$.
12.	For each set of the parameters from the grids defined in Steps 9 and 10, choose the trading volume (by changing the default volume v), the strikes of the put and call options to buy such that the maximum possible loss is smaller than $cL_{\max}$ but is bigger than $cL_{\min}$.
13.	For each trading day in the history where the value of the underlying at maturity is already known, calculate the profit from the trading portfolio constructed in Step 12.
14.	Repeat Steps 2-7,12,13 for each trading day in the history to calculate the expected risk and return for each set of parameters from the grids defined in Steps 9 and 10.
15.	Plot the expected risk and return on a graph for each set of parameters, yielding the efficient frontier.

Table 8.1.1: The algorithm for our trading strategy.

it is too short, the calibration procedure will not be accurate enough. One of the main challenges to our trading strategy is to predict the direction of movement of the underlying, which is discussed in detail in Section 8.4. Hence, we want to choose the length, τ, such that the prediction error is minimised:

$$\tau = \min_{\tau'} \frac{1}{T - \tau' + 1} \sum_{t=\tau'}^{T} \left(S_t - S_{t-\tau'}\right)^2, \tag{8.1}$$

where T is the total number of data points available and $T - \tau' + 1$ must be reasonably large for the result to be accurate.

Step 1: Given the historical data series of the underlying, find τ which is given by equation (8.1).

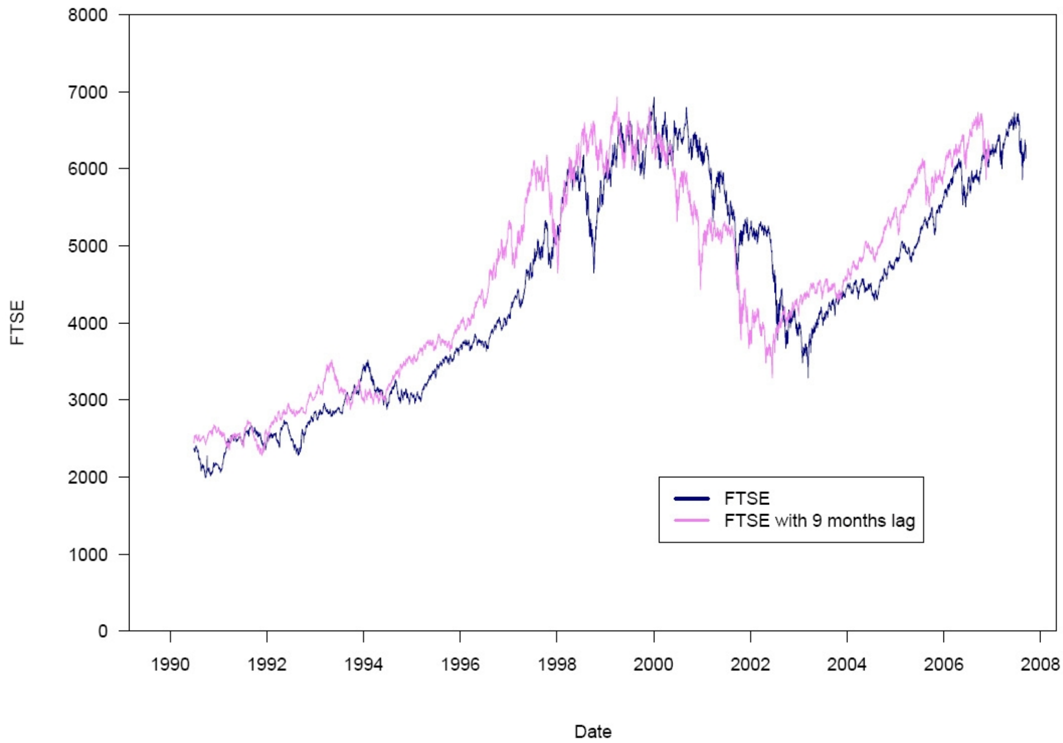

Figure 8.1.1: The time series of FTSE.

For FTSE 100, we find that τ is approximately equal to 9 months. Figure 8.1.1 gives the evolution of the FTSE index over 16 years and the data series with 9-month time lag. Since we need to predict the trend of the underlying for the next 3 months, the length of the historical data used should be equal to 3 months as well so as to correctly capture the statistical behaviour in a 3-month period. For example, if we perform the density comparison today, we use historical data of 9 months to 6 months old to calibrate the model. The whole 9 months period is not used because it is well known that the stock price process is not stationary. To minimize the effect of nonstationarity of the parameters, we calibrate the model on data of the same length as the time period we simulate.

Step 2: For each trading day D, calibrate the parameters of the model using historical data of τ to $\tau - M$ older than D, where M is the time length from D to maturity of the options under consideration. In our case, we use the VG model, discussed in Section

6.4.1.

Note that other stochastic models can be used to describe the dynamics of the underlying stock process. The VG model is used for its simplicity and the ability to handle skewness and kurtosis explicitly. For example, the stochastic volatility model using a VG process (VGSAM), introduced by Carr *et al.* (2003), can be used such that the effects due to stochastic volatility can be handled. In the VG model, the volatility of the underlying process is assumed to be deterministic and the process is assumed to be stationary. The VGSAM allows the modelling of the variability of the volatility as observed in market data. Therefore, the deviations of the distributions of the underlying at maturity implied by historical data and that implied by current option prices due to different views on the stochastic volatility can be captured. For example, the history might imply lower variability of the volatility than the investors in the market. By using VGSAM, this deviation can be captured and traded upon. Similarly, if our portfolio contains options on more than one underlying asset, we may want to employ a financial model that can capture the correlations between the different assets. However, this thesis focuses on the idea of trading on the deviations of the distributions in a Lévy market and the use of more complicated model is out of the scope of this thesis.

Step 3: For each trading day, simulate forward N times to get the points at which the underlying will be at maturity using equation (6.35), where N is a large positive number. Apply a change of measure method to get a risk-neutral distribution.

Note that for more accurate results, one should employ (Markov Chain Monte Carlo (MCMC) to estimate the parameters of the model and then simulate forward from the distributions of the parameters. We do not employ MCMC because it is more complicated to implement and we want to focus on the construction of the trading strategy. Note that our approach underestimates the variability of the parameters. We apply the change of measure method described in Section 7.3 because of its simplicity. Other ways of changing the measure can be used, see Miyahara (2005) for the different properties of a few common kinds of equivalent martingale measures for geometric Lévy processes. In our simulation, we choose $N = 10,000$. Since a VG process can be decomposed into a difference of two Gamma processes, see equation (6.32), and the parameters are constant over time, simulation is easily implemented, as we do not have to simulate the price series over daily interval. We can simply simulate with the time interval equal to the maturity.

Step 4: Using the N values of the underlying at maturity for each trading day D obtained from Step 3, estimate the density function using a kernel smoothing method.

We use the kernel smoothing method introduced by Bowman & Azzalini (1997). In MatLab 7.3.0, this can be done using the function `ksdensity()` in the statistics toolbox. The estimate is based on a normal kernel function under the assumption that the data sample is independent and identically distributed, which is satisfied in our case since in each of our simulations the VG model produces i.i.d. random variables. We use the default bandwidth of the `ksdensity()` function, which produces a smooth density estimate. Hence, we obtain an approximation of the density function implied by historical data series of the underlying. The next step is to obtain the historical data of the option prices. Recall that we perform this density comparison on each trading day in the history to obtain the expected return and risk for each set of parameters such that we can approximate the 'optimal parameters' using risk-return analysis. In general, it is difficult to obtain the history of option prices since the bid and ask prices of the options are generally not stored for over two years. Normally, only the implied volatilities, calculated using the Black-Scholes model, of the option prices are stored for a long period of time. To obtain the option prices, the Black-Scholes option pricing formula is used to calculate the prices again. The Black-Scholes model is used because only the implied volatilities calculated using the Black-Scholes model are stored in the database rather than the prices themselves. Therefore, we do not have the information about exactly what strikes of options were traded and the trading volume on a particular day. Since the whole volatility surface, that is, the volatilities for different strikes and maturities, is stored in the volatility database, the price of an option can be calculated given a strike and a maturity date. For calibrating the model using option prices, we use prices of options with maturity of 3 months and 80 different strikes around the 'at the money' strike, that is, the strike price equal to the current price of the underlying asset. When testing the performance of the strategy, we assume only certain strikes can be traded liquidly, as will be discussed in Section 8.6.

Step 5: Given the prices of options with the same maturity and different strikes, estimate the parameters of the VG model.

We use the calibration method introduced by Chourdakis (2005) using fractional Fractional Fast Fourier Transform (FRFT) to estimate the parameters of the VG model from the option price data. Choudakis (2005) showed how FRFT can be used to retrieve option prices from the corresponding characteristic function of the log return of the price process. In the case of VG model, the characteristic function is given by

$$\phi_S\left(u,t;r,\sigma_{RN},\nu_{RN},\theta_{RN}\right)=\frac{\exp\left(\mathrm{i}u\left(r+\omega_{RN}\right)t\right)}{\left(1-\mathrm{i}\theta_{RN}\nu_{RN}u+\frac{\sigma_{RN}^2\nu_{RN}}{2}u^2\right)^{\frac{t}{\nu_{RN}}}}. \tag{8.2}$$

It was shown in Chourdakis (2005) that in the case of VG model, the FRFT method can deliver option prices up to forty-five times faster than the well-known Fast Fourier Transform (FFT) by Carr & Madan (1999), without substantial loss of accuracy.

Step 6: For each trading day, simulate forward N times to get the points at which the underlying will be at maturity, where N is a large positive number.

Step 7: Using the N values of the underlying at maturity for each trading day D obtained from Step 6, estimate the density function using a kernel smoothing method as in Step 4.

8.2 Quantifying payoff using the comparison functions

We repeat Steps 1 to 8 for 1874 trading days between 29th June 2000 and 23rd November 2007 and show eight of the plots in Figure 8.2.1. We notice that the relative shapes of the two densities are similar. The densities implied by the options have much fatter tails and are more negatively skewed. Their peaks are generally on the right of that of the history implied densities. There is less risk aversion in the history implied deviations, and the distributions are substantially more peaked. Figure 8.2.2 gives the general shape of the distributions. In this section, we determine the range of strikes of the options which are overpriced by comparing the two density functions, based on the subjective belief that the distribution implied by historical data series of the underlying is more accurate than the one implied by current market prices due to that fact that investors are risk-averse. The put and call option pricing formulae are in terms of the density functions and are given by

$$P\left(S_t, \mathrm{P}, r, T-t\right) = e^{-r(T-t)} \int_0^{\mathrm{P}} (\mathrm{P}-s)\, f_{S_T}\left(s, S_t\right) \mathrm{d}s, \tag{8.3}$$

$$C\left(S_t, \mathrm{C}, r, T-t\right) = e^{-r(T-t)} \int_{\mathrm{C}}^{\infty} (s-\mathrm{C})\, f_{S_T}\left(s, S_t\right) \mathrm{d}s, \tag{8.4}$$

where P and C are the strikes of the put and call options, respectively, and $f_{S_T}\left(s, S_t\right)$ is the probability density function (pdf) of the underlying at maturity, depending on the value of the stock price today, S_t. In other words, the option price is the discounted expectation of payoff with respect to the probability distribution of the underlying at maturity with respect to a risk-neutral measure. Hence, we can find out the range of values of P and C such that the option prices implied from history are lower than those traded in the market. We can discuss this graphically with the help of Figure 8.2.2.

Let X_R and X_L be the intersections of the two density functions on the right and on the left respectively. Let E_R and E_L be the intersections of the option implied density with the x-axis. Let $f^o(\cdot)$ and $f^h(\cdot)$ be the density functions of the option implied and history implied densities respectively. To quantify the difference between the two, we define the *area difference function* to be:

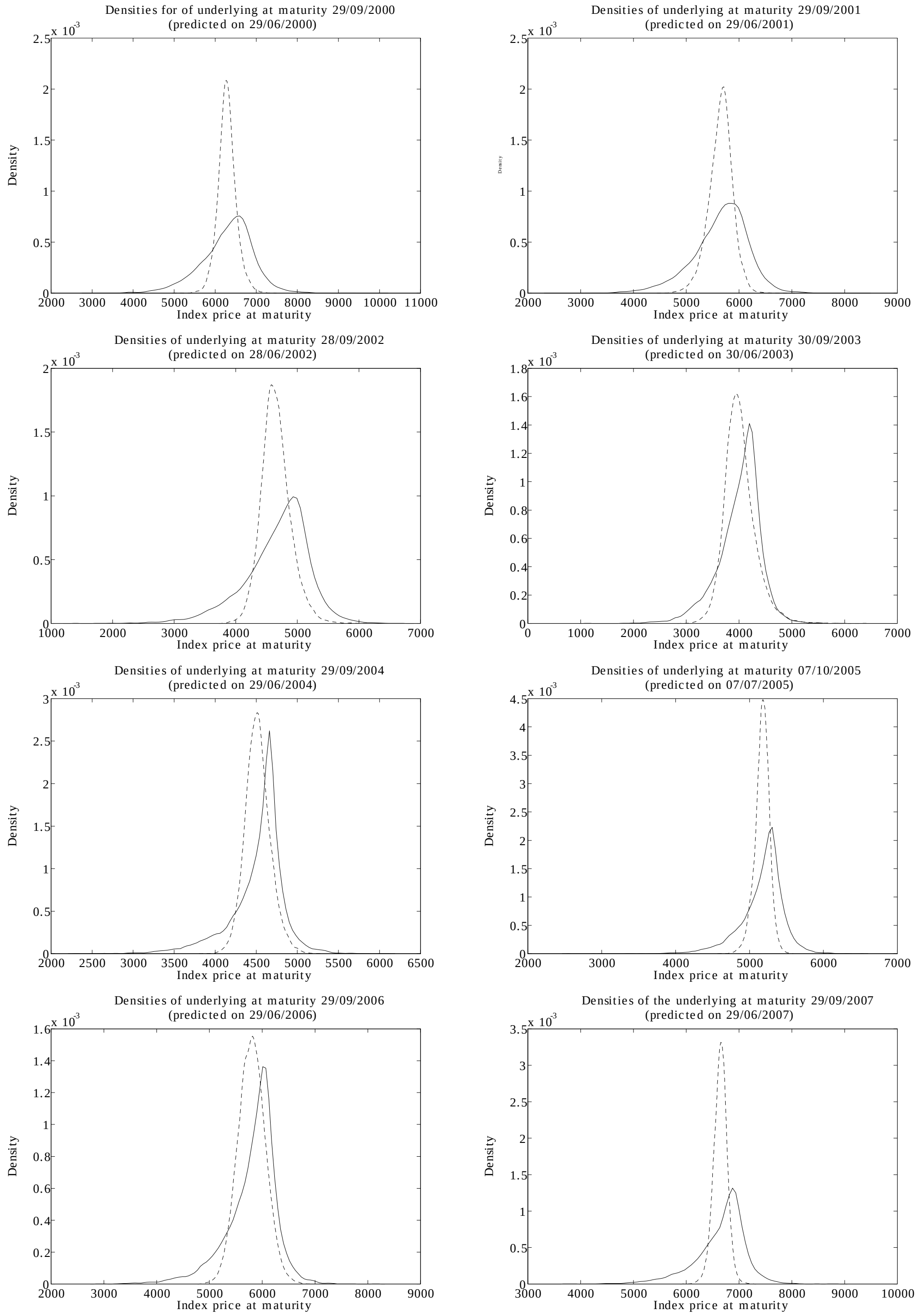

Figure 8.2.1: The densities implied by options (solid line) and historical data (dotted line) for FTSE. The VG models are calibrated using data available at 29/6/2000, 29/6/2001, 28/6/2002, 30/6/2003, 29/6/2004, 7/7/2005, 29/6/2006, 29/06/2007 with 3 months maturity. The plots give the density functions of the index at maturity.

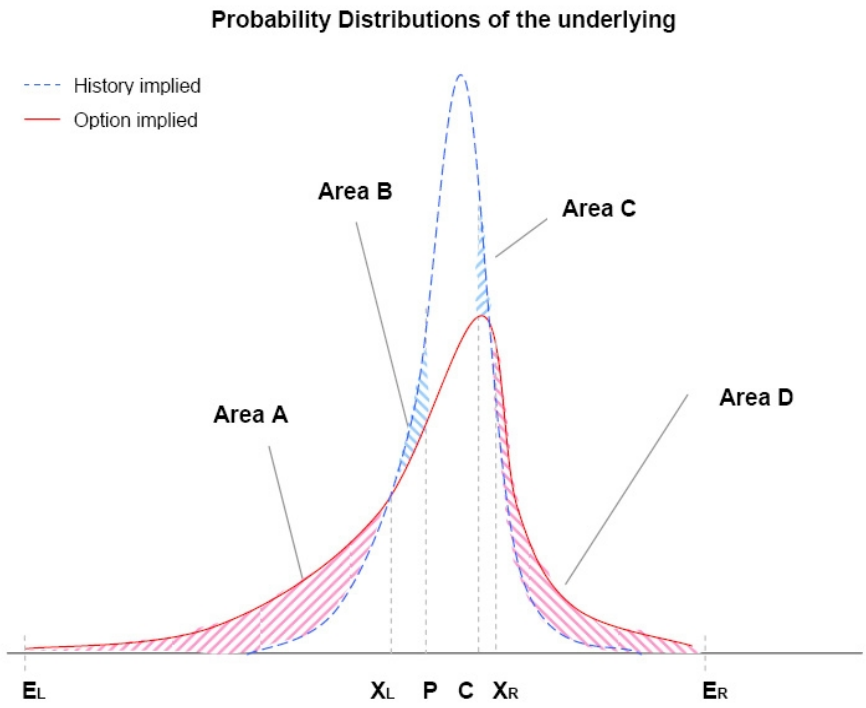

Figure 8.2.2: Density comparison with definitions of Areas A, B, C and D.

$$\Delta G(x_1, x_2) = \int_{x_1}^{x_2} \left[f^o(s) - f^h(s)\right] \mathrm{d}s. \tag{8.5}$$

This function has the fundamental quality as we will show later in Proposition 8.2.1 that the price differences of the options implied from the history and from the current option prices can be expressed as a function of this area difference function. The *comparison functions* are defined to be:

$$p(f_1(\cdot), f_2(\cdot), x_1, x_2, \mathrm{P}) = \int_{x_1}^{x_2} (\mathrm{P} - s)\,[f_1(s) - f_2(s)]\,\mathrm{d}s \tag{8.6}$$

and

$$c(f_1(\cdot), f_2(\cdot), x_1, x_2, \mathrm{C}) = \int_{x_1}^{x_2} (s - \mathrm{C})\,[f_1(s) - f_2(s)]\,\mathrm{d}s. \tag{8.7}$$

These functions facilitate the price comparison to obtain the range of strikes at which the options are overpriced.

For example, by comparing equation (8.3) with equation (8.6) for a given pair of densities $f^o(\cdot)$ and $f^h(\cdot)$, we can see that if $p\left(f^o(\cdot), f^h(\cdot), E_L, \mathrm{P}, \mathrm{P}\right) > 0$, then the put options with strike P are overpriced. Similarly, if $c\left(f^o(\cdot), f^h(\cdot), \mathrm{C}, E_R, \mathrm{C}\right) > 0$, then the call options with strike C are overpriced. In our analysis, we need to parametrise the choice of the strike to trade, which is in the range of strikes at which the options are overpriced, such that we can optimise the parameters through risk-return analysis. We note that $p\left(f^o(\cdot), f^h(\cdot), E_L, \mathrm{P}, \mathrm{P}\right)$ and $c\left(f^o(\cdot), f^h(\cdot), \mathrm{C}, E_R, \mathrm{C}\right)$ can be decomposed into a sum of two comparison functions in terms of the intersection points X_R and X_L :

$$\begin{aligned} & p\left(f^o(\cdot), f^h(\cdot), E_L, \mathrm{P}, \mathrm{P}\right) \\ & \quad = \begin{cases} p\left(f^o(\cdot), f^h(\cdot), E_L, X_L, \mathrm{P}\right) - p\left(f^h(\cdot), f^o(\cdot), X_L, \mathrm{P}, \mathrm{P}\right) & \text{if } X_L < \mathrm{P} \\ p\left(f^o(\cdot), f^h(\cdot), E_L, \mathrm{P}, \mathrm{P}\right) & \text{if } X_L \geq \mathrm{P} \end{cases} \\ & \quad = p\left(f^o(\cdot), f^h(\cdot), E_L, \min\left(X_L, \mathrm{P}\right), \mathrm{P}\right) - p\left(f^h(\cdot), f^o(\cdot), X_L, \max\left(\mathrm{P}, X_L\right), \mathrm{P}\right) \end{aligned}$$

and

$$\begin{aligned} & c\left(f^o(\cdot), f^h(\cdot), \mathrm{C}, E_R, \mathrm{C}\right) \\ & \quad = \begin{cases} c\left(f^o(\cdot), f^h(\cdot), X_R, E_R, \mathrm{C}\right) - c\left(f^h(\cdot), f^o(\cdot), \mathrm{C}, X_R, \mathrm{C}\right) & \text{if } X_R > \mathrm{C} \\ c\left(f^o(\cdot), f^h(\cdot), \mathrm{C}, E_R, \mathrm{C}\right) & \text{if } X_R \leq \mathrm{C} \end{cases} \\ & \quad = c\left(f^o(\cdot), f^h(\cdot), \max(X_R, \mathrm{C}), E_R, \mathrm{C}\right) - c\left(f^h(\cdot), f^o(\cdot), \min\left(X_R, \mathrm{C}\right), X_R, \mathrm{C}\right). \end{aligned}$$

Therefore, the put option with strike P is overpriced if

$$p\left(f^o(\cdot), f^h(\cdot), E_L, \min\left(X_L, \mathrm{P}\right), \mathrm{P}\right) > p\left(f^h(\cdot), f^o(\cdot), X_L, \max\left(\mathrm{P}, X_L\right), \mathrm{P}\right) \tag{8.8}$$

and the call option with strike C is overpriced if

$$c\left(f^o(\cdot), f^h(\cdot), \max(X_R, \mathrm{C}), E_R, \mathrm{C}\right) > c\left(f^h(\cdot), f^o(\cdot), \min\left(X_R, \mathrm{C}\right), X_R, \mathrm{C}\right). \tag{8.9}$$

Graphically, we can relate $p\left(f^o(\cdot), f^h(\cdot), E_L, X_L, \mathrm{P}\right)$ to Area A in Figure 8.2.2 by the fact that they are both related to the difference of the density functions $f^o(\cdot)$ and $f^h(\cdot)$ from range E_L to X_L. Note that in the case of $X_L < P$ and $X_R > C$, from (8.5),

$$\begin{aligned} \text{Area A} &= \int_{E_L}^{X_L} \left[f^o(s) - f^h(s)\right] \mathrm{d}s = \Delta G\left(E_L, X_L\right) \\ \text{Area B} &= \int_{X_L}^{\mathrm{P}} \left[f^h(s) - f^o(s)\right] \mathrm{d}s = -\Delta G\left(X_L, \mathrm{P}\right) \\ \text{Area C} &= \int_{\mathrm{C}}^{X_R} \left[f^h(s) - f^o(s)\right] \mathrm{d}s = -\Delta G\left(\mathrm{C}, X_R\right) \\ \text{Area D} &= \int_{X_R}^{E_R} \left[f^o(s) - f^h(s)\right] \mathrm{d}s = \Delta G\left(X_R, E_R\right). \end{aligned}$$

The following proposition gives the relationships of the comparison functions in (8.8) and (8.9) with the above areas, that is, the relationships of $p\left(f^o(\cdot), f^h(\cdot), E_L, X_L, \mathrm{P}\right)$, $p\left(f^h(\cdot), f^o(\cdot), X_L, \mathrm{P}, \mathrm{P}\right)$, $c\left(f^h(\cdot), f^o(\cdot), \mathrm{C}, X_R, \mathrm{C}\right)$ and $c\left(f^o(\cdot), f^h(\cdot), X_R, E_R, \mathrm{C}\right)$, with areas A, B, C and D through the area function, respectively.

Proposition 8.2.1 *i) The comparison function* $p\left(f^o(\cdot), f^h(\cdot), E_L, X_L, \mathrm{P}\right)$ *is an increasing function of* $\Delta G(E_L, X_L)$ *and hence of Area A.*
ii) The comparison function $p\left(f^h(\cdot), f^o(\cdot), X_L, \mathrm{P}, \mathrm{P}\right)$ *is an increasing function of* $-\Delta G(X_L, \mathrm{P})$ *and hence of Area B.*
iii) The comparison function $c\left(f^h(\cdot), f^o(\cdot), \mathrm{C}, X_R, \mathrm{C}\right)$ *is an increasing function of* $-\Delta G(\mathrm{C}, X_R)$ *and hence of Area C.*
iv) The comparison function $c\left(f^o(\cdot), f^h(\cdot), X_R, E_R, \mathrm{C}\right)$ *is an increasing function of* $\Delta G(X_R, E_R)$ *and hence of Area D.*

Proof. The proof is included in Appendix C.2. □

Equations (8.8) and (8.9) will be useful in constructing the parameters R_L and R_R introduced in Section 7.2. This is further discussed in next section.

8.3 Determining which options to sell

In last section, we see that there is a range of strikes within which we can sell the overpriced options. To maximise profit, we should predict the direction of the movement of the underlying, which is discussed in Section 8.4. Intuitively, if we predict an upward move, we would sell a put and call options with relatively higher strikes in their respective ranges. In fact, we should specify the left and right price ratios, R_L and R_R, with which we decide which options to sell. For example, if we expect a rise in the underlying, we sell the put options with strike at which

$$R_L = \frac{p\left(f^h(\cdot), f^o(\cdot), X_L, \max(\mathrm{P}, X_L), \mathrm{P}\right)}{p\left(f^o(\cdot), f^h(\cdot), E_L, \min(X_L, \mathrm{P}), \mathrm{P}\right)}, \tag{8.10}$$

and sell call options with strike at which

$$R_R = \frac{c\left(f^o(\cdot), f^h(\cdot), X_R, E_R, \mathrm{C}\right) - c\left(f^o(\cdot), f^h(\cdot), \mathrm{C}, E_R, \mathrm{C}\right)}{c\left(f^o(\cdot), f^h(\cdot), X_R, E_R, \mathrm{C}\right)} \text{ and } \mathrm{C} \geq X_R. \tag{8.11}$$

If $R_R = 0$, $X_R = \mathrm{C}$, so we sell call options with strike at the right hand side interaction. If $R_R = 1$, $c\left(f^o(\cdot), f^h(\cdot), \mathrm{C}, E_R, \mathrm{C}\right) = 0$, so we sell call options with strikes at the right hand side intersection of the option implied density and the x-axis. Similarly, if we expect a fall in the underlying, we sell call options with strike at which

$$R_R = \frac{c\left(f^h(\cdot), f^o(\cdot), \min(X_R, \mathrm{C}), X_R, \mathrm{C}\right)}{c\left(f^o(\cdot), f^h(\cdot), \max(X_R, \mathrm{C}), E_R, \mathrm{C}\right)}, \tag{8.12}$$

and sell the put options at which

$$R_L = \frac{p\left(f^o(\cdot), f^h(\cdot), E_L, X_L, \mathrm{P}\right) - p\left(f^o(\cdot), f^h(\cdot), E_L, \mathrm{P}, \mathrm{P}\right)}{p\left(f^o(\cdot), f^h(\cdot), E_L, X_L, \mathrm{P}\right)} \text{ and } \mathrm{P} \leq X_L. \tag{8.13}$$

Remark 8.3.1 *It may be tempting to buy options near the peaks of the distributions since Figure 8.2.2 implies that they are underpriced. However, we should not buy these underpriced options since both densities have heavy mass around the peaks, indicating that the options will expire worthless with a very high probability. Moreover, options with strikes near the spot price are much more expensive than those with strikes far away from the spot price. If we bought these options, the initial cashflow would be negative and it is very likely that all the options would expire worthless, thus leaving us a zero payoff at maturity and we would suffer a loss. Therefore, we conclude that we should* ***not*** *buy the options around the peaks of the densities even though the densities indicate that they are underpriced.*

To prevent infinite loss, whenever we sell a call (put) option, we buy back a call (put) option at a higher (lower) strike. This strike is chosen such that the initial cash flow is maximised while the maximum possible loss is smaller than a pre-chosen threshold, which is discussed in detail in Section 8.5. This strike follows from setting the monetary parameters.

8.4 Predicting the direction of movement of the underlying (Steps 8-9)

As mentioned in the last section, we need to predict the direction of movement of the underlying. Since the underlying is a random process, knowing the historical data is not sufficient to determine the direction of future movement. The direction of movement of the underlying may also depend on external data that cannot be deduced from historical data. Investors base their strategy on a combination of external data and historical data. Instead of using external data directly we could refine our strategy using the behaviour of the investors in the market.

Step 8 Estimate the optimal value of the prediction parameter, p.

If the investors are expecting an upwards movement, since they are risk averse, they generally will not anticipate an extreme rise and instead tend to expect moderate upwards movement. Hence the distribution implied by the investors' behaviour will be more pointed and has larger kurtosis. To see this, compare the plots in Figure 8.2.1, the heights of the peaks change noticeably in different market conditions shown in Figure 8.1.1. To determine whether the peak, the skewness or the kurtosis of the option implied distribution has the strongest relationship with the prediction of future direction of movement, we perform the following analysis.

Let $\Delta S_{t_i,t_i+\Delta t}$ be the change in price of the underlying from time t_i to $t_i+\Delta t$, where $i=\{1,2,...,N\}$. Let $x_i^{(1)}$, $x_i^{(2)}$ and $x_i^{(3)}$ be the value of the peak, skewness and kurtosis of the distribution estimated on t_i, respectively. Let ω_b and ω_s be the functions such that

$$\omega_b\left(\{x_1,x_2,...x_N\},x'\right)=\sum_{i=1}^{N}1_{\left\{\mathrm{sign}(\Delta S_{t_i,t_i+\Delta t})=\mathrm{sign}(x_i-x')\right\}}$$

and

$$\omega_s\left(\{x_1,x_2,...x_N\},x'\right)=\sum_{i=1}^{N}1_{\left\{\mathrm{sign}(\Delta S_{t_i,t_i+\Delta t})=\mathrm{sign}(x'-x_i)\right\}},$$

where

$$\mathrm{sign}\left(x\right)=\begin{cases}1 & \text{if } x\geq 0\\ -1 & \text{if } x<0\end{cases}.$$

The function ω_b counts the number of time such that $\Delta S_{t_i,t_i+\Delta t}\geq 0$ and $x_i\geq x'$ at the same time or $\Delta S_{t_i,t_i+\Delta t}<0$ and $x_i<x'$ at the same time. The function ω_s counts the number of time such that $\Delta S_{t_i,t_i+\Delta t}\geq 0$ and $x_i\leq x'$ at the same time or $\Delta S_{t_i,t_i+\Delta t}<0$ and $x_i>x'$ at the same time. Let $o_b^{(j)}$ and $o_s^{(j)}$ be the value such that

$$o_b^{(j)}=\max_{x'}\omega_b\left(\left\{x_1^{(j)},x_2^{(j)},...x_N^{(j)}\right\},x'\right),$$

and

$$o_s^{(j)}=\max_{x'}\omega_s\left(\left\{x_1^{(j)},x_2^{(j)},...x_N^{(j)}\right\},x'\right),$$

for $j=1,2$ or 3. That is, $o_b^{(j)}$ is a value that maximises the number of time $\Delta S_{t_i,t_i+\Delta t}$ and $x_i^{(j)}-o_b^{(j)}$ having the same sign. $o_s^{(j)}$ is a value that maximises the number of time $\Delta S_{t_i,t_i+\Delta t}$ and $o_s^{(j)}-x_i^{(j)}$ having the same sign.

In our analysis, there are 1301 trading days, that is, $N=1301$. The analysis result is given in Table 8.4.1. The result shows that the peak and the kurtosis are relatively better in predicting future direction movements. If we predict the movement will be positive if the kurtosis is greater than 4.904, we get 954 correct predictions out of 1301 days, that is, 73.33% accuracy. Therefore, we should use kurtosis as the prediction parameter, p, and we should choose $p = 4.904$.

	peak	skewness	kurtosis
j	1	2	3
$o_b^{(j)}$	1.238×10^{-3}	-1.430	4.904
$\omega_b\left(\left\{x_1^{(j)}, x_2^{(j)}, ...x_N^{(j)}\right\}, o_b^{(j)}\right)$	946	749	954
$o_s^{(j)}$	2.626×10^{-3}	-0.7009	7.811
$\omega_s\left(\{x_1, x_2, ...x_N\}, o_s^{(j)}\right)$	753	905	752

Table 8.4.1: The prediction results of the future direction of movement using the peak, the skewness and the kurtosis of the option implied distribution.

Recall in Section 8.3, the definitions of the parameters R_L and R_R depend on the prediction on the direction of movement of the underlying. The parameters R_L and R_R determine the strikes at which as should sell the put and call options. In the analysis of the performance of the trading strategy, we assume only options with certain strikes are available to be traded liquidly in the market. On day D, let P_p be the price of the underlying on the previous trading day. We assume the differences between the strikes and P_p must be in multiples of 50, as is the case for options on FTSE 100, see www.euronext.com (2007), where the interval between exercise prices is determined by the time to maturity of a particular expiry month and is either 50 or 100 index points. The minimum and maximum strikes of put options we can trade are assumed to be $P_p - 900$ and $P_p + 250$ respectively. Again this is chosen based on the FTSE 100. The minimum and maximum strikes of call options we can trade are assumed to be $P_p - 150$ and $P_p + 500$ respectively. Since we do not have the information about the bid-ask spreads, we assume that, if an option worth s, it costs us $s + 6$ to buy and we can sell it for $s - 6$. The conservative value of six is determined by comparing with real life option data. In order to find the optimal values of the parameters, R_R and R_L, using optimisation through grid search, we firstly specify a grid of values for them.

Step 9: Define the values of the grids for parameters R_R and R_L.

For each combination of the parameters, we calculate which strikes of put and call options to sell on each trading day D. If the prediction parameter of the option implied density is lower than p, we presume the market is more likely to fall. Hence, we sell call options with strike C such that (8.12) is satisfied, and we sell the put options with strike $\mathrm{P} \leq X_L$ such that (8.13) is satisfied. Similarly, from equations (8.10) and (8.11), if the prediction parameter of the option implied density is higher than p, we presume that the market is more likely to rise. Hence, we sell the put options with strike P such that (8.10) is satisfied and sell call options with strike $\mathrm{C} \geq X_R$ such that (8.11) is satisfied.

Therefore, for each combination of values of R_R, R_L and p, we obtain the strikes of the call and put options to sell.

8.5 Determining the dynamic trading volume to control maximum possible loss (Steps 10-13)

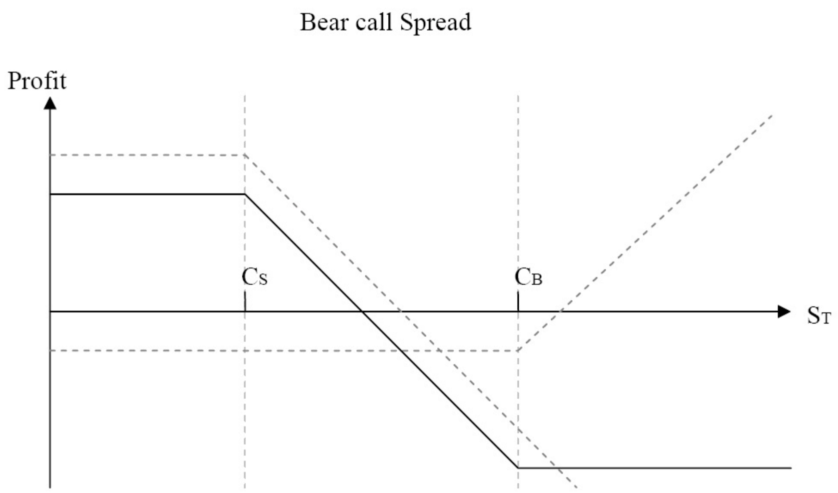

Figure 8.5.1: A bull put spread.

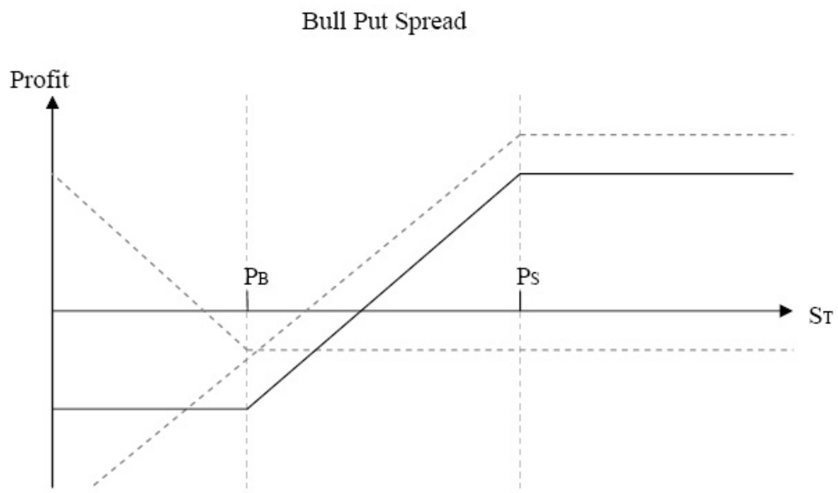

Figure 8.5.2: A bear call spread.

To limit our maximum possible loss at maturity, we buy put (call) options with strike lower (higher) than the ones we sell. In other words, we perform a bear call spread and a bull put spread, which are illustrated in Figures 8.5.1 and 8.5.2. For more information about derivatives strategies, we refer the reader to Hull (2003). For example, if we sell a call option, the payoff at maturity is equal to $\min(C_S - S_T, 0)$. Therefore, our maximum possible loss is unlimited since S_T is unbounded from above. However, if we buy a call option with strike C_B as well, where $C_B > C_S$, our maximum possible loss at maturity is equal to $C_B - C_S$, which is bounded. Similarly, if we sell a put option, the payoff at maturity is equal to $\min(S_T - P_S, 0)$, where S_T is the price of the underlying at maturity and P_S is the strike of our option. Therefore, our maximum possible loss is $-P_S$ since $S_T \geq 0$. However, if we buy a put option with strike P_B as well, where $P_B < P_S$, our maximum possible loss at maturity is equal to $P_B - P_S$.

To determine the strikes of the far out of money strikes and the volumes to be bought, we need four parameters: *Initial Capital*, c, which is the initial capital input to the trade; *Maximum Loss Percentage*, $L_{\max}$, which is the percentage of current capital we can afford to lose; *Minimum Loss Percentage*, $L_{\min}$, which is the percentage of current capital that represents an acceptable loss; *Default Trading Volume*, v, which is the default number of contracts of each type of option we sell. Note that $L_{\max}$, $L_{\min}$ and v are reflecting our level of risk aversion.

Step 10: Define the values of the grids for parameters c, $L_{\max}$, $L_{\min}$ and v.

Step 11: Let c_D be the current level of capital on trading day D. The maximum and minimum amounts we can lose are given by $c_D L_{\max}$ and $c_D L_{\min}$.

Step 12: The strikes of the put and call options to sell are given by Step 9, and we let them be P_S and C_S, respectively. For call options, for each buyable strike, C_B, we calculate the initial cash flow, denoted by $C_i^{(C)}$:

$$C_i^{(C)} = v_D \left(\text{price of call options with strike } C_S - \text{price of call options with strike } C_B\right),$$

where v_D is the trading volume on day D. We set $v_D = v$, the default trading volume and we may change the value of v_D later. The maximum possible loss is denoted by M_L and is the most negative possible payoff at maturity minus the initial cashflow on the trading day:

$$M_L = v_D \left(C_B - C_S\right) - C_i^{(C)} e^{r\tau} + \text{transaction cost involved} \times e^{r\tau},$$

where r is the risk-free interest rate and τ is the time to maturity of the options. The transaction cost involved is discussed later. If M_L is bigger than $cL_{\max}$, we reduce the value of v_D until $M_L \leq cL_{\max}$. Similarly, if M_L is smaller than $cL_{\min}$, we increase the value of v_D until $M_L \geq cL_{\min}$. In reality, the trading volume, v_D, should not be 'too big' as we require a counterparty to trade with us. Therefore, it is not realistic if v_D is unbounded. From the historical trading volumes of the options of FTSE 100, we assume $v_D \leq 1000$. We choose to buy options with strike C_B with its corresponding trading volume such that the initial cashflow, $C_i^{(C)}$ is maximised among all the buyable strikes. Similarly, for put options, for each buyable strike, P_B, the initial cash flow, denoted by $C_i^{(P)}$, is

$$C_i^{(P)} = v_D \,(\text{price of put options with strike } P_S - \text{price of put options with strike } P_B)$$

and the maximum possible loss is

$$M_L = v_D\,(P_S - P_B) - C_i^{(P)} e^{r\tau} + \text{transaction cost involved} \times e^{r\tau},$$

where r is the risk-free interest rate and τ is the time to maturity of the options. We adjust the value of v_D such that $cL_{\min} \leq M_L \leq cL_{\max}$. We choose to buy options with strike P_B with its corresponding trading volume such that the initial cashflow, $C_i^{(P)}$, is maximised among all the buyable strikes.

There are two different kinds of transaction costs for the options of FTSE 100 at the time of this thesis. The first kind is the trading fee, which is fixed at 25p and payable per side per lot. The second is a clearing fee, which is set at 22p per side per upper boundary for client business of £1200 (A 'lot' is defined as the shares in a single transaction).

Step 13: With the value of the underlying at maturity, calculate the profit from the trading portfolio constructed in Step 12. The gain/loss is added to the total capital available at maturity of the options. The amount of capital not used in the trading strategy is assumed to be invested in a risk-free bank account and the interest earned is subject to a tax rate of 40%. Note that the earning from the trading strategy is not subject to tax because of the possible loss at maturity. Interest after tax is added to the total capital. Repeat Steps 11-13 for all trading days under consideration.

8.6 Risk-return analysis to find the optimal set of parameters (Step 14)

In this section, we explain how to obtain the optimal values of the six parameters, R_L, R_R (Section 8.2), p (Section 8.4), c, $L_{\max}$, $L_{\min}$ and v (Section 8.5).

We run our strategy over the grids of values of the parameters and plot their expected returns against their CVaR. We can then obtain an efficient frontier defined in Section 7.4 and decide which parameters we should use. We can separate the parameters into two groups: $\{p, R_L, R_R\}$ relate to the density comparison, and $\{c, v, L_{\max}, L_{\min}\}$ relate to monetary values. Hence, we optimise these parameters separately using grids with predefined step sizes. We first fix some values for $\{c, v, L_{\max}, L_{\min}\}$ and find the efficient frontier for $\{p, R_L, R_R\}$. Then we use the optimal values for $\{p, R_L, R_R\}$ to find the efficient frontier for $\{c, v, L_{\max}, L_{\min}\}$. The reason for separating the two groups is that it reduces the number of iterations needed by a substantial factor.

Recall we conduct the trading strategy on the European call and put options of FTSE 100 to investigate its performance. All transaction costs are taken into account and tax, assumed to be 40%, is deducted from guaranteed profits. We use European option data of 1874 trading days between 29th June 2000 and 23rd November 2007. Since making use of the trends of the historical data series and the option prices is a crucial part of our trading strategy, we cannot resample from the data randomly. Hence, we take all the continuous data series of length 250 (approximate number of business days in a year) from the 1368 data points available. Hence, we get (1368-250+1) = 1119 data samples and over which we calculate the expected CAGRs and CVaRs. The idea behind this is for each trading day, we calculate the value of the portfolio one year later following our trading strategy. Note that the data used are not independent because the time periods are overlapping. However, this is necessary since the length of the data is limited. Nonetheless, we can make use of this dependence as an extra piece of information. For example, if the market is currently falling, we want to take into account the behavior of a falling market and we want to conduct the trade up to a relatively short period of time, say one year, and we predict the market will not have recovered yet, we should use data since the market started to fall until now to perform the risk-return analysis. In the next section, we show performance of the trading strategy over seven years and in different individual market conditions.

8.7 Performance of our trading strategy (Step 15)

In this section, we present and discuss the performance of the trading strategy. We first consider the performance of the strategy throughout different market conditions and then consider the performance during a falling market, a recovering market and a rising market individually. We do not consider a worsening market because the data set considered did not contain such a period. Although the trading strategy has a high earning potential, it is important to note that we assume we are able to trade in a 3-month option **everyday**. In practice, options which are liquidly traded on exchange has fixed maturities. For example, the FTSE options have maturities on the third Fridays of quarterly months. Therefore, we could not actually trade in FTSE options with 3-month maturities everyday in the year. Note that our trading strategy relies on the initial positive cashflow from selling the more expensive options and profit if the options expire worthlessly. However, if we do have to pay out for the options we sold at maturity, the amount would be much bigger than the earning we make from selling the options. In other words, the profit of the trading strategy replying on the fact that most of the time the options expire worthlessly and only occasionally do we result in a loss. Hence, we have to trade in a 3-month option everyday in order to have a bigger chance of making profits. In the current market where an exact 3-month option can only be traded on a few days a year, our trading strategy is not applicable. Nonetheless, over-the-counter (OTC) trading may be available and it is still an important finding. In the following, we first consider the performance of the strategy throughout different market conditions and then consider the performance during a falling market, a recovering market and a rising market individually. We do not consider a worsening market because the data set considered did not contain such a period. We investigate the effect of individual parameters to the risk and return of the strategy in order to have a better understanding of their roles and to facilitate the decision making process of choosing the parameter values. Since the trading strategy introduced in this paper is a new approach in the literature that the readers would not be familiar with, we investigate the effects of each parameter in hope of giving a better idea why we would want to introduce these seven parameters in the first place. The analysis below would show that the values of the parameters have a clear effect on the risk and return of the strategy and vary throughout different market conditions.

8.7.1 Efficient frontier analysis by varying the density parameters

Figure 8.7.1.1 gives the performance of the trading strategy from 2000 to 2007. The point on the far left of the efficient frontier has risk $= -0.2636$ and expected return $= 9.010$, meaning that the expected earning in the 5% worst case scenarios is a gain of 26.36% of the initial investment and the expected return is about 9.010 times the initial investment over a year. Despite the fact that some of the points of the efficient frontier have negative risks, this does not indicate an arbitrage opportunity since we only test the strategy on 7 years historical data and the performance of our strategy strongly depends on the market condition. We have bounded the maximum trading volume in a day to be 1000 units. If the trading volume is not bounded, the profit attainable could be much greater. Note that some points on the efficient frontier has a risk bigger than 1, which means that the worst 5% scenarios can result from losses more than the initial investment we input in our trading strategy. Note that this can be prevented by setting the maximum amount of money we could afford to lose on each day to be the value equal to the current amount of capital available divided by the number of business days until the maturity of the option.

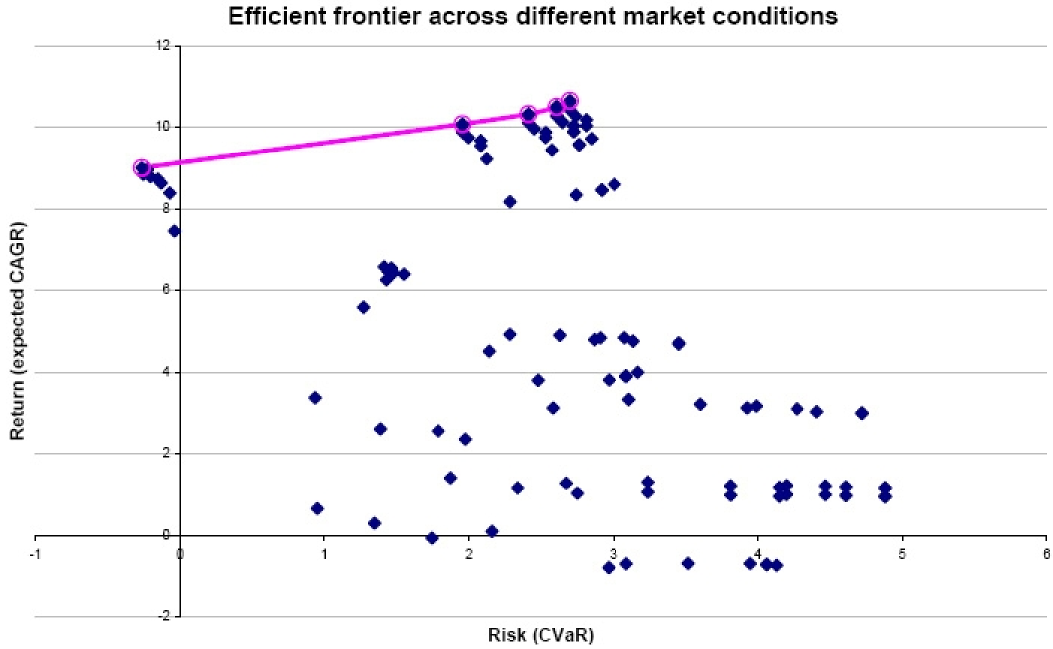

Figure 8.7.1.1: Efficient frontier obtained by fixing $c = 600,000$, $v = 100$, $L_{\min} = 0.033$, $L_{\max} = 0.13$ and varying R_L from 0 to 1 by 0.1 and R_R from 0 to 1 by 0.1.

To gain a better understanding of the contributions of the density parameters to the performance of the strategy, we include the colour-coded versions of Figure 8.7.1.1 in Appendix C.3 as Figures C.3.1-C.3.2. We summarise our observations as follows:

- Distribution of values of R_L in Figure C.3.1
  - As R_L increases from 0 to 0.5, the points move gradually along the efficient frontier from right to left.
  - When $R_L = 0.6$, the points suddenly move downwards and to the right, away from the efficient frontier.
  - As R_L increases from 0.6 to 1, the points move gradually downwards. The sudden jump can be due to the fact that option payoff function is a step function. The higher the value of R_L, the more risky the trading strategy is since it corresponds to selling put options with strikes relatively closer to the current stock price. While we could make a significant profit when $R_L = 0.5$, when $R_L = 0.6$, the strikes of the put options sold may become higher than the stock prices at maturity and we suffer losses.
  - It shows that when $R_L = \{0, 0.1, ..., 0.5\}$, p contributes more to the level of risk.

  - When $R_L = \{0.6, 0.7, ..., 1\}$, R_L contributes more to the level of expected return.

- Distribution of values of R_R in Figure C.3.2
  - As R_R increases from 0 to 1, there is a clear drift upwards until the points hit the efficient frontier at $R_R = 1$.
  - It shows that R_R contributes more to the level of expected return of the strategy.

8.7.2 Efficient frontier analysis by varying the monetary parameters

Figure 8.7.2.1 shows the efficient frontier obtained from varying the monetary parameters on data starting from year 2000 to 2007. The far left point has risk = −4.797 and expected return = 16.38. Note that a proportion of points in the plot have negative expected return and positive risk. This shows the calibration of the momentary parameters of the trading strategy is crucially important.

The colour-coded versions of Figure 8.7.2.1 are included as Figures C.3.3-C.3.6 in Appendix C.3. We summarise our observations as follows:

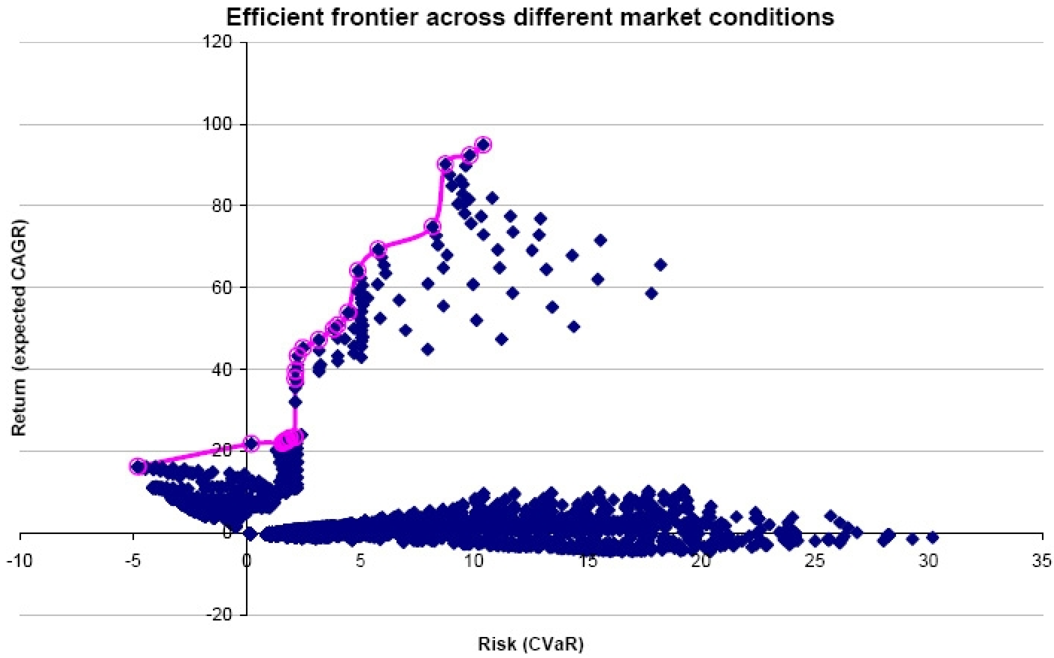

Figure 8.7.2.1: Efficient frontier obtained by fixing $R_L = 0.5$, $R_R = 0.8$ and varying v from 0 to 900 by 100, c from $100,000$ to $4,600,000$ by $500,000$, $L_{\min}$ from 0.001 to 0.091 by 0.01 and $L_{\max}$ from 0.1 to 0.28 by 0.02.

- Distribution of values of v in Figure C.3.3

  - When $v = 0$, both risk and expected return are very close to zero.
  - As v increases from 100 to 200, the risk increases even though the expected return only increases very slightly.
  - As $v = 300$, the points suddenly move up a great deal and to the left. In fact, all the points on the efficient frontier and all the points with negative risks have value $v = 300$.
  - As $v = 400$, the points move back down to near the x-axis and on the right of the y-axis, that is, the risk is positive and the expected return is nearly zero.
  - As v increases from 400 to 900, the points move generally to the right but still stay very close to the x-axis.
  - This shows that the optimal value of the default trading volume is $v = 300$. Further optimisation can be done to find the optimal value in the range between 250 and 350.

- Distribution of values of c in Figure C.3.4

  - When $c = 100,000$, the points distribute as two groups in the plot. The first group stays near to the right part of the efficient frontier, that is, they have high returns and high risks. The second group stays very close to the x-axis and on the right of the y-axis, meaning they have high risks but very small expected returns. All in all, the value $c = 100,000$ is not ideal since it leads to very high risk and the expected return is volatile.
  - When $c = 600,000$, the points move significantly downwards and slightly to the left with some points lying on the efficient frontier.
  - When $c = 1,100,000$, the points spread out a little bit more and some of them lie on the far left part of the efficient frontier.
  - As c increases from 1,100,000 to 4,600,000, the points move gradually downwards and towards the $(0, 0)$ coordinates.
  - This results show that if the initial capital is greater than 1,100,000, it has an adverse effect on the performance.

- Distribution of values of $L_{\min}$ in Figure C.3.5

  - The points can be considered to be distributed into 4 different groups. The first group stay around the left part of the efficient frontier, the second stay very closed together around the middle of the efficient frontier, the third spread out more and stay around the right of the efficient frontier and the fourth group stay around the x-axis on the right of the y-axis.

- In general, the points move gradually upwards and to the right as $L_{\min}$ increases.
- This shows that $L_{\min}$ contributes less to the performance of the strategy compared to other monetary parameters but in general, the risk and expected return increase as $L_{\min}$ increases.

- Distribution of values of $L_{\max}$ in Figure C.3.6
  - The points distribute as four groups in the plot as in the plot of $L_{\min}$ and the points move gradually upwards and to the right as $L_{\max}$ increases.
  - This shows that $L_{\max}$ does not contribute significantly to the risk and return level if we trade throughout different market conditions. Its effect will become more obvious as we discuss the performance in different market conditions.

It is important to note that during the simulation period, the equity market went through three stages: From 29th June 2000 to 1st November 2002, the market was falling; from 2nd November 2002 to 2nd December 2003, the market was recovering; from 3rd December 2003 to 29th January 2007, the market was rising (see Figure 8.1.1). In the following, we optimise the parameters' values separately under these three different market conditions in order to understand the dependency of the parameters on the condition of the market.

8.7.3 Efficient frontiers in a falling market

We test the trading strategy in a falling market on data samples of length 250 (number of business days in a year) with starting dates ranging from 29th June 2000 to 1st November 2002, totally 470 data samples. We repeat the efficient frontier analysis in the previous section and Figure 8.7.3.1 shows the efficient frontier obtained from varying the density parameters. In contrary to the performance of the strategy assessed throughout different market conditions, the strategy can no longer result in a negative risk. In fact, the far left point of the efficient frontier has risk = 0.3978 and expected return = 4.731. The colour-coded versions of Figure 8.7.3.1 are included as Figures C.4.1-C.4.2 in Appendix C.4.

- Distribution of values of R_L in Figure C.4.1
  - As R_L increases from 0 to 0.5, the points move from right to left along the efficient frontier.

  - When $R_L = 0.6$, the points jump downwards to the middle of the plot.
  - As R_L increases from 0.6 to 1, the points move slightly upwards.
  - It shows that R_L contributes more to the level of risk.

- Distribution of values of R_R in Figure C.4.2

  - The points move gradually upwards as R_R increases, showing that R_R contributes more to the level of expected return.

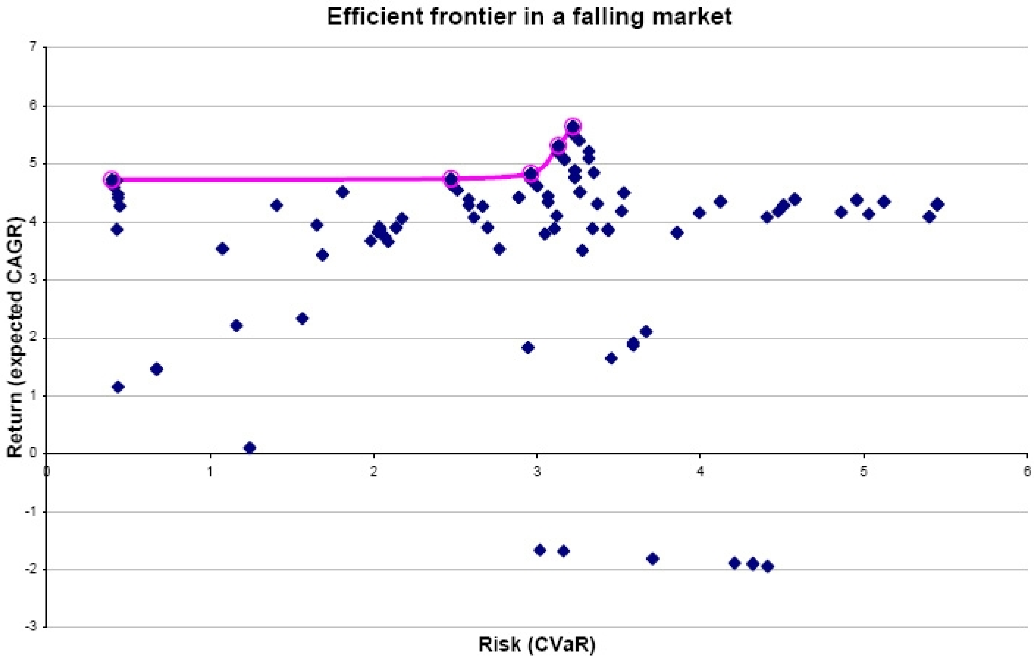

Figure 8.7.3.1: Efficient frontier in a falling market, obtained by fixing $c = 600,000$, $v = 100$, $L_{\min} = 0.033$, $L_{\max} = 0.13$ and varying R_L from 0 to 1 by 0.1 and R_R from 0 to 1 by 0.1.

Figure 8.7.3.2 shows the efficient frontier in a falling market. The shape of the efficient frontier is very similar to the case when we study the performance throughout different market conditions. However, the expected returns achievable are much lower in this case. In the previous case, the point on the far right of the efficient frontier has risk = 10.40 and expected return = 94.89. In a falling market, the point has risk = 14.35 and expected return = 48.54. The points on the far left of the efficient frontiers in the two cases have comparatively more similar values. This shows that if we are looking for high risk and high return investment, we should not only trade when the market is falling.

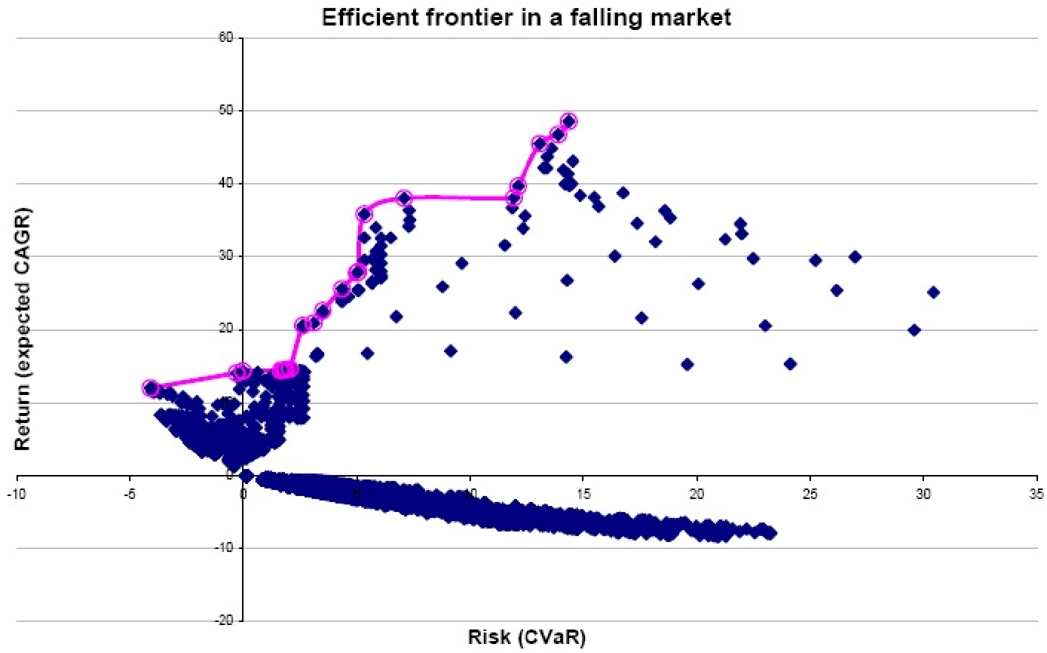

Figure 8.7.3.2: Efficient frontier in a falling market, obtained by fixing $R_L = 0.5$, $R_R = 0.8$ and varying v from 0 to 900 by 100, c from $100,000$ to $4,600,000$ by $500,000$, $L_{\min}$ from 0.001 to 0.091 by 0.01 and $L_{\max}$ from 0.1 to 0.28 by 0.02.

The colour-code versions of Figure 8.7.3.2 are included as Figures C.4.3-C.4.6 in Appendix C.4. The distributions of the points of all the parameters' values are very similar to the case when we trade throughout different market conditions. This shows that when we study the performance throughout different market conditions, the performance is dominated by the losses resulted from trading in a falling market. We will study the performance in a recovering and risking market in the next two sections and see a great improvement to the performance of the trading strategy.

8.7.4 Efficient frontiers in a recovering market

We test the trading strategy in a recovering market on data samples with starting dates ranging from 2nd November 2002 to 2nd December 2003, 200 data samples in total. Figure 8.7.4.1 shows the efficient frontier in a recovering market. The distributions of the points are very different from the results so far. There are only two points on the efficient frontier and they are very close to each other. The risk of the point on the left $= -11.89$ and expected return $= 19.67$. It shows that our trading strategy can lead to great profit if the appropriate parameters are chosen in a recovering market. However, it may be difficult to determine where the market in three month time will be a recovering market or already a rising market. Contrary to the last two analyses, all the points on the efficient frontier have negative risks and high expected returns.

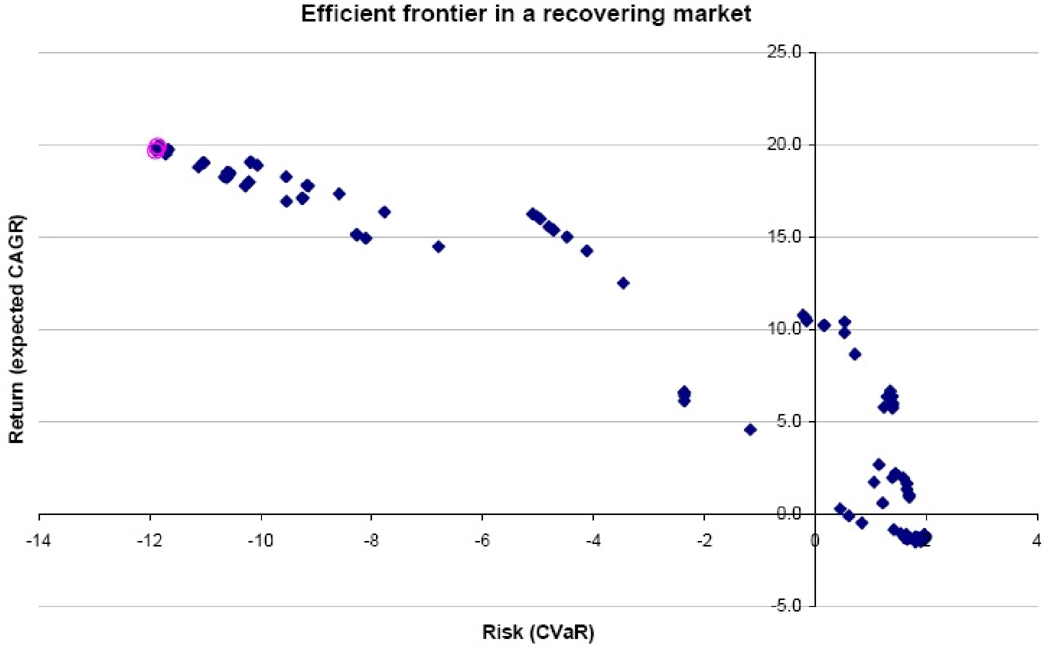

Figure 8.7.4.1: Efficient frontier in a recovering market, obtained by fixing $c = 600,000$, $v = 100$, $L_{\min} = 0.033$, $L_{\max} = 0.13$ and varying R_L from 0 to 1 by 0.1 and R_R from 0 to 1 by 0.1.

Figures C.5.1-C.5.2 show the distributions of the values of the density parameters R_L and R_R, respectively.

- Distribution of values of R_L in Figure C.5.1
  - The movement of the points is different from the analyses throughout different market conditions and in a falling market.
  - The points stay very closed together and near to the efficient frontier when $R_L = \{0, 0.1, 0.2, 0.3\}$.
  - As R_L increases from 0.4 to 1, the points move gradually downwards and to the right.
  - In the last two analyses, the value $R_L = 0.5$ lies on the far left point on the efficient frontier. In this analysis, $R_L = 0.3$ on the left point and $R_L = 0$ on the right point.
  - The three analyses so far show that we should not choose $R_L > 0.5$.

- Distribution of values of R_R in Figure C.5.2
  - As R_R increases from 0 to 1, the points move gradually upwards and to the left until touching the efficient frontier.
  - As in the last two analyses, the point on the efficient frontier has large value of R_R. In the analysis throughout different market conditions, $R_R = 0.8$ or 1. In the analysis for a falling market, $R_R = 0.8$. In this analysis, $R_R = 1$.

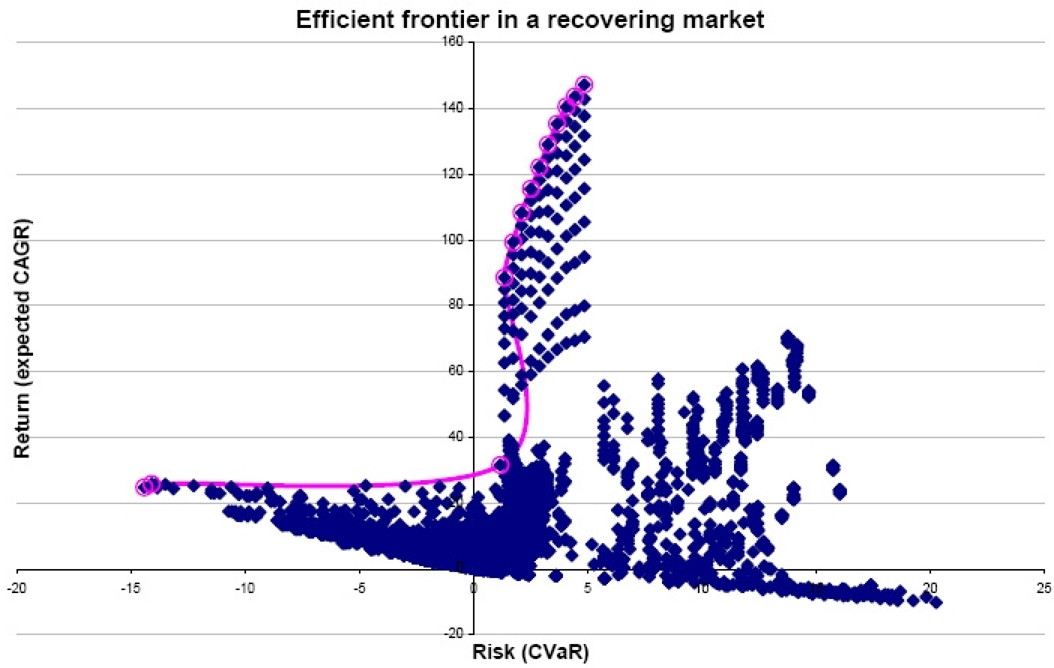

Figure 8.7.4.2: Efficient frontier in a recovering market, obtained by fixing $R_L = 0.5$, $R_R = 0.8$ and varying v from 0 to 900 by 100, c from 100,000 to 4,600,000 by 500,000, $L_{\min}$ from 0.001 to 0.091 by 0.01 and $L_{\max}$ from 0.1 to 0.28 by 0.02.

Figure 8.7.4.2 shows the efficient frontier by varying the monetary parameters in a recovering market. The shape of the efficient frontier and the pattern of the points on the plot are significantly different from the previous two cases. The points on the efficient frontier have much less risks and higher expected returns than before. It shows that trading in a recovering market is much more profitable than trading in a falling market. However, the difficulty is that we have to predict whether we will be in a recovering market 3 months later if the current market is falling.

To gain a better understand of the contributions of the different monetary parameters to the performance of the trading strategy in a recovering market, we include the colour-code versions of Figure 8.7.4.2 in Appendix C.5 as Figures C.5.3-C.5.6.

- Distribution of values of v in Figure C.5.3
  - All the points on the efficient frontier have value $v = 300$, which is the optimal value of the parameter, as in the previous cases.
- Distribution of values of c in Figure C.5.4
  - The points lying on the efficient frontier have three different values for c. The left part of the efficient frontier has value $c = 1,600,000$. The middle has value $c = 600,000$ and the right part has value $c = 100,000$. Since the points with negative risks on the efficient frontier all have $c = 1,600,000$, while the expected return is still very high, we should choose $c = 1,600,000$ in a recovering market.
- Distribution of values of $L_{\min}$ in Figure C.5.5
  - As $L_{\min}$ increases from 0.001 to 0.091, the points generally move upwards. This shows that the parameter $L_{\min}$ contributes more to the level of expected return of the strategy in a recovering market.
- Distribution of values of $L_{\max}$ in Figure C.5.6
  - As $L_{\max}$ increases from 0.1 to 0.28, the points generally move upwards and to the right. This shows that the parameter $L_{\max}$ has a significant contribution to both the levels of expected return and risk in a recovering market.

8.7.5 Efficient frontiers in a rising market

We test the trading strategy on data samples with starting dates ranging from 2nd December 2003 to 23rd November 2007, totally 1007 data samples. We repeat the efficient frontier analysis by varying the density parameters and the monetary parameters separately. Figure 8.7.5.1 shows the efficient frontier in a rising market. The plot look very similar to the one in the analysis for a recovering market, but with a less negative risk and a lower expected return: risk $= -8.616$ and expected return $= 12.09$ for the far left point on the efficient frontier. Similar to the analysis for a recovering market, all the points on the efficient frontier have negative risks and high expected returns.

Figures C.6.1-C.6.2 show the distributions of the values of the density parameters R_L and R_R, respectively.

- Distribution of values of R_L in Figure C.6.1

- The points stay very close together and near the efficient frontier as $R_L = \{0, 0.1, 0.2, 0.3\}$.
- As R_L increases from 0.4 to 1, the points move gradually downwards and to the right.
- It is interesting to note that the point $R_L = 0.3$ lies on the efficient frontier in all of the analyses.

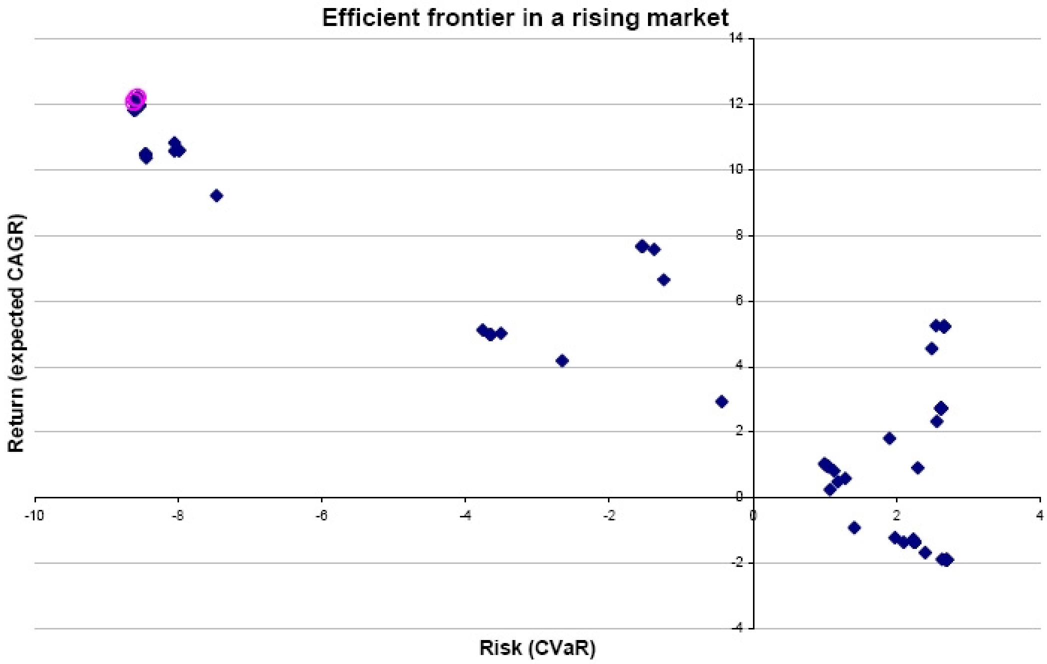

Figure 8.7.5.1: Efficient frontier in a rising market, obtained by fixing $c = 600,000$, $v = 100$, $L_{\min} = 0.033$, $L_{\max} = 0.13$ and varying R_L from 0 to 1 by 0.1 and R_R from 0 to 1 by 0.1.

- Distribution of values of R_R in Figure C.6.2
  - The pattern is very different from the last three analyses so far.
  - It is hard to determine a pattern of movement of points from the plot.
  - The points on the efficient frontier have values $R_R = 0.3$, 0.4 and 0.8, from left to right, contrary to the last three analyses, where the values of R_R on the efficient frontier are always large.
  - It seems that the points overlap each other quite often.

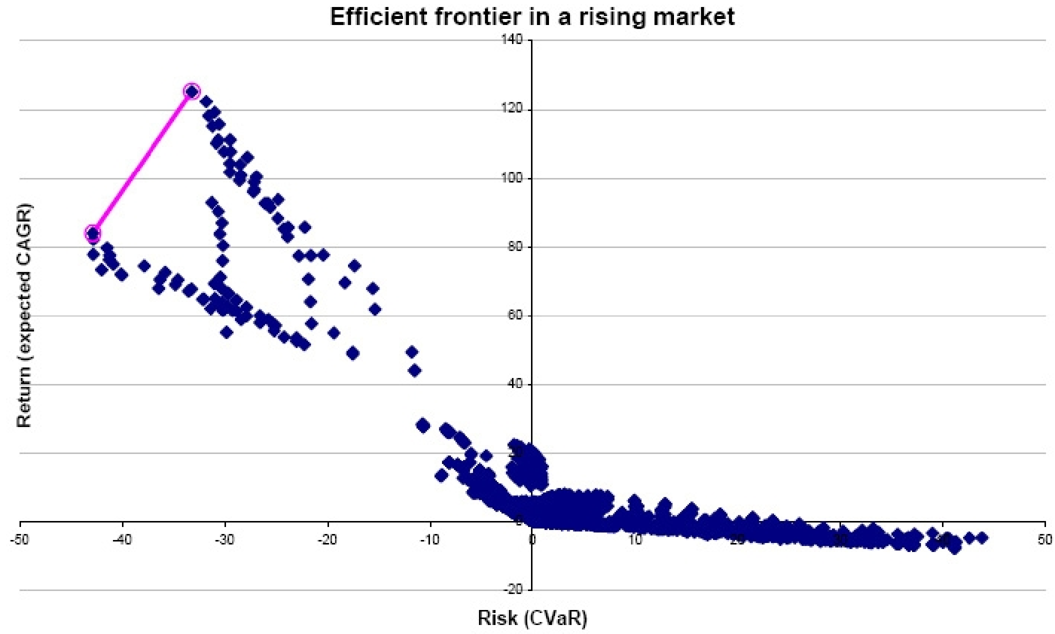

Figure 8.7.5.2: Efficient frontier in a rising market, obtained by fixing $R_L = 0.5$, $R_R = 0.8$ and varying v from 0 to 900 by 100, c from $100,000$ to $4,600,000$ by $500,000$, $L_{\min}$ from 0.001 to 0.091 by 0.01 and $L_{\max}$ from 0.1 to 0.28 by 0.02.

Figure 8.7.5.2 shows the efficient frontier in a rising market by varying the monetary parameters. The shape of the efficient frontier and the pattern of the points in the plot are quite different from the previous cases. There are only two points on the efficient frontier and they have high returns and very negative risks. This shows that it is very profitable to use the trading strategy in a risking market if we choose the right parameters.

To gain a better understanding of the contribution of different monetary parameters to the levels of risk and expected return, we include the colour-coded versions of Figure 8.7.5.2 in Appendix C.6 as Figure C.6.3-C.6.6.

- Distribution of values of v in Figure C.6.3
  - As in the previous cases, it is obvious that $v = 300$ is the optimal value of the trading strategy.
- Distribution of values of c in Figure C.6.4
  - When $c = 100,000$, the points either have positive expected return and negative risk, or stay very close to the x-axis and on the right of the y-axis.
  - As c increases from 100,000 to 4,600,000, the points quickly move towards the coordinates $(0, 0)$ on the plot.

  - The plot shows that the lower the initial capital, the more volatile is the performance of the trading strategy.

- Distribution of values of $L_{\min}$ in Figure C.6.5
  - The points can be considered to be distributed into two groups. The first group stay on the left of the y-axis and spread across the plot. The second group stay on the right of the y-axis and stay very close to the x-axis.
  - The first group of points move upwards and to the right as $L_{\min}$ increases.
  - The second group of points move a little bit downwards and to the right as $L_{\min}$ increases.
  - This shows that $L_{\min}$ is not the dominant parameter in a rising market.

- Distribution of values of $L_{\max}$ in Figure C.6.6
  - As $L_{\max}$ increases from 0.1 to 0.28, the points move gradually upwards and to the left, until hitting the efficient frontier.
  - This shows that the parameter $L_{\max}$ contributes to both the levels of risk and expected return in a rising market.
  - However, there are a large proportions of points with $L_{\max} = 0.28$ staying around the x-axis, showing that $L_{\max}$ is not the dominant parameter in a rising market.

8.7.6 Suggestion for choices of parameters

If we are looking for negative risks and high returns, from the analysis performed in this chapter so far, we should choose:

Market condition	R_L	R_R	Risk	Expected return
General	0.5	0.8	−0.26	9.01
Falling	0.5	0.8	0.40	4.73
Recovering	0	1	−11.86	19.91
Rising	0.3	0.8	−8.57	12.21

Market condition	v	c	$L_{\min}$	$L_{\max}$	Risk	Expected return
General	300	1,100,000	0.091	0.28	−4.797	16.38
Falling	300	600,000	0.071	0.16	−0.02097	14.34
Recovering	300	1,100,000	0.081	0.28	−14.09	25.79
Rising	300	100,000	0.091	0.28	−33.22	125.3

We see that some parameters' values are more sensitive than others. If we are very risk averse and want to invest with the most negative risk, we should choose:

Market condition	R_L	R_R	Risk	Expected return
General	0.5	0.8	−0.26	9.01
Falling	0.5	0.8	0.40	4.73
Recovering	0.3	1	−11.89	19.67
Rising	0.3	0.3	−8.62	12.09

Market condition	v	c	$L_{\min}$	$L_{\max}$	Risk	Expected return
General	300	1,100,000	0.091	0.28	−4.797	16.38
Falling	300	1,100,000	0.091	0.28	−4.100	11.98
Recovering	300	1,100,000	0.071	0.28	−14.41	24.78
Rising	300	100,000	0.041	0.28	−42.85	84.03

8.8 Combining with portfolio insurance

The analysis so far shows that while the trading strategy has a high earning potential, it can also lead to losses. This kind of strategy may not appeal to investors who are looking for safe capital returns. In this section we aim to combine our trading strategy with portfolio insurance strategies to ensure guaranteed returns. Portfolio insurance is designed to give investors the ability to limit downside risk while allowing some participation in upside markets.

The traditional portfolio insurance strategies invest in risk-free bonds and a risky asset. In this chapter, we give a brief introduction of the strategies and comment on the proper choices of parameters' values from the efficient frontiers generated in Section 8.7.

8.8.1 Guaranteed capital return

Let $\{P_t, 0 \leq t \leq T\}$ be the value of our capital. We want to invest in risk-free bonds and a risky asset such that at time T, we have $P_T \geq P_t$. In other words, the capital return is guaranteed. The most basic strategy is to invest in a zero coupon bond (costing Z_t, say) which will payout P_t at time T and invest the amount $(P_t - Z_t)$ in a risky asset. Therefore, even if the entire investment in the risky asset is lost, we will still get P_t at time T from the zero coupon bond. In our case, the risky asset will be the portfolio determined by our trading strategy. Note that in our case, the loss from the risky asset can be more than 100% of the initial investment into the strategy. However, this is unlikely if we choose the correct parameters' values. If we want to eliminate the risk of losing more than 100% of the initial investment, we can set the maximum amount of money we could afford to lose on each day to be the value equal to the current amount of capital available divided by the number of business days until the maturity of the option.

Given this guaranteed capital return, we can be risk-seeking when choosing the parameters' values from the efficient frontiers. So far, we have chosen parameters' values with negative risks and moderate returns. Combined with guaranteed capital return, we can choose parameters' values with moderate risks and very large returns. It is sensible since the amount to be invested in the trading portfolio, $(P_t - Z_t)$, will be comparatively small. We should make use of the high earning potential of the trading strategy and huge profits may be returned using such small investment.

8.8.2 Constant proportion portfolio insurance (CPPI)

The constant proportion portfolio insurance (CPPI) was introduced by Perold (1986) (see also Perold & Sharpe (1988)) for fixed-income instruments and by Black & Jones (1987) for equity instruments. CPPI is a capital guaranteed derivative security that embeds a dynamic trading strategy in order to provide participation to the performance of a certain underlying. In order to be able to guarantee the capital invested, the option writer (option seller) needs to buy a zero coupon bond and use the proceeds to get the exposure he wants. While in the case of Section 8.8.1, the client would only get the remaining proceeds (or initial cushion) invested in a risky asset, bought once and for all, the CPPI provides leverage through a multiplier. For example, say an investor has a P_t portfolio, a floor of Z_t (price of the bond to guarantee his P_t at maturity) and a multiple of $\alpha \in \mathbb{R}^+$. Then on day one, the writer will allocate $\alpha\,(P_t - Z_t)$ to the risky asset and the remaining $P_t - \alpha\,(P_t - Z_t)$ to the risk-free asset (the bond). The exposure will be revised as the portfolio value changes, that is, when the risky asset performs and with leverage multiplies by α the performance (or vice versa). These rules are predefined and agreed once and for all during the life of the product. In our case, the risky asset will be the portfolio of our trading strategy.

The higher the multiple, α, the more we will earn when the trading strategy results in a gain. However, the higher the multiple, the faster the portfolio will approach the floor when there are sustained losses resulting from the trading strategy. The exposure approaches zero as the cushion approaches zero. This keeps the portfolio value from falling below the floor in continuous time. However, it will fall below the floor when there is a very sharp loss before the investor has a chance to trade.

Given this extra protection, we can be less risk averse when choosing the parameters' values from the efficient frontiers. In Section 8.7, we have chosen parameters' values with negative risks and moderate returns. Combined with CPPI, we should choose the parameters' values of the points with moderate risks (risk ≤ 1) and high returns on the efficient frontiers:

Market condition	R_L	R_R	Risk	Expected return
General	0.5	0.8	−0.26	9.01
Falling	0.5	0.8	0.40	4.73
Recovering	0	1	−11.86	19.91
Rising	0.3	0.8	−8.57	12.21

Market condition	v	c	$L_{\min}$	$L_{\max}$	Risk	Expected return
General	300	600,000	0.091	0.1	0.1861	21.88
Falling	300	600,000	0.071	0.16	−0.02097	14.34
Recovering	300	1,100,000	0.081	0.28	−14.09	25.79
Rising	300	100,000	0.091	0.28	−33.22	125.3

Summary of Part III

The skewness and kurtosis trades based on the Black-Scholes model in the literature were motivated by the fact that the density function of the underlying at maturity implied by historical data is different from that implied by current option prices. However, according to the Black-Scholes model, the two density functions should be identical, as in the Black-Scholes world, there is an unique martingale measure. This problem can be solved by applying a Lévy market model. When the underlying asset is driven by a Lévy process, there are infinitely many martingale measures and hence we would expect deviations of the density functions. In fact, from our analysis, we observed that the densities implied from current option prices always have fatter tails than those implied from historical data. This can be caused by the risk aversion of the investors in the market. While the skewness and kurtosis trades in the literature only involve calculating the values of the skewness and kurtosis of the two distributions, we extend these ideas forward to develop a quantitative and dynamic trading strategy which makes use of the deviations of the distributions to obtain the optimal strikes of the options to be traded and at the same time takes into account the practical issues such as the maximum and minimum capital allowed to use in a single trade and the trading volume.

We used the Variance Gamma (VG) market model and calibrated two sets of parameters: the first set was obtained from analysing the historical data series of the underlying and the second set was obtained from applying fractional Fast Fourier Transform methods on option prices. With these two sets of parameters, we then simulated forward to get the density functions of the underlying at maturity. Based on the deviation of the two density functions, we constructed a dynamic trading strategy and reported its performance on European options of FTSE 100. The performance of the trading strategy under different market conditions were reported. It is important to note that in our performance analysis, the historical option prices used are implied from a volatility database. The actual prices and availability of the options traded in our performance checking routines may not be accurate due to lack of actual data on historical bid-ask spreads and trading volumes. However, we have introduced bid-ask spreads and maximum trading volume assumptions which are set to be more strict than real life circumstances. In practical situations, the trading strategy can be adjusted to real life data and due to the risk aversion of investors, overall profits are still expected. Simulation results showed that the trading strategy has a high earning potential.

Chapter 9

Conclusion of the thesis

This thesis contributes to the development of the theoretical representation and financial applications of Lévy processes. The thesis is composed of three parts. The first part presents an explicit formula for the chaotic representation of powers of increments of Lévy processes. The second part provides a prefect hedging strategy for European and exotic options in a Lévy market. The third part gives a speculating option trading strategy in a Lévy market.

In the first part, we presented the explicit formulae for the chaotic representation in terms of orthogonalised compensated power jump processes and the chaotic representation in terms of the Poisson random measure. The result is important since it enables the calculation of the chaotic representations of some common kinds of contingent claims traded in the market. Unfortunately, the results are constricted to the kind of Lévy functionals which are powers of increments of the underlying Lévy process. The derivation of an explicit formula for general Lévy functions directly is difficult. In the derivation of our result, we applied the Itô formula to the powers of increments of the underlying Lévy process, $(X_{t+\Delta t} - X_t)^n$. Since the derivatives of the $(X_{t+\Delta t} - X_t)^n$ can be calculated readily, we could write out the expression of the CRP of $(X_{t+\Delta t} - X_t)^n$ for $n = 2, 3, 4$ and find the general pattern of the CRP. We then proved the result using induction. However, for general Lévy functionals, F, the form of the derivatives of F is unknown and hence we could not write out the CRP of F using Itô's formula. To solve this problem in the case where F is a general Lévy functional which is an analytic function of the Lévy process, we suggested using a Taylor expansion to express F in terms of an infinite sum of the powers of increments of the underlying Lévy process such that our formula can be applied. This approach firstly decomposes F into a sum of terms contributed by $(X_{t+\Delta t} - X_t)^n$ for $n = 1, 2, 3, ...$ The unknown values of the derivatives of F only come through the derivative terms of the Taylor expansion. We could then apply our explicit formula to $(X_{t+\Delta t} - X_t)^n$ for $n = 1, 2, 3, ...$ and obtain the CRP of F.

The second part of the thesis gives the perfect hedging strategies of European and exotic options in a Lévy market using Taylor expansions. Firstly, we derived the hedging strategies which invest in some higher moments derivatives, namely, the power jump assets introduced by Corcuera *et al.* (2005) and the moment swaps introduced by Schoutens (2005). Note that a variance swap is a special case of a moment swap and is frequently traded in the (over-the-counter) OTC market. The power jump assets and moment swaps are not always traded liquidly in the market, this fact limiting the application of our results. To remedy this problem, we extended the delta and gamma hedges commonly used by traders in the market to higher moment hedges. By doing this, we could hedge by investing in other traded derivatives depending on the same underlying asset. We also derived the minimal variance portfolios, corresponding to making an investment in a risk-free bank account, the underlying asset and (potentially) variance swaps. Numerical procedures were provided and difficulties with the numerical implementation were discussed. The performance of the trading strategies was investigated. We found that perfect hedging is possible as long as we could invest an exact specified amount in as many higher moment financial derivatives as required to hedge. However, it is time-consuming to calculate the derivatives of the option pricing function. We suggested building a look-up table which can be used to calculate the derivatives of all functions and such table would speed up the calculation a great deal. Future work is required to design more efficient implementations of the strategies and to compute the results in a reasonable amount of time for real time application.

In the current market condition, we suspect there might not be much interest in trading in moment swaps and power jump assets. It is because they are extremely risky and investors are not willing to take on such risk. This thesis provides a motivation for trading in such moment swaps and power jump assets. Traditionally, options are only hedged by investing in a risk-free bank account and the underlying stock or by investing in a portfolio of simpler options. However, when we move away from the Black-Scholes world into a market driven by Lévy processes, there are an infinite number of higher moments terms in the expansions of the option pricing formulae. In the Black-Scholes model, the option pricing formula of an European option does not involve any moments higher than the first one. However, in reality, traders use not only the delta of the portfolio but also the gamma, which corresponds to the second moment term in the Taylor expansion of the price of the portfolio. The Lévy market model allows the price formula of the portfolio to have an infinite number of higher moment terms in the Taylor expansion. This motivates the usage of higher moment assets, such as moment swaps and power jump assets to be used in option hedging. These financial derivatives are of practical importance since the

i-th moment swap and the i-th power jump asset depend only on the i-th moment of the underlying asset. They allow the decomposition of the risk of a portfolio of options into different order of moments, which can then be hedged separately.

The third part of the thesis proposes an option trading strategy based on the deviations of the density implied by the historical data of the underlying and that implied by the current option prices in the market. Numerical implementation was discussed in details and the performance of the trading strategy under different market conditions was investigated. Simulation results showed that the trading strategy has a high earning potential. Nonetheless, in the current market, we cannot benefit from our trading strategy. It is because the strategy requires us to be able to trade in some particular three-month options everyday in a year. Options liquidly traded on the exchange have fixed maturities. For FTSE 100 European options, the maturity dates are the third Fridays of the delivery months. Therefore, we cannot trade in a three-month options liquidly on exchange everyday as the strategy requires. The reason why trading in three-month options everyday is necessary to make profit is that the strategy makes a profit by selling over-priced options and buying far out of money options to avoid infinite loss. If the options expire worthlessly, we profit from the sale of the over-priced options. However, if the options we sold were in the money, we would suffer losses much larger than the amount we earned from selling the options. From the analysis on the performance of the strategy, we could see that if we choose the correct parameters' values, the strategy can lead to huge profits. It is because we rely on the fact that the options expire worthless most of the time. Even though we would lose money occasionally, the losses would be covered by the sum of the small profits earned most of the time throughout the year. This is the reason why we have to invest in the same kind of options everyday. Although we cannot trade in three-month options everyday on exchange, such options maybe available in the OTC market and our trading strategy can then be applied.

From our back-testing of the performance of the trading strategy, we could see that high profits are attainable if the correct parameter's values are chosen. The analyses for individual market conditions showed that the performance of the strategy throughout different market conditions is dominated by the performance in a falling market, where the performance is the worst compared to a recovering or a rising market. Therefore, if the current market is recovering or rising, the parameter's values obtained from the analysis throughout different market conditions do not actually lead to the best performance of the strategy. Unless we would like to invest in the strategy for a long time such that the market goes through all the market conditions (seven years in our case), we should choose the parameters' values according to the predicted market condition in three months time. However, this is difficult and we often suffer from losses during unexpected changes of market conditions three months later. Note that the strategy can lead to losses bigger than 100% of the initial investment, that is, we may have to pay more than we invest. Therefore, we should be extremely careful when choosing the parameters' values. We suggested choosing the values which result in profits with minimal risk in all the analyses provided in this part. Those points may not lie on any of the efficient frontiers of the plots but they increase the chance that the strategy would lead to moderate returns and avoid losses. Alternatively, we could set the maximum amount of money we could afford to lose on each day to be the value equal to the current amount of capital available divided by the number of business days until the maturity of the option.

At the time of the production of this thesis, the Black-Scholes model is still the most popular model among practitioners in spite of its well-known drawbacks, for example, the volatility smile. Lévy models are gaining popularity since they can handle the skewness and kurtosis of the market data. However, there are still practitioners who would avoid the use of Lévy models and prefer considering extensions of the Black-Scholes model in other directions, such as using stochastic volatility driven by Brownian motion. The reluctance to move from the Black-Scholes world is partly due to the fact that some of the nice properties of the model are lost when we try to generalise it to include jumps. For example, if the stock price process is driven by a Brownian motion, the log returns of the stock price are normally distributed. Suppose we have simulated a set of realisations of the stock price processes, with parameters μ_1 and σ_1. If later we want to change the parameters to μ_2, σ_2, instead of having to simulate a new set of realisations, we could use the scaling property of normal random variables: If $X \sim N\left(\mu_X, \sigma_X^2\right)$,

$$Y = \left(\frac{X - \mu_X}{\sigma_X}\right) \sigma_Y + \mu_Y$$

is also normally distributed with parameters, $N\left(\mu_Y, \sigma_Y^2\right)$. These nice scaling properties are generally lost in the Lévy market model.

Another nice property of normally distributed random variables is that if X and Y are normally distributed and independent of each other, $Z = X + Y$ is also a normal random variable. This property is very useful in the calibration of the model on options data across time. Suppose we have options data maturing in 1 year and 2 years. We assume the log returns of the stock price process from t_0 to t_1 is normally distributed with parameters $\mu_{0,1}$, $\sigma_{0,1}$ and those from t_1 to t_2 has distribution $N\left(\mu_{1,2}, \sigma_{1,2}\right)$. Assume the log returns of the stock price process from t_0 to t_2 has distribution $N\left(\mu_{0,2}, \sigma_{0,2}\right)$. We can calibrate the values of $\left(\mu_{0,1}, \sigma_{0,1}\right)$ and $\left(\mu_{0,2}, \sigma_{0,2}\right)$ using option data with maturities of 1 year and 2 year, respectively. The values of $\left(\mu_{1,2}, \sigma_{1,2}\right)$ are obtained by $\mu_{1,2} = \mu_{0,2} - \mu_{0,1}$

and $\sigma_{1,2} = \sqrt{\left(\sigma_{0,2}^2 - \sigma_{0,1}^2\right)/2}$. Without the additive property of the random variable, we would have to input the parameter values for time t_0 to time t_1 into the calibration of parameters for time t_1 to time t_2. It is because the option data with maturities of 2 years can only be used to obtain parameters for time t_0 to time t_2. The log return of the stock price process from time t_0 to time t_2 is the sum of the log return from time t_0 to time t_1 and the log return from time t_1 to time t_2. Therefore the parameter values calibrated for time t_0 to time t_1 must be input to the calibration of parameters for time t_1 to time t_2. As a result, any approximation errors in the calibration of time t_0 to time t_1 will be propagated into the calibration of time t_1 to time t_2.

Among all the nice properties of the Black-Scholes model, the closed forms of the option pricing formulae probably contribute greatly to its popularity. Calculation of the option price is quick and easily implemented. In the Lévy market models, option pricing formulae can no longer be expressed in such an elegant closed form. Carr & Madan (1999) developed an option pricing method using the Fast Fourier Transform (FFT) given the characteristic function of the distribution of the log return random variables. Chourdakis (2005) further extended it using the Fractional Fast Fourier Transform (FRFT) and the computation time has been much reduced. These developments in option pricing under Lévy market models are extremely important and contribute to the increasing popularity of the Lévy model. With these option pricing methods, Lévy market models are no longer abstract mathematical objects but practical financial models that provide accurate calibration tools. Although, the option pricing methods for Lévy market models are much more difficult to implement than the Black-Scholes model, the computation time required is still reasonably short.

Though some of the nice properties of the Black-Scholes model are lost if we employ more advanced financial models, it is inevitable since problems like the volatility smile and skewness and kurtosis of the log returns of the price processes have to be resolved in order to get more accurate pricing and financial modelling. Lévy driven models are a natural choice since they are simple extensions to the Black-Scholes model obtained by replacing the Brownian motions with its generalisation with jumps, that is, Lévy processes. Lévy models can fit the market option prices across different strikes and the skewness and kurtosis of the log return distributions are controlled by extra parameters of the models. In this thesis, we investigated the chaotic representations of the Lévy processes and presented some hedging and trading strategies of options under the Lévy market model. This thesis contributed both to the theoretical aspect and practical financial applications of the Lévy market models.

Part IV

Appendices

Appendix A

Part I

A.1 A note on the Nualart and Schoutens representation

Nualart & Schoutens (2000) derived the basic result for representing $(X_{t+t_0} - X_{t_0})^k$ when $t_0 = 0$ and $k = 2$. In the proof of the CRP, Nualart & Schoutens (2000) made use of Proposition 2 in their paper, given in Proposition 3.0.1 in this thesis, and the following equation derived from the Itô formula (equation (5) in Nualart & Schoutens (2000)):

$$
\begin{aligned}
&(X_{t+t_0} - X_{t_0})^k \\
&= \frac{\sigma^2}{2} k (k-1) \left((X_{t+t_0} - X_{t_0})^{k-2} t - \int_0^t s \ \mathrm{d} (X_{s+t_0} - X_{t_0})^{k-2} \right) && \text{(A.1)} \\
&\quad + \sum_{j=1}^{k} \binom{k}{j} \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \,\mathrm{d}Y_s^{(j)} + \sum_{j=1}^{k-1} \binom{k}{j} m_j t \,(X_{t+t_0} - X_{t_0})^{k-j} && \text{(A.2)} \\
&\quad - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_{t_0}^{t+t_0} s \ \mathrm{d} (X_s - X_{t_0})^{k-j} + m_k t. && \text{(A.3)}
\end{aligned}
$$

There is a small inaccuracy in this equation and we provide the corrected one necessary for the derivation of the explicit formula. The second term in (A.2) should be

$$
\sum_{j=1}^{k-1} \binom{k}{j} m_j \,(t + t_0)\, (X_{t+t_0} - X_{t_0})^{k-j}
$$

rather than $\sum_{j=1}^{k-1} \binom{k}{j} m_j t \,(X_{t+t_0} - X_{t_0})^{k-j}$. The error propagates from equation (4) in Nualart & Schoutens (2000). By integration by parts, $\sum_{j=1}^{k} \binom{k}{j} m_j \int_{t_0}^{t+t_0} (X_{s-} - X_{t_0})^{k-j} \,\mathrm{d}s$ should give

$$
\sum_{j=1}^{k-1} \binom{k}{j} m_j \,(t + t_0)\, (X_{t+t_0} - X_{t_0})^{k-j} - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_{t_0}^{t+t_0} s \ \mathrm{d} (X_s - X_{t_0})^{k-j} + m_k t
$$

rather than the term

$$\sum_{j=1}^{k-1}\binom{k}{j}m_j t\left(X_{t+t_0}-X_{t_0}\right)^{k-j}-\sum_{j=1}^{k-1}\binom{k}{j}m_j\int_{t_0}^{t+t_0} s\ \mathrm{d}\left(X_s-X_{t_0}\right)^{k-j}+m_k t$$

stated in Nualart & Schoutens (2000, p.114). Omitting t_0 makes the constant term of the representation not equal to the expectation of $\left(X_{t+t_0}-X_{t_0}\right)^k$ since it depends on t_0. Equation (5) in Nualart & Schoutens (2000) should in fact be:

$$\begin{aligned}
&\left(X_{t+t_0}-X_{t_0}\right)^k \\
&= \frac{\sigma^2}{2}k\left(k-1\right)\left(\left(X_{t+t_0}-X_{t_0}\right)^{k-2}t-\int_0^t s\ \mathrm{d}\left(X_{s+t_0}-X_{t_0}\right)^{k-2}\right) && \text{(A.4)}\\
&\quad+\sum_{j=1}^{k}\binom{k}{j}\int_{t_0}^{t+t_0}\left(X_{s-}-X_{t_0}\right)^{k-j}\mathrm{d}Y_s^{(j)}+\sum_{j=1}^{k-1}\binom{k}{j}m_j\left(t+t_0\right)\left(X_{t+t_0}-X_{t_0}\right)^{k-j} && \text{(A.5)}\\
&\quad-\sum_{j=1}^{k-1}\binom{k}{j}m_j\int_{t_0}^{t+t_0} s\ \mathrm{d}\left(X_s-X_{t_0}\right)^{k-j}+m_k t. && \text{(A.6)}
\end{aligned}$$

Let $G=\{G_t, t\geq 0\}$ be a pure jump Lévy process with no Brownian part (that is, $\sigma^2=0$), $G^{(i)}=\left\{G_t^{(i)}, t\geq 0\right\}$ be its i-th power jump process and $\widehat{G}^{(i)}=\left\{\widehat{G}_t^{(i)}, t\geq 0\right\}$ be its i-th compensated power jump process. As an illustration of this representation, we derive $\left(G_{t+t_0}-G_{t_0}\right)^2$ using (A.1)-(A.3) to inspect the constant terms. Since $\sigma^2=0$, the terms in (A.1) are equal to zero. We have

$$\begin{aligned}
\left(G_{t+t_0}-G_{t_0}\right)^2 &= 2\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+2m_1 t\left(G_{t+t_0}-G_{t_0}\right)\\
&\quad -2m_1\int_{t_0}^{t+t_0} t_1\ \mathrm{d}\left(G_{t_1}-G_{t_0}\right)+m_2 t\\
&= 2\int_{t_0}^{t+t_0}\left[\left(\widehat{G}_{t_1-}^{(1)}-\widehat{G}_{t_0}^{(1)}\right)+m_1\left(t_1-t_0\right)\right]\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}\\
&\quad +2m_1 t\left[\left(\widehat{G}_{t+t_0}^{(1)}-\widehat{G}_{t_0}^{(1)}\right)+m_1 t\right]-2m_1\int_{t_0}^{t+t_0} t_1\ \mathrm{d}\left[\widehat{G}_{t_1}^{(1)}+m_1 t_1\right]+m_2 t\\
&= 2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+2m_1\left(t-t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +m_1^2t^2+m_2 t-2m_1^2 t t_0.
\end{aligned}$$

The expectation of $2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+2m_1\left(t-t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}$ is zero since the compensated processes $\widehat{G}_t^{(1)}$ and $\widehat{G}_t^{(2)}$ have zero means. We see that $m_1^2t^2+m_2t-2m_1^2tt_0$ depends on t_0 which in fact cannot be the expectation of $\left(G_{t+t_0}-G_{t_0}\right)^2$ since the increments of G_t are stationary. Starting from (A.4)-(A.6), we can find that

$$\left(G_{t+t_0}-G_{t_0}\right)^2=2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+2m_1 t\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+m_1^2t^2+m_2 t, \quad \text{(A.7)}$$

where the detailed derivation is given in Appendix A.2.1.

A.2 Calculation of $(G_{t+t_0} - G_{t_0})^k$ for $k = 2, 3, 4$ when $\sigma = 0$

A.2.1 $(G_{t+t_0} - G_{t_0})^2$

Starting from (A.4)-(A.6), we have

$$
\begin{aligned}
&(G_{t+t_0} - G_{t_0})^2 \\
&\quad = 2\int_{t_0}^{t+t_0} (G_{t_1-} - G_{t_0})\, \mathrm{d}\widehat{G}_{t_1}^{(1)} + \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(2)} + 2m_1 (t+t_0)(G_{t+t_0} - G_{t_0}) \\
&\quad\quad -2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}(G_{t_1} - G_{t_0}) + m_2 t \\
&\quad = 2\int_{t_0}^{t+t_0} \left[\widehat{G}_{t_1-}^{(1)} - \widehat{G}_{t_0}^{(1)} + m_1 (t_1 - t_0)\right] \mathrm{d}\widehat{G}_{t_1}^{(1)} + \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(2)} + m_2 t \\
&\quad\quad +2m_1 (t+t_0)\left[\widehat{G}_{t+t_0}^{(1)} - \widehat{G}_{t_0}^{(1)} + m_1 t\right] - 2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[\widehat{G}_{t_1}^{(1)} + m_1 t_1\right] \\
&= 2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} + 2m_1 \int_{t_0}^{t+t_0} (t_1 - t_0)\, \mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad + \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(2)} + 2m_1 (t+t_0)\left[\widehat{G}_{t+t_0}^{(1)} - \widehat{G}_{t_0}^{(1)}\right] + 2m_1^2 (t+t_0)\, t \\
&\quad -2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} - m_1^2 \left[t_1^2\right]_{t_0}^{t+t_0} + m_2 t \\
&= 2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} + 2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} - 2m_1 t_0 \left[\widehat{G}_{t+t_0}^{(1)} - \widehat{G}_{t_0}^{(1)}\right] \\
&\quad + \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(2)} + 2m_1 (t+t_0)\, \widehat{G}_{t+t_0}^{(1)} - 2m_1 (t+t_0)\, \widehat{G}_{t_0}^{(1)} \\
&\quad +2m_1^2 (t+t_0)\, t - 2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} - m_1^2 t^2 - 2m_1^2 t t_0 + m_2 t \\
&= 2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} + 2m_1 t \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(1)} + \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(2)} + m_1^2 t^2 + m_2 t.
\end{aligned}
$$

A.2.2 $(G_{t+t_0} - G_{t_0})^3$

For $k = 3$, we can start from (A.4)-(A.6)

$$
\begin{aligned}
(G_{t+t_0} - G_{t_0})^3 \quad = \quad & 3\int_{t_0}^{t+t_0} (G_{t_1-} - G_{t_0})^2 \mathrm{d}\widehat{G}_{t_1}^{(1)} + 3\int_{t_0}^{t+t_0} (G_{t_1-} - G_{t_0})\, \mathrm{d}\widehat{G}_{t_1}^{(2)} \\
& + \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}_{t_1}^{(3)} + 3m_1 (t+t_0)(G_{t+t_0} - G_{t_0})^2 \\
& +3m_2 (t+t_0)(G_{t+t_0} - G_{t_0}) - 3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}(G_{t_1} - G_{t_0})^2
\end{aligned}
$$

$$-3m_2\int_{t_0}^{t+t_0} t_1 \mathrm{d}\left(G_{t_1}-G_{t_0}\right)+m_3 t.$$

Accordingly, let

$$\left(G_{t+t_0}-G_{t_0}\right)^3 = I_1+I_2+I_3+I_4+I_5+I_6+I_7+m_3 t.$$

Firstly,

$$\begin{aligned}
I_1 &= 3\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)^2 \mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= 3\int_{t_0}^{t+t_0}\left[2\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}+2m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right. \\
&\quad \left.+\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}+m_1^2\left(t_1-t_0\right)^2+m_2\left(t_1-t_0\right)\right]\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= 6\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+6m_1\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\left(t_1-t_0\right)\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad +3\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+3m_1^2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)^2\mathrm{d}\widehat{G}_{t_1}^{(1)}+3m_2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\mathrm{d}\widehat{G}_{t_1}^{(1)}. \\
I_2 &= 3\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)\mathrm{d}\widehat{G}_{t_1}^{(2)} = 3\int_{t_0}^{t+t_0}\left(\widehat{G}_{t_1-}^{(1)}-\widehat{G}_{t_0}^{(1)}+m_1\left(t_1-t_0\right)\right)\mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&= 3\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)}+3m_1\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\mathrm{d}\widehat{G}_{t_1}^{(2)}. \\
I_4 &= 3m_1\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)^2 \\
&= 3m_1\left(t+t_0\right)\left[2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+2m_1 t\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+m_1^2 t^2+m_2 t\right] \\
&= 6m_1\left(t+t_0\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+6m_1^2 t\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad +3m_1\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+3m_1^3 t^2\left(t+t_0\right)+3m_1 m_2 t\left(t+t_0\right). \\
I_5 &= 3m_2\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right) = 3m_2\left(t+t_0\right)\left(\widehat{G}_{t+t_0}^{(1)}-\widehat{G}_{t_0}^{(1)}+m_1 t\right) \\
&= 3m_2\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+3m_1 m_2 t\left(t+t_0\right). \\
I_6 &= -3m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)^2 \\
&= -3m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[2\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}+2m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right. \\
&\quad \left.+\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}+m_1^2\left(t_1-t_0\right)^2+m_2\left(t_1-t_0\right)\right] \\
&= -6m_1\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}-3m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[2m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&\quad -3m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(2)}-3m_1^3\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[t_1^2-2t_1 t_0\right]-3m_1 m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}t_1,
\end{aligned}$$

$$
\begin{aligned}
I_6 &= -6m_1 \int_{t_0}^{t+t_0} t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} - 3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[2m_1 \left(t_1 - t_0\right) \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&\quad -3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(2)} - 6m_1^3 \int_{t_0}^{t+t_0} t_1^2 \mathrm{d}t_1 + 6m_1^3 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}t_1 - \frac{3}{2} m_1 m_2 \left[t_1^2\right]_{t_0}^{t+t_0} \\
&= -6m_1 \int_{t_0}^{t+t_0} t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} - 3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[2m_1 \left(t_1 - t_0\right) \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&\quad -3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(2)} - 2m_1^3 \left[t_1^3\right]_{t_0}^{t+t_0} + 3m_1^3 t_0 \left[t_1^2\right]_{t_0}^{t+t_0} - \frac{3}{2} m_1 m_2 \left[t^2 + 2tt_0\right] \\
&= -6m_1 \int_{t_0}^{t+t_0} t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} - 3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[2m_1 \left(t_1 - t_0\right) \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&\quad -3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(2)} - 2m_1^3 \left[t^3 + 3t^2 t_0 + 3tt_0^2\right] \\
&\quad +3m_1^3 t_0 \left[t^2 + 2tt_0\right] - \frac{3}{2} m_1 m_2 \left[t^2 + 2tt_0\right] \\
&= -6m_1 \int_{t_0}^{t+t_0} t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)} \mathrm{d}\widehat{G}_{t_1}^{(1)} - 3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[2m_1 \left(t_1 - t_0\right) \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&\quad -3m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(2)} - 2m_1^3 t^3 - 3m_1^3 t^2 t_0 - \frac{3}{2} m_1 m_2 \left[t^2 + 2tt_0\right].
\end{aligned}
$$

We need to evaluate the term $\int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[2m_1 \left(t_1 - t_0\right) \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)}\right]$ very carefully. Using chain rule or simply taking away the integral sign is not correct. Rather we need to proceed as follows:

$$
\begin{aligned}
& 2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[\left(t_1 - t_0\right) \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&= 2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[t_1 \left(\widehat{G}_{t_1}^{(1)} - \widehat{G}_{t_0}^{(1)}\right)\right] - 2m_1 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[\widehat{G}_{t_1}^{(1)} - \widehat{G}_{t_0}^{(1)}\right] \\
&= 2m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[t_1 \widehat{G}_{t_1}^{(1)}\right] - 2m_1 \widehat{G}_{t_0}^{(1)} \int_{t_0}^{t+t_0} t_1 \mathrm{d}t_1 - 2m_1 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= 2m_1 \left[t_1^2 \widehat{G}_{t_1}^{(1)}\right]_{t_0}^{t+t_0} - 2m_1 \int_{t_0}^{t+t_0} t_1 \widehat{G}_{t_1}^{(1)} \mathrm{d}t_1 - m_1 \widehat{G}_{t_0}^{(1)} \left[t_1^2\right]_{t_0}^{t+t_0} - 2m_1 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= 2m_1 \left[\left(t + t_0\right)^2 \widehat{G}_{t+t_0}^{(1)} - t_0^2 \widehat{G}_{t_0}^{(1)}\right] - m_1 \int_{t_0}^{t+t_0} \widehat{G}_{t_1}^{(1)} \mathrm{d}t_1^2 \\
&\qquad - m_1 \widehat{G}_{t_0}^{(1)} \left[t^2 + 2tt_0\right] - 2m_1 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= 2m_1 \left(t + t_0\right)^2 \widehat{G}_{t+t_0}^{(1)} - 2m_1 t_0^2 \widehat{G}_{t_0}^{(1)} - m_1 \left[\widehat{G}_{t_1}^{(1)} t_1^2\right]_{t_0}^{t+t_0} \\
&\quad +m_1 \int_{t_0}^{t+t_0} t_1^2 \mathrm{d}\widehat{G}_{t_1}^{(1)} - m_1 \widehat{G}_{t_0}^{(1)} \left[t^2 + 2tt_0\right] - 2m_1 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= 2m_1 \left(t + t_0\right)^2 \widehat{G}_{t+t_0}^{(1)} - 2m_1 t_0^2 \widehat{G}_{t_0}^{(1)} - m_1 \left[\widehat{G}_{t+t_0}^{(1)} \left(t + t_0\right)^2 - \widehat{G}_{t_0}^{(1)} t_0^2\right] \\
&\quad +m_1 \int_{t_0}^{t+t_0} t_1^2 \mathrm{d}\widehat{G}_{t_1}^{(1)} - m_1 \widehat{G}_{t_0}^{(1)} \left[t^2 + 2tt_0\right] - 2m_1 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}_{t_1}^{(1)},
\end{aligned}
$$

$$\begin{aligned}
&2m_1\int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[(t_1-t_0)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]\\
&= m_1\left(t+t_0\right)^2\widehat{G}_{t+t_0}^{(1)} - m_1 t_0^2\widehat{G}_{t_0}^{(1)} + m_1\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad -m_1\widehat{G}_{t_0}^{(1)}\left[t^2+2tt_0\right] - 2m_1t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= m_1\left(t+t_0\right)^2\widehat{G}_{t+t_0}^{(1)} - m_1\left(t+t_0\right)^2\widehat{G}_{t_0}^{(1)} + m_1\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)} - 2m_1t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= m_1\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} + m_1\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)} - 2m_1t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}.
\end{aligned}$$

Hence,

$$\begin{aligned}
I_6 &= -6m_1\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} - 3m_1^2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad -3m_1^2\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)} + 6m_1^2t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad -3m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(2)} - 2m_1^3t^3 - 3m_1^3t^2t_0 - \frac{3}{2}m_1m_2\left[t^2+2tt_0\right].\\
I_7 &= -3m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right) = -3m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[\widehat{G}_{t_1}^{(1)} + m_1t_1\right]\\
&= -3m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)} - 3m_1m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}t_1 = -3m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)} - \frac{3}{2}m_1m_2\left[t^2+2tt_0\right].
\end{aligned}$$

Altogether, we have

$$\begin{aligned}
\left(G_{t+t_0}-G_{t_0}\right)^3 &= 6\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} + 3\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +3\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)} + 6m_1t\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(3)} + 3m_1t\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)} + \left(3m_1^2t^2+3m_2t\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +m_1^3t^3 + 3m_1m_2t^2 + m_3t.
\end{aligned}$$

A.2.3 $\left(G_{t+t_0}-G_{t_0}\right)^4$

Starting from (A.4)-(A.6),

$$\begin{aligned}
\left(G_{t+t_0}-G_{t_0}\right)^4 &= \sum_{j=1}^{4}\binom{4}{j}\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)^{4-j}\mathrm{d}\widehat{G}_{t_1}^{(j)} + \sum_{j=1}^{3}\binom{4}{j}m_j\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)^{4-j}\\
&\quad -\sum_{j=1}^{3}\binom{4}{j}m_j\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)^{4-j} + m_4t,
\end{aligned}$$

$$\begin{aligned}
\left(G_{t+t_0}-G_{t_0}\right)^4 &= 4\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)^3 \mathrm{d}\widehat{G}_{t_1}^{(1)}+6\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)^2 \mathrm{d}\widehat{G}_{t_1}^{(2)}\\
&\quad +4\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)\mathrm{d}\widehat{G}_{t_1}^{(3)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(4)}+4m_1\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)^3\\
&\quad +6m_2\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)^2+4m_3\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)\\
&\quad -4m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)^3-6m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)^2\\
&\quad -4m_3\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)+m_4 t\\
&= J_1+J_2+J_3+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(4)}+J_4+J_5+J_6+J_7+J_8+J_9+m_4 t.
\end{aligned}$$

$$\begin{aligned}
J_1 &= 4\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)^3 \mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= 4\int_{t_0}^{t+t_0}\Bigg[6\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\int_{t_0}^{t_3-}\mathrm{d}\widehat{G}_{t_4}^{(1)}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}+3\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(2)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\\
&\quad +3\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(2)}+6m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\\
&\quad +\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(3)}+3m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\\
&\quad +\Big(3m_1^2\left(t_1-t_0\right)^2+3m_2\left(t_1-t_0\right)\Big)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}+m_1^3\left(t_1-t_0\right)^3\\
&\quad +3m_1m_2\left(t_1-t_0\right)^2+m_3\left(t_1-t_0\right)\Bigg]\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= 24\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\int_{t_0}^{t_3-}\mathrm{d}\widehat{G}_{t_4}^{(1)}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(2)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +24m_1\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(3)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12m_1^2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)^2\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12m_2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+4m_1^3\int_{t_0}^{t+t_0}\left(t_1-t_0\right)^3\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12m_1m_2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)^2\mathrm{d}\widehat{G}_{t_1}^{(1)}+4m_3\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\mathrm{d}\widehat{G}_{t_1}^{(1)}.
\end{aligned}$$

$$
\begin{aligned}
J_2 &= 6\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)^2 \mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&= 6\int_{t_0}^{t+t_0}\left[2\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}+2m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right.\\
&\left.+\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}+m_1^2\left(t_1-t_0\right)^2+m_2\left(t_1-t_0\right)\right]\mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&= 12\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)}+12m_1\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&+6\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(2)}+6m_1^2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)^2\mathrm{d}\widehat{G}_{t_1}^{(2)}+6m_2\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\mathrm{d}\widehat{G}_{t_1}^{(2)}. \\
J_3 &= 4\int_{t_0}^{t+t_0}\left(G_{t_1-}-G_{t_0}\right)\mathrm{d}\widehat{G}_{t_1}^{(3)} = 4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}G_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(3)} \\
&= 4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\left[\widehat{G}_{t_2}^{(1)}+m_1t_2\right]\mathrm{d}\widehat{G}_{t_1}^{(3)} \\
&= 4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(3)}+4m_1\int_{t_0}^{t+t_0}\left(t_1-t_0\right)\mathrm{d}\widehat{G}_{t_1}^{(3)} \\
&= 4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(3)}+4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(3)}-4m_1t_0\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(3)}. \\
J_4 &= 4m_1\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)^3 \\
&= 4m_1\left(t+t_0\right)\left[6\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+3\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)}\right.\\
&+3\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)}+6m_1t\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(3)} \\
&\left.+3m_1t\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+\left(3m_1^2t^2+3m_2t\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+m_1^3t^3+3m_1m_2t^2+m_3t\right] \\
&= 24m_1\left(t+t_0\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&+12m_1\left(t+t_0\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1\left(t+t_0\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&+24m_1^2t\left(t+t_0\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+4m_1\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(3)} \\
&+12m_1^2t\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+12m_1^3t^2\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1m_2t\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&+4m_1^4t^3\left(t+t_0\right)+12m_1^2m_2t^2\left(t+t_0\right)+4m_1m_3t\left(t+t_0\right). \\
J_5 &= 6m_2\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)^2 \\
&= 6m_2\left(t+t_0\right)\left[2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+2m_1t\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+m_1^2t^2+m_2t\right] \\
&= 12m_2\left(t+t_0\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1m_2t\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&+6m_2\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)}+6m_1^2m_2t^2\left(t+t_0\right)+6m_2^2t\left(t+t_0\right).
\end{aligned}
$$

$$\begin{aligned}
J_6 &= 4m_3\left(t+t_0\right)\left(G_{t+t_0}-G_{t_0}\right)=4m_3\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}G_{t_1}^{(1)} \\
&= 4m_3\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\left[\widehat{G}_{t_1}^{(1)}+m_1t_1\right]=4m_3\left(t+t_0\right)\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+4m_1m_3t\left(t+t_0\right). \\
J_7 &= -4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)^3 \\
&= -4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[6\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\int_{t_0}^{t_3-}\mathrm{d}\widehat{G}_{t_4}^{(1)}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}+3\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(2)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right. \\
&\quad +3\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(2)}+6m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}+\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(3)} \\
&\quad +3m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}+\left(3m_1^2\left(t_1-t_0\right)^2+3m_2\left(t_1-t_0\right)\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)} \\
&\quad \left.+m_1^3\left(t_1-t_0\right)^3+3m_1m_2\left(t_1-t_0\right)^2+m_3\left(t_1-t_0\right)\right] \\
&= -24m_1\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad -12m_1\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)}+K_1-4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(3)}+K_2+K_3+K_4 \\
&\quad +K_5+K_6+K_7. \\
K_1 &= -4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[6m_1\left(t_1-t_0\right)\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&= -24m_1^2\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[\left(t_1-t_0\right)\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right] \\
&= -24m_1^2\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[t_1\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]+24m_1^2t_0\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= -24m_1^2\left[t_1^2\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]_{t_0}^{t+t_0}+24m_1^2\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}t_1 \\
&\quad +24m_1^2t_0\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= -24m_1^2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1^2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}t_1^2 \\
&\quad +24m_1^2t_0\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= -24m_1^2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1^2\left[t_1^2\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]_{t_0}^{t+t_0} \\
&\quad -12m_1^2\int_{t_0}^{t+t_0}t_1^2\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}+24m_1^2t_0\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&= -12m_1^2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1^2\int_{t_0}^{t+t_0}t_1^2\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad +24m_1^2t_0\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)}.
\end{aligned}$$

$$
\begin{aligned}
K_2 &= -4m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[3m_1(t_1-t_0)\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\right] \\
&= -12m_1^2 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\right] + 12m_1^2 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(2)}_{t_1} \\
&= -12m_1^2 \left[t_1^2 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\right]_{t_0}^{t+t_0} + 12m_1^2 \int_{t_0}^{t+t_0} t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\mathrm{d}t_1 + 12m_1^2 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(2)}_{t_1} \\
&= -12m_1^2 (t+t_0)^2 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(2)}_{t_1} + 6m_1^2 \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\mathrm{d}t_1^2 + 12m_1^2 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(2)}_{t_1} \\
&= -12m_1^2 (t+t_0)^2 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(2)}_{t_1} + 6m_1^2 \left[t_1^2 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\right]_{t_0}^{t+t_0} - 6m_1^2 \int_{t_0}^{t+t_0} t_1^2 \mathrm{d}\widehat{G}^{(2)}_{t_1} \\
&\quad +12m_1^2 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(2)}_{t_1} \\
&= -6m_1^2 (t+t_0)^2 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(2)}_{t_1} - 6m_1^2 \int_{t_0}^{t+t_0} t_1^2 \mathrm{d}\widehat{G}^{(2)}_{t_1} + 12m_1^2 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(2)}_{t_1}. \\
K_3 &= -4m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[3m_1^2(t_1-t_0)^2\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right] \\
&= -4m_1 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[3m_1^2\left(t_1^2-2t_1t_0+t_0^2\right)\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right] \\
&= -12m_1^3 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[t_1^2 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right] + 24m_1^3 t_0 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\left[t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right] \\
&\quad -12m_1^3 t_0^2 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&= -12m_1^3 \left[t_1^3 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right]_{t_0}^{t+t_0} + 12m_1^3 \int_{t_0}^{t+t_0} t_1^2 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}t_1 - 12m_1^3 t_0^2 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +24m_1^3 t_0 \left[t_1^2 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right]_{t_0}^{t+t_0} - 24m_1^3 t_0 \int_{t_0}^{t+t_0} t_1 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}t_1 \\
&= -12m_1^3 (t+t_0)^3 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(1)}_{t_1} + 4m_1^3 \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}t_1^3 - 12m_1^3 t_0^2 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +24m_1^3 t_0 (t+t_0)^2 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(1)}_{t_1} - 12m_1^3 t_0 \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}t_1^2 \\
&= -12m_1^3 (t+t_0)^3 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(1)}_{t_1} + 4m_1^3 \left[t_1^3 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right]_{t_0}^{t+t_0} - 4m_1^3 \int_{t_0}^{t+t_0} t_1^3 \mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +24m_1^3 t_0 (t+t_0)^2 \int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(1)}_{t_1} - 12m_1^3 t_0 \left[t_1^2 \int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\right]_{t_0}^{t+t_0} \\
&\quad +12m_1^3 t_0 \int_{t_0}^{t+t_0} t_1^2 \mathrm{d}\widehat{G}^{(1)}_{t_1} - 12m_1^3 t_0^2 \int_{t_0}^{t+t_0} t_1 \mathrm{d}\widehat{G}^{(1)}_{t_1},
\end{aligned}
$$

$$
\begin{aligned}
K_3 &= -12m_1^3\left(t+t_0\right)^3\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+4m_1^3\left(t+t_0\right)^3\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}-4m_1^3\int_{t_0}^{t+t_0}t_1^3\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +24m_1^3t_0\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1^3t_0\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12m_1^3t_0\int_{t_0}^{t+t_0}t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1^3t_0^2\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= -8m_1^3\left(t+t_0\right)^3\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}-4m_1^3\int_{t_0}^{t+t_0}t_1^3\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1^3t_0\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12m_1^3t_0\int_{t_0}^{t+t_0}t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1^3t_0^2\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= 4m_1^3\left[-2\left(t^3+3t^2t_0+3tt_0^2+t_0^3\right)+3t_0\left(t^2+2tt_0+t_0^2\right)\right]\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad -4m_1^3\int_{t_0}^{t+t_0}t_1^3\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1^3t_0\int_{t_0}^{t+t_0}t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1^3t_0^2\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= 4m_1^3\left[-2t^3-3t^2t_0+t_0^3\right]\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}-4m_1^3\int_{t_0}^{t+t_0}t_1^3\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&\quad +12m_1^3t_0\int_{t_0}^{t+t_0}t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}-12m_1^3t_0^2\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}.\\
K_4 &= -4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[3m_2\left(t_1-t_0\right)\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]\\
&= -12m_1m_2\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]+12m_1m_2t_0\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= -12m_1m_2\left[t_1^2\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]_{t_0}^{t+t_0}+12m_1m_2\int_{t_0}^{t+t_0}t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}t_1\\
&\quad +12m_1m_2t_0\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= -12m_1m_2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+6m_1m_2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}t_1^2\\
&\quad +12m_1m_2t_0\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= -12m_1m_2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+6m_1m_2\left[t_1^2\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\right]_{t_0}^{t+t_0}\\
&\quad -6m_1m_2\int_{t_0}^{t+t_0}t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1m_2t_0\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}\\
&= -6m_1m_2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)}+12m_1m_2t_0\int_{t_0}^{t+t_0}t_1\mathrm{d}\widehat{G}_{t_1}^{(1)}-6m_1m_2\int_{t_0}^{t+t_0}t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)}.\\
K_5 &= -4m_1\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[m_1^3\left(t_1-t_0\right)^3\right]=-4m_1^4\int_{t_0}^{t+t_0}t_1\mathrm{d}\left[t_1^3-3t_1^2t_0+3t_1t_0^2-t_0^3\right]\\
&= -12m_1^4\int_{t_0}^{t+t_0}t_1^3\mathrm{d}t_1+24m_1^4t_0\int_{t_0}^{t+t_0}t_1^2\mathrm{d}t_1-12m_1^4t_0^2\int_{t_0}^{t+t_0}t_1\mathrm{d}t_1\\
&= -3m_1^4\left[t_1^4\right]_{t_0}^{t+t_0}+8m_1^4t_0\left[t_1^3\right]_{t_0}^{t+t_0}-6m_1^4t_0^2\left[t_1^2\right]_{t_0}^{t+t_0}\\
&= -3m_1^4\left[t^4+4t^3t_0+6t^2t_0^2+4tt_0^3\right]+8m_1^4t_0\left[t^3+3t^2t_0+3tt_0^2\right]-6m_1^4t_0^2\left[t^2+2tt_0\right]\\
&= -3m_1^4t^4-4m_1^4t^3t_0.
\end{aligned}
$$

$$\begin{aligned}
K_6 &= -4m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[3m_1m_2\left(t_1-t_0\right)^2\right] = -12m_1^2m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[t_1^2-2t_1t_0+t_0^2\right] \\
&= -24m_1^2m_2\int_{t_0}^{t+t_0} t_1^2\mathrm{d}t_1 + 24m_1^2m_2t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}t_1 = -8m_1^2m_2\left[t_1^3\right]_{t_0}^{t+t_0} + 12m_1^2m_2t_0\left[t_1^2\right]_{t_0}^{t+t_0} \\
&= -8m_1^2m_2\left[t^3+3t^2t_0+3tt_0^2\right] + 12m_1^2m_2t_0\left[t^2+2tt_0\right] = -8m_1^2m_2t^3 - 12m_1^2m_2t^2t_0. \\
K_7 &= -4m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\left[m_3\left(t_1-t_0\right)\right] = -4m_1m_3\int_{t_0}^{t+t_0} t_1\mathrm{d}t_1 \\
&= -2m_1m_3\left[t_1^2\right]_{t_0}^{t+t_0} = -2m_1m_3\left[t^2+2tt_0\right] = -2m_1m_3t^2 - 4m_1m_3tt_0.
\end{aligned}$$

Hence,

$$\begin{aligned}
J_7 &= -3m_1^4t^4 - 24m_1\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\mathrm{d}\widehat{G}_{t_3}^{(1)}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad -12m_1\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(2)}\mathrm{d}\widehat{G}_{t_1}^{(1)} - 12m_1\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&\quad -12m_1^2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} - 12m_1^2\int_{t_0}^{t+t_0} t_1^2\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad +24m_1^2t_0\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} - 4m_1\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(3)} \\
&\quad -6m_1^2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(2)} - 6m_1^2\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(2)} + 12m_1^2t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(2)} \\
&\quad +4m_1^3\left[-2t^3-3t^2t_0+t_0^3\right]\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} - 4m_1^3\int_{t_0}^{t+t_0} t_1^3\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad +12m_1^3t_0\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)} - 12m_1^3t_0^2\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad -6m_1m_2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} + 12m_1m_2t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad -6m_1m_2\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)} - 4m_1^4t^3t_0 - 8m_1^2m_2t^3 \\
&\quad -12m_1^2m_2t^2t_0 - 2m_1m_3t^2 - 4m_1m_3tt_0. \\
J_8 &= -6m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right)^2 = \left(\frac{-6m_2}{-3m_1}\right)I_6 = \frac{2m_2}{m_1}I_6 \\
&= -12m_2\int_{t_0}^{t+t_0} t_1\int_{t_0}^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(1)} - 6m_1m_2\left(t+t_0\right)^2\int_{t_0}^{t+t_0}\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad -6m_1m_2\int_{t_0}^{t+t_0} t_1^2\mathrm{d}\widehat{G}_{t_1}^{(1)} + 12m_1m_2t_0\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)} \\
&\quad -6m_2\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(2)} - 4m_1^2m_2t^3 - 6m_1^2m_2t^2t_0 - 3m_2^2\left[t^2+2tt_0\right]. \\
J_9 &= -4m_3\int_{t_0}^{t+t_0} t_1\mathrm{d}\left(G_{t_1}-G_{t_0}\right) = \left(\frac{-4m_3}{-3m_2}\right)I_7 \\
&= -4m_3\int_{t_0}^{t+t_0} t_1\mathrm{d}\widehat{G}_{t_1}^{(1)} - 2m_1m_3\left[t^2+2tt_0\right].
\end{aligned}$$

$$
\begin{aligned}
&(G_{t+t_0}-G_{t_0})^4 \\
&= m_4 t + 24\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-}\int_{t_0}^{t_3-} \mathrm{d}\widehat{G}^{(1)}_{t_4}\mathrm{d}\widehat{G}^{(1)}_{t_3}\mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +12\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-} \mathrm{d}\widehat{G}^{(2)}_{t_3}\mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} + 12\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-} \mathrm{d}\widehat{G}^{(1)}_{t_3}\mathrm{d}\widehat{G}^{(2)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +12\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-} \mathrm{d}\widehat{G}^{(1)}_{t_3}\mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(2)}_{t_1} + 24m_1 t\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-}\int_{t_0}^{t_2-} \mathrm{d}\widehat{G}^{(1)}_{t_3}\mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(3)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} + 4\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(3)}_{t_1} \\
&\quad +6\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\mathrm{d}\widehat{G}^{(2)}_{t_1} + 12m_1 t\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +12m_1 t\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(2)}_{t_1} + \left(12m_1^2t^2+12m_2t\right)\int_{t_0}^{t+t_0}\int_{t_0}^{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_2}\mathrm{d}\widehat{G}^{(1)}_{t_1} \\
&\quad +\int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(4)}_{t_1} + 4m_1 t\int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(3)}_{t_1} + \left(6m_1^2t^2+6m_2t\right)\int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(2)}_{t_1} \\
&\quad +\left(4m_1^3t^3+12m_1m_2t^2+4m_3t\right)\int_{t_0}^{t+t_0} \mathrm{d}\widehat{G}^{(1)}_{t_1} + m_1^4t^4 + 6m_1^2m_2t^3 + \left(4m_1m_3+3m_2^2\right)t^2.
\end{aligned}
$$

A.3 Proof of Proposition 3.1.2

We prove this result using strong induction. Clearly, the proposition is true for $k=1$ and 2. Assume the proposition is true for $k=n$, where n is an integer ≥ 1. Then for $k=n+1$, firstly we prove that the sum of the indices of all the m_q's appearing in each of the terms of $C_t^{(n+1)}$ are equal to $n+1$. By Proposition 3.1.1, we have

$$
C_t^{(n+1)} = \sum_{j=1}^{n}\binom{n+1}{j} m_j t C_t^{(n+1-j)} - \sum_{j=1}^{n}\binom{n+1}{j} m_j \int_0^t t_1 \, \mathrm{d}C_{t_1}^{(n+1-j)} + m_{n+1}t. \tag{A.8}
$$

By the induction step, the tuples of the indices of all the m_q's appearing in each of the terms of $C_t^{(n+1-j)}$ are elements of $\mathcal{L}_{n+1-j}$ defined in (3.9). Since we have $m_j C_t^{(n+1-j)}$ appearing in the first term of (A.8), $m_j C_{t_1}^{(n+1-j)}$ in the second term and m_{n+1} in the last term, it is clear that the tuples of the indices of all the m_q's appearing in each of the terms of $C_t^{(n+1)}$ are elements of $\mathcal{L}_{n+1}$. By the induction step, the first term of (A.8) is given by

$$
\begin{aligned}
&\sum_{j=1}^{n}\binom{n+1}{j} m_j t C_t^{(n+1-j)} \\
&= \sum_{j=1}^{n} \sum_{\phi_{n+1-j}=\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \dots, i_l^{(n+1-j)}\right)\in\mathcal{L}_{n+1-j}} \frac{1}{l!}\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \dots, i_l^{(n+1-j)}, j\right)! \\
&\quad \times \left(p_1^{\phi_{n+1-j}}, p_2^{\phi_{n+1-j}}, \dots, p_{n+1-j}^{\phi_{n+1-j}}\right)! \left[\prod_{q\in\phi_{n+1-j}\cup\{j\}} m_q\right] t^{l+1}
\end{aligned}
$$

and the second term is given by

$$
\begin{aligned}
&-\sum_{j=1}^{n}\binom{n+1}{j} m_j \int_0^t t_1 \, \mathrm{d}C_{t_1}^{(n+1-j)} \\
&= -\sum_{j=1}^{n} \sum_{\phi_{n+1-j}=\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}\right) \in \mathcal{L}_{n+1-j}} \frac{1}{l!}\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}, j\right)! \\
&\quad \times \left(p_1^{\phi_{n+1-j}}, p_2^{\phi_{n+1-j}}, \ldots, p_{n+1-j}^{\phi_{n+1-j}}\right)! \left[\prod_{q \in \phi_{n+1-j} \cup \{j\}} m_q\right] \frac{l}{l+1} t^{l+1}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
C_t^{(n+1)} &= \sum_{j=1}^{n} \sum_{\phi_{n+1-j}=\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}\right) \in \mathcal{L}_{n+1-j}} \frac{1}{l!}\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}, j\right)! \\
&\quad \times \left(p_1^{\phi_{n+1-j}}, p_2^{\phi_{n+1-j}}, \ldots, p_{n+1-j}^{\phi_{n+1-j}}\right)! \left[\prod_{q \in \phi_{n+1-j} \cup \{j\}} m_q\right] t^{l+1} \frac{1}{l+1} + m_{n+1} t.
\end{aligned}
$$

Next we are going to prove that

$$
\begin{aligned}
&\sum_{\phi_{n+1}=\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right) \in \mathcal{L}_{n+1}} \frac{1}{(l+1)!}\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)! \\
&\times \left(p_1^{\phi_{n+1}}, p_2^{\phi_{n+1}}, \ldots, p_{n+1}^{\phi_{n+1}}\right)! \left[\prod_{q \in \phi_{n+1}} m_q\right] t^{l+1} \\
&= \sum_{j=1}^{n} \sum_{\phi_{n+1-j}=\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}\right) \in \mathcal{L}_{n+1-j}} \frac{1}{(l+1)!}\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}, j\right)! \\
&\quad \times \left(p_1^{\phi^{n+1-j}}, p_2^{\phi^{n+1-j}}, \ldots, p_{n+1-j}^{\phi^{n+1-j}}\right)! \left[\prod_{q \in \phi_{n+1-j} \cup \{j\}} m_q\right] t^{l+1} + m_{n+1} t. \qquad \text{(A.9)}
\end{aligned}
$$

On the R.H.S., we are adding a j to each tuple $\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}\right)$ such that we have $\sum_{q=1}^{l} i_q^{(n+1-j)} + j = n+1$. Suppose $\phi_{n+1} = \left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)$ has one extra element compared to the tuple $\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}\right)$ and otherwise they are the same. Since $\sum_{q=1}^{l+1} i_q^{(n+1)} = n+1$, to obtain $\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)$ from $\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, \ldots, i_l^{(n+1-j)}\right)$, we are adding an element j to the latter such that the sum of the tuple is equal to $n+1$. Suppose there are r distinct value(s) in $\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)$. Let $x_1, x_2, \ldots, x_r$ be the distinct values in $\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)$ and let $f_i, i = 1, \ldots, r$ be the number of times x_i appears in $\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)$. Note that $\sum_{q=1}^{r} f_q$ is equal to the length of $\left(i_1^{(n+1)}, i_2^{(n+1)}, \ldots, i_{l+1}^{(n+1)}\right)$, that is, $\sum_{q=1}^{r} f_q = l+1$. Since $\left(i_1^{(n+1)}, i_2^{(n+1)}, ..., i_{l+1}^{(n+1)}\right)$ can be obtained by adding an element j to a tuple $\left(i_1^{(n+1-j)}, i_2^{(n+1-j)}, ..., i_l^{(n+1-j)}\right)$ whose elements add up to $n+1-j$, j can take one of the r distinct value(s): $x_1, x_2, ..., x_r$. For example, suppose $j = x_i$, then the corresponding term on the R.H.S. of (A.9) is

$$\frac{(n+1)!}{(x_1!)^{f_1}(x_2!)^{f_2}\cdots(x_i!)^{f_i-1}\cdots(x_r!)^{f_r}\,x_i!f_1!\cdots(f_i-1)!\cdots f_r!}\left[\prod_{q\in\phi_{n+1}} m_q\right] t^{l+1}\frac{1}{\sum_{q=1}^{r} f_q}.$$

Summing up all the possible $j \in \{x_1, x_2, ..., x_r\}$,

$$\begin{aligned}
&\sum_{i=1}^{r}\frac{(n+1)!}{(x_1!)^{f_1}(x_2!)^{f_2}\cdots(x_i!)^{f_i-1}\cdots(x_r!)^{f_r}\,x_i!f_1!\cdots(f_i-1)!\cdots f_r!}\left[\prod_{q\in\phi_{n+1}} m_q\right] t^{l+1}\frac{1}{\sum_{q=1}^{r} f_q}\\
&= \frac{(n+1)!}{(x_1!)^{f_1}\cdots(x_r!)^{f_r}\,f_1!\cdots f_r!}\frac{1}{\sum_{q=1}^{r} f_q}\left[\prod_{q\in\phi_{n+1}} m_q\right] t^{l+1}\sum_{i=1}^{r} f_i\\
&= \frac{(n+1)!}{(x_1!)^{f_1}\cdots(x_r!)^{f_r}\,f_1!\cdots f_r!}\left[\prod_{q\in\phi_{n+1}} m_q\right] t^{l+1}.
\end{aligned}$$

For the case $\phi_{n+1} = \left(i_1^{(n+1)}\right)$, it is clear that the L.H.S. of (A.9) is equal to $m_{n+1}t$. Hence, by applying the same argument to each possible tuple $\left(i_1^{(n+1)}, i_2^{(n+1)}, ..., i_{l+1}^{(n+1)}\right) \in \mathcal{L}_{n+1}$, we have proven (A.9) and therefore

$$\begin{aligned}
C_t^{(n+1)} &= \sum_{\phi_{n+1}=\left(i_1^{(n+1)}, i_2^{(n+1)}, ..., i_l^{(n+1)}\right)\in\mathcal{L}_{n+1}} \frac{1}{l!}\left(i_1^{(n+1)}, i_2^{(n+1)}, ..., i_l^{(n+1)}\right)!\\
&\quad\times\left(p_1^{\phi_{n+1}}, p_2^{\phi_{n+1}}, ..., p_{n+1}^{\phi_{n+1}}\right)!\left[\prod_{q\in\phi_{n+1}} m_q\right] t^l.
\end{aligned}$$

A.4 Proof of Theorem 3.1.4

We prove the result using strong induction. Firstly we need to consider $\mathcal{I}_k$ defined by equation (3.7). We need to know what tuples are in $\mathcal{I}_{k+1}$ but not in $\mathcal{I}_k$, and these correspond to those elements adding up exactly to $k+1$. Let $\mathcal{J}_{k+1}$ be the collection of these tuples, that is, $\mathcal{I}_{k+1} \equiv \mathcal{I}_k \cup \mathcal{J}_{k+1}$. We have

$$\mathcal{J}_{k+1} = \left\{(i_1, i_2, ..., i_j) \mid j \in \{1, 2, ..., k+1\},\ i_p \in \{1, 2, ..., k+1\} \text{ and } \sum_{p=1}^{j} i_p = k+1\right\}.$$

To construct $\mathcal{J}_{k+1}$ from $\mathcal{I}_k$, we can simply add an element to the end of each tuple in $\mathcal{I}_k$ so that the elements of each new tuple add up exactly to $k+1$, and finally including the tuple $(k+1)$ in $\mathcal{J}_{k+1}$.

We are going to prove by strong induction that $(G_{t+t_0} - G_{t_0})^k = \sum_{\theta_k \in \mathcal{I}_k} \Pi^{(k)}_{\theta_k,t} \mathcal{S}_{\theta_k,t,t_0} + C^{(k)}_t$ for any non-negative integer k. For $k = 0$, clearly both sides equal 1. For $k = 1$ and 2, it can be checked easily that the proposition is true using derivation given in Appendix A.2.1. Assume the proposition is true for $k = 0, 1, 2, ..., n$, where n is a positive integer. Note that it is sufficient to prove the representation for G^{n+1}_t only since $\{G_{t+t_0} - G_{t_0}, t \geq 0\}$ and $\{G_t, t \geq 0\}$ have the same distribution. For $i = 1, 2, 3, ...,$

$$G^{(i)}_{t+t_0} - G^{(i)}_{t_0} = \sum_{0<s\leq t+t_0} (\Delta G_s)^i - \sum_{0<s\leq t_0} (\Delta G_s)^i = \sum_{t_0<s\leq t+t_0} (\Delta G_s)^i \,,$$

which has the same distribution as $\sum_{0<s\leq t} (\Delta G_s)^i$. Since both $\{G_{t+t_0} - G_{t_0}, t \geq 0\}$ and $\{G_t, t \geq 0\}$ are created by the same infinitely divisible distribution, the compensators for their i-th power jump processes are both equal to $m_i t$. Hence, we have the i-th compensated power jump process of $\{G_t, t \geq 0\}\,, \left\{\widehat{G}^{(i)}_t\right\}$ has the same distribution as $\left\{\widehat{G}^{(i)}_{t+t_0} - \widehat{G}^{(i)}_{t_0}\right\}$. For $k = n + 1$, by (3.3)-(3.6),

$$\begin{aligned} G^{n+1}_t \;&=\; \sum_{j=1}^{n+1} \binom{n+1}{j} \int_0^t G^{n+1-j}_{t_1-} \;\mathrm{d}\widehat{G}^{(j)}_{t_1} + \sum_{j=1}^{n} \binom{n+1}{j} m_j t G^{n+1-j}_t \\ &\quad - \sum_{j=1}^{n} \binom{n+1}{j} m_j \int_0^t t_1 \;\mathrm{d}G^{n+1-j}_{t_1} + m_{n+1} t. \end{aligned} \tag{A.10}$$

Firstly, we want to prove that all the stochastic integrals in G^{n+1}_t are of the form $\mathcal{S}_{\theta_{n+1},t,0}$, where $\theta_{n+1} \in \mathcal{I}_{n+1}$. From (A.10), it is clear that the first term is the only term introducing new stochastic integrals which are not in $\mathcal{I}_n$. The general term of the stochastic integrals in the first term is

$$\int_0^t G^{n+1-j}_{t_1-} \;\mathrm{d}\widehat{G}^{(j)}_{t_1}, \quad j = 1, 2, ..., n + 1. \tag{A.11}$$

By assumption,

$$G^{n+1-j}_{t_1-} = \sum_{\theta_{n+1-j} \in \mathcal{I}_{n+1-j}} \Pi^{(n+1-j)}_{\theta_{n+1-j},t_1} \mathcal{S}_{\theta_{n+1-j},t_1,0} + C^{(n+1-j)}_{t_1},\; j = 1, 2, ..., n + 1.$$

When $j = 1$ in (A.11), we have $\int_0^t G^n_{t_1-} \mathrm{d}\widehat{G}^{(1)}_{t_1}$, meaning that we are adding a 1 to the end of all tuples in $\mathcal{I}_n$. Since by definition

$$\mathcal{I}_n = \left\{ (i_1, i_2, ..., i_j) \,|\, j \in \{1, 2, ..., n\}\,,\; i_p \in \{1, 2, ..., n\} \text{ and } \sum_{p=1}^{j} i_p \leq n \right\},$$

we know that the sums of the elements of the new tuples we get from adding a 1 to the end of each tuple of $\mathcal{I}_n$ are less than or equal to $n + 1$. Similarly, when $j = 2$, we have $\int_0^t G^{n-1}_{t_1-} \mathrm{d}\widehat{G}^{(2)}_{t_1}$, meaning that we are adding a 2 to the end of all tuples in $\mathcal{I}_{n-1}$ and since by definition

$$\mathcal{I}_{n-1} = \left\{ (i_1, i_2, ..., i_j) \,|\, j \in \{1, 2, ..., n-1\}\,,\; i_p \in \{1, 2, ..., n-1\} \text{ and } \sum_{p=1}^{j} i_p \leq n-1 \right\},$$

we know that the sums of the elements of the new tuples we get from adding a 2 to the end of each tuple of $\mathcal{I}_{n-1}$ are less than or equal to $n+1$. We can continue the same argument until $j = n$. When $j = n+1$, we have $\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(n+1)}$. Since $\mathcal{I}_n \supset \mathcal{I}_{n-1} \supset ... \supset \mathcal{I}_2 \supset \mathcal{I}_1$, the above way of introducing new stochastic integrals is the same as adding an element to the end of each tuple in $\mathcal{I}_n$ so that the elements of each new tuple add up exactly to $n+1$. Hence all the elements in $\mathcal{J}_{n+1}$ have been created and since $\mathcal{I}_{n+1} \equiv \mathcal{I}_n \cup \mathcal{J}_{n+1}$, we have proved that all the stochastic integrals in G_t^{n+1} have the form $\mathcal{S}_{\theta_{n+1},t,0}$, where $\theta_{n+1} \in \mathcal{I}_{n+1}$. By definition, $C_t^{(n+1)}$ is the term in G_t^{n+1} not containing any stochastic integral. Hence it is correct to write $C_t^{(n+1)}$ as the final term. Finally, we want to consider the coefficients of the stochastic integrals, that is, we are going to prove Proposition 3.1.3. By assumption of the induction step and from (A.10), we have

$$
\begin{aligned}
G_t^{n+1} &= \sum_{j=1}^{n} \binom{n+1}{j} \sum_{\theta_{n+1-j} \in \mathcal{I}_{n+1-j}} \int_0^t \Pi_{\theta_{n+1-j},t_1}^{(n+1-j)} \mathcal{S}_{\theta_{n+1-j},t_1,0} \, \mathrm{d}\widehat{G}_{t_1}^{(j)} \\
&\quad + \sum_{j=1}^{n} \binom{n+1}{j} m_j t \sum_{\theta_{n+1-j} \in \mathcal{I}_{n+1-j}} \Pi_{\theta_{n+1-j},t}^{(n+1-j)} \mathcal{S}_{\theta_{n+1-j},t,0} \\
&\quad - \sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{\theta_{n+1-j} \in \mathcal{I}_{n+1-j}} \int_0^t t_1 \, \mathrm{d}\left[\Pi_{\theta_{n+1-j},t_1}^{(n+1-j)} \mathcal{S}_{\theta_{n+1-j},t_1,0}\right] \\
&\quad + \sum_{j=1}^{n} \binom{n+1}{j} \int_0^t C_{t_1}^{(n+1-j)} \, \mathrm{d}\widehat{G}_{t_1}^{(j)} + \sum_{j=1}^{n} \binom{n+1}{j} m_j t C_t^{(n+1-j)} \\
&\quad - \sum_{j=1}^{n} \binom{n+1}{j} m_j \int_0^t t_1 \, \mathrm{d}\left[C_{t_1}^{(n+1-j)}\right] + \int_0^t \mathrm{d}\widehat{G}_{t_1}^{(n+1)} + m_{n+1} t \\
&= L_1 + L_2 + L_3 + L_4 + L_5 + L_6 + \int_0^t \mathrm{d}\widehat{G}_{t_1}^{(n+1)} + m_{n+1} t. \qquad \text{(A.12)}
\end{aligned}
$$

Let $\mathcal{K}_{l,s} = \left\{ (i_1, ..., i_l) \, | i_j \in \{1, 2, ..., s\} \text{ and } \sum_{j=1}^{l} i_j = s \right\}$. Since the length of a tuple must not be greater than the sum of all the elements in the tuple (because an element must be at least 1), $l \le s$. By definition, we have $\mathcal{I}_n = \bigcup_{s=1}^{n} \bigcup_{l=1}^{s} \mathcal{K}_{l,s}$. For any $\theta_{l,s} \in \mathcal{K}_{l,s}$, let $\theta_{l,s} = \left(i_1^{\theta_{l,s}}, i_2^{\theta_{l,s}}, ..., i_l^{\theta_{l,s}} \right)$. It is obvious from Proposition 3.1.1 that $C_t^{(k)}$ has the form $C_t^{(k)} = q_0^{(k)} + q_1^{(k)} t + q_2^{(k)} t^2 + ... + q_k^{(k)} t^k$. Note that $q_0^{(k)}$ is non-zero only when $k = 0$. When $k = 0$, by definition $C_t^{(k)} = 1$, so we have $q_0^{(0)} = 1$. We need to find out the recursive relationships between the $q_r^{(k)}$'s. From (3.8), for $k > 1$,

$$
\begin{aligned}
& q_1^{(k)} t + q_2^{(k)} t^2 + \cdots + q_k^{(k)} t^k \\
& \quad = m_k t + \sum_{j=1}^{k-1} \binom{k}{j} m_j t \left[q_1^{(k-j)} t + q_2^{(k-j)} t^2 + \cdots + q_{k-j}^{(k-j)} t^{k-j} \right] \\
& \qquad - \sum_{j=1}^{k-1} \binom{k}{j} m_j \int_0^t t_1 \, \mathrm{d}\left[q_1^{(k-j)} t_1 + q_2^{(k-j)} t_1^2 + \cdots + q_{k-j}^{(k-j)} t_1^{k-j} \right]
\end{aligned}
$$

$$
\begin{aligned}
&= m_k t + \sum_{j=1}^{k-1} \binom{k}{j} m_j \left[q_1^{(k-j)} t^2 + q_2^{(k-j)} t^3 + \cdots + q_{k-j}^{(k-j)} t^{k-j+1} \right] \\
&\quad - \sum_{j=1}^{k-1} \binom{k}{j} m_j \left[\frac{1}{2} q_1^{(k-j)} t^2 + \frac{2}{3} q_2^{(k-j)} t^3 + \cdots + \frac{k-j}{k-j+1} q_{k-j}^{(k-j)} t^{k-j+1} \right].
\end{aligned}
$$

By comparing the coefficients of t, $q_1^{(k)} = m_k$. By comparing the coefficients of t^r, $r = 2, ..., k$,

$$
q_r^{(k)} = \sum_{j=1}^{k+1-r} \binom{k}{j} m_j q_{r-1}^{(k-j)} - \sum_{j=1}^{k+1-r} \binom{k}{j} m_j \frac{r-1}{r} q_{r-1}^{(k-j)} = \frac{1}{r} \sum_{j=1}^{k+1-r} \binom{k}{j} m_j q_{r-1}^{(k-j)}. \tag{A.13}
$$

From (A.12), we have

$$
\begin{aligned}
L_1 &= \sum_{j=1}^{n} \left\{ \binom{n+1}{j} \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \int_0^t \Pi_{\theta_{l,s},t_1}^{(n+1-j)} \mathcal{S}_{\theta_{l,s},t_1,0} \mathrm{d}\widehat{G}_{t_1}^{(j)} \right\} \\
&= \sum_{j=1}^{n} \left\{ \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \int_0^t \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! j!} \frac{1}{(n+1-j-s)!} \right. \\
&\quad \left. \times \sum_{w=0}^{n+1-j-s} q_w^{(n+1-j-s)} t_1^w \int_0^{t_1-} \cdots \int_0^{t_l-} \mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_1}^{(j)} \right\}. \\
L_3 &= -\sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \int_0^t t_1 \mathrm{d} \left[\Pi_{\theta_{l,s},t_1}^{(n+1-j)} \mathcal{S}_{\theta_{l,s},t_1,0} \right] \\
&= -\sum_{j=1}^{n} m_j \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! j!} \times \frac{1}{(n+1-j-s)!} \\
&\quad \times \int_0^t t_1 \mathrm{d} \left[\left\{ q_0^{(n+1-j-s)} + q_1^{(n+1-j-s)} t_1 + \sum_{w=2}^{n+1-j-s} q_w^{(n+1-j-s)} t_1^w \right\} \right. \\
&\quad \left. \times \int_0^{t_1-} \int_0^{t_2-} \cdots \int_0^{t_l-} \mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)} \right].
\end{aligned}
$$

By integration by parts,

$$
\begin{aligned}
L_3 &= -\sum_{j=1}^{n} m_j \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! j!} \times \frac{1}{(n+1-j-s)!} \\
&\quad \times \left\{ q_0^{(n+1-j-s)} \int_0^t t_1 \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)} \right. \\
&\quad + q_1^{(n+1-j-s)} \left[t_1^2 \int_0^{t_1-} \int_0^{t_2-} \cdots \int_0^{t_l-} \mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)} \right]_0^t
\end{aligned}
$$

$$
\begin{aligned}
& -q_1^{(n+1-j-s)}\int_0^t t_1\int_0^{t_1-}\int_0^{t_2-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\mathrm{d}t_1 \\
& +\sum_{w=2}^{n+1-j-s}q_w^{(n+1-j-s)}\Bigg\{\left[t_1^{w+1}\int_0^{t_1-}\int_0^{t_2-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\right]_0^t \\
& -\int_0^t t_1^w\int_0^{t_1-}\int_0^{t_2-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\mathrm{d}t_1\Bigg\}\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
L_3 \;=\; & -\sum_{j=1}^{n}m_j\sum_{s=1}^{n+1-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}\times\frac{1}{(n+1-j-s)!} \\
& \times\Bigg\{q_0^{(n+1-j-s)}\int_0^t t_1\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)} \\
& +q_1^{(n+1-j-s)}\left[t^2\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\right] \\
& -q_1^{(n+1-j-s)}\frac{1}{2}\int_0^t\int_0^{t_1-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\mathrm{d}t_1^2 \\
& +\sum_{w=2}^{n+1-j-s}q_w^{(n+1-j-s)}\Bigg\{\left[t^{w+1}\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\right] \\
& -\frac{1}{w+1}\int_0^t\int_0^{t_1-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\mathrm{d}t_1^{w+1}\Bigg\}\Bigg\}.
\end{aligned}
$$

By integration by parts again,

$$
\begin{aligned}
L_3 \;=\; & -\sum_{j=1}^{n}m_j\sum_{s=1}^{n+1-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}\times\frac{1}{(n+1-j-s)!} \\
& \times\Bigg\{q_0^{(n+1-j-s)}\int_0^t t_1\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)} \\
& +q_1^{(n+1-j-s)}\left[t^2\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\right] \\
& -\frac{1}{2}q_1^{(n+1-j-s)}\left[t_1^2\int_0^{t_1-}\int_0^{t_2-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\right]_0^t \\
& +\frac{1}{2}q_1^{(n+1-j-s)}\int_0^t t_1^2\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)} \\
& +\sum_{w=2}^{n+1-j-s}q_w^{(n+1-j-s)}\Bigg\{\left[t^{w+1}\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\right] \\
& -\frac{1}{w+1}\left[t_1^{w+1}\int_0^{t_1-}\int_0^{t_2-}\cdots\int_{t0}^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_3}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)}\right]_0^t
\end{aligned}
$$

$$+\frac{1}{w+1}\int_0^t t_1^{w+1}\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\Bigg\}\Bigg\}.$$

Finally, we have

$$\begin{aligned}
L_3 &= -\sum_{j=1}^{n} m_j \sum_{s=1}^{n+1-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}\times\frac{1}{(n+1-j-s)!}\\
&\times\Bigg\{q_0^{(n+1-j-s)}\int_0^t t_1\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\\
&+\frac{1}{2}q_1^{(n+1-j-s)}\left[t^2\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\right]\\
&+\frac{1}{2}q_1^{(n+1-j-s)}\int_0^t t_1^2\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\\
&+\sum_{w=2}^{n+1-j-s} q_w^{(n+1-j-s)}\Bigg\{\frac{w}{w+1}t^{w+1}\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\\
&+\frac{1}{w+1}\int_0^t t_1^{w+1}\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\Bigg\}\Bigg\}.
\end{aligned}$$

Next, consider

$$\begin{aligned}
L_2 &= \sum_{j=1}^{n}\binom{n+1}{j}m_j t\sum_{s=1}^{n+1-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\Pi_{\theta_{l,s},t}^{(n+1-j)}\mathcal{S}_{\theta_{l,s},t,0}\\
&= \sum_{j=1}^{n}\binom{n+1}{j}m_j t\sum_{s=1}^{n+1-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\left\{\frac{(n+1-j)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!}\times\frac{C_t^{(n+1-j-s)}\mathcal{S}_{\theta_{l,s},t,0}}{(n+1-j-s)!}\right\}\\
&= \sum_{j=1}^{n}m_j\sum_{s=1}^{n+1-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\Bigg\{\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}\frac{1}{(n+1-j-s)!}\\
&\times\sum_{w=0}^{n+1-j-s}q_w^{(n+1-j-s)}t^{w+1}\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\Bigg\}.
\end{aligned}$$

Since $q_0^{(k)}$ is non-zero only when $k=0$ and $C_t^{(0)}=1$,

$$\begin{aligned}
L_2 &= \sum_{j=1}^{n}m_j\sum_{s=1}^{n+1-j}\Bigg\{1_{\{s=n+1-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}t\Bigg\{\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}\\
&\times\int_0^t\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)}\mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}\Bigg\}\\
&+1_{\{s\le n-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\Bigg\{\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}\frac{1}{(n+1-j-s)!}
\end{aligned}$$

$$\times \sum_{w=1}^{n+1-j-s} q_w^{(n+1-j-s)} t^{w+1} \int_0^t \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\binom{\theta_{l,s}}{i_1}} \cdots \mathrm{d}\widehat{G}_{t_2}^{\binom{\theta_{l,s}}{i_{l-1}}} \mathrm{d}\widehat{G}_{t_1}^{\binom{\theta_{l,s}}{i_l}} \Bigg\}\Bigg\}.$$

$$L_4 = \sum_{j=1}^{n} \binom{n+1}{j} \int_0^t \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} t_1^w \mathrm{d}\widehat{G}_{t_1}^{(j)}.$$

$$L_5 = \sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} t^{w+1}.$$

since $n+1-j \geq 1$ gives $q_0^{(n+1-j)} = 0$.

$$\begin{aligned} L_6 &= -\sum_{j=1}^{n} \binom{n+1}{j} m_j \left\{ \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} \int_0^t t_1 \mathrm{d}t_1^w \right\} \\ &= -\sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} \left\{ \left[t_1^{w+1}\right]_0^t - \int_0^t t_1^w \mathrm{d}t_1 \right\} \\ &= -\sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} \left\{ t^{w+1} - \frac{1}{w+1} t^{w+1} \right\} \\ &= -\sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} \left\{ \frac{w}{w+1} t^{w+1} \right\}. \end{aligned}$$

To ease notation, we let $\mathbf{F}_1 = \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! j!}$, $\mathbf{F}_2 = \frac{\mathbf{F}_1}{(n+1-j-s)!}$, $\mathbf{G}_i^j = \widehat{G}_{t_i}^{\binom{\theta_{l,s}}{i_j}}$, $\mathbf{I}_1 = \int_0^{t_1-} \cdots \int_0^{t_l-} \mathrm{d}\mathbf{G}_{l+1}^1 \cdots \mathrm{d}\mathbf{G}_2^l$, $\mathbf{I}_2 = \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\mathbf{G}_l^1 \cdots \mathrm{d}\mathbf{G}_2^{l-1}$, $\mathbf{I}_3 = \int_0^t \mathbf{I}_2 \ \mathrm{d}\mathbf{G}_1^l$, $\tilde{\mathbf{q}}_i = q_i^{(n+1-j-s)}$. Note that these definitions are only for simplicity in writing out equations. When doing calculation, we should always use the long but clear notation. Altogether, we have

$$L_1 = \sum_{j=1}^{n} \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \int_0^t \mathbf{F}_2 \sum_{w=0}^{n+1-j-s} \tilde{\mathbf{q}}_w t_1^w \mathbf{I}_1 \ \mathrm{d}\widehat{G}_{t_1}^{(j)}.$$

$$\begin{aligned} L_2 = \sum_{j=1}^{n} m_j \sum_{s=1}^{n+1-j} \Bigg\{ & 1_{\{s=n+1-j\}} t \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \mathbf{F}_1 \mathbf{I}_3 \\ & + 1_{\{s \leq n-j\}} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \mathbf{F}_2 \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w t^{w+1} \mathbf{I}_3 \Bigg\}. \end{aligned}$$

$$\begin{aligned} L_3 = -\sum_{j=1}^{n} m_j \sum_{s=1}^{n+1-j} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \mathbf{F}_2 \Bigg\{ & \tilde{\mathbf{q}}_0 \int_0^t t_1 \mathbf{I}_2 \ \mathrm{d}\mathbf{G}_1^l + \frac{1}{2} \tilde{\mathbf{q}}_1 t^2 \mathbf{I}_3 + \frac{1}{2} \tilde{\mathbf{q}}_1 \int_0^t t_1^2 \mathbf{I}_2 \ \mathrm{d}\mathbf{G}_1^l \\ & + \sum_{w=2}^{n+1-j-s} \tilde{\mathbf{q}}_w \left\{ \frac{w}{w+1} t^{w+1} \mathbf{I}_3 + \frac{1}{w+1} \int_0^t t_1^{w+1} \mathbf{I}_2 \ \mathrm{d}\mathbf{G}_1^l \right\} \Bigg\}. \end{aligned}$$

$$L_4 = \sum_{j=1}^{n} \binom{n+1}{j} \int_0^t \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} t_1^w \ \mathrm{d}\widehat{G}_{t_1}^{(j)}.$$

$$\begin{aligned} L_5 &= \sum_{j=1}^{n}\binom{n+1}{j}m_j\sum_{w=1}^{n+1-j}q_w^{(n+1-j)}t^{w+1}. \\ L_6 &= -\sum_{j=1}^{n}\binom{n+1}{j}m_j\sum_{w=1}^{n+1-j}q_w^{(n+1-j)}\left\{\frac{w}{w+1}t^{w+1}\right\}. \end{aligned}$$

Next, consider L_1 and L_3. Let $u,v\in\{1,2,...,n-1\}$ and $u+v\le n$. In L_1, when $j=u$, $s=v$ (hence $s\le n-j$),

$$\begin{aligned} L_1 &= \sum_{l=1}^{v}\sum_{\theta_{l,v}\in\mathcal{K}_{l,v}}\int_0^t\frac{(n+1)!}{i_1^{\theta_{l,v}}!i_2^{\theta_{l,v}}!\cdots i_l^{\theta_{l,v}}!u!}\frac{1}{(n+1-u-v)!} \\ &\quad\times\left\{m_{n+1-u-v}t_1+\sum_{w=2}^{n+1-u-v}q_w^{(n+1-u-v)}t_1^w\right\} \\ &\quad\times\int_0^{t_1-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,v}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,v}}\right)}\mathrm{d}\widehat{G}_{t_1}^{(u)}. \end{aligned}$$

Since $s=v$, $l\in\{1,2,...,v\}$, we have by definition $\left(i_1^{\theta_{l,v}},i_2^{\theta_{l,v}},...,i_l^{\theta_{l,v}}\right)\in\mathcal{J}_v$. In L_3, when $j=n+1-u-v$ (hence $j\in\{1,...,n-1\}$), $s=u+v$ (hence $s=n+1-j$) and $i_l^{\theta_{l,s}}=u$ (hence $i_l^{\theta_{l,s}}<s$),

$$\begin{aligned} L_3 &= -m_{n+1-u-v}\sum_{l=1}^{u+v}\sum_{\theta_{l,u+v}\in\mathcal{K}_{l,u+v}}\frac{(n+1)!}{i_1^{\theta_{l,u+v}}!i_2^{\theta_{l,u+v}}!\cdots i_{l-1}^{\theta_{l,u+v}}!u!\,(n+1-u-v)!} \\ &\quad\times\int_0^t t_1\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,u+v}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,u+v}}\right)}\mathrm{d}\widehat{G}_{t_1}^{(u)}. \end{aligned}$$

Since $s=u+v$ and $i_l^{\theta_{l,s}}=u$, $\sum_{p=1}^{l-1}i_p^{\theta_{l,u+v}}=v$, we have by definition $\left(i_1^{\theta_{l,u+v}},i_2^{\theta_{l,u+v}},...,i_{l-1}^{\theta_{l,u+v}}\right)\in\mathcal{J}_v$. Hence the terms

$$\begin{aligned} &\sum_{l=1}^{v}\sum_{\theta_{l,v}\in\mathcal{K}_{l,v}}\int_0^t\frac{(n+1)!}{i_1^{\theta_{l,v}}!i_2^{\theta_{l,v}}!\cdots i_l^{\theta_{l,v}}!u!}\frac{1}{(n+1-u-v)!}m_{n+1-u-v}t_1 \\ &\times\int_0^{t_1-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,v}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,v}}\right)}\mathrm{d}\widehat{G}_1^{(u)} \end{aligned}$$

in L_1 and

$$\begin{aligned} &-m_{n+1-u-v}\sum_{l=1}^{u+v}\sum_{\theta_{l,u+v}\in\mathcal{K}_{l,u+v}}\frac{(n+1)!}{i_1^{\theta_{l,u+v}}!i_2^{\theta_{l,u+v}}!\cdots i_{l-1}^{\theta_{l,u+v}}!u!\,(n+1-u-v)!} \\ &\times\int_0^t t_1\int_0^{t_1-}\cdots\int_0^{t_{l-1}-}\mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,u+v}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,u+v}}\right)}\mathrm{d}\widehat{G}_{t_1}^{(u)} \end{aligned}$$

cancel each other. So we now have

$$L_1 = \sum_{j=1}^{n}\left\{1_{\{j\le n-1\}}\sum_{s=1}^{n+1-j}\left\{1_{\{s\le n-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\int_0^t \mathbf{F}_2\sum_{w=2}^{n+1-j-s}\tilde{\mathbf{q}}_w t_1^w \mathbf{I}_1\ \mathrm{d}\widehat{G}_{t_1}^{(j)}\right.\right.$$
$$\left.\left.+1_{\{s=n+1-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\int_0^t \mathbf{F}_1\mathbf{I}_1\ \mathrm{d}\widehat{G}_{t_1}^{(j)}\right\}+1_{\{j=n\}}(n+1)\int_0^t\int_0^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(n)}\right\}.$$

Since $q_0^{(k)}=0$ for $k>0$,

$$L_3 = -\sum_{j=1}^{n} m_j\left\{1_{\{j\le n-1\}}\sum_{s=1}^{n+1-j}\left\{1_{\{s\le n-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\mathbf{F}_2\left\{\frac{1}{2}\tilde{\mathbf{q}}_1 t^2\mathbf{I}_3+\frac{1}{2}\tilde{\mathbf{q}}_1\int_0^t t_1^2\mathbf{I}_2\ \mathrm{d}\mathbf{G}_1^l\right.\right.\right.$$
$$\left.+\sum_{w=2}^{n+1-j-s}\tilde{\mathbf{q}}_w\left\{\frac{w}{w+1}t^{w+1}\mathbf{I}_3+\frac{1}{w+1}\int_0^t t_1^{w+1}\mathbf{I}_2\ \mathrm{d}\mathbf{G}_1^l\right\}\right\}$$
$$\left.\left.+1_{\{s=n+1-j\}}\frac{(n+1)!}{(n+1-j)!j!}\int_0^t t_1\ \mathrm{d}\widehat{G}_{t_1}^{(n+1-j)}\right\}+1_{\{j=n\}}m_n(n+1)\int_0^t t_1\ \mathrm{d}\widehat{G}_{t_1}^{(1)}\right\}.$$

Next, consider L_3 and L_4. Let $u\in\{1,2,...,n\}$. In L_3, when $j=n+1-u$ (hence $j\in\{1,...,n\}$), $s=u$ (hence $s=n+1-j$) and $i_l^{\theta_{l,s}}=u$ (hence $i_l^{\theta_{l,s}}=s$), we have

$$L_3 = m_{n+1-u}\frac{(n+1)!}{u!\,(n+1-u)!}\int_0^t t_1\ \mathrm{d}\widehat{G}_{t_1}^{(u)}.$$

In L_4, when $j=u$, we have

$$L_4 = \binom{n+1}{u}\int_0^t\left\{m_{n+1-u}t_1+\sum_{w=2}^{n+1-u}q_w^{(n+1-u)}t_1^w\right\}\mathrm{d}\widehat{G}_{t_1}^{(u)}.$$

Hence the terms $m_{n+1-u}\frac{(n+1)!}{u!(n+1-u)!}\int_0^t t_1\ \mathrm{d}\widehat{G}_{t_1}^{(u)}$ in L_3 and $\binom{n+1}{u}\int_0^t m_{n+1-u}t_1\ \mathrm{d}\widehat{G}_{t_1}^{(u)}$ cancel each other. In L_3, since the terms where $(j=n)$ and $(j\le n-1,\ s=n+1-j)$ get cancelled, we can sum j from 1 to $n-1$ and sum s from 1 to $n-j$.

$$L_3 = -\sum_{j=1}^{n-1}m_j\sum_{s=1}^{n-j}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\mathbf{F}_2\sum_{w=1}^{n+1-j-s}\tilde{\mathbf{q}}_w\left\{\frac{w}{w+1}t^{w+1}\mathbf{I}_3+\frac{1}{w+1}\int_0^t t_1^{w+1}\mathbf{I}_2\ \mathrm{d}\mathbf{G}_1^l\right\}.$$

$$L_4 = \sum_{j=1}^{n-1}\left\{\binom{n+1}{j}\int_0^t\sum_{w=2}^{n+1-j}\frac{1}{w}\left[\sum_{z=1}^{n+2-j-w}\binom{n+1-j}{z}m_z q_{w-1}^{(n+1-j-z)}\right]t_1^w\ \mathrm{d}\widehat{G}_{t_1}^{(j)}\right\}$$

by (A.13). Next, consider L_4 and L_3 again. Let $u\in\{1,2,...,n-1\}$, $v\in\{1,2,...,n-1\}$, $u+v\le n$, $x\in\{1,2,...,v\}$ and hence $x+u\le n$. In L_4, when $j=u$, $w=n+2-u-v$ (hence $w\in\{2,...,n+1-j\}$), $z=x$ (hence $z\in\{1,...,n+2-j-w\}$),

$$L_4 = \binom{n+1}{u}\int_0^t\frac{1}{n+2-u-v}\binom{n+1-u}{x}m_x q_{n+1-u-v}^{(n+1-u-x)}t_1^{n+2-u-v}\ \mathrm{d}\widehat{G}_{t_1}^{(u)}. \tag{A.14}$$

In L_3, when $j=x$ (hence $j\in\{1,...,n-1\}$), $s=u$ (hence $s\le n-j$), $i_l^{\theta_{l,s}}=u$ (hence $i_l^{\theta_{l,s}}=s$), $w=n+1-u-v$ (hence $w\le n+1-s-j$ because $j=x\le v$),

$$\begin{aligned}L_3 &= m_x\frac{(n+1)!}{u!x!}\frac{1}{(n+1-x-u)!}q_{n+1-u-v}^{(n+1-u-x)}\\&\quad\times\left\{\frac{n+1-u-v}{n+2-u-v}t^{n+2-u-v}\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(u)}+\frac{1}{n+2-u-v}\int_0^t t_1^{n+2-u-v}\,\mathrm{d}\widehat{G}_{t_1}^{(u)}\right\},\end{aligned}$$

where the second term cancels (A.14). So now we have

$$\begin{aligned}L_4 &= 0.\\L_3 &= -\sum_{j=1}^{n-1}m_j\sum_{s=1}^{n-j}\left\{1_{\{s=1\}}\frac{(n+1)!}{j!}\frac{1}{(n-j)!}\sum_{w=1}^{n-j}q_w^{(n-j)}\frac{w}{w+1}t^{w+1}\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(1)}\right.\\&\quad+1_{\{2\le s\le n-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\left\{1_{\left\{i_l^{\theta_{l,s}}<s\right\}}\mathbf{F}_2\right.\\&\quad\times\sum_{w=1}^{n+1-j-s}\tilde{\mathbf{q}}_w\left\{\frac{w}{w+1}t^{w+1}\mathbf{I}_3+\frac{1}{w+1}\int_0^t t_1^{w+1}\mathbf{I}_2\,\mathrm{d}\mathbf{G}_1^l\right\}\\&\quad\left.\left.+1_{\left\{i_l^{\theta_{l,s}}=s\right\}}\frac{(n+1)!}{s!j!}\frac{1}{(n+1-j-s)!}\sum_{w=1}^{n+1-j-s}\tilde{\mathbf{q}}_w\frac{w}{w+1}t^{w+1}\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(s)}\right\}\right\}.\end{aligned}$$

Next, consider L_1 and L_3. By the equation for $q_w^{(n+1-j-s)}$ given in (A.13), we have

$$\begin{aligned}L_1 &= \sum_{j=1}^{n}\left\{1_{\{j\le n-1\}}\sum_{s=1}^{n+1-j}\left\{1_{\{s\le n-1-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\int_0^t\mathbf{F}_2\sum_{w=2}^{n+1-j-s}\frac{1}{w}\right.\right.\\&\quad\times\sum_{z=1}^{n+2-j-s-w}\binom{n+1-j-s}{z}m_zq_{w-1}^{(n+1-j-s-z)}t_1^w\mathbf{I}_1\,\mathrm{d}\widehat{G}_{t_1}^{(j)}\\&\quad\left.\left.+1_{\{s=n+1-j\}}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\int_0^t\mathbf{F}_1\mathbf{I}_1\,\mathrm{d}\widehat{G}_{t_1}^{(j)}\right\}+1_{\{j=n\}}\,(n+1)\int_0^t\int_0^{t_1-}\mathrm{d}\widehat{G}_{t_2}^{(1)}\mathrm{d}\widehat{G}_{t_1}^{(n)}\right\}.\end{aligned}$$

Let $u\in\{1,2,...,n-2\}$, $v\in\{1,2,...,n-2\}$, $u+v\le n-1$, $x\in\{1,2,...,v\}$, $\beta\in\{1,2,...,v+1-x\}$. In L_1, when $j=u$, $s=n-u-v$ (hence $s\in\{1,...,n-1-j\}$), $w=x+1$ (hence $w\in\{2,...,n+1-j-s\}$), $z=\beta$ (hence $z\in\{1,...,n+2-j-s-w\}$),

$$\begin{aligned}L_1 &= \sum_{l=1}^{n-u-v}\sum_{\theta_{l,n-u-v}\in\mathcal{K}_{l,n-u-v}}\int_0^t\frac{(n+1)!}{i_1^{\theta_{l,n-u-v}}!i_2^{\theta_{l,n-u-v}}!\cdots i_l^{\theta_{l,n-u-v}}!u!}\frac{1}{(v+1)!}\frac{1}{x+1}\binom{v+1}{\beta}\\&\quad\times m_\beta q_x^{(v+1-\beta)}t_1^{x+1}\int_0^{t_1-}\cdots\int_0^{t_l-}\mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,n-u-v}}\right)}\cdots\mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,n-u-v}}\right)}\mathrm{d}\widehat{G}_{t_1}^{(u)}.\end{aligned}\tag{A.15}$$

By definition, since $s=n-u-v$ and $l\in\{1,2,...,n-u-v\}$, $\left(i_1^{\theta_{l,n-u-v}},i_2^{\theta_{l,n-u-v}},...,i_l^{\theta_{l,n-u-v}}\right)\in$

$\mathcal{J}_{n-u-v}$. In L_3, when $j=\beta$ (hence $j \in \{1,...,n-2\}$), $s=n-v$ (hence $s \in \{2,...,n-j\}$), $i_l^{\theta_{l,s}}=u$ (hence $i_l^{\theta_{l,s}}<s$), $w=x$ (hence $w \in \{1,...,n+1-j-s\}$),

$$\begin{aligned} L_3 &= -m_\beta \sum_{l=1}^{n-v} \sum_{\theta_{l,n-v}\in\mathcal{K}_{l,n-v}} \frac{(n+1)!}{i_1^{\theta_{l,n-v}}! i_2^{\theta_{l,n-v}}! \cdots i_{l-1}^{\theta_{l,n-v}}! u! \beta!} \frac{1}{(v+1-\beta)!} \\ &\times q_x^{(v+1-\beta)} \Bigg\{ \frac{x}{x+1} t^{x+1} \int_0^t \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,n-v}}\right)} \cdots \mathrm{d}\widehat{G}_{t2}^{\left(i_{l-1}^{\theta_{l,n-v}}\right)} \mathrm{d}\widehat{G}_{t_1}^{(u)} \\ &+ \frac{1}{x+1} \int_0^t t_1^{x+1} \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,n-v}}\right)} \cdots \mathrm{d}\widehat{G}_{t2}^{\left(i_{l-1}^{\theta_{l,n-v}}\right)} \mathrm{d}\widehat{G}_{t_1}^{(u)} \Bigg\}. \end{aligned}$$

The final term in L_3

$$\begin{aligned} &-m_\beta \sum_{l=1}^{n-v} \sum_{\theta_{l,n-v}\in\mathcal{K}_{l,n-v}} \frac{(n+1)!}{i_1^{\theta_{l,n-v}}! i_2^{\theta_{l,n-v}}! \cdots i_{l-1}^{\theta_{l,n-v}}! u! \beta!} \frac{1}{(v+1-\beta)!} \\ &\times q_x^{(v+1-\beta)} \frac{1}{x+1} \int_0^t t_1^{x+1} \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,n-v}}\right)} \cdots \mathrm{d}\widehat{G}_{t2}^{\left(i_{l-1}^{\theta_{l,n-v}}\right)} \mathrm{d}\widehat{G}_{t_1}^{(u)} \end{aligned}$$

clearly cancels (A.15) in L_1. So now we can write

$$\begin{aligned} L_1 &= \sum_{j=1}^{n} 1_{\{j\le n\}} \sum_{s=1}^{n+1-j} 1_{\{s=n+1-j\}} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} \int_0^t \mathbf{F}_1 \mathbf{I}_1 \,\mathrm{d}\widehat{G}_{t_1}^{(j)}. \\ L_3 &= -\sum_{j=1}^{n-1} m_j \sum_{s=1}^{n-j} \Bigg\{ 1_{\{s=1\}} \frac{(n+1)!}{j!} \frac{1}{(n-j)!} \sum_{w=1}^{n-j} q_w^{(n-j)} \frac{w}{w+1} t^{w+1} \int_0^t \mathrm{d}\widehat{G}_{t_1}^{(1)} \\ &+ 1_{\{2\le s\le n-j\}} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} \Bigg\{ 1_{\left\{i_l^{\theta_{l,s}}<s\right\}} \mathbf{F}_2 \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w \frac{w}{w+1} t^{w+1} \mathbf{I}_3 \\ &+ 1_{\left\{i_l^{\theta_{l,s}}=s\right\}} \frac{(n+1)!}{s!j!} \frac{1}{(n+1-j-s)!} \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w \frac{w}{w+1} t^{w+1} \int_0^t \mathrm{d}\widehat{G}_{t_1}^{(s)} \Bigg\} \Bigg\}. \end{aligned}$$

We can now simplify it as

$$L_3 = -\sum_{j=1}^{n-1} m_j \sum_{s=1}^{n-j} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} \mathbf{F}_2 \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w \frac{w}{w+1} t^{w+1}\, \mathbf{I}_3.$$

Altogether, we have

$$L_1 = \sum_{j=1}^{n} 1_{\{j\le n\}} \left\{ \sum_{s=1}^{n+1-j} 1_{\{s=n+1-j\}} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} \int_0^t \mathbf{F}_1 \mathbf{I}_1 \,\mathrm{d}\widehat{G}_{t_1}^{(j)} \right\}.$$

$$L_2 = \sum_{j=1}^{n} m_j \sum_{s=1}^{n+1-j} \left\{ 1_{\{s=n+1-j\}} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} t\mathbf{F}_1\mathbf{I}_3 + 1_{\{s\le n-j\}} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} \mathbf{F}_2 \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w t^{w+1} \mathbf{I}_3 \right\}.$$

$$L_3 = -\sum_{j=1}^{n-1} m_j \sum_{s=1}^{n-j} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} \mathbf{F}_2 \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w \frac{w}{w+1} t^{w+1} \mathbf{I}_3.$$

$$L_4 = 0 \quad \text{and} \quad L_5 = \sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} t^{w+1}.$$

$$L_6 = -\sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} \frac{w}{w+1} t^{w+1}.$$

Since at the beginning of the proof, we have already showed that the stochastic integrals of G_t^{n+1} are of the form $\mathcal{S}_{\theta_{n+1},t,0}$ where $\theta_{n+1,t} \in \mathcal{I}_{n+1}$. We are now going to show that the coefficient of each $\mathcal{S}_{\theta_{n+1},t,0}$ is $\Pi^{(n+1)}_{\theta_{n+1,t}}$. Consider $\int_0^t \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_{l+1}}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_2}^{\left(i_l^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_1}^{(j)}$, where $\theta_{l,s} \in \mathcal{K}_{l,s},\ j \in \{1,2,...,n\},\ s = n+1-j$. This stochastic integral only appears in L_1 and its coefficient is $\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}$. And from (3.11), since $n+1-s-j = n+1-(n+1-j)-j = 0$,

$$\Pi^{(n+1)}_{\left(i_1^{\theta_{l,s}},i_2^{\theta_{l,s}},\ldots,i_l^{\theta_{l,s}},j\right)} = \frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!0!} C_t^{(0)} = \frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!j!}$$

since $C_t^{(0)} = 1$ by definition (3.8). Hence we have proved that the coefficient is given by $\Pi^{(n+1)}_{\left(i_1^{\theta_{l,s}},i_2^{\theta_{l,s}},\ldots,i_l^{\theta_{l,s}},j\right)}$. Next, we change the summation sign of j and s in L_2 to obtain:

$$L_2 = \sum_{s=1}^{n} \left\{ \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} m_{n+1-s} t \frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!\,(n+1-s)!} \mathbf{I}_3 + \sum_{j=1}^{n-s} \sum_{l=1}^{s} \sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} m_j \mathbf{F}_2 \sum_{w=1}^{n+1-j-s} \tilde{\mathbf{q}}_w t^{w+1} \mathbf{I}_3 \right\}.$$

Similarly, by changing the summation sign of j and w, we have

$$\begin{aligned} L_2 &= \sum_{s=1}^{n}\left\{\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} m_{n+1-s}t\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!\,(n+1-s)!}\mathbf{I}_3\right.\\ &\left.+\sum_{w=1}^{n-s}\sum_{j=1}^{n+1-w-s}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} m_j\mathbf{F}_2\tilde{\mathbf{q}}_w t^{w+1}\mathbf{I}_3\right\}. \end{aligned}$$

By (A.13), $\frac{1}{w+1}\sum_{j=1}^{n+1-w-s}\frac{(n+1-s)!}{j!(n+1-j-s)!}m_j q_w^{(n+1-s-j)} = q_{w+1}^{(n+1-s)}$, so we have

$$\begin{aligned} L_2 &= \sum_{s=1}^{n}\left\{\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} m_{n+1-s}t\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!\,(n+1-s)!}\mathbf{I}_3\right.\\ &\left.+\sum_{w=1}^{n-s}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!}\,(w+1)\,\frac{1}{(n+1-s)!}q_{w+1}^{(n+1-s)}t^{w+1}\mathbf{I}_3\right\}. \end{aligned}$$

Changing $\sum_{w=1}^{n-s}$ to $\sum_{w=2}^{n+1-s}$, we have

$$\begin{aligned} L_2 &= \sum_{s=1}^{n}\left\{\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}} m_{n+1-s}t\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!\,(n+1-s)!}\mathbf{I}_3\right.\\ &\left.+\sum_{w=2}^{n+1-s}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!}\frac{w}{(n+1-s)!}q_{w}^{(n+1-s)}t^{w}\mathbf{I}_3\right\}. \end{aligned}$$

Similarly,

$$L_3 = -\sum_{s=1}^{n-1}\sum_{w=1}^{n-s}\sum_{j=1}^{n+1-w-s} m_j\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\mathbf{F}_2\tilde{\mathbf{q}}_w\frac{w}{w+1}t^{w+1}\mathbf{I}_3.$$

By (A.13), $\frac{1}{w+1}\sum_{j=1}^{n+1-w-s}\frac{(n+1-s)!}{j!(n+1-j-s)!}m_j q_w^{(n+1-s-j)} = q_{w+1}^{(n+1-s)}$, so we have

$$L_3 = -\sum_{s=1}^{n-1}\sum_{w=2}^{n+1-s}\sum_{l=1}^{s}\sum_{\theta_{l,s}\in\mathcal{K}_{l,s}}\frac{(n+1)!}{i_1^{\theta_{l,s}}!i_2^{\theta_{l,s}}!\cdots i_l^{\theta_{l,s}}!}\frac{w-1}{(n+1-s)!}q_w^{(n+1-s)}t^w\mathbf{I}_3.$$

For $s=1$, the stochastic integral $\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(1)}$ appears in both L_2 and L_3. Its coefficient is given by

$$\begin{aligned} &\sum_{w=2}^{n}(n+1)\,wq_w^{(n)}t^w + m_n\,(n+1)\,t - (n+1)\sum_{w=2}^{n}(w-1)\,q_w^{(n)}t^w\\ &= (n+1)\left[m_n t + \sum_{w=2}^{n} q_w^{(n)}t^w\right] = (n+1)\,C_t^{(n)}. \end{aligned}$$

By (3.11),

$$\Pi_{(1)}^{(n+1)} = \frac{(n+1)!}{(n+1-1)!}C_t^{(n+1-1)} = (n+1)\,C_t^{(n)}.$$

For $s \in \{2, 3, ..., n-1\}$, the coefficient of the stochastic integral

$$\int_0^t \int_0^{t_1-} \cdots \int_0^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)}$$

is given by

$$\begin{aligned}
& m_{n+1-s} t \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! \,(n+1-s)!} + \sum_{w=2}^{n+1-s} \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}!} \frac{w}{(n+1-s)!} q_w^{(n+1-s)} t^w \\
& - \sum_{w=2}^{n+1-s} \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}!} \frac{(w-1)}{(n+1-s)!} q_w^{(n+1-s)} t^w \\
& = \sum_{w=1}^{n+1-s} \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}!} \frac{1}{(n+1-s)!} q_w^{(n+1-s)} t^w \\
& = \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}!} \frac{1}{(n+1-s)!} C_t^{(n+1-s)} = \Pi^{(n+1)}_{\left(i_1^{\theta_{l,s}}, i_2^{\theta_{l,s}}, ..., i_l^{\theta_{l,s}}\right)}
\end{aligned}$$

by (3.11). For $s = n$, the stochastic integral appears in L_2 only and its coefficient is given by

$$m_1 t \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}!} = \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}!} C_t^{(1)} = \Pi^{(n+1)}_{\left(i_1^{\theta_{l,s}}, i_2^{\theta_{l,s}}, ..., i_l^{\theta_{l,s}}\right)}.$$

The stochastic integral $\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(n+1)}$ appears only once in G_t^{n+1} and its coefficient is equal to one. By (3.11),

$$\Pi^{(n+1)}_{(n+1)} = \frac{(n+1)!}{(n+1)!} C_t^{(0)} = 1.$$

For stochastic integrals where the indices add up exactly to $n+1$, other than $\int_0^t \mathrm{d}\widehat{G}_{t_1}^{(n+1)}$, they only appear in L_1 and their coefficients are given by:

$$\frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! j!} = \frac{(n+1)!}{i_1^{\theta_{l,s}}! i_2^{\theta_{l,s}}! \cdots i_l^{\theta_{l,s}}! j!} C_t^{(0)} = \Pi^{(n+1)}_{\left(i_1^{\theta_{l,s}}, i_2^{\theta_{l,s}}, ..., i_l^{\theta_{l,s}}, j\right)},$$

where the sum of $i_1^{\theta_{l,s}}, i_2^{\theta_{l,s}}, ..., i_l^{\theta_{l,s}}, j$ is equal to $n+1$. Finally, we have to show that $L_5 + L_6 + m_{n+1} t = C_t^{(n+1)}$. By (A.13),

$$\frac{1}{w+1} \sum_{j=1}^{n+1-w} \binom{n+1}{j} m_j q_w^{(n+1-j)} = q_{w+1}^{(n+1)},$$

we have

$$\begin{aligned} L_5 &= \sum_{j=1}^{n} \binom{n+1}{j} m_j \sum_{w=1}^{n+1-j} q_w^{(n+1-j)} t^{w+1} \\ &= \sum_{w=1}^{n} \sum_{j=1}^{n+1-w} \binom{n+1}{j} m_j q_w^{(n+1-j)} t^{w+1} = \sum_{w=1}^{n} (w+1)\, q_{w+1}^{(n+1)} t^{w+1}. \\ L_6 &= -\sum_{w=1}^{n} \sum_{j=1}^{n+1-w} \binom{n+1}{j} m_j q_w^{(n+1-j)} \frac{w}{w+1} t^{w+1} = -\sum_{w=1}^{n} w q_{w+1}^{(n+1)} t^{w+1}. \end{aligned}$$

Hence

$$L_5 + L_6 + m_{n+1} t = \sum_{w=1}^{n} q_{w+1}^{(n+1)} t^{w+1} + m_{n+1} t = \sum_{w=2}^{n+1} q_w^{(n+1)} t^w + m_{n+1} t = \sum_{w=1}^{n+1} q_w^{(n+1)} t^w = C_{n+1}^{(k)}.$$

Thus, we have proved that

$$G_t^{n+1} = \sum_{\theta_{n+1} \in \mathcal{I}_{n+1}} \Pi_{\theta_{n+1},t}^{(n+1)} \mathcal{S}_{\theta_{n+1},t,0} + C_t^{(n+1)}.$$

As explained in the beginning of the proof, since $\{G_t, t \geq 0\}$ and $\{G_{t+t_0} - G_{t_0}, t \geq 0\}$ have the same distribution and since $\mathrm{d}\left(\widehat{G}_{t+t_0}^{(i)} - \widehat{G}_{t_0}^{(i)}\right) = \mathrm{d}\widehat{G}_{t+t_0}^{(i)}$, we have

$$\begin{aligned} &(G_{t+t_0} - G_{t_0})^{n+1} \\ &= \sum_{s=1}^{n+1} \sum_{l=1}^{s} \sum_{\theta_{l,s} \in \mathcal{K}_{l,s}} \Pi_{\theta_{l,s},t}^{(n+1)} \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \cdots \int_{t_0}^{t_{l-1}-} \mathrm{d}\widehat{G}_{t_l}^{\left(i_1^{\theta_{l,s}}\right)} \cdots \mathrm{d}\widehat{G}_{t_2}^{\left(i_{l-1}^{\theta_{l,s}}\right)} \mathrm{d}\widehat{G}_{t_1}^{\left(i_l^{\theta_{l,s}}\right)} \\ &= \sum_{\theta_{n+1} \in \mathcal{I}_{n+1}} \Pi_{\theta_{n+1},t}^{(n+1)} \mathcal{S}_{\theta_{n+1},t,t_0} + C_t^{(n+1)}. \end{aligned}$$

Therefore, by the principle of strong induction,

$$(G_{t+t_0} - G_{t_0})^k = \sum_{\theta_k \in \mathcal{I}_k} \Pi_{\theta_k,t}^{(k)} \mathcal{S}_{\theta_k,t,t_0} + C_t^{(k)}$$

for all non-negative integers k.

A.5 Proof of Proposition 3.1.5

We prove by induction. For $n = 1$, $Y_t^{(1)} = H_t^{(1)}$ by definition and since $\mathcal{M}^{n,n} = \{1\}$, the proposition is true. Assume the proposition is true for all $k \geq n$. Now, consider $n+1$,

$$\begin{aligned} Y_t^{(n+1)} &= H_t^{(n+1)} - \sum_{l=1}^{n} a_{n+1,l} Y_t^{(l)} = H_t^{(n+1)} - \sum_{l=1}^{n} a_{n+1,l} \left\{ H_t^{(l)} + \sum_{k=1}^{l-1} b_{l,k} H_t^{(k)} \right\} \\ &= H_t^{(n+1)} - \sum_{l=1}^{n} a_{n+1,l} \sum_{k=1}^{l} b_{l,k} H_t^{(k)} = H_t^{(n+1)} + \sum_{k=1}^{n} b_{n+1,k} H_t^{(k)}, \end{aligned}$$

which completes the proof.

A.6 Calculation of $(X_{t+t_0} - X_{t_0})^k$ for $k = 3, 4, 5$ when $\sigma \neq 0$

A.6.1 $(X_{t+t_0} - X_{t_0})^3$

When $k = 3$, from (3.14),

$$\begin{aligned} &(X_{t+t_0} - X_{t_0})^3 \\ &= \frac{\sigma^2}{2} 3 \cdot 2 (X_{t+t_0} - X_{t_0}) t - \frac{\sigma^2}{2} 3 \cdot 2 \int_{t_0}^{t+t_0} (s - t_0) \,\mathrm{d} (X_s - X_{t_0}) \\ &\quad + 3 \int_{t_0}^{t+t_0} \sigma^2 (s - t_0) \,\mathrm{d} Y_s^{(1)} + 3 m_1 (t + t_0) \sigma^2 t \\ &\quad - 3 m_1 \int_{t_0}^{t+t_0} s \mathrm{d} \left[\sigma^2 (s - t_0) \right] + A_1 (X_{t+t_0}, X_{t_0}; 3) \\ &= 3\sigma^2 (X_{t+t_0} - X_{t_0}) t - 3\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \,\mathrm{d} \left[\int_{t_0}^{s} \mathrm{d} \left[Y_{t_1}^{(1)} + m_1 t_1 \right] \right] \\ &\quad + 3\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \,\mathrm{d} Y_s^{(1)} + 3 m_1 (t + t_0) \sigma^2 t - \frac{3}{2} m_1 \sigma^2 \left[s^2 \right]_{t_0}^{t+t_0} + A_1 (X_{t+t_0}, X_{t_0}; 3) \\ &= 3\sigma^2 (X_{t+t_0} - X_{t_0}) t - 3\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \,\mathrm{d} Y_s^{(1)} - 3\sigma^2 m_1 \int_{t_0}^{t+t_0} (s - t_0) \,\mathrm{d} (s - t_0) \\ &\quad + 3\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \,\mathrm{d} Y_s^{(1)} + 3 m_1 (t + t_0) \sigma^2 t - \frac{3}{2} m_1 \sigma^2 \left(t^2 + 2 t t_0 \right) \\ &\quad + A_1 (X_{t+t_0}, X_{t_0}; 3) \\ &= 3\sigma^2 (X_{t+t_0} - X_{t_0}) t - \frac{3}{2} \sigma^2 m_1 \left[(s - t_0)^2 \right]_{t_0}^{t+t_0} + \frac{3}{2} m_1 \sigma^2 t^2 + A_1 (X_{t+t_0}, X_{t_0}; 3) \\ &= 3\sigma^2 (X_{t+t_0} - X_{t_0}) t - \frac{3}{2} \sigma^2 m_1 t^2 + \frac{3}{2} m_1 \sigma^2 t^2 + A_1 (X_{t+t_0}, X_{t_0}; 3) \\ &= 3\sigma^2 A_1 (X_{t+t_0}, X_{t_0}; 1) t + A_1 (X_{t+t_0}, X_{t_0}; 3). \end{aligned}$$

A.6.2 $(X_{t+t_0} - X_{t_0})^4$

When $k = 4$, we are going to have a closer look at the cancellation pattern.

$$
\begin{aligned}
(X_{t+t_0} - X_{t_0})^4 &= \frac{\sigma^2}{2} 4 \cdot 3 \left(X_{t+t_0} - X_{t_0}\right)^2 t - \frac{\sigma^2}{2} 4 \cdot 3 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left(X_s - X_{t_0}\right)^2 \\
&\quad + \binom{4}{1} \int_{t_0}^{t+t_0} A_2 \left(X_{t+t_0}, X_{t_0}; 3\right) \mathrm{d}Y_s^{(1)} + \binom{4}{2} \int_{t_0}^{t+t_0} A_2 \left(X_{t+t_0}, X_{t_0}; 2\right) \mathrm{d}Y_s^{(2)} \\
&\quad + \binom{4}{1} m_1 \left(t + t_0\right) A_2 \left(X_{t+t_0}, X_{t_0}; 3\right) + \binom{4}{2} m_2 \left(t + t_0\right) A_2 \left(X_{t+t_0}, X_{t_0}; 2\right) \\
&\quad - \binom{4}{1} m_1 \int_{t_0}^{t+t_0} s \mathrm{d} A_2 \left(X_s, X_{t_0}; 3\right) - \binom{4}{2} m_2 \int_{t_0}^{t+t_0} s \mathrm{d} A_2 \left(X_s, X_{t_0}; 2\right) \\
&\quad + A_1 \left(X_{t+t_0}, X_{t_0}; 4\right).
\end{aligned}
$$

Let

$$
\begin{aligned}
(X_{t+t_0} - X_{t_0})^4 &= \frac{\sigma^2}{2} 4 \cdot 3 \left(X_{t+t_0} - X_{t_0}\right)^2 t + M_1 + M_2 + M_3 \\
&\quad + M_4 + M_5 + M_6 + M_7 + A_1 \left(X_{t+t_0}, X_{t_0}; 4\right). \\
M_1 &= -6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left(X_s - X_{t_0}\right)^2 \\
&= -6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left[\sigma^2 (s - t)\right] \\
&\quad -6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left[2 \int_{t_0}^{s} \int_{t_0}^{t_1 -} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)}\right] \\
&\quad -6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left[2 m_1 (s - t_0) \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right] \\
&\quad -6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left[\int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(2)}\right] \\
&\quad -6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left[m_1^2 (s - t_0)^2 + m_2 (s - t_0)\right] \\
&= -3\sigma^4 \left[(s - t_0)^2\right]_{t_0}^{t+t_0} - 12\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_{t_2}^{(1)} \\
&\quad -12\sigma^2 m_1 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d} \left[(s - t_0) \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right] - 6\sigma^2 \int_{t_0}^{t+t_0} (s - t_0) \, \mathrm{d}Y_s^{(2)} \\
&\quad -12\sigma^2 m_1^2 \int_{t_0}^{t+t_0} (s - t_0)^2 \, \mathrm{d} (s - t_0) - 3\sigma^2 m_2 \left[(s - t_0)^2\right]_{t_0}^{t+t_0}.
\end{aligned}
$$

$$
\begin{aligned}
M_2 &= 4\int_{t_0}^{t+t_0} 3\sigma^2 \left(X_{s-} - X_{t_0}\right)(s-t_0)\,\mathrm{d}Y_s^{(1)} \\
&= 12\sigma^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s} \mathrm{d}\left[Y_{t_1}^{(1)} + m_1 t_1\right](s-t_0)\,\mathrm{d}Y_s^{(1)}. \\
M_3 &= 6\int_{t_0}^{t+t_0} \sigma^2 (s-t_0)\,\mathrm{d}Y_s^{(2)}.
\end{aligned}
$$

In M_1, the second term is cancelled with the first term in M_2 and the fourth term is cancelled with M_3. Hence we now have

$$
\begin{aligned}
M_1 &= -3\sigma^4 t^2 - 12\sigma^2 m_1 \left[(s-t_0)^2 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} - 3\sigma^2 m_2 t^2 \\
&\quad +12\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\mathrm{d}(s-t_0) - 4\sigma^2 m_1^2 \left[(s-t_0)^3\right]_{t_0}^{t+t_0} \\
&= -3\sigma^4 t^2 - 12\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} + 6\sigma^2 m_1 \int_{t_0}^{t+t_0} \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\mathrm{d}(s-t_0)^2 \\
&\quad -4\sigma^2 m_1^2 t^3 - 3\sigma^2 m_2 t^2. \\
M_2 &= 12\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2\,\mathrm{d}Y_s^{(1)} \text{ and } M_3 = 0. \\
M_4 &= 4m_1 (t+t_0)\, 3\sigma^2 \left(X_{t+t_0} - X_{t_0}\right) t = 12 m_1 \sigma^2 (t+t_0) \int_{t_0}^{t+t_0} \mathrm{d}\left[Y_s^{(1)} + m_1 s\right] t \\
&= 12 m_1 \sigma^2 t (t+t_0) \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} + 12 m_1^2 \sigma^2 (t+t_0)\, t^2. \\
M_5 &= 6 m_2 (t+t_0)\,\sigma^2 t. \\
M_6 &= -4m_1 \int_{t_0}^{t+t_0} s\,\mathrm{d}\left[3\sigma^2 \left(X_{s-} - X_{t_0}\right)(s-t_0)\right] \\
&= -12 m_1 \sigma^2 \int_{t_0}^{t+t_0} s\,\mathrm{d}\left[\int_{t_0}^{s} \mathrm{d}\left[Y_{t_1}^{(1)} + m_1 t_1\right](s-t_0)\right] \\
&= -12 m_1 \sigma^2 \int_{t_0}^{t+t_0} s\,\mathrm{d}\left[(s-t_0) \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right] - 12 m_1^2 \sigma^2 \int_{t_0}^{t+t_0} s\,\mathrm{d}\left[(s-t_0)^2\right] \\
&= -12 m_1 \sigma^2 \int_{t_0}^{t+t_0} s\,\mathrm{d}\left[s \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right] + 12 m_1 \sigma^2 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} \\
&\quad -12 m_1^2 \sigma^2 \int_{t_0}^{t+t_0} s\mathrm{d}s^2 + 24 m_1^2 \sigma^2 \int_{t_0}^{t+t_0} s\,\mathrm{d}\left[t_0 s\right] \\
&= -12 m_1 \sigma^2 \left[s^2 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} + 12 m_1 \sigma^2 \int_{t_0}^{t+t_0} s \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s \\
&\quad +12 m_1 \sigma^2 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} - 24 m_1^2 \sigma^2 \int_{t_0}^{t+t_0} s^2\mathrm{d}s + 12 m_1^2 \sigma^2 t_0 \left[s^2\right]_{t_0}^{t+t_0}
\end{aligned}
$$

$$
\begin{aligned}
M_6 &= -12m_1\sigma^2\left(t^2+2tt_0+t_0^2\right)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+6m_1\sigma^2\int_{t_0}^{t+t_0}\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s^2\\
&\quad +12m_1\sigma^2t_0\int_{t_0}^{t+t_0}s\mathrm{d}Y_s^{(1)}-8m_1^2\sigma^2\left[s^3\right]_{t_0}^{t+t_0}+12m_1^2\sigma^2t_0\left[t^2+2tt_0\right]\\
&= -12m_1\sigma^2\left(t^2+2tt_0+t_0^2\right)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+6m_1\sigma^2\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}-6m_1\sigma^2\int_{t_0}^{t+t_0}s^2\mathrm{d}Y_s^{(1)}\\
&\quad +12m_1\sigma^2t_0\int_{t_0}^{t+t_0}s\mathrm{d}Y_s^{(1)}-8m_1^2\sigma^2\left[t^3+3t^2t_0+3tt_0^2\right]+12m_1^2\sigma^2t_0\left[t^2+2tt_0\right],\\
&= -6m_1\sigma^2\left(t^2+2tt_0+t_0^2\right)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}-6m_1\sigma^2\int_{t_0}^{t+t_0}s^2\mathrm{d}Y_s^{(1)}\\
&\quad +12m_1\sigma^2t_0\int_{t_0}^{t+t_0}s\mathrm{d}Y_s^{(1)}-8m_1^2\sigma^2\left[t^3+3t^2t_0+3tt_0^2\right]+12m_1^2\sigma^2t_0\left[t^2+2tt_0\right].\\
M_7 &= -6m_2\int_{t_0}^{t+t_0}s\mathrm{d}\left[\sigma^2\left(s-t_0\right)\right]=3m_2\sigma^2\left[s^2\right]_{t_0}^{t+t_0}=-3m_2\sigma^2\left[t^2+2tt_0\right].
\end{aligned}
$$

The second terms in M_5 and M_7 cancel each other. In M_1, the second term is cancelled with the first term in M_4, the last term is cancelled with the first term in M_5 and the first term in M_7. Now we have

$$
\begin{aligned}
M_1 &= -3\sigma^4t^2+6\sigma^2m_1\left[(s-t_0)^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}-6\sigma^2m_1\int_{t_0}^{t+t_0}(s-t_0)^2\mathrm{d}Y_s^{(1)}\\
&\quad -4\sigma^2m_1^2t^3\\
&= -3\sigma^4t^2+6\sigma^2m_1t^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}-6\sigma^2m_1\int_{t_0}^{t+t_0}(s-t_0)^2\mathrm{d}Y_s^{(1)}-4\sigma^2m_1^2t^3.\\
M_4 &= 12m_1\sigma^2tt_0\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+12m_1^2\sigma^2t^3+12m_1^2\sigma^2t^2t_0.\\
M_5 &= M_7=0.\\
M_1+M_2+M_4 &= -3\sigma^4t^2+6\sigma^2m_1t^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+6\sigma^2m_1\int_{t_0}^{t+t_0}(s-t_0)^2\mathrm{d}Y_s^{(1)}\\
&\quad +8\sigma^2m_1^2t^3+12m_1\sigma^2tt_0\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+12m_1^2\sigma^2t^2t_0.
\end{aligned}
$$

Hence, $M_1+M_2+M_4+M_6=-3\sigma^4t^2$ and we have

$$
\left(X_{t+t_0}-X_{t_0}\right)^4=6\sigma^2tA_1\left(X_{t+t_0},X_{t_0};2\right)+3\sigma^4t^2+A_1\left(X_{t+t_0},X_{t_0};4\right).
$$

A.6.3 $(X_{t+t_0}-X_{t_0})^5$

When $k=5$, from (A.4)-(A.6),

$$\begin{aligned}
(X_{t+t_0}-X_{t_0})^5 &= \frac{\sigma^2}{2}5\cdot 4\,(X_{t+t_0}-X_{t_0})^3\,t-\frac{\sigma^2}{2}5\cdot 4\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\,(X_s-X_{t_0})^3\\
&\quad +5\int_{t_0}^{t+t_0}A_2\,(X_{s-},X_{t_0};4)\,\mathrm{d}Y_s^{(1)}+10\int_{t_0}^{t+t_0}A_2\,(X_{s-},X_{t_0};3)\,\mathrm{d}Y_s^{(2)}\\
&\quad +10\int_{t_0}^{t+t_0}A_2\,(X_{s-},X_{t_0};2)\,\mathrm{d}Y_s^{(3)}+5m_1\,(t+t_0)\,A_2\,(X_{t+t_0},X_{t_0};4)\\
&\quad +10m_2\,(t+t_0)\,A_2\,(X_{t+t_0},X_{t_0};3)+10m_3\,(t+t_0)\,A_2\,(X_{t+t_0},X_{t_0};2)\\
&\quad -5m_1\int_{t_0}^{t+t_0}s\mathrm{d}A_2\,(X_s,X_{t_0};4)-10m_2\int_{t_0}^{t+t_0}s\mathrm{d}A_2\,(X_s,X_{t_0};3)\\
&\quad -10m_3\int_{t_0}^{t+t_0}s\mathrm{d}A_2\,(X_s,X_{t_0};2)+A_1\,(X_{t+t_0},X_{t_0};5)\\
&= 10\sigma^2\,(X_{t+t_0}-X_{t_0})^3\,t+N_1+N_2+N_3+N_4+N_5\\
&\quad +N_6+N_7+N_8+N_9+N_{10}+A_1\,(X_{t+t_0},X_{t_0};5)\,.\\
N_1 &= -\frac{\sigma^2}{2}5\cdot 4\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\,(X_s-X_{t_0})^3\\
&= -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[3\sigma^2A_1\,(X_s,X_{t_0};1)\,(s-t_0)+A_1\,(X_s,X_{t_0};3)\right]\\
&= -30\sigma^4\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[(X_s-X_{t_0})\,(s-t_0)\right]\\
&\quad -60\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)}\\
&\quad -30\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}Y_s^{(1)}-30\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(2)}\\
&\quad -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[6m_1\,(s-t_0)\int_{t_0}^{s}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\right]\\
&\quad -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[m_1^3\,(s-t_0)^3\right]\\
&\quad -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}Y_s^{(3)}-30\sigma^2m_1\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[(s-t_0)\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(2)}\right]\\
&\quad -30\sigma^2m_1^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[(s-t_0)^2\int_{t_0}^{s}dY_{t_1}^{(1)}\right]\\
&\quad -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[3m_2\,(s-t_0)\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]\\
&\quad -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[3m_1m_2\,(s-t_0)^2\right]-10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}\left[m_3\,(s-t_0)\right]
\end{aligned}$$

$$\begin{aligned}
N_1 &= -30\sigma^4\left[(s-t_0)^2\left(X_s-X_{t_0}\right)\right]_{t_0}^{t+t_0}+30\sigma^4\int_{t_0}^{t+t_0}(s-t_0)\left(X_{s-}-X_{t_0}\right)\mathrm{d}\left(s-t_0\right)\\
&\quad -60\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)}\\
&\quad -30\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}Y_s^{(1)}-30\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(2)}\\
&\quad -60\sigma^2 m_1\left[(s-t_0)^2\int_{t_0}^{s}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}\\
&\quad +60\sigma^2 m_1\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}\left(s-t_0\right)\\
&\quad -10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}Y_s^{(3)}-30\sigma^2 m_1\left[(s-t_0)^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(2)}\right]_{t_0}^{t+t_0}\\
&\quad +30\sigma^2 m_1\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}\left(s-t_0\right)-30\sigma^2 m_1^2\left[(s-t_0)^3\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}\\
&\quad +30\sigma^2 m_1^2\int_{t_0}^{t+t_0}(s-t_0)^2\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}\left(s-t_0\right)-30\sigma^2 m_2\left[(s-t_0)^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}\\
&\quad +30\sigma^2 m_2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}\left(s-t_0\right)-30\sigma^2 m_1^3\int_{t_0}^{t+t_0}(s-t_0)^3\,\mathrm{d}\left(s-t_0\right)\\
&\quad -60\sigma^2 m_1 m_2\int_{t_0}^{t+t_0}(s-t_0)^2\,\mathrm{d}\left(s-t_0\right)-5\sigma^2 m_3\left[(s-t_0)^2\right]_{t_0}^{t+t_0}\\
&= -30\sigma^4 t^2\left(X_{t+t_0}-X_{t_0}\right)+15\sigma^4\int_{t_0}^{t+t_0}\left(X_{s-}-X_{t_0}\right)\mathrm{d}\left(s-t_0\right)^2\\
&\quad -60\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)}-30\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}Y_s^{(1)}\\
&\quad -30\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(2)}-60\sigma^2 m_1 t^2\int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)}\\
&\quad +30\sigma^2 m_1\int_{t_0}^{t+t_0}\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}\left(s-t_0\right)^2-10\sigma^2\int_{t_0}^{t+t_0}(s-t_0)\,\mathrm{d}Y_s^{(3)}\\
&\quad -30\sigma^2 m_1 t^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(2)}+15\sigma^2 m_1\int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}\left(s-t_0\right)^2-30\sigma^2 m_1^2 t^3\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}\\
&\quad +10\sigma^2 m_1^2\int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}\left(s-t_0\right)^3-30\sigma^2 m_2 t^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}\\
&\quad +15\sigma^2 m_2\int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}\left(s-t_0\right)^2-\frac{15}{2}\sigma^2 m_1^3 t^4-20\sigma^2 m_1 m_2 t^3-5\sigma^2 m_3 t^2.
\end{aligned}$$

By integration by parts again,

$$
\begin{aligned}
N_1 \quad = \quad & -30\sigma^4 t^2 \int_{t_0}^{t+t_0} \mathrm{d}\left[Y_s^{(1)} + m_1 s\right] + 15\sigma^4 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}\left[Y_{t_1}^{(1)} + m_1 t_1\right] \mathrm{d}\left(s-t_0\right)^2 \\
& -60\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(2)} \mathrm{d}Y_s^{(1)} - 30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(2)} \\
& -60\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} + 30\sigma^2 m_1 \left[(s-t_0)^2 \int_{t_0}^{s} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} \\
& -30\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} - 10\sigma^2 \int_{t_0}^{t+t_0} (s-t_0)\, \mathrm{d}Y_s^{(3)} \\
& -30\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(2)} + 15\sigma^2 m_1 \left[(s-t_0)^2 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(2)}\right]_{t_0}^{t+t_0} \\
& -15\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2\, \mathrm{d}Y_s^{(2)} - 30\sigma^2 m_1^2 t^3 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} \\
& +10\sigma^2 m_1^2 \left[(s-t_0)^3 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} - 10\sigma^2 m_1^2 \int_{t_0}^{t+t_0} (s-t_0)^3\, \mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 m_2 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} + 15\sigma^2 m_2 \left[(s-t_0)^2 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} \\
& -15\sigma^2 m_2 \int_{t_0}^{t+t_0} (s-t_0)^2\, \mathrm{d}Y_s^{(1)} - \frac{15}{2}\sigma^2 m_1^3 t^4 - 20\sigma^2 m_1 m_2 t^3 - 5\sigma^2 m_3 t^2 \\
= \quad & -30\sigma^4 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} - 30\sigma^4 m_1 t^3 + 15\sigma^4 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}\left(s-t_0\right)^2 \\
& +15\sigma^4 m_1 \int_{t_0}^{t+t_0} (s-t_0)\, \mathrm{d}\left(s-t_0\right)^2 - 60\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(2)} \mathrm{d}Y_s^{(1)} - 30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(2)} \\
& -60\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} + 30\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} - 10\sigma^2 \int_{t_0}^{t+t_0} (s-t_0)\, \mathrm{d}Y_s^{(3)} \\
& -30\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(2)} + 15\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(2)} - 15\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2\, \mathrm{d}Y_s^{(2)} \\
& -30\sigma^2 m_1^2 t^3 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} + 10\sigma^2 m_1^2 t^3 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} - 10\sigma^2 m_1^2 \int_{t_0}^{t+t_0} (s-t_0)^3\, \mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 m_2 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} + 15\sigma^2 m_2 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} \\
& -15\sigma^2 m_2 \int_{t_0}^{t+t_0} (s-t_0)^2\, \mathrm{d}Y_s^{(1)} - \frac{15}{2}\sigma^2 m_1^3 t^4 - 20\sigma^2 m_1 m_2 t^3 - 5\sigma^2 m_3 t^2
\end{aligned}
$$

$$
\begin{aligned}
N_1 &= -30\sigma^4 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} - 30\sigma^4 m_1 t^3 + 15\sigma^4 \left[(s-t_0)^2 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)} \right]_{t_0}^{t+t_0} \\
&\quad -15\sigma^4 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(1)} + 30\sigma^4 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}(s-t_0) \\
&\quad -60\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} - 30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(2)} \mathrm{d}Y_s^{(1)} \\
&\quad -30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(2)} - 30\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} \\
&\quad -30\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} - 10\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \,\mathrm{d}Y_s^{(3)} \\
&\quad -15\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(2)} - 15\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(2)} - 20\sigma^2 m_1^2 t^3 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} \\
&\quad -10\sigma^2 m_1^2 \int_{t_0}^{t+t_0} (s-t_0)^3 \,\mathrm{d}Y_s^{(1)} - 15\sigma^2 m_2 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} \\
&\quad -15\sigma^2 m_2 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(1)} - \frac{15}{2}\sigma^2 m_1^3 t^4 - 20\sigma^2 m_1 m_2 t^3 - 5\sigma^2 m_3 t^2 \\
&= -30\sigma^4 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} - 30\sigma^4 m_1 t^3 + 15\sigma^4 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} - 15\sigma^4 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(1)} \\
&\quad +10\sigma^4 m_1 t^3 - 60\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} \\
&\quad -30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(2)} \mathrm{d}Y_s^{(1)} - 30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(2)} \\
&\quad -30\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} - 30\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} \\
&\quad -10\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \,\mathrm{d}Y_s^{(3)} - 15\sigma^2 m_1 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(2)} - 15\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(2)} \\
&\quad -20\sigma^2 m_1^2 t^3 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} - 10\sigma^2 m_1^2 \int_{t_0}^{t+t_0} (s-t_0)^3 \,\mathrm{d}Y_s^{(1)} - 15\sigma^2 m_2 t^2 \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} \\
&\quad -15\sigma^2 m_2 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(1)} - \frac{15}{2}\sigma^2 m_1^3 t^4 - 20\sigma^2 m_1 m_2 t^3 - 5\sigma^2 m_3 t^2.
\end{aligned}
$$

$$
\begin{aligned}
N_2 &= 5 \int_{t_0}^{t+t_0} A_2\left(X_{s-}, X_{t_0}; 4\right) \mathrm{d}Y_s^{(1)} \\
&= 5 \int_{t_0}^{t+t_0} \left[6\sigma^2 (s-t_0) A_1\left(X_{s-}, X_{t_0}; 2\right) + 3\sigma^4 (s-t_0)^2 \right] \mathrm{d}Y_s^{(1)} \\
&= 5 \int_{t_0}^{t+t_0} \left\{ 6\sigma^2 (s-t_0) \left[2 \int_{t_0}^{s-} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} + 2m_1 (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} + \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(2)} \right] \right\} \\
&\qquad \mathrm{d}Y_s^{(1)} + 5 \int_{t_0}^{t+t_0} \left\{ 6\sigma^2 (s-t_0) \left[m_1^2 (s-t_0)^2 + m_2 (s-t_0) \right] \right\} \mathrm{d}Y_s^{(1)} \\
&\quad +15\sigma^4 \int_{t_0}^{t+t_0} (s-t_0)^2 \,\mathrm{d}Y_s^{(1)}
\end{aligned}
$$

$$
\begin{aligned}
N_2 &= 60\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} + 60\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \\
&\quad \times \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(1)} + 30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(2)} \mathrm{d}Y_s^{(1)} \\
&\quad +30\sigma^2 m_1^2 \int_{t_0}^{t+t_0} (s-t_0)^3 \mathrm{d}Y_s^{(1)} + 30\sigma^2 m_2 \int_{t_0}^{t+t_0} (s-t_0)^2 \mathrm{d}Y_s^{(1)} \\
&\quad +15\sigma^4 \int_{t_0}^{t+t_0} (s-t_0)^2 \mathrm{d}Y_s^{(1)}. \\
N_3 &= 10 \int_{t_0}^{t+t_0} A_2(X_{s-}, X_{t_0}; 3) \mathrm{d}Y_s^{(2)} = 10 \int_{t_0}^{t+t_0} \left[3\sigma^2 A_1(X_{s-}, X_{t_0}; 1)(s-t_0)\right] \mathrm{d}Y_s^{(2)} \\
&= 30\sigma^2 \int_{t_0}^{t+t_0} \int_{t_0}^{s-} \mathrm{d}\left[Y_{t_1}^{(1)} + m_1 t_1\right] (s-t_0) \mathrm{d}Y_s^{(2)} \\
&= 30\sigma^2 \int_{t_0}^{t+t_0} (s-t_0) \int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)} \mathrm{d}Y_s^{(2)} + 30\sigma^2 m_1 \int_{t_0}^{t+t_0} (s-t_0)^2 \mathrm{d}Y_s^{(2)}. \\
N_4 &= 10 \int_{t_0}^{t+t_0} A_2(X_{s-}, X_{t_0}; 2) \mathrm{d}Y_s^{(3)} = 10 \int_{t_0}^{t+t_0} \sigma^2 (s-t_0) \mathrm{d}Y_s^{(3)}. \\
N_5 &= 5m_1 (t+t_0) A_2(X_{t+t_0}, X_{t_0}; 4) = 5m_1 (t+t_0) \left[6\sigma^2 t A_1(X_{t+t_0}, X_{t_0}; 2) + 3\sigma^4 t^2\right] \\
&= 30\sigma^2 m_1 t (t+t_0) \left[2 \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} + 2m_1 t \int_{t_0}^{t+t_0} \mathrm{d}Y_{t_1}^{(1)} + \int_{t_0}^{t+t_0} \mathrm{d}Y_{t_1}^{(2)}\right] \\
&\quad +30\sigma^2 m_1 t (t+t_0) \left[m_1^2 t^2 + m_2 t\right] + 15\sigma^4 m_1 t^2 (t+t_0) \\
&= 60\sigma^2 m_1 t (t+t_0) \int_{t_0}^{t+t_0} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)} + 60\sigma^2 m_1^2 t^2 (t+t_0) \int_{t_0}^{t+t_0} \mathrm{d}Y_{t_1}^{(1)} \\
&\quad +30\sigma^2 m_1 m_2 t^2 (t+t_0) + 30\sigma^2 m_1 t (t+t_0) \int_{t_0}^{t+t_0} \mathrm{d}Y_{t_1}^{(2)} \\
&\quad +30\sigma^2 m_1^3 t^3 (t+t_0) + 15\sigma^4 m_1 t^2 (t+t_0). \\
N_6 &= 10m_2 (t+t_0) A_2(X_{t+t_0}, X_{t_0}; 3) = 10m_2 (t+t_0) \left[3\sigma^2 A_1(X_{t+t_0}, X_{t_0}; 1) t\right] \\
&= 30\sigma^2 m_2 t (t+t_0) \int_{t_0}^{t+t_0} \mathrm{d}\left[Y_s^{(1)} + m_1 s\right] \\
&= 30\sigma^2 m_2 t (t+t_0) \int_{t_0}^{t+t_0} \mathrm{d}Y_s^{(1)} + 30\sigma^2 m_1 m_2 t^2 (t+t_0). \\
N_7 &= 10m_3 (t+t_0) A_2(X_{t+t_0}, X_{t_0}; 2) = 10\sigma^2 m_3 t (t+t_0). \\
N_8 &= -5m_1 \int_{t_0}^{t+t_0} s \mathrm{d}A_2(X_s, X_{t_0}; 4) \\
&= -5m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[6\sigma^2 (s-t_0) A_1(X_s, X_{t_0}; 2) + 3\sigma^4 (s-t_0)^2\right] \\
&= -30\sigma^2 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[(s-t_0) A_1(X_s, X_{t_0}; 2)\right] - 15\sigma^4 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[s^2 - 2t_0 s\right] \\
&= -30\sigma^2 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[2(s-t_0) \int_{t_0}^{s} \int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)} \mathrm{d}Y_{t_1}^{(1)}\right] \\
&\quad -30\sigma^2 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[2m_1 (s-t_0)^2 \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(1)}\right] \\
&\quad -30\sigma^2 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[(s-t_0) \int_{t_0}^{s} \mathrm{d}Y_{t_1}^{(2)}\right] - 30\sigma^2 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[m_1^2 (s-t_0)^3\right] \\
&\quad -30\sigma^2 m_1 \int_{t_0}^{t+t_0} s \mathrm{d}\left[m_2 (s-t_0)^2\right] - 30\sigma^4 m_1 \int_{t_0}^{t+t_0} s^2 \mathrm{d}s + 15\sigma^4 m_1 t_0 \left(t^2 + 2t_0 t\right)
\end{aligned}
$$

$$
\begin{aligned}
N_8 \;=\;& -60\sigma^2 m_1 \int_{t_0}^{t+t_0} s\mathrm{d}\left[s\int_{t_0}^{s}\int_{t_0}^{t_1-} \mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\right] + 60\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} \\
& -60\sigma^2 m_1^2 \int_{t_0}^{t+t_0} s\mathrm{d}\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right] + 120\sigma^2 m_1^2 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}\left[s\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right] \\
& -60\sigma^2 m_1^2 t_0^2 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} - 30\sigma^2 m_1 \int_{t_0}^{t+t_0} s\mathrm{d}\left[s\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(2)}\right] \\
& +30\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(2)} - 30\sigma^2 m_1^3 \int_{t_0}^{t+t_0} s\mathrm{d}s^3 + 90\sigma^2 m_1^3 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}s^2 \\
& -90\sigma^2 m_1^3 t_0^2 \int_{t_0}^{t+t_0} s\mathrm{d}s - 30\sigma^2 m_1 m_2 \int_{t_0}^{t+t_0} s\mathrm{d}s^2 + 60\sigma^2 m_1 m_2 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}s \\
& -10\sigma^4 m_1\left(t^3+3t_0t^2+3t_0^2t\right) + 15\sigma^4 m_1 t_0 t^2 + 30\sigma^4 m_1 t_0^2 t \\
=\;& -60\sigma^2 m_1 \left[s^2\int_{t_0}^{s}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} + 60\sigma^2 m_1 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s \\
& +60\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} - 60\sigma^2 m_1^2\left[s^3\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} \\
& +60\sigma^2 m_1^2 \int_{t_0}^{t+t_0} s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s + 120\sigma^2 m_1^2 t_0\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} \\
& -120\sigma^2 m_1^2 t_0 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s - 60\sigma^2 m_1^2 t_0^2 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 m_1\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(2)}\right]_{t_0}^{t+t_0} + 30\sigma^2 m_1 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}s \\
& +30\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(2)} - 90\sigma^2 m_1^3 \int_{t_0}^{t+t_0} s^3\mathrm{d}s + 180\sigma^2 m_1^3 t_0 \int_{t_0}^{t+t_0} s^2\mathrm{d}s \\
& -45\sigma^2 m_1^3 t_0^2\left(t^2+2t_0t\right) - 60\sigma^2 m_1 m_2 \int_{t_0}^{t+t_0} s^2\mathrm{d}s \\
& +30\sigma^2 m_1 m_2 t_0\left(t^2+2t_0t\right) - 10\sigma^4 m_1\left(t^3+3t_0t^2+3t_0^2t\right) + 15\sigma^4 m_1 t_0 t^2 + 30\sigma^4 m_1 t_0^2 t \\
=\;& -60\sigma^2 m_1 (t+t_0)^2 \int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} + 30\sigma^2 m_1 \int_{t_0}^{t+t_0}\int_{t_0}^{s-}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s^2 \\
& +60\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} - 60\sigma^2 m_1^2 (t+t_0)^3 \int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} \\
& +20\sigma^2 m_1^2 \int_{t_0}^{t+t_0}\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s^3 + 120\sigma^2 m_1^2 t_0 (t+t_0)^2 \int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} \\
& -60\sigma^2 m_1^2 t_0 \int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s^2 - 60\sigma^2 m_1^2 t_0^2 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 m_1 (t+t_0)^2 \int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(2)} + 15\sigma^2 m_1 \int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(2)}\mathrm{d}s^2 + 30\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(2)} \\
& -\frac{45}{2}\sigma^2 m_1^3\left(t^4+4t_0t^3+6t_0^2t^2+4t_0^3t\right) + 60\sigma^2 m_1^3 t_0\left(t^3+3t_0^2t+3t_0t^2\right) \\
& -45\sigma^2 m_1^3 t_0^2\left(t^2+2t_0t\right) - 20\sigma^2 m_1 m_2\left(t^3+3t_0t^2+3t_0^2t\right) + 30\sigma^2 m_1 m_2 t_0\left(t^2+2t_0t\right) \\
& -10\sigma^4 m_1\left(t^3+3t_0t^2+3t_0^2t\right) + 15\sigma^4 m_1 t_0 t^2 + 30\sigma^4 m_1 t_0^2 t
\end{aligned}
$$

$$
\begin{aligned}
N_8 \;=\;& -60\sigma^2 m_1 (t+t_0)^2 \int_{t_0}^{t+t_0}\int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} + 30\sigma^2 m_1 \left[s^2\int_{t_0}^{s}\int_{t_0}^{t_1-}\mathrm{d}Y_{t_2}^{(1)}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} \\
& -30\sigma^2 m_1 \int_{t_0}^{t+t_0} s^2 \int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} + 60\sigma^2 m_1 t_0 \int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} \\
& -60\sigma^2 m_1^2 (t+t_0)^3 \int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} + 20\sigma^2 m_1^2\left[s^3\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} - 20\sigma^2 m_1^2\int_{t_0}^{t+t_0} s^3\mathrm{d}Y_s^{(1)} \\
& +120\sigma^2 m_1^2 t_0 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} - 60\sigma^2 m_1^2 t_0\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0} \\
& +60\sigma^2 m_1^2 t_0\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(1)} - 60\sigma^2 m_1^2 t_0^2\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} - 30\sigma^2 m_1 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(2)} \\
& +15\sigma^2 m_1\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(2)}\right]_{t_0}^{t+t_0} - 15\sigma^2 m_1\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(2)} + 30\sigma^2 m_1 t_0\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(2)} \\
& -\frac{45}{2}\sigma^2 m_1^3 t^4 - 90\sigma^2 m_1^3 t_0 t^3 - 180\sigma^2 m_1^3 t_0^2 t^2 - 180\sigma^2 m_1^3 t_0^3 t - 20\sigma^2 m_1 m_2 t^3 \\
& +60\sigma^2 m_1^3 t_0\left(t^3 + 3t_0^2 t + 3t_0 t^2\right) - 30\sigma^2 m_1 m_2 t_0 t^2 - 10\sigma^4 m_1 t^3 - 15\sigma^4 m_1 t_0 t^2 \\
=\;& -60\sigma^2 m_1 (t+t_0)^2 \int_{t_0}^{t+t_0}\int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} + 30\sigma^2 m_1 (t+t_0)^2\int_{t_0}^{t+t_0}\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} \\
& -90\sigma^2 m_1^3 t_0 t^3 - 30\sigma^2 m_1\int_{t_0}^{t+t_0} s^2\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} + 60\sigma^2 m_1 t_0\int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} \\
& -180\sigma^2 m_1^3 t_0^2 t^2 - 60\sigma^2 m_1^2 (t+t_0)^3\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} + 20\sigma^2 m_1^2 (t+t_0)^3\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} \\
& -20\sigma^2 m_1^2\int_{t_0}^{t+t_0} s^3\mathrm{d}Y_s^{(1)} + 120\sigma^2 m_1^2 t_0 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} \\
& +60\sigma^2 m_1^2 t_0 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} + 60\sigma^2 m_1^2 t_0\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(1)} - 60\sigma^2 m_1^2 t_0^2\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} \\
& -30\sigma^2 m_1 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(2)} + 15\sigma^2 m_1 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(2)} - 15\sigma^2 m_1\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(2)} \\
& +30\sigma^2 m_1 t_0\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(2)} - \frac{45}{2}\sigma^2 m_1^3 t^4 - 180\sigma^2 m_1^3 t_0^3 t - 15\sigma^4 m_1 t_0 t^2 \\
& +60\sigma^2 m_1^3 t_0\left(t^3 + 3t_0^2 t + 3t_0 t^2\right) - 20\sigma^2 m_1 m_2 t^3 - 30\sigma^2 m_1 m_2 t_0 t^2 - 10\sigma^4 m_1 t^3 \\
=\;& -30\sigma^2 m_1 (t+t_0)^2 \int_{t_0}^{t+t_0}\int_{t_0}^{s-} \mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} - 30\sigma^2 m_1\int_{t_0}^{t+t_0} s^2\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} \\
& +60\sigma^2 m_1 t_0\int_{t_0}^{t+t_0} s\int_{t_0}^{s-}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}Y_s^{(1)} - 40\sigma^2 m_1^2 (t+t_0)^3\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} \\
& -20\sigma^2 m_1^2\int_{t_0}^{t+t_0} s^3\mathrm{d}Y_s^{(1)} + 60\sigma^2 m_1^2 t_0 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)} \\
& +60\sigma^2 m_1^2 t_0\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(1)} - 60\sigma^2 m_1^2 t_0^2\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)} \\
& -15\sigma^2 m_1 (t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(2)} - 15\sigma^2 m_1\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(2)} + 30\sigma^2 m_1 t_0\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(2)} \\
& -\frac{45}{2}\sigma^2 m_1^3 t^4 - 90\sigma^2 m_1^3 t_0 t^3 - 180\sigma^2 m_1^3 t_0^2 t^2 - 180\sigma^2 m_1^3 t_0^3 t - 20\sigma^2 m_1 m_2 t^3 \\
& +60\sigma^2 m_1^3 t_0\left(t^3 + 3t_0^2 t + 3t_0 t^2\right) - 30\sigma^2 m_1 m_2 t_0 t^2 - 10\sigma^4 m_1 t^3 - 15\sigma^4 m_1 t_0 t^2.
\end{aligned}
$$

$$\begin{aligned}
N_9 &= -10m_2\int_{t_0}^{t+t_0} s\mathrm{d}A_2\left(X_s,X_{t_0};3\right) = -10m_2\int_{t_0}^{t+t_0} s\mathrm{d}\left[3\sigma^2\left(X_s-X_{t_0}\right)(s-t_0)\right]\\
&= -30\sigma^2 m_2\int_{t_0}^{t+t_0} s\mathrm{d}\left[(s-t_0)\int_{t_0}^{s}\mathrm{d}\left[Y_{t_1}^{(1)}+m_1 t\right]\right]\\
&= -30\sigma^2 m_2\int_{t_0}^{t+t_0} s\mathrm{d}\left[(s-t_0)\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]-30\sigma^2 m_1 m_2\int_{t_0}^{t+t_0} s\mathrm{d}\left[(s-t_0)^2\right]\\
&= -30\sigma^2 m_2\left[s(s-t_0)\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}+30\sigma^2 m_2\int_{t_0}^{t+t_0}(s-t_0)\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s\\
&\quad -30\sigma^2 m_1 m_2\int_{t_0}^{t+t_0} s\mathrm{d}\left[s^2-2t_0 s\right]\\
&= -30\sigma^2 m_2 t(t+t_0)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+30\sigma^2 m_2\int_{t_0}^{t+t_0} s\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s\\
&\quad -30\sigma^2 m_2 t_0\int_{t_0}^{t+t_0}\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s-60\sigma^2 m_1 m_2\int_{t_0}^{t+t_0} s^2\mathrm{d}s+30\sigma^2 m_1 m_2 t_0\left(t^2+2t_0 t\right)\\
&= -30\sigma^2 m_2 t(t+t_0)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+15\sigma^2 m_2\int_{t_0}^{t+t_0}\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\mathrm{d}s^2\\
&\quad -30\sigma^2 m_2 t_0\left[s\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}+30\sigma^2 m_2 t_0\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)}\\
&\quad -20\sigma^2 m_1 m_2\left(t^3+3t_0 t^2+3t_0^2 t\right)+30\sigma^2 m_1 m_2 t_0\left(t^2+2t_0 t\right)\\
&= -30\sigma^2 m_2 t(t+t_0)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+15\sigma^2 m_2\left[s^2\int_{t_0}^{s}\mathrm{d}Y_{t_1}^{(1)}\right]_{t_0}^{t+t_0}\\
&\quad -15\sigma^2 m_2\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(1)}-30\sigma^2 m_2 t_0(t+t_0)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+30\sigma^2 m_2 t_0\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)}\\
&\quad -20\sigma^2 m_1 m_2 t^3-30\sigma^2 m_1 m_2 t_0 t^2\\
&= -30\sigma^2 m_2 t(t+t_0)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+15\sigma^2 m_2(t+t_0)^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}-15\sigma^2 m_2\int_{t_0}^{t+t_0} s^2\mathrm{d}Y_s^{(1)}\\
&\quad -30\sigma^2 m_2 t_0(t+t_0)\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}+30\sigma^2 m_2 t_0\int_{t_0}^{t+t_0} s\mathrm{d}Y_s^{(1)}\\
&\quad -20\sigma^2 m_1 m_2 t^3-30\sigma^2 m_1 m_2 t_0 t^2.\\
N_{10} &= -10m_3\int_{t_0}^{t+t_0} s\mathrm{d}A_2\left(X_s,X_{t_0};2\right) = -10m_3\int_{t_0}^{t+t_0} s\mathrm{d}\left[\sigma^2(s-t_0)\right] = -5\sigma^2 m_3\left(t^2+2t_0 t\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\left(X_{t+t_0}-X_{t_0}\right)^5 &= A_1\left(X_{t+t_0},X_{t_0};5\right)+10\sigma^2\left(X_{t+t_0}-X_{t_0}\right)^3 t-15\sigma^4 t^2\int_{t_0}^{t+t_0}\mathrm{d}Y_s^{(1)}-15\sigma^4 m_1 t^3\\
&= A_1\left(X_{t+t_0},X_{t_0};5\right)+10\sigma^2 A_1\left(X_{t+t_0},X_{t_0};3\right)t\\
&\quad +30\sigma^4 A_1\left(X_{t+t_0},X_{t_0};1\right)t^2-15\sigma^4 A_1\left(X_{t+t_0},X_{t_0};1\right)t^2\\
&= A_1\left(X_{t+t_0},X_{t_0};5\right)+10\sigma^2 A_1\left(X_{t+t_0},X_{t_0};3\right)t+15\sigma^4 A_1\left(X_{t+t_0},X_{t_0};1\right)t^2.
\end{aligned}$$

A.7 Simulation algorithms for stochastic integrals with respect to Lévy processes

In this appendix we summarize the numerical methods used in our simulation to evaluate multiple integrals. We first describe the stochastic Euler scheme for the stochastic differential equations (SDE) of general Lévy processes. By the results of Protter & Talay (1997) the approximation is satisfactory. For an up-to-date introduction to numerical solutions of SDE we refer the reader to Higham & Kloeden (2002), Higham (2001) and Kloeden (2002). For more theoretical details, we recommend Kloeden & Platen (1999).

A.7.1 Euler scheme for Wiener processes W_t

The simplest numerical method for SDEs is the *stochastic Euler*, or *Euler-Maruyama*, scheme, see Higham & Kloeden (2002), which forms a natural generalisation of the deterministic Euler scheme. For a scalar Itô's SDE

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + b(t, X_t)\mathrm{d}Z_t,$$

where Z_t is a general Lévy process, the scheme has the form

$$X_{n+1} = X_n + a\left(t_n, X_n\right)\Delta_n + b\left(t_n, X_n\right)\Delta Z_n, \tag{A.16}$$

where

$$\Delta_n = t_{n+1} - t_n = \int_{t_n}^{t_{n+1}} \mathrm{d}s, \quad \Delta Z_n = Z_{t_{n+1}} - Z_{t_n} = \int_{t_n}^{t_{n+1}} \mathrm{d}Z_s.$$

The scheme computes discrete approximations $X_n \approx X\left(t_n\right)$, at times $t_n = \sum_{r=0}^{n-1} \Delta_r$. In practice, it is common to use a single pre-chosen value for the step size Δ_r.

Convergence for numerical schemes may be defined in many ways. It is usual to distinguish between strong and weak convergences depending on whether the realisations or in general only their probability distributions are required to be close, respectively. Under suitable conditions on the SDE, for a fixed T, letting $t_{N_T} = T$ and $\Delta = \max_{0 \leq n \leq N_T - 1} \Delta_n$, a numerical scheme is said to converge with *strong order* γ if, for sufficiently small Δ,

$$E\left(\left|X_T - X_{N_T}\right|\right) \leq K_T \Delta^{\gamma},$$

for some constant K_T. Similarly, we have *weak order* β if, for each polynomial g,

$$\left|E\left(g\left(X_T\right)\right) - E\left(g\left(X_{N_T}\right)\right)\right| \leq K_{g,T} \Delta^{\beta}$$

for some constant $K_{g,T}$.

A.7.2 Double stochastic integrals

In general, a multiple stochastic integral can be approximated by applying a suitable numerical method. For example, to evaluate the double integral

$$I_n = I\left[t_n;t_{n+1}\right] = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \mathrm{d}Z_s^{(2)} \mathrm{d}Z_t^{(1)}$$

we consider the 2-dimensional Itô's SDE

$$\mathrm{d}X_t^{(1)} = X_t^{(2)} \mathrm{d}Z_t^{(1)}, \;\; \mathrm{d}X_t^{(2)} = \mathrm{d}Z_t^{(2)}, \tag{A.17}$$

with initial conditions $X_{t_n}^{(1)} = 0$, $X_{t_n}^{(2)} = Z_{t_n}^{(2)}$, for which the solution at time $t = t_{n+1}$ satisfies $X_{t_{n+1}}^{(2)} = \Delta Z_n^{(2)}$ and $X_{t_{n+1}}^{(1)} = I\left[t_n;t_{n+1}\right]$. We may apply the stochastic Euler scheme (A.16) to (A.17) over the discretization subinterval $[t_n, t_{n+1}]$ with a suitable step size $\delta = \left(t_{n+1} - t_n\right)/K$. If we let $t'_k = t_n + k\delta$ and $\delta Z_k^{(j)} = Z_{t'_{k+1}}^{(j)} - Z_{t'_k}^{(j)}$ then the Euler scheme gives $Y_0^{(1)} = 0, Y_0^{(2)} = Z_{t_n}^{(2)}$, and

$$Y_{k+1}^{(1)} = Y_k^{(1)} + Y_k^{(2)} \delta Z_k^{(1)}, \;\; Y_{k+1}^{(2)} = Y_k^{(2)} + \delta Z_k^{(2)}, \;\; \text{for } 0 \le k \le K-1.$$

The strong order of convergence $\gamma = \frac{1}{2}$ of the stochastic Euler scheme ensures that, see Higham & Kloeden (2002),

$$E\left(\left|Y_K^{(1)} - I\left[t_n;t_{n+1}\right]\right|\right) \le C\sqrt{\delta}.$$

that is, its strong order of convergence is $\frac{1}{2}$.

A.7.3 Multiple stochastic integrals

Similarly, to evaluate $I_n = I\left[t_n;t_{n+1}\right] = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t_1} \cdots \int_{t_n}^{t_m} \mathrm{d}Z_{t_{m+1}}^{(m+1)} \cdots \mathrm{d}Z_{t_2}^{(2)} \mathrm{d}Z_{t_1}^{(1)}$, the Euler scheme gives $Y_0^{(1)} = 0$, $Y_0^{(2)} = 0$, ..., $Y_0^{(m)} = 0$, $Y_0^{(m+1)} = Z_{t_n}^{(m+1)}$ and

$$\begin{aligned} Y_{k+1}^{(1)} &= Y_k^{(1)} + Y_k^{(2)} \delta Z_k^{(1)}, \qquad Y_{k+1}^{(2)} = Y_k^{(2)} + Y_k^{(3)} \delta Z_k^{(2)}, \\ &\vdots \\ Y_{k+1}^{(m)} &= Y_k^{(m)} + Y_k^{(m+1)} \delta Z_k^{(m)}, \;\; Y_{k+1}^{(m+1)} = Y_k^{(m+1)} + \delta Z_k^{(m+1)}, \end{aligned}$$

for $0 \le k \le K-1$ using the same notations in the last subsection.

A.7.4 Rate of convergence of the Euler scheme for Lévy processes

Protter & Talay (1997) derived the rate of convergence of the Euler scheme for general Lévy processes. Here we quote one of their results which shows that the error of our simulation due to the use of Euler scheme is bounded. Let X be the solution of

$$X_t = X_0 + \int_0^t f\left(X_{s-}\right) \mathrm{d}Z_s \tag{A.18}$$

for a given and fixed Lévy process Z. This corresponds to the Itô's and not the Stratonovich integral. Note that in finance the usage of the Itô's integral makes physical sense in terms of non-anticipatory behaviour of the option. We denote by ΔZ_s the jump of $\{Z_t, 0 \leq t \leq T\}$ at time s: $\Delta Z_s = Z_s - Z_{s-}$. The Lévy decomposition of Z is given by

$$Z_t = \sigma W_t + \beta t + \int_{\|x\|<1} x\left(N_t\left(\omega, \mathrm{d}x\right) - t\nu\left(\mathrm{d}x\right)\right) + \sum_{0<s\leq t} \Delta Z_s \mathbf{1}_{[\|\Delta Z_s\| \geq 1]}.$$

In general, the law of the random variable X_T is unknown. The Euler scheme discretises (A.18) in time. Let $\frac{T}{n}$ be the discretisation step of the time interval $[0, T]$ and let $\left(\bar{X}_t^n\right)$ be the piecewise constant process defined by $\bar{X}_0^n = X_0$ and $\bar{X}_{(p+1)T/n}^n = \bar{X}_{pT/n}^n + f\left(\bar{X}_{pT/n}^n\right)\left(Z_{(p+1)T/n} - Z_{pT/n}\right)$. Actually this is just a direct application of (A.16) with the Wiener process W_t replaced by the Lévy process Z_t. Next we quoted the result for the rate of convergence by Protter & Talay (1997).

For $K > 0, m > 0$ and $p \in \mathbb{N} - \{0\}$, set

$$\begin{aligned} \rho_p\left(m\right) &= 1 + \|\beta\|^2 + \|\sigma\|^2 + \int_{-m}^{m} \|z\|^2 \nu(\mathrm{d}z) + \|\beta\|^p + \|\sigma\|^p \\ &\quad + \left(\int_{-m}^{m} \|z\|^2 \nu\left(\mathrm{d}z\right)\right)^{p/2} + \int_{-m}^{m} \|z\|^p \nu\left(\mathrm{d}z\right) \end{aligned}$$

where ν is the Lévy measure and $\eta_{K,p}\left(m\right) := \exp\left(K\rho_p\left(m\right)\right)$. Define $h\left(m\right) := \nu\left(\{x; \|x\| \geq m\}\right)$ for $m > 0$. With conditions on $f\left(\cdot\right), g\left(\cdot\right)$ and X_0, Protter & Talay (1997) proved

$$\left|Eg\left(X_T\right) - Eg\left(\bar{X}_T^n\right)\right| \leq 4\left\|g\right\|_{L^\infty(\mathbb{R}^d)}\left(1 - \exp\left(-h\left(m\right)T\right)\right) + \frac{\eta_{K(T),8}\left(m\right)}{n}.$$

Thus the convergence rate is governed by the rate of increase to infinity of the functions $h\left(\cdot\right)$ and $\eta_{K(T),8}\left(\cdot\right)$. In our case, we set $g\left(X_t\right) = X_t$. Hence we can see that the order of (weak) convergence of our simulation is bounded.

A.8 Plots of CRP

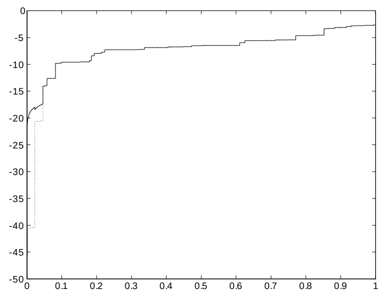

Figure A.8.1: G_t^4 generated using CRP and directly from the Gamma process in log scale.

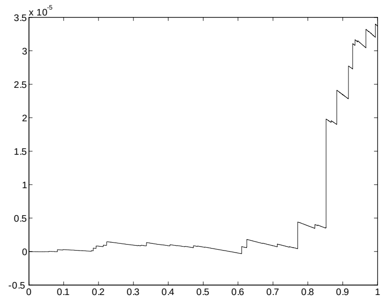

Figure A.8.2: The difference of the two series in Figure A.8.1.

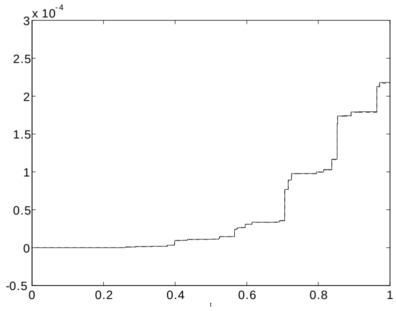

Figure A.8.3: $(G_{t+t_0} - G_{t_0})^9$ generated using CRP and also calculated directly from the Gamma process.

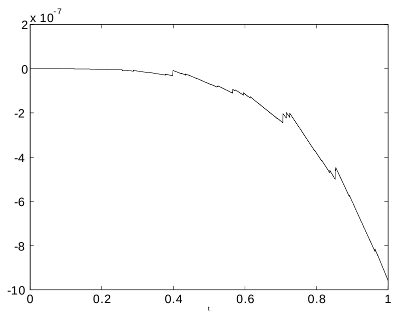

Figure A.8.4: The difference of the two series in Figure A.8.3.

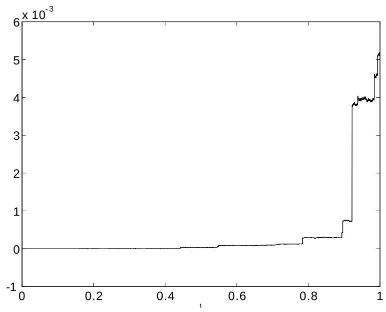

Figure A.8.5: X_t^5 generated using CRP and also calculated directly from the Wiener and Gamma processes.

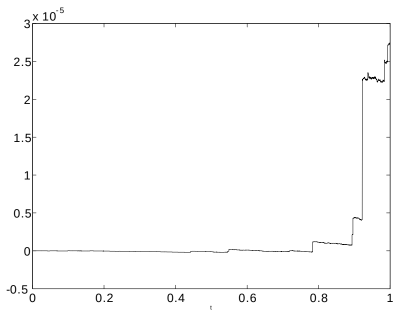

Figure A.8.6: The difference of the two series in Figure A.8.5.

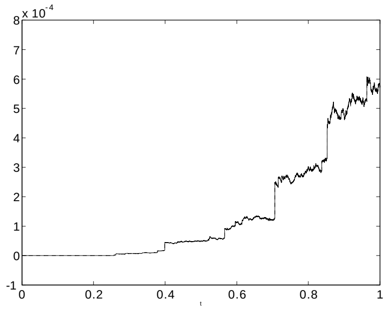

Figure A.8.7: $(X_{t+t_0} - X_{t_0})^8$ generated using CRP and also calculated directly from the Wiener and Gamma processes.

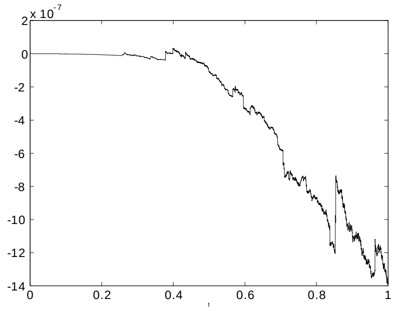

Figure A.8.8: The difference of the two series in Figure A.8.7.

Figures A.8.1-A.8.8: Solid line is generated using the CRP and the dotted line is generated by the Wiener and Gamma processes. Time step $= \frac{1}{10000}$, $a = 10$, $b = 20$. In Figure A.8.3, $t_0 = 0.0099$; in Figure A.8.5, $\sigma = 0.01$; in Figure A.8.7, $t_0 = 0.0019$ and $\sigma = 0.02$.

A.9 Proof of Lemma 5.4.1

We need the following propositions and lemma before we can prove Lemma 5.4.1. Recall $H_t^{(i)} = \sum_{j=1}^{i} a_{i,j} Y_t^{(j)}$, $i \geq 1$ and define $q_i = \sum_{j,j'=1,...,i} a_{i,j} a_{i,j'} m_{j+j'} + a_{i,1}^2 \sigma^2$ as well as

$$J_n^{(i_1,...,i_n)}(f,t) = \int_{t_0}^{t} \left(\int_{t_0}^{t_n -} \cdots \left(\int_{t_0}^{t_2 -} f(t_1,...,t_n) \, \mathrm{d}H_{t_1}^{(i_1)} \right) \cdots \mathrm{d}H_{t_{n-1}}^{(i_{n-1})} \right) \mathrm{d}H_{t_n}^{(i_n)}.$$

Léon *et al.* (2002, Proposition 1.1) proved that

$$E\left[J_n^{(i_1,...,i_n)}(f,t) \, J_m^{(j_1,...,j_m)}(g,t) \right] = \begin{cases} q_{i_1} \cdots q_{i_n} \int_{t_0}^{t} \cdots \int_{t_0}^{t_2 -} f(t_1,...,t_n) \, g(t_1,...,t_n) \, \mathrm{d}t_1 \cdots \mathrm{d}t_n, \\ \qquad \text{if } n = m \text{ and } (i_1,...,i_n) = (j_1,...,j_n), \\ 0, \qquad \text{otherwise.} \end{cases} \tag{A.19}$$

Define $\tilde{m}_2 = m_2 + \sigma^2$ and $\tilde{m}_n = m_n$ for $n = 1$ and $n = 3, 4, ...$

Proposition A.9.1 *For* $i, j = 1, 2, ...$, $E\left[Y_t^{(i)} Y_t^{(j)} \right] = \left(m_{i+j} + 1_{\{i=j=1\}} \sigma^2 \right) t = \tilde{m}_{i+j} t.$

Proof. Léon *et al.* (2002, P. 3) noted that the predictable quadratic variation process of $\left\{ H_t^{(i)}, t \geq 0 \right\}$ is given by

$$\left\langle H^{(i)}, H^{(i)} \right\rangle_t = q_i t$$

and the predictable quadratic covariation process of $Y^{(i)}$ and $Y^{(j)}$ $(i, j \geq 2)$ is given by

$$\left\langle Y^{(i)}, Y^{(j)} \right\rangle_t = m_{i+j} t.$$

Since $Y_t^{(1)} = H_t^{(1)}$, we have $\left\langle Y^{(1)}, Y^{(1)} \right\rangle_t = m_2 + \sigma^2 t$ and thus

$$\left\langle Y^{(i)}, Y^{(j)} \right\rangle_t = \left(m_{i+j} + 1_{\{i=j=1\}} \sigma^2 \right) t.$$

Hence,

$$E\left[Y_t^{(i)} Y_t^{(j)} \right] = E\left[\left\langle Y^{(i)}, Y^{(j)} \right\rangle_t \right] = \left(m_{i+j} + 1_{\{i=j=1\}} \sigma^2 \right) t = \tilde{m}_{i+j} t.$$

□

Proposition A.9.2 *In the Gamma(a, b) case, $\tilde{m}_n = m_n$ for all n and $m_n = \frac{(n-1)!a}{b^n}$.*

Proof. In the Gamma case, $\sigma^2 = 0$. We prove $m_n = \frac{(n-1)!a}{b^n}$ by induction. Assume true for n.

$$\begin{aligned} m_{n+1} &= \int_0^\infty x^n a e^{-bx} \mathrm{d}x = -\frac{a}{b} \int_0^\infty x^n \mathrm{d}e^{-bx} = -\frac{a}{b} \left[x^n e^{-bx} \right]_0^\infty + \frac{a}{b} \int_0^\infty e^{-bx} n x^{n-1} \mathrm{d}x \\ &= \frac{n}{b} \int_0^\infty x^{n-1} a e^{-bx} \mathrm{d}x = \frac{n}{b} \frac{(n-1)!a}{b^n} = \frac{n!a}{b^{n+1}}. \end{aligned}$$

□

Lemma A.9.3 *For any integer $k \geq 1$, $n = 1, 2, ..., k$ and $r = 1, 2, ..., n$,*

$$\frac{\mathrm{d}^r}{\mathrm{d}x^r} x^n (x-1)^k = \sum_{i=0}^{r} \binom{r}{i} \frac{n!}{(n-r+i)!} x^{n-r+i} \frac{k!}{(k-i)!} (x-1)^{k-i}.$$

Proof. We prove by induction on r. When $r = 1$,

$$L.H.S. = n x^{n-1} (x-1)^k + k x^n (x-1)^{k-1} = R.H.S.$$

Assume it is true for $r = p$, where p is an integer ≥ 1. Then, for $r = p + 1$,

$$\begin{aligned} \frac{\mathrm{d}^{p+1}}{\mathrm{d}x^{p+1}} x^n (x-1)^k &= \frac{\mathrm{d}}{\mathrm{d}x} \left[\sum_{i=0}^{p} \binom{p}{i} \frac{n!}{(n-p+i)!} x^{n-p+i} \frac{k!}{(k-i)!} (x-1)^{k-i} \right] \\ &= \sum_{i=0}^{p} \binom{p}{i} \frac{n!}{(n-p+i)!} \frac{k!}{(k-i)!} \left[(n-p+i) x^{n-p+i-1} (x-1)^{k-i} \right. \\ &\quad \left. + (k-i) x^{n-p+i} (x-1)^{k-i-1} \right] \\ &= \sum_{i=0}^{p} \binom{p}{i} \left[\frac{n!}{(n-p+i-1)!} x^{n-p+i-1} \frac{k!}{(k-i)!} (x-1)^{k-i} \right. \\ &\quad \left. + \frac{n!}{(n-p+i)!} x^{n-p+i} \frac{k!}{(k-i-1)!} (x-1)^{k-i-1} \right] \end{aligned}$$

$$= \sum_{i=0}^{p} \binom{p}{i} \frac{n!}{(n-p+i-1)!} x^{n-p+i-1} \frac{k!}{(k-i)!} (x-1)^{k-i}$$
$$+ \sum_{i=1}^{p+1} \binom{p}{i-1} \frac{n!}{(n-p+i-1)!} x^{n-p+i-1} \frac{k!}{(k-i)!} (x-1)^{k-i}$$
$$= \sum_{i=1}^{p} \left[\binom{p}{i} + \binom{p}{i-1} \right] \frac{n!}{(n-p+i-1)!} x^{n-p+i-1} \frac{k!}{(k-i)!} (x-1)^{k-i}$$
$$+ \frac{n!}{(n-p-1)!} x^{n-p-1} (x-1)^{k} + x^{n} \frac{k!}{(k-p-1)!} (x-1)^{k-p-1}.$$

Therefore, we have

$$\frac{\mathrm{d}^{p+1}}{\mathrm{d}x^{p+1}} x^n (x-1)^k = \sum_{i=1}^{p} \left[\frac{1}{i} + \frac{1}{p-i+1} \right] \frac{p!}{(i-1)!\,(p-i)!} \frac{n! x^{n-p+i-1}}{(n-p+i-1)!} \frac{k!\,(x-1)^{k-i}}{(k-i)!}$$
$$+ \frac{n!}{(n-p-1)!} x^{n-p-1} (x-1)^{k} + x^{n} \frac{k!}{(k-p-1)!} (x-1)^{k-p-1}$$
$$= \sum_{i=0}^{p+1} \binom{p+1}{i} \frac{n!}{(n-p+i-1)!} x^{n-p+i-1} \frac{k!}{(k-i)!} (x-1)^{k-i}.$$

□

Proof of Lemma 5.4.1. We prove the result by strong induction on i. Using the Laguerre polynomials, we obtain

$$H_t^{(2)} = \widehat{G}_t^{(2)} - \frac{2}{b} \widehat{G}_t^{(1)}$$

and since

$$m_2 = \frac{a}{b^2} \text{ and } m_3 = \frac{2a}{b^3},$$

it is clear that the proposition is true for $i = 2$. Assume it is true for all integers $k \geq 2$. Then for $i = k + 1$, by the orthogonality of H's, we have

$$E\left[H_t^{(n)} H_t^{(k+1)} \right] = 0 \quad \text{for all } n = 1, 2, ...k$$

or we can write

$$E\left[\left\{ a_{n,1} Y_t^{(1)} + \cdots + a_{n,n-1} Y_t^{(n-1)} + Y_t^{(n)} \right\} \right.$$
$$\left. \times \left\{ a_{k+1,1}\, Y_t^{(1)} + a_{k+1,2} Y_t^{(2)} + \cdots + a_{k+1,k} Y_t^{(k)} + Y_t^{(k+1)} \right\} \right]$$
$$= 0$$

for all $n = 1, 2, ..., k$. By Proposition A.9.1 and since $\tilde{m}_n = m_n$ in the Gamma case, we have

$$a_{k+1,1} \left(a_{n,1} m_2 + \cdots + m_{n+1} \right) + a_{k+1,2} \left(a_{n,1} m_3 + \cdots + m_{n+2} \right)$$
$$+ \cdots + \left(a_{n,1} m_{k+2} + \cdots + m_{n+k+1} \right)$$
$$= 0$$

for all $n = 1, 2, ..., k$, that is

$$\begin{bmatrix} m_2 & \cdots & m_{k+1} & m_{k+2} \\ a_{2,1}m_2 + m_3 & \cdots & a_{2,1}m_{k+1} + m_{k+2} & a_{2,1}m_{k+2} + m_{k+3} \\ & \vdots & & \\ a_{n,1}m_2 + \cdots + m_{n+1} & \cdots & a_{n,1}m_{k+1} + \cdots + m_{n+k} & a_{n,1}m_{k+2} + \cdots + m_{n+k+1} \\ & \vdots & & \\ a_{k,1}m_2 + \cdots + m_{k+1} & \cdots & a_{k,1}m_{k+1} + \cdots + m_{2k} & a_{k,1}m_{k+2} + m_{2k+1} \end{bmatrix}$$

$$\times \begin{bmatrix} a_{k+1,1} \\ a_{k+1,2} \\ \vdots \\ a_{k+1,n} \\ \vdots \\ a_{k+1,k} \\ a_{k+1,k+1} \end{bmatrix} = 0,$$

and by matrix operations, we have

$$\begin{bmatrix} m_2 & m_3 & m_4 & \cdots & m_{k+1} & m_{k+2} \\ m_3 & m_4 & m_5 & \cdots & m_{k+2} & m_{k+3} \\ & & & \vdots & & \\ m_{n+1} & m_{n+2} & m_{n+3} & \cdots & m_{n+k} & m_{n+k+1} \\ & & & \vdots & & \\ m_{k+1} & m_{k+2} & m_{k+3} & \cdots & m_{2k} & m_{2k+1} \end{bmatrix} \begin{bmatrix} a_{k+1,1} \\ a_{k+1,2} \\ \vdots \\ a_{k+1,n} \\ \vdots \\ a_{k+1,k} \\ a_{k+1,k+1} \end{bmatrix} = 0. \tag{A.20}$$

It remains to prove that $a_{k+1,j} = (-1)^{k+1-j} \binom{k}{j-1} \frac{m_{k+2}}{m_{j+1}}$ for $j = 1, 2, ..., k$ solve (A.20). In other words, we need to prove for $n = 1, 2, ..., k$,

$$\sum_{j=1}^{k+1} m_{n+j} a_{k+1,j} = \sum_{j=1}^{k+1} m_{n+j} (-1)^{k+1-j} \binom{k}{j-1} \frac{m_{k+2}}{m_{j+1}} = \sum_{j=0}^{k} m_{n+j+1} (-1)^{k-j} \binom{k}{j} \frac{m_{k+2}}{m_{j+2}} = 0.$$

Note that this proof so far can be applied to all Lévy processes by replacing m with $\tilde{m}$, given that the general form of $\tilde{m}$ is known. Since we have only proven a simple formula of m for Gamma in Proposition A.9.2, we restrict our results to the Gamma case here. As

$$\sum_{j=0}^{k} m_{n+j+1} (-1)^{k-j} \binom{k}{j} \frac{m_{k+2}}{m_{j+2}} = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \frac{a}{b^{n+k+1}} \frac{(n+j)!\,(k+1)!}{(j+1)!},$$

we have to prove

$$\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\frac{a}{b^{n+k+1}}\frac{(n+j)!\,(k+1)!}{(j+1)!}=0. \tag{A.21}$$

Consider $x^n\,(x-1)^k=\sum_{j=0}^{k}\binom{k}{j}x^{n+j}\,(-1)^{k-j}$. Differentiating both sides with respect to x by $n-1$ times,

$$\begin{aligned}\text{L.H.S.} &= \sum_{j=0}^{n-1}\binom{n-1}{j}\frac{n!}{(j+1)!}x^{j+1}\frac{k!}{(k-j)!}\,(x-1)^{k-j},\\ \text{R.H.S.} &= \sum_{j=0}^{k}\binom{k}{j}\frac{(n+j)!}{(j+1)!}x^{j+1}\,(-1)^{k-j},\end{aligned}$$

where L.H.S. is obtained using Lemma A.9.3. Put $x=1$, $\sum_{j=0}^{k}\binom{k}{j}\frac{(n+j)!}{(j+1)!}\,(-1)^{k-j}=0$ and hence (A.21) is proven and this concludes the proof. □

A.10 Proof of Proposition 5.4.2

The Lévy measure of *Gamma* $(a,b)\,,\nu_G\,(\mathrm{d}x)$ is given by, see Schoutens (2003),

$$\nu_G\,(\mathrm{d}x)=\frac{ae^{-bx}}{x}1_{(x>0)}\mathrm{d}x. \tag{A.22}$$

Following Nualart & Schoutens (2000, p. 119), we use (5.10) by taking $f\,(x)=\mathrm{i}\theta x^j$ and put $z=x^j$ to obtain for $j\geq 2$,

$$\begin{aligned}E\left[\exp\left(\mathrm{i}\theta G_1^{(j)}\right)\right] &= \exp\left(\mathrm{i}\theta\int_{-1}^{1}\frac{a\exp\left(-bz^{\frac{1}{j}}\right)}{j}1_{(z>0)}\mathrm{d}z\right.\\ &\quad\left.+\int_{-\infty}^{\infty}\left(e^{\mathrm{i}\theta z}-1-\mathrm{i}\theta z1_{\{|z|<1\}}\right)\frac{a\exp\left(-bz^{\frac{1}{j}}\right)}{jz}1_{(z>0)}\mathrm{d}z\right).\end{aligned}$$

The Brownian part σ for $G^{(j)}$ is 0 because the original Gamma process has no Brownian part and we have

$$\int_{-1}^{1}\frac{a\exp\left(-bz^{\frac{1}{j}}\right)}{j}1_{(z>0)}\mathrm{d}z=\frac{a}{j}\int_{0}^{1}\exp\left(-bz^{\frac{1}{j}}\right)\mathrm{d}z=\frac{a}{b^j}\int_{0}^{b}\exp\left(-z\right)z^{j-1}\mathrm{d}z,$$

where the integral is the lower incomplete Gamma function $\gamma\,(j,b)$, see Weisstein (1999a), and can be calculated by integration by parts since j is an integer ≥ 1. So its value can be found in closed form when j is known. Hence the Lévy triplet of $G^{(j)}$ is given by

$$\left(\frac{a}{b^j}\int_0^b \exp\left(-z\right) z^{j-1}\mathrm{d}z, 0, \frac{a\exp\left(-bz^{\frac{1}{j}}\right)}{jz}1_{(z>0)}\mathrm{d}z\right) \tag{A.23}$$

using the Lévy-Khintchine formula given in equation (2.1). Note that (A.22) is a special case of (A.23) when $j = 1$.

A.11 Proof of Proposition 5.4.3

We have 2 cases:

Case I: k is even and so p_1 is a local minimum.

$$\begin{aligned}
& E\left\{\exp\left(i\theta\left[G_t^{(j)} + a_{j,j-1}G_t^{(j-1)} + a_{j,j-2}G_t^{(j-2)} + ... + a_{j,1}G_t^{(1)}\right]\right)\right\} \\
= \; & \exp\left\{t\int_{-\infty}^{\infty}\left(e^{i\theta z}-1\right)a\left[-\frac{e^{-bh_1^{(j)}(z)}}{h_1^{(j)}\left(z\right)}\mathrm{d}h_1^{(j)}\left(z\right)1_{\left(h^{(j)}(p_1)<z<h^{(j)}(0)\right)}\right.\right. \\
& +\frac{e^{-bh_2^{(j)}(z)}}{h_2^{(j)}\left(z\right)}\mathrm{d}h_2^{(j)}\left(z\right)1_{\left(h^{(j)}(p_1)<z<h^{(j)}(p_2)\right)} - \frac{e^{-bh_3^{(j)}(z)}}{h_3^{(j)}\left(z\right)}\mathrm{d}h_3^{(j)}\left(z\right)1_{\left(h^{(j)}(p_3)<z<h^{(j)}(p_2)\right)} + ... \\
& \left.\left. -\frac{e^{-bh_{k-1}^{(j)}(z)}}{h_{k-1}^{(j)}\left(z\right)}\mathrm{d}h_{k-1}^{(j)}\left(z\right)1_{\left(h^{(j)}(p_{k-1})<z<h^{(j)}(p_{k-2})\right)} + \frac{e^{-bh_k^{(j)}(z)}}{h_k^{(j)}\left(z\right)}\mathrm{d}h_k^{(j)}\left(z\right)1_{\left(h^{(j)}(p_{k-1})<z<\infty\right)}\right]\right\}\mathrm{d}z.
\end{aligned}$$

Case II: k is odd and so p_1 is a local maximum.

$$\begin{aligned}
& E\left\{\exp\left(i\theta\left[G_t^{(j)} + a_{j,j-1}G_t^{(j-1)} + a_{j,j-2}G_t^{(j-2)} + ... + a_{j,1}G_t^{(1)}\right]\right)\right\} \\
= \; & \exp\left\{t\int_{-\infty}^{\infty}\left(e^{i\theta z}-1\right)a\left[\frac{e^{-bh_1^{(j)}(z)}}{h_1^{(j)}\left(z\right)}\mathrm{d}h_1^{(j)}\left(z\right)1_{\left(h^{(j)}(0)<z<h^{(j)}(p_1)\right)}\right.\right. \\
& -\frac{e^{-bh_2^{(j)}(z)}}{h_2^{(j)}\left(z\right)}\mathrm{d}h_2^{(j)}\left(z\right)1_{\left(h^{(j)}(p_2)<z<h^{(j)}(p_1)\right)} + \frac{e^{-bh_3^{(j)}(z)}}{h_3^{(j)}\left(z\right)}\mathrm{d}h_3^{(j)}\left(z\right)1_{\left(h^{(j)}(p_2)<z<h^{(j)}(p_3)\right)} + ... \\
& \left.\left. -\frac{e^{-bh_{k-1}^{(j)}(z)}}{h_{k-1}^{(j)}\left(z\right)}\mathrm{d}h_{k-1}^{(j)}\left(z\right)1_{\left(h^{(j)}(p_{k-1})<z<h^{(j)}(p_{k-2})\right)} + \frac{e^{-bh_k^{(j)}(z)}}{h_k^{(j)}\left(z\right)}\mathrm{d}h_k^{(j)}\left(z\right)1_{\left(h^{(j)}(p_{k-1})<z<\infty\right)}\right]\right\}\mathrm{d}z.
\end{aligned}$$

This means that the Lévy measure of $H_t^{(j)}$ is given by:

Case I: k is even.

$$
\begin{aligned}
&\nu_H(\mathrm{d}z)\\
&= a\Bigg[-\frac{e^{-bh_1^{(j)}(z)}}{h_1^{(j)}(z)}\mathrm{d}h_1^{(j)}(z)\,1_{\left(h^{(j)}(p_1)<z<h^{(j)}(0)\right)} + \frac{e^{-bh_2^{(j)}(z)}}{h_2^{(j)}(z)}\mathrm{d}h_2^{(j)}(z)\,1_{\left(h^{(j)}(p_1)<z<h^{(j)}(p_2)\right)}\\
&-\frac{e^{-bh_3^{(j)}(z)}}{h_3^{(j)}(z)}\mathrm{d}h_3^{(j)}(z)\,1_{\left(h^{(j)}(p_3)<z<h^{(j)}(p_2)\right)} + ... - \frac{e^{-bh_{k-1}^{(j)}(z)}}{h_{k-1}^{(j)}(z)}\mathrm{d}h_{k-1}^{(j)}(z)\,1_{\left(h^{(j)}(p_{k-1})<z<h^{(j)}(p_{k-2})\right)}\\
&+\frac{e^{-bh_k^{(j)}(z)}}{h_k^{(j)}(z)}\mathrm{d}h_k^{(j)}(z)\,1_{\left(h^{(j)}(p_{k-1})<z<\infty\right)}\Bigg]\mathrm{d}z.
\end{aligned}
$$

Case II: k is odd.

$$
\begin{aligned}
&\nu_H(\mathrm{d}z)\\
&= a\Bigg[\frac{e^{-bh_1^{(j)}(z)}}{h_1^{(j)}(z)}\mathrm{d}h_1^{(j)}(z)\,1_{\left(h^{(j)}(0)<z<h^{(j)}(p_1)\right)} - \frac{e^{-bh_2^{(j)}(z)}}{h_2^{(j)}(z)}\mathrm{d}h_2^{(j)}(z)\,1_{\left(h^{(j)}(p_2)<z<h^{(j)}(p_1)\right)}\\
&+\frac{e^{-bh_3^{(j)}(z)}}{h_3^{(j)}(z)}\mathrm{d}h_3^{(j)}(z)\,1_{\left(h^{(j)}(p_2)<z<h^{(j)}(p_3)\right)} + ... - \frac{e^{-bh_{k-1}^{(j)}(z)}}{h_{k-1}^{(j)}(z)}\mathrm{d}h_{k-1}^{(j)}(z)\,1_{\left(h^{(j)}(p_{k-1})<z<h^{(j)}(p_{k-2})\right)}\\
&+\frac{e^{-bh_k^{(j)}(z)}}{h_k^{(j)}(z)}\mathrm{d}h_k^{(j)}(z)\,1_{\left(h^{(j)}(p_{k-1})<z<\infty\right)}\Bigg]\mathrm{d}z.
\end{aligned}
$$

Altogether, we have

$$
\nu_H^{(j)}(\mathrm{d}z) = a\left[\sum_{i=1}^{k} g\left(h^{(j)}(p_{i-1}), h^{(j)}(p_i); z\right)\frac{e^{-bh_i^{(j)}(z)}}{h_i^{(j)}(z)}\mathrm{d}h_i^{(j)}(z)\right]\mathrm{d}z.
$$

Hence, the Lévy triplet of $H_t^{(j)}$ is given by

$$
\left(-\left[m_j + a_{j,j-1}m_{j-1} + a_{j,j-2}m_{j-2} + ... + a_{j,1}m_1\right] + \int_{-1}^{1} z\nu_H(\mathrm{d}z), 0, \nu_H(\mathrm{d}z)\right).
$$

A.12 Validity of $\nu_H^{(2)}(\mathrm{d}z)$ as a Lévy measure

We want to prove that the Lévy measure $\nu_H^{(2)}(\mathrm{d}z)$ given in (5.14) is valid, that is,

$$
\int_{-\infty}^{+\infty}\left(1\wedge z^2\right)\nu_H^{(2)}(\mathrm{d}z) < \infty,
$$

which means $\int_{-\infty}^{-1}\nu_H^{(2)}(\mathrm{d}z) + \int_{-1}^{1} z^2\nu_H^{(2)}(\mathrm{d}z) + \int_1^{\infty}\nu_H^{(2)}(\mathrm{d}z) < \infty$. Since $-\frac{1}{b^2}$ can be bigger or smaller than -1, we have 2 cases:

Case I: $b < 1$, we have to prove

$$\int_{-\frac{1}{b^2}}^{-1} \nu_H^{(2)}(\mathrm{d}z) + \int_{-1}^{1} z^2 \nu_H^{(2)}(\mathrm{d}z) + \int_{1}^{\infty} \nu_H^{(2)}(\mathrm{d}z) = I_1 + I_2 + I_3 < \infty.$$

We need the following lemma:

Lemma A.12.1 *Let $-1 \le c < 0$ be a constant. Then $\int_c^1 z^2 \nu_H^{(2)}(\mathrm{d}z) < \infty$.*

Proof. By putting $u = \sqrt{1 + zb^2}$, the integral can be calculated easily. □

We have

$$\begin{aligned} I_1 &= \int_{-\frac{1}{b^2}}^{-1} \nu_H^{(2)}(\mathrm{d}z) \\ &= \int_{-\frac{1}{b^2}}^{-1} \frac{e^{-\left[1-\sqrt{1+zb^2}\right]}}{\left[1-\sqrt{1+zb^2}\right]} \frac{ab^2}{2\sqrt{1+zb^2}} \mathrm{d}z + \int_{-\frac{1}{b^2}}^{-1} \frac{e^{-\left[1+\sqrt{1+zb^2}\right]}}{\left[1+\sqrt{1+zb^2}\right]} \frac{ab^2}{2\sqrt{1+zb^2}} \mathrm{d}z. \end{aligned}$$

Put $y = \sqrt{1 + zb^2}$, so $z = \frac{y^2-1}{b^2}$ and $\mathrm{d}z = \frac{2y}{b^2}\mathrm{d}y$.

$$I_1 = \int_0^{\sqrt{1-b^2}} a \frac{e^{-1+y}}{1-y} \frac{b^2}{2} \frac{1}{y} \frac{2y}{b^2} \mathrm{d}y + \int_0^{\sqrt{1-b^2}} a \frac{e^{-1-y}}{1+y} \frac{b^2}{2} \frac{1}{y} \frac{2y}{b^2} \mathrm{d}y = I_{11} + I_{12}.$$

Put $x = 1 - y$, so $y = 1 - x$ and $\mathrm{d}y = -\mathrm{d}x$, $I_{11} = \int_{1-\sqrt{1-b^2}}^{1} a \frac{e^{-x}}{x} \mathrm{d}x$. Hence, we get $I_{11} = a\left[-E_1(1) + E_1\left(1 - \sqrt{1-b^2}\right)\right]$, where $E_1(q)$ is the exponential integral defined by, see Abramowitz & Stegun (1964),

$$E_n(q) = \int_1^{\infty} e^{-xq} x^{-n} \mathrm{d}x. \tag{A.24}$$

Since $E_1(q) < \infty$ for $q \neq 0$ and $0 < b < 1$, we have $I_{11} < \infty$. Similarly, by letting $x = 1 + y$, we have $I_{12} = \int_1^{1+\sqrt{1-b^2}} a \frac{e^{-x}}{x} \mathrm{d}x$ so that $I_{12} = a\left[-E_1\left(1 + \sqrt{1-b^2}\right) + E_1(1)\right]$. Hence, $I_{12} < \infty$ and $I_1 < \infty$. Secondly, for I_2 we make use of Lemma A.12.1 by putting $c = -1$. Hence we have proven that $I_2 < \infty$. Finally, we have to consider

$$I_3 = \int_1^{\infty} \nu_H^{(2)}(\mathrm{d}z) = \int_1^{\infty} \frac{ae^{-\left[1+\sqrt{1+zb^2}\right]}}{\left[1+\sqrt{1+zb^2}\right]} \frac{b^2}{2} \frac{1}{\sqrt{1+zb^2}} \mathrm{d}z.$$

Put $y = 1 + \sqrt{1 + zb^2}$, so $z = \frac{(y-1)^2-1}{b^2}$ and $\mathrm{d}z = \frac{2(y-1)}{b^2}\mathrm{d}y$.

$$I_3 = \int_{1+\sqrt{1+b^2}}^{\infty} \frac{ae^{-y}}{y} \frac{b^2}{2} \frac{1}{(y-1)} \frac{2(y-1)}{b^2} \mathrm{d}y = \int_{1+\sqrt{1+b^2}}^{\infty} \frac{ae^{-y}}{y} \mathrm{d}y.$$

Hence, we get $I_3 = aE_1\left(1 + \sqrt{1+b^2}\right)$ and $I_3 < \infty$. Combining with the above results, we have proven that in case I where $b < 1$, the Lévy measure of H is valid.

Case II: $b \ge 1$, we have to prove $\int_{-\frac{1}{b^2}}^{1} z^2 \nu_H^{(2)}(\mathrm{d}z) + \int_1^{\infty} \nu_H^{(2)}(\mathrm{d}z) < \infty$. We have proven in case I that $\int_1^{\infty} \nu_H^{(2)}(\mathrm{d}z) < \infty$. For $\int_{-\frac{1}{b^2}}^{1} z^2 \nu_H^{(2)}(\mathrm{d}z)$, we can make use of Lemma A.12.1 by putting

$c = -\frac{1}{b^2}$ and conclude that $\int_{-\frac{1}{b^2}}^{1} z^2 \nu_H^{(2)}(\mathrm{d}z) < \infty$. Hence in the case where $b \geq 1$ we have proven that the Lévy measure of H is valid. Combining the above two cases, we have shown that for all $b > 0$ the Lévy measure $\nu_H^{(2)}(\mathrm{d}z)$ is valid.

Appendix B

Part II

B.1 Exotic options

In this appendix, we list the types of options to which our hedging strategies given in Part II of this thesis can be applied. We use Monte-Carlo simulation and finite-difference methods to obtain the option prices and their derivatives.

European call options

The option price function before maturity of the European call option with strike K, maturity T is given at time t by:

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[(S_T - K)^+ | \mathcal{F}_t \right],$$

where Q is the risk-neutral measure and $\mathcal{F}_t$ is the filtration of S.

'Up-and-out' barrier call options

Let the maximum process of the stock price, $S = \{S_t, 0 \leq t \leq T\}$ be

$$M_t^S = \sup\{S_u; 0 \leq u \leq t\}, \ 0 \leq t \leq T.$$

The 'up-and-out' barrier call is worthless unless its maximum remains below some high barrier H, in which case it retains the structure of an European call with strike K. The price is given by

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[(S_T - K)^+ 1_{\{M_T^S < H\}} \right].$$

The option price is differentiable in the stock price in the range $S_t < H$ and $M_t^S < H$.

'Up-and-in' barrier call options

The 'up-and-in' barrier call is worthless unless its maximum crossed some high barrier H, in which case it retains the structure of an European call with strike K. The price is given by

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[(S_T - K)^+ 1_{\{M_T^S \geq H\}} \right].$$

'Down-and-out' barrier call options

Let the minimum process of the stock price, $S = \{S_t, 0 \leq t \leq T\}$ be

$$m_t^S = \inf\{S_u; 0 \leq u \leq t\}, \ 0 \leq t \leq T.$$

The 'down-and-out' barrier call is worthless unless its minimum retains above some low barrier H, in which case it retains the structure of an European call with strike K. The price is given by

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[(S_T - K)^+ 1_{\{m_T^S > H\}} \right].$$

The option price is differentiable in the stock price in the range $S_t > H$ and $m_t^S > H$.

'Down-and-in' barrier call options

The 'down-and-in' barrier call is worthless unless its minimum crossed some low barrier H, in which case it retains the structure of an European call with strike K. The price is given by

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[(S_T - K)^+ 1_{\{m_T^S \leq H\}} \right].$$

The option price is differentiable in the stock price in the range $S_t > H$ and $m_t^S > H$; or $m_t^S < H$. In the latter case, the option becomes a standard European call option. In our simulation for the 'down-and-in' barrier call options, we consider only the first case.

Lookback options with a floating strike

The price of a floating strike lookback option is given by

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[M_T^S - S_T \right].$$

Lookback options with a fixed strike

The price of a fixed strike lookback option is given by

$$F(t, S_t) = \exp(-r(T-t)) E_Q \left[\left(M_T^S - K \right)^+ \right].$$

Asian call options

The option price function before maturity of an European-style arithmetic average call option with strike K, maturity T is given at time t by:

$$F(t, S_t) = \frac{\exp(-r(T-t))}{n} E_Q \left[\left(\sum_{k=1}^{n} S_{t_k} - nK \right)^+ \middle| \mathcal{F}_t \right].$$

B.2 Proof of Proposition 6.2.3

The initial investment at time t is

$$C_i \left\{ S_t^i \exp\left(-r\left(t+\Delta t\right)\right) T_t^{(i)} + \frac{S_t^i \exp\left(-r\left(t+\Delta t\right)\right) T_t^{(i)}}{\exp\left(r\Delta t\right)-1} + \frac{S_t^i \left[-\exp\left(-rt\right) T_t^{(i)} + m_i \Delta t\right]}{\exp\left(r\Delta t\right)-1} \right\}.$$

At maturity, the value of the portfolio is equal to

$$C_i S_t^i \left\{ \exp\left(-r\left(t+\Delta t\right)\right) T_{t+\Delta t}^{(i)} + \frac{\exp\left(r\Delta t\right)}{\exp\left(r\Delta t\right)-1} \left\{ \exp\left(-r\left(t+\Delta t\right)\right) T_t^{(i)} - \exp\left(-rt\right) T_t^{(i)} + m_i \Delta t \right\} \right\}.$$

Hence, the change of value of the portfolio is equal to

$$C_i \left\{ S_t^i \exp\left(-r\left(t+\Delta t\right)\right) T_{t+\Delta t}^{(i)} + S_t^i \left[-\exp\left(-rt\right) T_t^{(i)} + m_i \Delta t\right] \right\} = C_i \left(\Delta S_t\right)^i,$$

by equation (6.15).

B.3 Proof of Proposition 6.2.4

The initial investment at time t is

$$C_2 \left\{ S_t^2 \exp\left(2b\Delta t\right) \exp\left(-r\left(t+\Delta t\right)\right) T_t^{(2)} + \frac{S_t^2 \exp\left(2b\Delta t\right) \exp\left(-r\left(t+\Delta t\right)\right) T_t^{(2)}}{\exp\left(r\Delta t\right)-1} + \frac{1}{\exp\left(r\Delta t\right)-1} \left\{ -S_t^2 \left[\exp\left(b\Delta t\right)-1\right]^2 + 2S_t \left[\exp\left(b\Delta t\right)-1\right] \Delta S_t + S_t^2 \exp\left(2b\Delta t\right) \left[-\exp\left(-rt\right) T_t^{(2)} + m_2 \Delta t\right] \right\} \right\}.$$

At maturity, the portfolio is worth

$$C_2 \left\{ S_t^2 \exp\left(2b\Delta t\right) \exp\left(-r\left(t+\Delta t\right)\right) T_{t+\Delta t}^{(2)} + \frac{\exp\left(r\Delta t\right)}{\exp\left(r\Delta t\right)-1} \left\{ S_t^2 \exp\left(2b\Delta t\right) \exp\left(-r\left(t+\Delta t\right)\right) T_t^{(2)} - S_t^2 \left[\exp\left(b\Delta t\right)-1\right]^2 + 2S_t \left[\exp\left(b\Delta t\right)-1\right] \Delta S_t + S_t^2 \exp\left(2b\Delta t\right) \left[-\exp\left(-rt\right) T_t^{(2)} + m_2 \Delta t\right] \right\} \right\}.$$

Hence, the change of value of the portfolio is

$$\begin{aligned}
& C_2 \Big\{ S_t^2 \exp\left(2b\Delta t\right) \exp\left(-r\left(t+\Delta t\right)\right) T_{t+\Delta t}^{(2)} - S_t^2 \left[\exp\left(b\Delta t\right)-1\right]^2 \\
& \quad +2S_t \left[\exp\left(b\Delta t\right)-1\right] \Delta S_t + S_t^2 \exp\left(2b\Delta t\right) \Big[-\exp\left(-rt\right) T_t^{(2)} + m_2 \Delta t\Big] \Big\} \\
= \; & C_2 \left(\Delta S\right)^2
\end{aligned}$$

by equation (6.17).

B.4 Proof of Proposition 6.3.1

Let

$$\xi = \xi^0 + \sum_{j=1}^{k} \int_0^T \varphi_j\left(s\right) \mathrm{d}S_j\left(s\right), \tag{B.1}$$

where ξ^0 denotes the difference of value between ξ and $\sum_{j=1}^{k} \int_0^T \varphi_j\left(s\right) \mathrm{d}S_j\left(s\right)$ for the portfolio $\varphi = \left(\varphi_1, ..., \varphi_k\right)$. By the results of Monat & Stricker (1995, Section 4.2), the Hilbert space argument in Benth *et al.* (2003, Theorem 2.3) and equation (6.25), the following orthogonality condition is satisfied:

$$E\left[\left(\xi - \widehat{\xi}\right)\Theta\right] = E\left[\left\{\xi^0 - E\left[\xi\right]\right\}\Theta\right] = E\left[\xi^0\Theta\right] - E\left[\xi\right]E\left[\Theta\right] = 0,$$

where

$$\Theta = \sum_{j=1}^{k} \int_0^T \theta_j\left(s\right) \sigma_j S_j\left(s_-\right) \mathrm{d}W_j\left(s\right) + \sum_{j=1}^{k} \int_0^T \int_{\mathbb{R}} x\theta_j\left(s\right) S_j\left(s_-\right) \tilde{N}_j\left(\mathrm{d}s, \mathrm{d}x\right) \tag{B.2}$$

for all $\theta = \left(\theta_1, ..., \theta_k\right) \in \mathcal{A}$, the set of all admissible portfolios. Since $E\left[\Theta\right] = 0$, we have

$$E\left[\xi^0\Theta\right] = 0.$$

From (6.25) and (6.26),

$$\begin{aligned}
\sum_{j=1}^{k} \int_0^T \varphi_j\left(s\right) \mathrm{d}S_j\left(s\right) \;=\; & \sum_{j=1}^{k} \int_0^T \varphi_j\left(s\right) S_j\left(s_-\right) b_j \mathrm{d}s + \sum_{j=1}^{k} \int_0^T \varphi_j\left(s\right) \sigma_j S_j\left(s_-\right) \mathrm{d}W_j\left(s\right) \\
& + \sum_{j=1}^{k} \int_0^T \int_{\mathbb{R}} x\varphi_j\left(s\right) S_j\left(s_-\right) \tilde{N}_j\left(\mathrm{d}s, \mathrm{d}x\right).
\end{aligned}$$

Hence, from (6.29) and (B.1),

$$\xi^0 = \xi - \sum_{j=1}^{k} \int_0^T \varphi_j\left(s\right) \mathrm{d}S_j\left(s\right),$$

$$\begin{aligned}
\xi^0 &= \sum_{j=1}^{k}\int_0^T f_1(\xi;s,j)\,\mathrm{d}W_j(s) + \sum_{j=1}^{k}\int_0^T\int_{\mathbb{R}} f_2(\xi;s,x,j)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x) \\
&\quad -\sum_{j=1}^{k}\int_0^T \varphi_j(s)\,S_j(s_-)\,b_j\mathrm{d}s - \sum_{j=1}^{k}\int_0^T \varphi_j(s)\,\sigma_j S_j(s_-)\,\mathrm{d}W_j(s) \\
&\quad -\sum_{j=1}^{k}\int_0^T\int_{\mathbb{R}} x\varphi_j(s)\,S_j(s_-)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x) + E[\xi] \\
&= -\sum_{j=1}^{k}\int_0^T \varphi_j(s)\,S_j(s_-)\,b_j\mathrm{d}s + \sum_{j=1}^{k}\int_0^T \left(\frac{1}{\sigma_j}f_1(\xi;s,j) - \varphi_j(s)\,S_j(s_-)\right)\sigma_j\mathrm{d}W_j(s) \\
&\quad +\sum_{j=1}^{k}\int_0^T\int_{\mathbb{R}} \big(f_2(\xi;s,x,j) - x\varphi_j(s)\,S_j(s_-)\big)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x) + E[\xi].
\end{aligned}$$

Hence, from (B.2),

$$\begin{aligned}
&E\left[\xi^0\Theta\right] \\
&= E[\xi]\times\sum_{j=1}^{k} E\left[\int_0^T \theta_j(s)\,\sigma_j S_j(s_-)\,\mathrm{d}W_j(s) + \int_0^T\int_{\mathbb{R}} x\theta_j(s)\,S_j(s_-)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x)\right] \\
&\quad -\sum_{j=1}^{k}\int_0^T \varphi_j(s)\,S_j(s_-)\,b_j\mathrm{d}s \\
&\quad \times\sum_{j=1}^{k} E\left[\int_0^T \theta_j(s)\,\sigma_j S_j(s_-)\,\mathrm{d}W_j(s) + \int_0^T\int_{\mathbb{R}} x\theta_j(s)\,S_j(s_-)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x)\right] \\
&\quad +\sum_{i,j=1}^{k} E\left[\int_0^T \theta_i(s)\,\sigma_i S_i(s_-)\,\mathrm{d}W_i(s)\cdot\int_0^T \left(\frac{1}{\sigma_j}f_1(\xi;s,j) - \varphi_j(s)\,S_j(s_-)\right)\sigma_j\mathrm{d}W_j(s)\right] \\
&\quad +\sum_{i,j=1}^{k} E\left[\int_0^T \theta_i(s)\,\sigma_i S_i(s_-)\,\mathrm{d}W_i(s)\cdot\int_0^T\int_{\mathbb{R}} \big(f_2(\xi;s,x,j) - x\varphi_j(s)\,S_j(s_-)\big)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x)\right] \\
&\quad +\sum_{i,j=1}^{k} E\left[\int_0^T\int_{\mathbb{R}} x\theta_i(s)\,S_i(s_-)\,\tilde{N}_i(\mathrm{d}s,\mathrm{d}x)\right. \\
&\quad \left.\times\int_0^T \left(\frac{1}{\sigma_j}f_1(\xi;s,j) - \varphi_j(s)\,S_j(s_-)\right)\sigma_j\mathrm{d}W_j(s)\right] \\
&\quad +\sum_{i,j=1}^{k} E\left[\int_0^T\int_{\mathbb{R}} x\theta_i(s)\,S_i(s_-)\,\tilde{N}_i(\mathrm{d}s,\mathrm{d}x)\right. \\
&\quad \left.\times\int_0^T\int_{\mathbb{R}} \big(f_2(\xi;s,x,j) - x\varphi_j(s)\,S_j(s_-)\big)\,\tilde{N}_j(\mathrm{d}s,\mathrm{d}x)\right].
\end{aligned}$$

By the well-known isometry, see Ikeda & Watanabe (1989), we have

$$\begin{aligned} E\left[\xi^0\Theta\right] &= \sum_{j=1}^{k} E\left[\int_0^T \theta_j(s)\, S_j(s_-)\left(\frac{1}{\sigma_j} f_1(\xi; s, j) - \varphi_j(s)\, S_j(s_-)\right)\sigma_j^2 \mathrm{d}s\right] \\ &+ \sum_{j=1}^{k} E\left[\int_0^T \int_{\mathbb{R}} x\theta_j(s)\, S_j(s_-)\left(f_2(\xi; s, x, j) - x\varphi_j(s)\, S_j(s_-)\right)\nu_j(\mathrm{d}x)\, \mathrm{d}s\right], \end{aligned}$$

$$\begin{aligned} E\left[\xi^0\Theta\right] &= \sum_{j=1}^{k} E\left[\int_0^T \theta_j(s)\, S_j(s_-)\left\{\left(f_1(\xi; s, j) - \sigma_j\varphi_j(s)\, S_j(s_-)\right)\sigma_j\right.\right. \\ &\left.\left. + \int_{\mathbb{R}} x\left(f_2(\xi; s, x, j) - x\varphi_j(s)\, S_j(s_-)\right)\nu_j(\mathrm{d}x)\right\}\mathrm{d}s\right] = 0. \end{aligned}$$

Thus,

$$f_1(\xi; s, j)\,\sigma_j + \int_{\mathbb{R}} x f_2(\xi; s, x, j)\,\nu_j(\mathrm{d}x) = \varphi_j(s)\, S_j(s)\left\{\sigma_j^2 + \int_{\mathbb{R}} x^2 \nu_j(\mathrm{d}x)\right\}$$

$$\varphi_j(s) = \frac{f_1(\xi; s, j)\,\sigma_j + \int_{\mathbb{R}} x f_2(\xi; s, x, j)\,\nu_j(\mathrm{d}x)}{\left\{\sigma_j^2 + \int_{\mathbb{R}} x^2\nu_j(\mathrm{d}x)\right\} S_j(s)}.$$

B.5 Proof of Proposition 6.3.2

From equations (6.9) and (6.30), the term $\sum_{i=2}^{q} C_i S_t^i m_i \Delta t$ can be hedged by investing

$$\sum_{i=2}^{q} \frac{C_i}{(\exp(r\Delta t) - 1)} S_t^i m_i \Delta t$$

in a risk-free bank account. To hedge the term

$$\sum_{i=2}^{q} C_i S_t^i \int_t^{t+\Delta t} \mathrm{d}Y_s^{(i)},$$

we let

$$\xi = \sum_{i=2}^{q} \int_t^{t+\Delta t} C_i S_t^i \mathrm{d}Y_s^{(i)} = \sum_{i=2}^{q} \int_t^{t+\Delta t} \int_{\mathbb{R}} C_i S_t^i x^i \tilde{N}(\mathrm{d}s, \mathrm{d}x)$$

by (4.4) and let the minimal variance portfolio to hedge ξ be

$$\widehat{\xi} = E[\xi] + \int_t^{t+\Delta t} \varphi_s \mathrm{d}S_s = \int_t^{t+\Delta t} \varphi_s \mathrm{d}S_s$$

since $E[\xi] = 0$. Hence, using Proposition 6.3.1 and equation (6.29) by putting $f_1(\xi; s, j) = 0$ and

$$f_2(\xi; s, x, j) = \sum_{i=2}^{q} C_i S_t^i x^i,$$

we have

$$\varphi_s = \frac{\int_{\mathbb{R}} \sum_{i=2}^{q} C_i S_t^i x^{i+1} \nu(\mathrm{d}x)}{\left[\sigma^2 + \int_{\mathbb{R}} x^2 \nu(\mathrm{d}x)\right] S_s}.$$

Hence, to hedge $\sum_{i=2}^{q} C_i (\Delta S_t)^i$ by minimal variance portfolio, we need to invest

$$\sum_{i=2}^{q} \frac{C_i}{(\exp(r\Delta t) - 1)} S_t^i m_i \Delta t$$

in a risk-free bank account and buy

$$\frac{\int_{\mathbb{R}} \sum_{i=2}^{q} C_i S_t^i x^{i+1} \nu(\mathrm{d}x)}{\left[\sigma^2 + \int_{\mathbb{R}} x^2 \nu(\mathrm{d}x)\right] S_t} = \frac{\sum_{i=2}^{q} C_i S_t^{i-1} m_{i+1}}{[\sigma^2 + m_2]}$$

amount of the underlying stock, S_t, where m_i are defined in (1.8) and for VG process, it is given by Lemma 6.4.1.

B.6 Proof of Proposition 6.3.3

If Δt is negligible compared to ΔS_t, from (6.9) and (6.31), the term $\sum_{i=3}^{q} C_i S_t^i m_i \Delta t$ can be hedged by investing

$$\sum_{i=3}^{q} \frac{C_i}{(\exp(r\Delta t) - 1)} S_t^i m_i \Delta t$$

in a risk-free bank account. Let

$$\begin{aligned} \xi &= \sum_{i=3}^{q} \int_t^{t+\Delta t} C_i S_t^i \mathrm{d}Y_s^{(i)} \\ &= \sum_{i=3}^{q} \int_t^{t+\Delta t} \int_{\mathbb{R}} C_i S_t^i x^i \tilde{N}(\mathrm{d}s, \mathrm{d}x) \end{aligned}$$

by (4.4) and let the minimal variance portfolio to hedge ξ be

$$\begin{aligned} \widehat{\xi} &= \int_t^{t+\Delta t} \varphi_s^{(1)} \mathrm{d}S_s + \int_t^{t+\Delta t} \varphi_s^{(2)} S_{s-} \mathrm{d}Y_s^{(2)} \\ &= \int_t^{t+\Delta t} \varphi_s^{(1)} \mathrm{d}S_s + \int_t^{t+\Delta t} \int_{\mathbb{R}} \varphi_s^{(2)} x^2 S_{s-} \tilde{N}(\mathrm{d}s, \mathrm{d}x). \end{aligned} \quad \text{(B.3)}$$

Similar to Proposition 6.3.1, we have the orthogonal condition

$$\begin{aligned} E\left[\Theta\left(\xi - \widehat{\xi}\right)\right] &= E\left[\Theta\left(\left(\int_t^{t+\Delta t} \int_{\mathbb{R}} \left[\sum_{i=3}^{q} C_i S_t^i x^i - \varphi_s^{(2)} x^2 S_{s-}\right] \tilde{N}(\mathrm{d}s, \mathrm{d}x)\right) \right.\right. \\ &\quad \left.\left. - \int_t^{t+\Delta t} \varphi_s^{(1)} \mathrm{d}S_s\right)\right] \\ &= 0, \end{aligned}$$

where

$$\Theta = \int_t^{t+\Delta t} \theta_s^{(1)} \sigma S_{s-} \mathrm{d}W_s + \int_t^{t+\Delta t} \int_{\mathbb{R}} \left[x\theta_s^{(1)} S_{s-} + \theta_s^{(2)} x^2 S_{s-} \right] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \tag{B.4}$$

for all $\theta \in \mathcal{A}$, the set of all admissible portfolios. From (6.25) and (6.26),

$$\int_t^{t+\Delta t} \varphi_s^{(1)} \mathrm{d}S_s = \int_t^{t+\Delta t} \varphi_s^{(1)} S_{s-} b \mathrm{d}s + \int_t^{t+\Delta t} \varphi_s^{(1)} \sigma S_{s-} \mathrm{d}W_s + \int_t^{t+\Delta t} \int_{\mathbb{R}} x\varphi_s^{(1)} S_{s-} \tilde{N}(\mathrm{d}s, \mathrm{d}x).$$

Hence

$$\begin{aligned}
& E\left[\Theta\left(\xi - \widehat{\xi}\right)\right] \\
= \; & E\Bigg[\Theta\Bigg(-\int_t^{t+\Delta t} \varphi_s^{(1)} S_{s-} b \mathrm{d}s - \int_t^{t+\Delta t} \varphi_s^{(1)} \sigma S_{s-} \mathrm{d}W_s \\
& + \int_t^{t+\Delta t} \int_{\mathbb{R}} \left[-x\varphi_s^{(1)} S_{s-} - \varphi_s^{(2)} x^2 S_{s-} + \sum_{i=3}^q C_i S_t^i x^i \right] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \Bigg)\Bigg] \\
= \; & E\left[-\int_t^{t+\Delta t} \theta_s^{(1)} \sigma S_{s-} \mathrm{d}W_s \cdot \int_t^{t+\Delta t} \varphi_s^{(1)} S_{s-} b \mathrm{d}s \right] \\
& + E\left[-\int_t^{t+\Delta t} \theta_s^{(1)} \sigma S_{s-} \mathrm{d}W_s \cdot \int_t^{t+\Delta t} \varphi_s^{(1)} \sigma S_{s-} \mathrm{d}W_s \right] \\
& + E\Bigg[\int_t^{t+\Delta t} \theta_s^{(1)} \sigma S_{s-} \mathrm{d}W_s \cdot \int_t^{t+\Delta t} \int_{\mathbb{R}} \Big[-x\varphi_s^{(1)} S_{s-} - \varphi_s^{(2)} x^2 S_{s-} \\
& + \sum_{i=3}^q C_i S_t^i x^i \Big] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \Bigg] \\
& + E\left[-\int_t^{t+\Delta t} \int_{\mathbb{R}} \left[x\theta_s^{(1)} S_{s-} + \theta_s^{(2)} x^2 S_{s-} \right] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \cdot \int_t^{t+\Delta t} \varphi_s^{(1)} S_{s-} b \mathrm{d}s \right] \\
& + E\left[-\int_t^{t+\Delta t} \int_{\mathbb{R}} \left[x\theta_s^{(1)} S_{s-} + \theta_s^{(2)} x^2 S_{s-} \right] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \cdot \int_t^{t+\Delta t} \varphi_s^{(1)} \sigma S_{s-} \mathrm{d}W_s \right] \\
& + E\Bigg[\int_t^{t+\Delta t} \int_{\mathbb{R}} \left[x\theta_s^{(1)} S_{s-} + \theta_s^{(2)} x^2 S_{s-} \right] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \\
& \times \int_t^{t+\Delta t} \int_{\mathbb{R}} \left[-x\varphi_s^{(1)} S_{s-} - \varphi_s^{(2)} x^2 S_{s-} + \sum_{i=3}^q C_i S_t^i x^i \right] \tilde{N}(\mathrm{d}s, \mathrm{d}x) \Bigg].
\end{aligned}$$

By the well-known isometry, see Ikeda & Watanabe (1989), we have

$$\begin{aligned}
E\left[\Theta\left(\xi - \widehat{\xi}\right)\right] = \; & -\int_t^{t+\Delta t} \theta_s^{(1)} \varphi_s^{(1)} \sigma^2 S_{s-}^2 \mathrm{d}s + \int_t^{t+\Delta t} \int_{\mathbb{R}} \left[x\theta_s^{(1)} S_{s-} + \theta_s^{(2)} x^2 S_{s-} \right] \\
& \times \left[-x\varphi_s^{(1)} S_{s-} - \varphi_s^{(2)} x^2 S_{s-} + \sum_{i=3}^q C_i S_t^i x^i \right] \nu(\mathrm{d}x) \mathrm{d}s,
\end{aligned}$$

where ν is the Lévy measure of the underlying Lévy process. Since $E\left[\Theta\left(\xi-\widehat{\xi}\right)\right]=0$ for all $\theta_s^{(1)}$ and $\theta_s^{(2)}$, we have

$$\begin{cases} \varphi_s^{(1)}\sigma^2 S_s + x\int_{\mathbb{R}}\left[-x\varphi_s^{(1)}S_s - \varphi_s^{(2)}x^2 S_s + \sum_{i=3}^q C_i S_t^i x^i\right]\nu\left(\mathrm{d}x\right)=0 \\ -\int_{\mathbb{R}} x\varphi_s^{(1)}S_s\nu\left(\mathrm{d}x\right) - \varphi_s^{(2)}\int_{\mathbb{R}} x^2 S_s\nu\left(\mathrm{d}x\right) + \sum_{i=3}^q C_i S_t^i\int_{\mathbb{R}} x^i\nu\left(\mathrm{d}x\right)=0 \end{cases}$$

$$\Rightarrow \begin{cases} \varphi_s^{(1)}=0 \\ \varphi_s^{(2)}=\sum_{i=3}^q C_i S_t^i\int_{\mathbb{R}} x^i\nu\left(\mathrm{d}x\right)/\left[\int_{\mathbb{R}} x^2 S_s\nu\left(\mathrm{d}x\right)\right] \end{cases}. \tag{B.5}$$

From (4.4) and (B.3), our hedging portfolio requests us to invest in $\int_t^{t+\Delta t}\varphi_s^{(2)}S_{s-}\mathrm{d}Y_s^{(2)}$. By (6.30) and (B.5),

$$\begin{aligned} \int_t^{t+\Delta t}\varphi_s^{(2)}S_{s-}\mathrm{d}Y_s^{(2)} &= \sum_{i=3}^q \frac{C_i S_t^i\int_{\mathbb{R}} x^i\nu\left(\mathrm{d}x\right)}{\int_{\mathbb{R}} x^2\nu\left(\mathrm{d}x\right)}\left[Y_{t+\Delta t}^{(2)}-Y_t^{(2)}\right] \\ &= \sum_{i=3}^q \frac{C_i S_t^{i-2}\int_{\mathbb{R}} x^i\nu\left(\mathrm{d}x\right)}{\int_{\mathbb{R}} x^2\nu\left(\mathrm{d}x\right)}\left\{\left(\Delta S_t\right)^2 - S_t^2 m_2\Delta t\right\}. \end{aligned} \tag{B.6}$$

We can hedge the terms $\sum_{i=3}^q C_i S_t^{i-2}\int_{\mathbb{R}} x^i\nu\left(\mathrm{d}x\right)\left(\Delta S_t\right)^2/\left[\int_{\mathbb{R}} x^2\nu\left(\mathrm{d}x\right)\right]$ using variance swaps. Let

$$\phi=\sum_{i=3}^q \frac{C_i S_t^{i-2}\int_{\mathbb{R}} x^i\nu\left(\mathrm{d}x\right)}{\int_{\mathbb{R}} x^2\nu\left(\mathrm{d}x\right)}=\sum_{i=3}^q \frac{C_i S_t^{i-2}m_i}{m_2}, \tag{B.7}$$

where m_i are defined in (1.8) and for VG process, it is given by Lemma 6.4.1. By Proposition 6.2.1, to hedge the term $\phi\left(\Delta S_t\right)^2$, we invest

$$\phi\left\{\frac{S_t^2\Delta s\left(n-2\right)}{\left[\exp\left(r\Delta t\right)-1\right]}\left[\sigma_{\text{strike}}^2-\overline{S}_{n,2}\right]+\frac{P_V\Delta s\left(n-2\right)S_t^2}{\left[\exp\left(r\Delta t\right)-1\right]}\right\}$$

in a risk-free bank account and buy $\phi\Delta s\left(n-2\right)S_t^2$ units of the variance swap with sampling points $\left\{..., s_{n-1}=t, s_n=t+\Delta t\right\}$, maturity $t+\Delta t$ and strike σ_{strike}^2, where P_V is the price of one unit of the variance swap. To hedge the term $-\phi S_t^2 m_2\Delta t$, by (6.9), we borrow $\phi S_t^2 m_2\Delta t/\left[\exp\left(r\Delta t\right)-1\right]$ from a risk-free bank account. Hence, altogether we should invest

$$\frac{1}{e^{r\Delta t}-1}\left\{\sum_{i=3}^q C_i S_t^i m_i\Delta t+\phi S_t^2\left\{\Delta s\left(n-2\right)\left[\sigma_{\text{strike}}^2-\overline{S}_{n,2}\right]+P_V\Delta s\left(n-2\right)-m_2\Delta t\right\}\right\}$$

in a risk-free bank account and buy

$$\phi\Delta s\left(n-2\right)S_t^2$$

units of variance swaps with sampling points $\left\{..., s_{n-1}=t, s_n=t+\Delta t\right\}$ and maturity $t+\Delta t$.

B.7 Proof of Lemma 6.4.1

Recall the Lévy measure of a VG process given in equation (6.33):

$$\nu(x)\,\mathrm{d}x = \begin{cases} \frac{\mu_n^2}{\nu_n}\frac{\exp\left(-\frac{\mu_n}{\nu_n}|x|\right)}{|x|}\mathrm{d}x & \text{for } x<0 \\ \frac{\mu_p^2}{\nu_p}\frac{\exp\left(-\frac{\mu_p}{\nu_p}x\right)}{x}\mathrm{d}x & \text{for } x>0. \end{cases}$$

Firstly, we should find out the general expression for

$$\int_{-\infty}^{0} x^n \frac{\exp(cx)}{-x}\mathrm{d}x \text{ and } \int_{0}^{\infty} x^n \frac{\exp(-cx)}{x}\mathrm{d}x \text{ for } n = 1, 2, 3, ...,$$

where c is a constant. We are going to calculate the values of of these expressions when $n = 2$ and $n = 3$ in order to observe a general pattern. We firstly evaluate the two integrals $\int_{-\infty}^{0} x^2 \frac{\exp(cx)}{-x}\mathrm{d}x$ and $\int_{0}^{\infty} x^2 \frac{\exp(-cx)}{x}\mathrm{d}x$.

$$\begin{aligned} \int_{-\infty}^{0} x^2 \frac{\exp(cx)}{-x}\mathrm{d}x &= \frac{1}{c}\int_{-\infty}^{0} -x\mathrm{d}\exp(cx) = \frac{1}{c}\left[-x\exp(cx)\right]_{-\infty}^{0} + \frac{1}{c}\int_{-\infty}^{0}\exp(cx)\,\mathrm{d}x \\ &= \frac{1}{c^2}\left[\exp(cx)\right]_{-\infty}^{0} = \frac{1}{c^2} \end{aligned}$$

$$\begin{aligned} \int_{0}^{\infty} x^2 \frac{\exp(-cx)}{x}\mathrm{d}x &= -\frac{1}{c}\int_{0}^{\infty} x\mathrm{d}\exp(-cx) = -\frac{1}{c}\left[x\exp(-cx)\right]_{0}^{\infty} + \frac{1}{c}\int_{0}^{\infty}\exp(-cx)\,\mathrm{d}x \\ &= -\frac{1}{c^2}\left[\exp(-cx)\right]_{0}^{\infty} = \frac{1}{c^2}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \int_{-\infty}^{\infty} x^2\nu(\mathrm{d}x) &= \frac{\mu_n^2}{\nu_n}\int_{-\infty}^{0} x^2\frac{\exp\left(\frac{\mu_n}{\nu_n}x\right)}{-x}\mathrm{d}x + \frac{\mu_p^2}{\nu_p}\int_{0}^{\infty} x^2\frac{\exp\left(-\frac{\mu_p}{\nu_p}x\right)}{x}\mathrm{d}x \\ &= \frac{\mu_n^2}{\nu_n}\frac{\nu_n^2}{\mu_n^2} + \frac{\mu_p^2}{\nu_p}\frac{\nu_p^2}{\mu_p^2} = \nu_n + \nu_p = \nu\left(\mu_n^2 + \mu_p^2\right), \end{aligned}$$

since

$$\nu_p = \mu_p^2\nu \text{ and } \nu_n = \mu_n^2\nu.$$

Similarly,

$$\begin{aligned} \int_{-\infty}^{0} x^3 \frac{\exp(cx)}{-x}\mathrm{d}x &= \frac{1}{c}\int_{-\infty}^{0} -x^2\mathrm{d}\exp(cx) = \frac{1}{c}\left[-x^2\exp(cx)\right]_{-\infty}^{0} + \frac{1}{c}\int_{-\infty}^{0}\exp(cx)\,\mathrm{d}x^2 \\ &= \frac{2}{c}\int_{-\infty}^{0} x\exp(cx)\,\mathrm{d}x = -\frac{2}{c^3}. \end{aligned}$$

$$\begin{aligned}\int_0^\infty x^3 \frac{\exp(-cx)}{x} \mathrm{d}x &= -\frac{1}{c}\int_0^\infty x^2 \mathrm{d}\exp(-cx) = -\frac{1}{c}\left[x^2\exp(-cx)\right]_0^\infty + \frac{1}{c}\int_0^\infty \exp(-cx)\,\mathrm{d}x^2 \\ &= \frac{2}{c}\int_0^\infty x\exp(-cx)\,\mathrm{d}x = \frac{2}{c^3}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\int_{-\infty}^\infty x^3\nu(\mathrm{d}x) &= \frac{\mu_n^2}{\nu_n}\int_{-\infty}^0 x^3\frac{\exp\left(\frac{\mu_n}{\nu_n}x\right)}{-x}\mathrm{d}x + \frac{\mu_p^2}{\nu_p}\int_0^\infty x^3\frac{\exp\left(-\frac{\mu_p}{\nu_p}x\right)}{x}\mathrm{d}x \\ &= -\frac{2\mu_n^2}{\nu_n}\frac{\nu_n^3}{\mu_n^3} + \frac{2\mu_p^2}{\nu_p}\frac{\nu_p^3}{\mu_p^3} = 2\left(-\frac{\nu_n^2}{\mu_n} + \frac{\nu_p^2}{\mu_p}\right) = 2\nu^2\left(-\mu_n^3 + \mu_p^3\right).\end{aligned}$$

Hence, we guess the general form of $\int_{-\infty}^0 x^n \frac{\exp(cx)}{-x}\mathrm{d}x$ is given by $(-1)^n \frac{(n-1)!}{c^n}$ and we proof it in the following lemma.

Lemma B.7.1

$$\int_{-\infty}^0 x^n \frac{\exp(cx)}{-x}\mathrm{d}x = (-1)^n\frac{(n-1)!}{c^n}.$$

Proof. For $k = n+1$,

$$\begin{aligned}\int_{-\infty}^0 x^{n+1}\frac{\exp(cx)}{-x}\mathrm{d}x &= -\frac{1}{c}\int_{-\infty}^0 x^n\mathrm{d}\exp(cx) = -\frac{1}{c}\left[x^n\exp(cx)\right]_{-\infty}^0 + \frac{1}{c}\int_{-\infty}^0 \exp(cx)\,\mathrm{d}x^n \\ &= \frac{n}{c}\int_{-\infty}^0 x^{n-1}\exp(cx)\,\mathrm{d}x = (-1)^{n+1}\frac{n!}{c^{n+1}}.\end{aligned}$$

□

Similarly, from the calculation above, we guess the general form of $\int_0^\infty x^n\frac{\exp(-cx)}{x}\mathrm{d}x$ is given by $\frac{(n-1)!}{c^n}$ and we proof it in the following lemma.

Lemma B.7.2

$$\int_0^\infty x^n\frac{\exp(-cx)}{x}\mathrm{d}x = \frac{(n-1)!}{c^n}.$$

Proof. For $k = n+1$,

$$\begin{aligned}\int_0^\infty x^{n+1}\frac{\exp(-cx)}{x}\mathrm{d}x &= -\frac{1}{c}\int_0^\infty x^n\mathrm{d}\exp(-cx) = -\frac{1}{c}\left[x^n\exp(-cx)\right]_0^\infty + \frac{1}{c}\int_0^\infty\exp(-cx)\,\mathrm{d}x^n \\ &= \frac{n}{c}\int_0^\infty x^{n-1}\exp(-cx)\,\mathrm{d}x = \frac{n!}{c^{n+1}}.\end{aligned}$$

Finally, we have the general form of $\int_{-\infty}^\infty x^n\nu(\mathrm{d}x)$, given in the following lemma. □

Lemma B.7.3

$$\int_{-\infty}^\infty x^n\nu(\mathrm{d}x) = (n-1)!\nu^{n-1}\left[(-1)^n\mu_n^n + \mu_p^n\right].$$

Proof.

$$\begin{aligned}
\int_{-\infty}^{\infty} x^n \nu\left(\mathrm{d}x\right) &= \frac{\mu_n^2}{\nu_n}\int_{-\infty}^{0} x^n \frac{\exp\left(\frac{\mu_n}{\nu_n}x\right)}{-x}\mathrm{d}x + \frac{\mu_p^2}{\nu_p}\int_{0}^{\infty} x^n \frac{\exp\left(-\frac{\mu_p}{\nu_p}x\right)}{x}\mathrm{d}x \\
&= \frac{\mu_n^2}{\nu_n}\left(-1\right)^n\left(n-1\right)!\frac{\nu_n^n}{\mu_n^n} + \frac{\mu_p^2}{\nu_p}\left(n-1\right)!\frac{\nu_p^n}{\mu_p^n} = \left(n-1\right)!\left[\left(-1\right)^n\frac{\nu_n^{n-1}}{\mu_n^{n-2}} + \frac{\nu_p^{n-1}}{\mu_p^{n-2}}\right] \\
&= \left(n-1\right)!\nu^{n-1}\left[\left(-1\right)^n\mu_n^n + \mu_p^n\right].
\end{aligned}$$

□

B.8 Taylor expansion of $\exp\left(x\right)$

By direct calculation, $\exp\left(\Delta x\right) = 1.001$. Approximation results is given in Table B.8.1. The second column gives the p-th derivative of $\exp\left(\Delta x\right)$, $D^i \exp\left(x\right)$, and the third column gives the approximation given by Taylor expansion using terms up to $i = p : \sum_{i=0}^{p}\frac{D^i\exp(x)}{i!}\left(\Delta x\right)^i$.

p	pth derivative	Up to i=p	p	pth derivative	Up to i=p	p	pth derivative	Up to i=p
1	1	1.001	8	1.51E+11	1.001	15	-3.08E+33	1.001
2	1	1.001	9	-1.05E+13	1.001	16	2.27E+37	1.001
3	1	1.001	10	1.63E+16	1.001	17	3.53E+39	1.001
4	0.998297	1.001	11	-2.80E+20	1.001	18	-2.67E+43	1.001
5	4.04659	1.001	12	3.90E+23	1.001	19	-2.69E+45	1.001
6	35989.7	1.001	13	9.44E+26	1.001	20	3.05E+49	1.001
7	-1.46E+07	1.001	14	8.73E+30	1.001			

Table B.8.1: Approximation of $\exp\left(\Delta x\right)$ using Taylor expansions.

Appendix C

Part III

C.1 Skewness and kurtosis trades

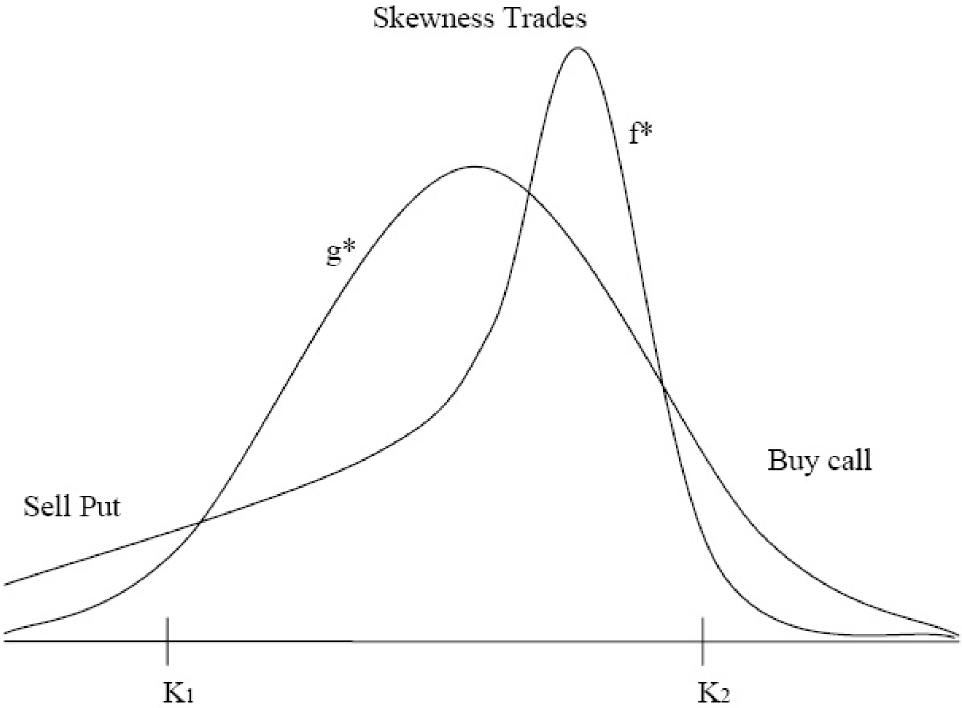

Figure C.1.1: Skewness trades, where f^* is the option implied density and g^* is the history implied density.

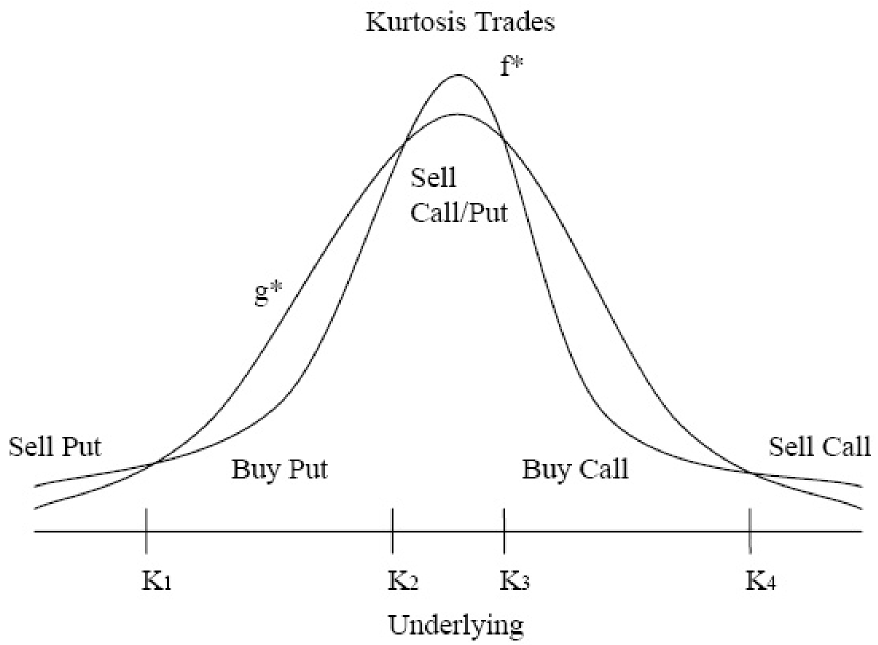

Figure C.1.2: Kurtosis trades, where f^* is the option implied density and g^* is the history implied density.

Skewness Trade	(S1)	$\text{skew}(f^*) > \text{skew}(g^*)$	Sell OTM put and buy OTM call
	(S2)	$\text{skew}(f^*) < \text{skew}(g^*)$	Buy OTM put and sell OTM call
Kurtosis Trade	(K1)	$\text{kurt}(f^*) > \text{kurt}(g^*)$	Sell FOTM, ATM and buy NOTM options
	(K2)	$\text{kurt}(f^*) < \text{kurt}(g^*)$	Buy FOTM, ATM and sell NOTM options

Table C.1.1: Strategies for skewness and kurtosis trades, where OTM, FOTM, NOTM and ATM stand for out-of-money, far-out-of-money, near-out-of-money and at-the-money options, $f^*$is the option implied density and g^* is the history implied density.

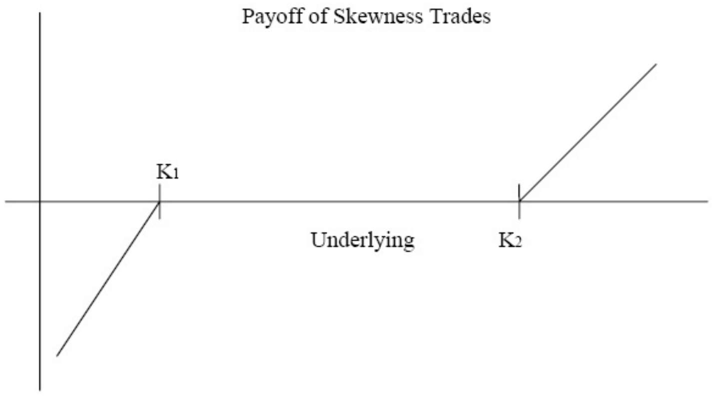

Figure C.1.3: Payoff of S1 trade.

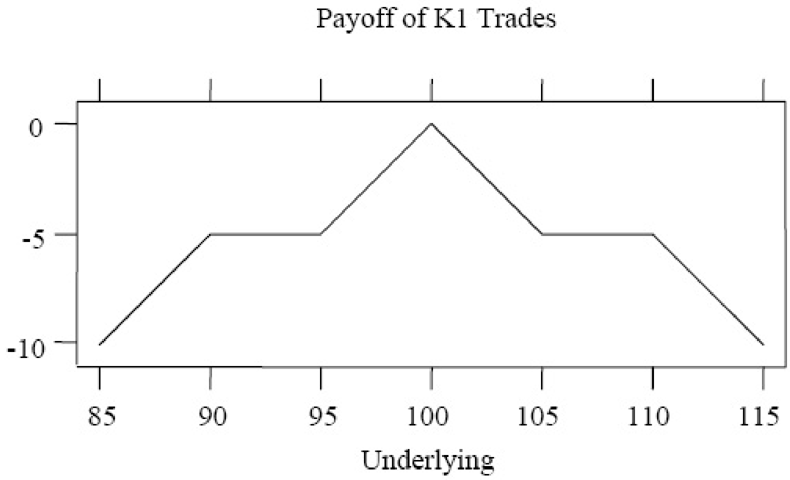

Figure C.1.4: Payoff of K1 trade.

C.2 Proof of Proposition 8.2.1

By integration by parts, we have,
i)

$$
\begin{aligned}
p\left(f^{o}(\cdot), f^{h}(\cdot), E_L, X_L, \mathrm{P}\right) &= \int_{E_L}^{X_L} (\mathrm{P}-s)\left[f^{o}(s)-f^{h}(s)\right] \mathrm{d}s \\
&= \mathrm{P}\Delta G\left(E_L, X_L\right) - \int_{E_L}^{X_L} s\left[f^{o}(s)-f^{h}(s)\right] \mathrm{d}s \\
&= \mathrm{P}\Delta G\left(E_L, X_L\right) - \int_{E_L}^{X_L} s \mathrm{d}\left\{\int_{E_L}^{s}\left[f^{o}(u)-f^{h}(u)\right] \mathrm{d}u\right\} \\
&= \mathrm{P}\Delta G\left(E_L, X_L\right) - \left[s\left\{\int_{E_L}^{s}\left[f^{o}(u)-f^{h}(u)\right] \mathrm{d}u\right\}\right]_{E_L}^{X_L} \\
&\quad + \int_{E_L}^{X_L} \int_{E_L}^{s}\left[f^{o}(u)-f^{h}(u)\right] \mathrm{d}u \mathrm{d}s \\
&= \left(\mathrm{P}-X_L\right) \Delta G\left(E_L, X_L\right) + \int_{E_L}^{X_L} \Delta G\left(E_L, s\right) \mathrm{d}s.
\end{aligned}
$$

ii)

$$
\begin{aligned}
p\left(f^h(\cdot), f^o(\cdot), X_L, \mathrm{P}, \mathrm{P}\right) &= \int_{X_L}^{\mathrm{P}} (\mathrm{P}-s)\left[f^h(s)-f^o(s)\right] \mathrm{d}s \\
&= -\mathrm{P}\Delta G(X_L,\mathrm{P}) - \int_{X_L}^{\mathrm{P}} s\left[f^h(s)-f^o(s)\right] \mathrm{d}s \\
&= -\mathrm{P}\Delta G(X_L,\mathrm{P}) - \int_{X_L}^{\mathrm{P}} s\mathrm{d}\left\{\int_{X_L}^{s} \left[f^h(s)-f^o(s)\right] \mathrm{d}u\right\} \\
&= -\mathrm{P}\Delta G(X_L,\mathrm{P}) - \left[s\left\{\int_{X_L}^{s} \left[f^h(s)-f^o(s)\right] \mathrm{d}u\right\}\right]_{X_L}^{\mathrm{P}} \\
&\quad + \int_{X_L}^{\mathrm{P}}\int_{X_L}^{s} \left[f^h(s)-f^o(s)\right] \mathrm{d}u\mathrm{d}s \\
&= -\mathrm{P}\Delta G(X_L,\mathrm{P}) - \left[-\mathrm{P}\Delta G(X_L,\mathrm{P})\right] - \int_{X_L}^{\mathrm{P}} \Delta G(X_L,s)\,\mathrm{d}s \\
&= -\int_{X_L}^{\mathrm{P}} \Delta G(X_L,s)\,\mathrm{d}s.
\end{aligned}
$$

iii)

$$
\begin{aligned}
c\left(f^h(\cdot), f^o(\cdot), \mathrm{C}, X_R, \mathrm{C}\right) &= \int_{\mathrm{C}}^{X_R} (s-\mathrm{C})\left[f^h(s)-f^o(s)\right] \mathrm{d}s \\
&= \int_{\mathrm{C}}^{X_R} s\left[f^h(s)-f^o(s)\right] \mathrm{d}s + \mathrm{C}\Delta G(\mathrm{C},X_R) \\
&= -\int_{\mathrm{C}}^{X_R} s\mathrm{d}\left\{\int_{\mathrm{C}}^{s} \left[f^o(u)-f^h(u)\right] \mathrm{d}u\right\} \mathrm{d}s \\
&\quad + \mathrm{C}\Delta G(\mathrm{C},X_R) \\
&= -\left[s\int_{\mathrm{C}}^{s} \left[f^o(u)-f^h(u)\right] \mathrm{d}u\right]_{\mathrm{C}}^{X_R} \\
&\quad + \int_{\mathrm{C}}^{X_R}\int_{\mathrm{C}}^{s} \left[f^o(u)-f^h(u)\right] \mathrm{d}u\mathrm{d}s + \mathrm{C}\Delta G(\mathrm{C},X_R) \\
&= -X_R\Delta G(\mathrm{C},X_R) + \int_{\mathrm{C}}^{X_R} \Delta G(\mathrm{C},s)\,\mathrm{d}s + \mathrm{C}\Delta G(\mathrm{C},X_R) \\
&= -(X_R-\mathrm{C})\,\Delta G(\mathrm{C},X_R) + \int_{\mathrm{C}}^{X_R} \Delta G(\mathrm{C},s)\,\mathrm{d}s.
\end{aligned}
$$

iv)

$$
\begin{aligned}
c\left(f^h(\cdot), f^o(\cdot), X_R, \infty, \mathrm{C}\right) &= \int_{X_R}^{E_R} (s-\mathrm{C})\left[f^o(s) - f^h(s)\right] \mathrm{d}s \\
&= \int_{X_R}^{E_R} s\left[f^o(s) - f^h(s)\right] \mathrm{d}s - \mathrm{C}\Delta G(X_R, E_R) \\
&= \int_{X_R}^{E_R} s\mathrm{d}\left\{\int_{X_R}^{s} \left[f^o(u) - f^h(u)\right] \mathrm{d}u\right\} \\
&\quad -\mathrm{C}\Delta G(X_R, E_R) \\
&= \left[s\int_{X_R}^{s} \left[f^o(u) - f^h(u)\right] \mathrm{d}u\right]_{X_R}^{E_R} - \mathrm{C}\Delta G(X_R, E_R) \\
&\quad - \int_{X_R}^{E_R}\int_{X_R}^{s} \left[f^o(u) - f^h(u)\right] \mathrm{d}u\mathrm{d}s \\
&= E_R\Delta G(X_R, E_R) - \int_{X_R}^{E_R} \Delta G(X_R, s)\, \mathrm{d}s \\
&\quad -\mathrm{C}\Delta G(X_R, E_R) \\
&= (E_R - \mathrm{C})\,\Delta G(X_R, E_R) - \int_{X_R}^{E_R} \Delta G(X_R, s)\, \mathrm{d}s.
\end{aligned}
$$

C.3 Risk-return analysis across different market conditions

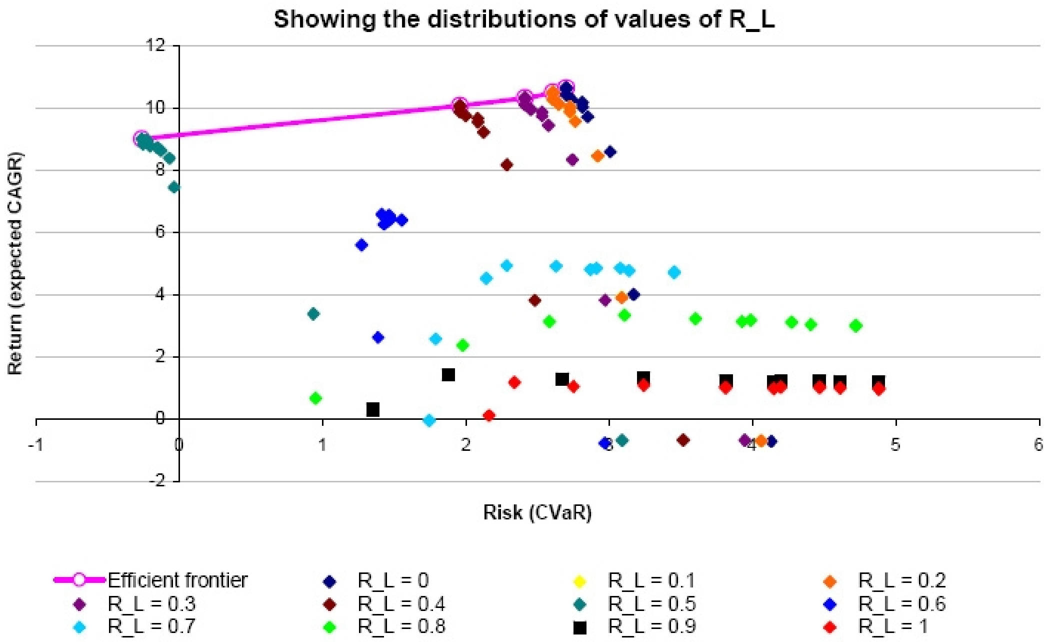

Figure C.3.1: The colour-coded version of Figure 8.7.1.1, illustrating the effects of varying R_L.

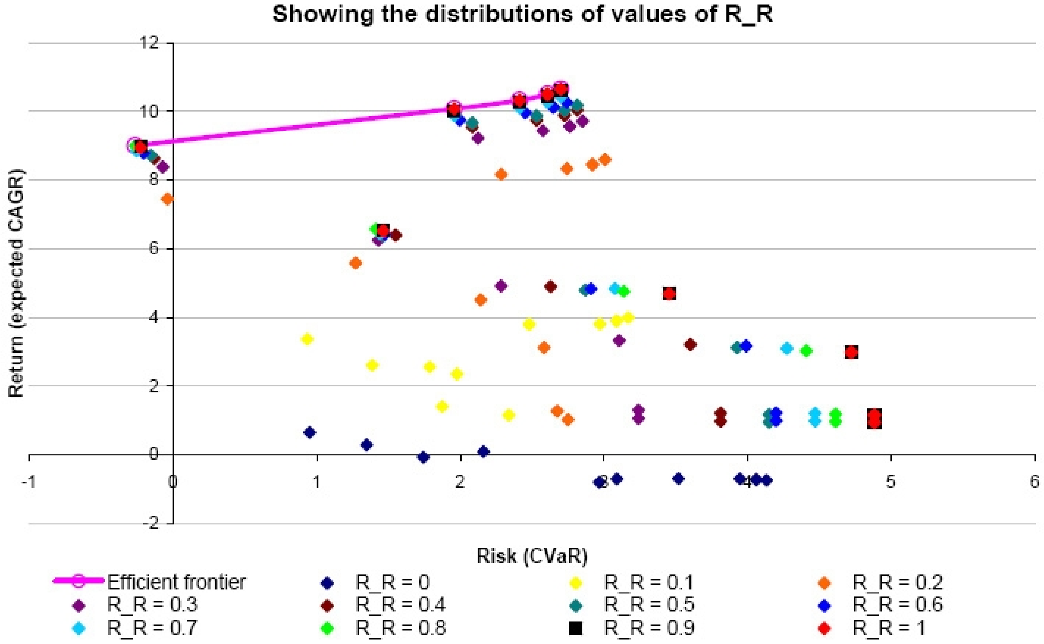

Figure C.3.2: The colour-coded version of Figure 8.7.1.1, illustrating the effects of varying R_R.

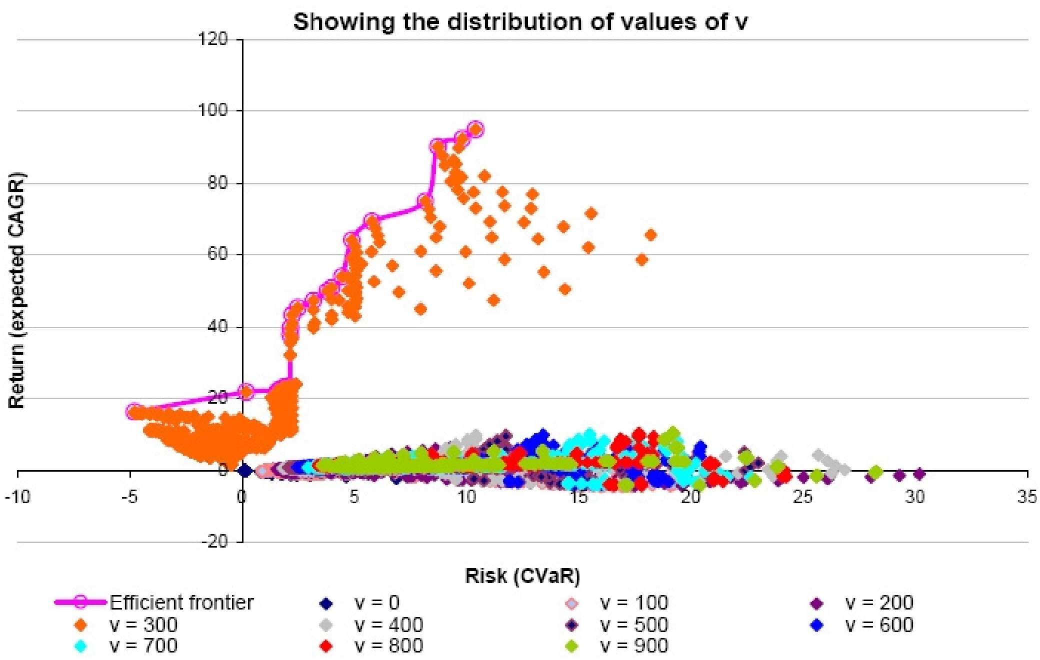

Figure C.3.3: The colour-coded version of Figure 8.7.2.2, illustrating the effects of varying v.

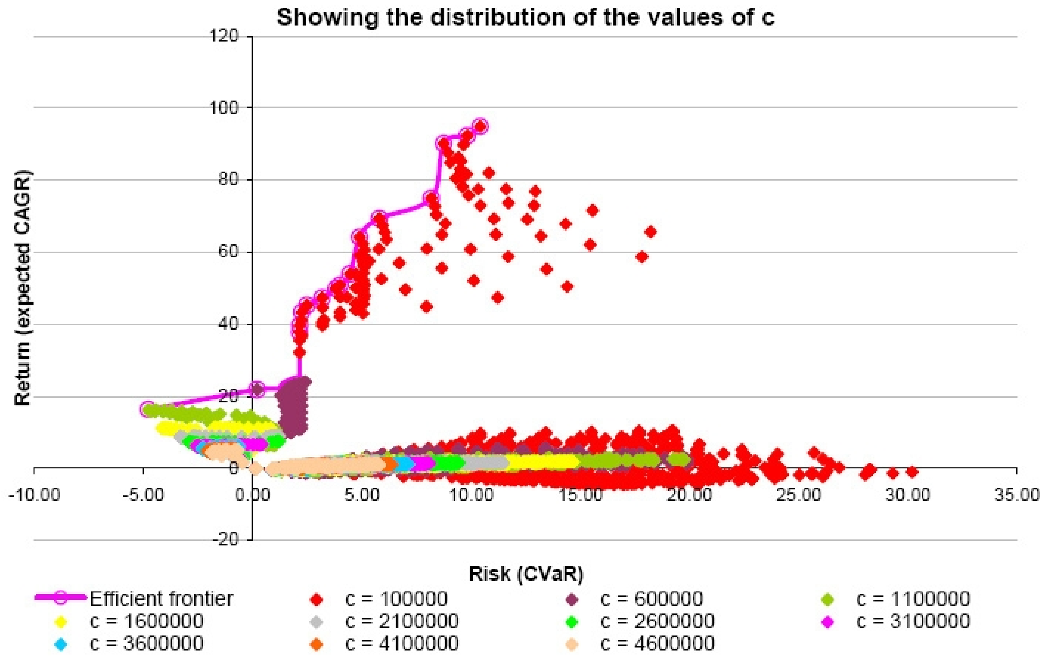

Figure C.3.4: The colour-coded version of Figure 8.7.2.2, illustrating the effects of varying c.

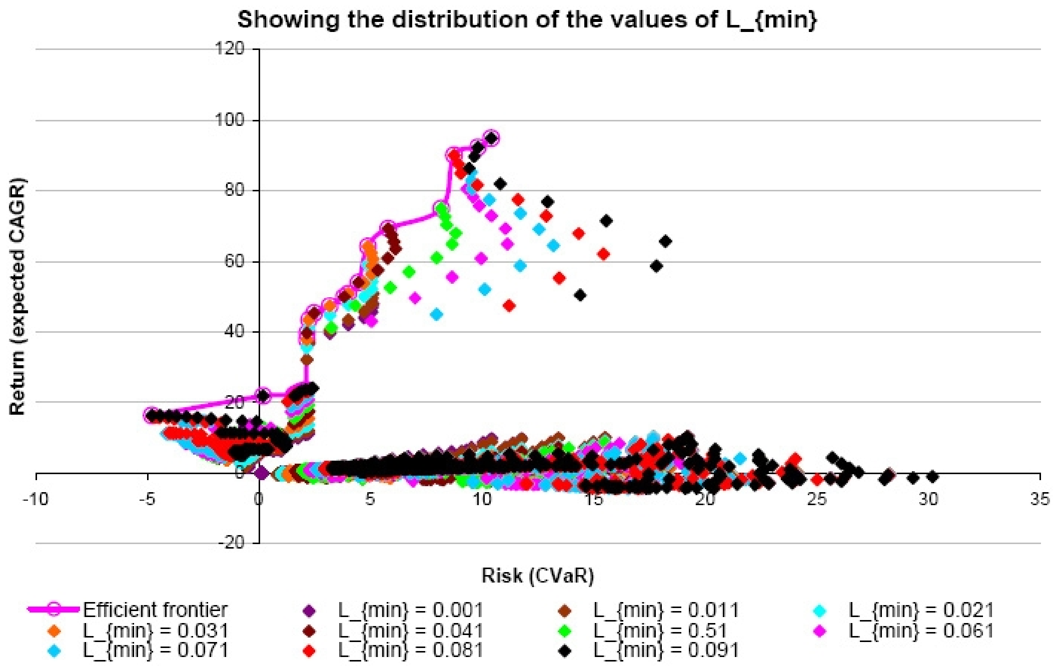

Figure C.3.5: The colour-coded version of Figure 8.7.2.2, illustrating the effects of varying $L_{\min}$.

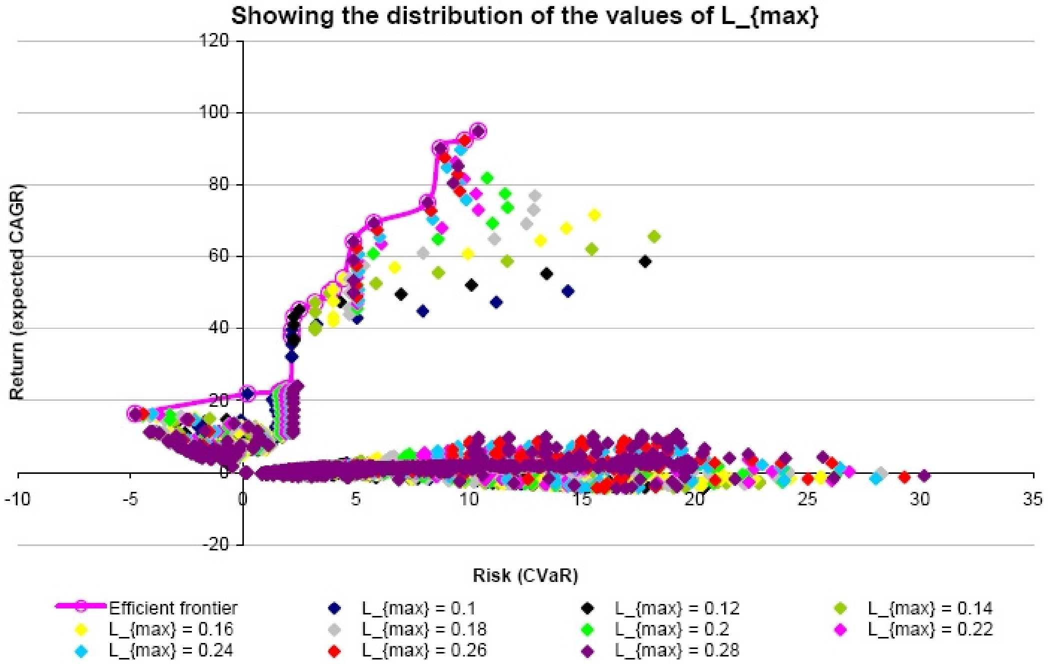

Figure C.3.6: The colour-coded version of Figure 8.7.2.2, illustrating the effects of varying $L_{\max}$.

C.4 Risk-return analysis in a falling market

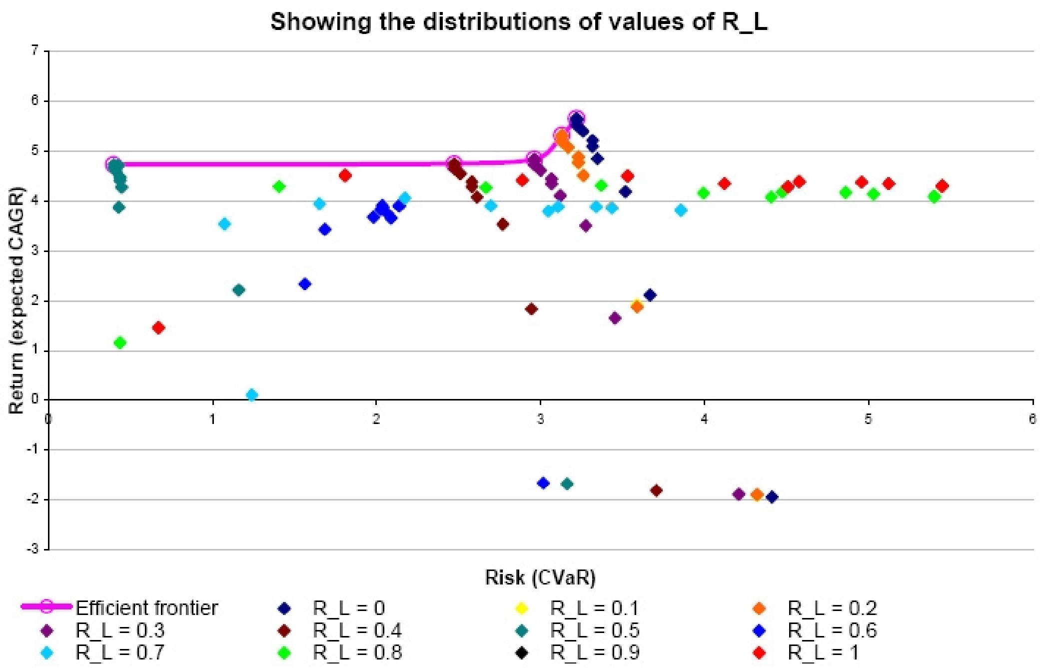

Figure C.4.1: The colour-coded version of Figure 8.7.3.1, illustrating the effects of varying R_L.

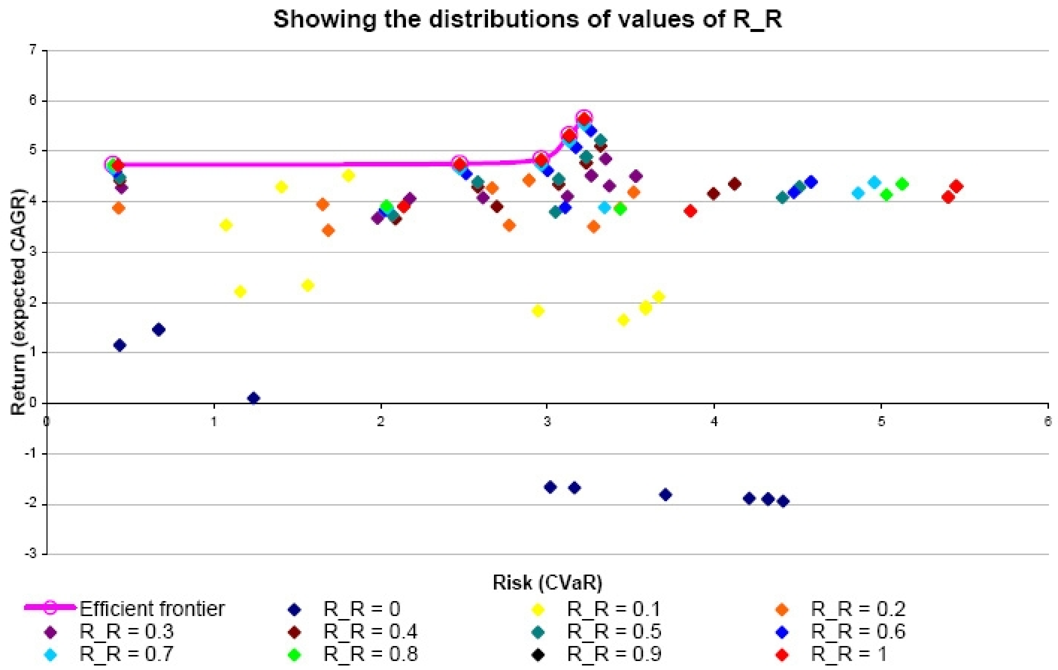

Figure C.4.2: The colour-coded version of Figure 8.7.3.1, illustrating the effects of varying R_R.

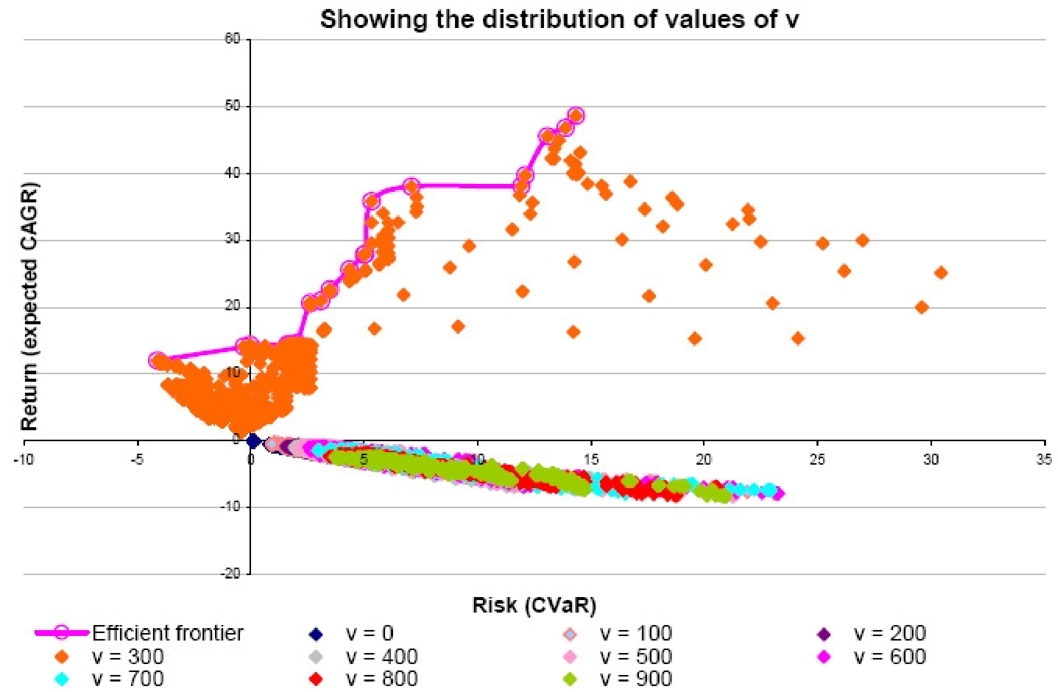

Figure C.4.3: The colour-coded version of Figure 8.7.3.2, illustrating the effects of varying v.

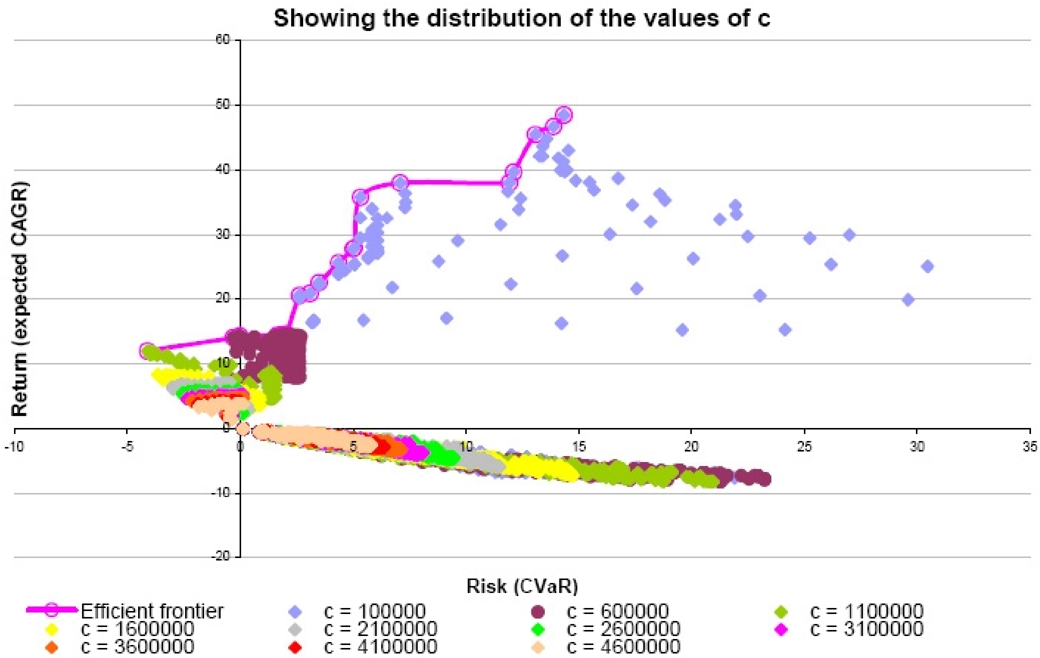

Figure C.4.4: The colour-coded version of Figure 8.7.3.2, illustrating the effects of varying c.

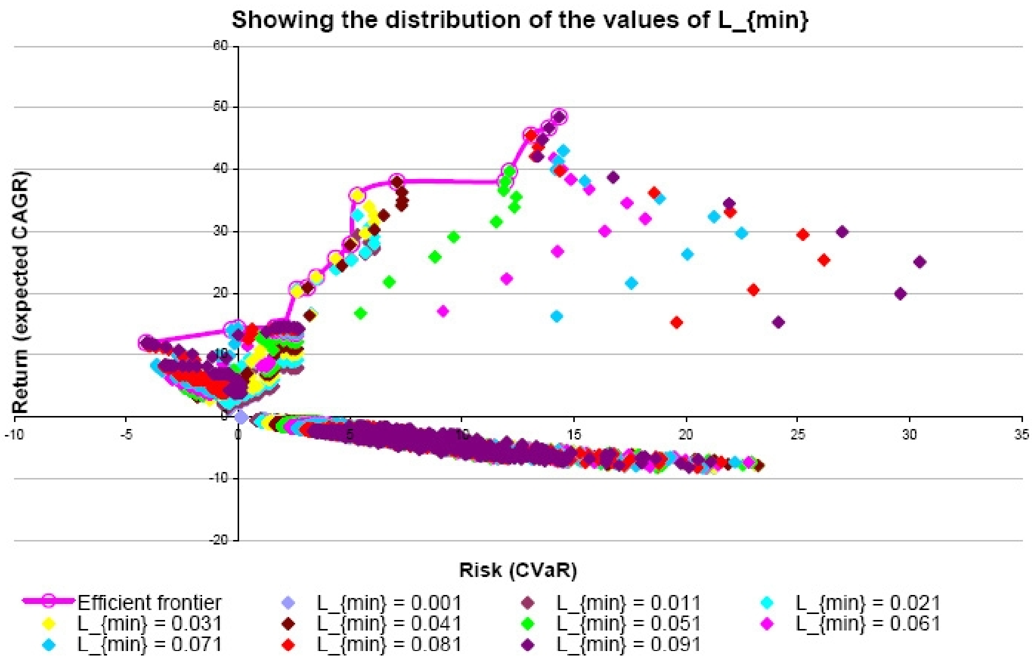

Figure C.4.5: The colour-coded version of Figure 8.7.3.2, illustrating the effects of varying $L_{\min}$.

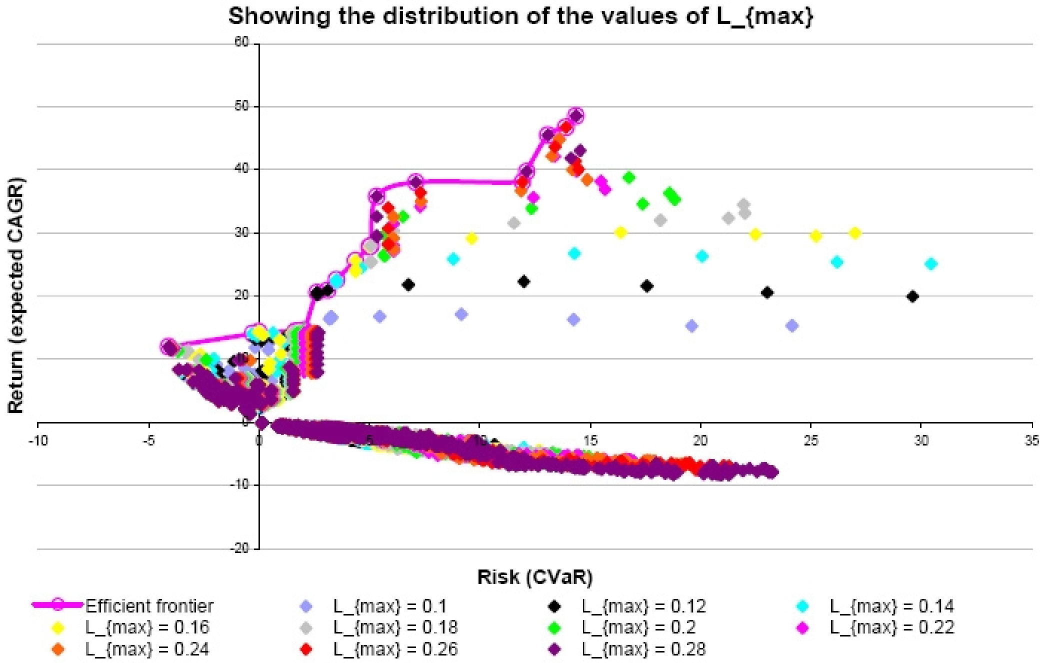

Figure C.4.6: The colour-coded version of Figure 8.7.3.2, illustrating the effects of varying $L_{\max}$.

C.5 Risk-return analysis in a recovering market

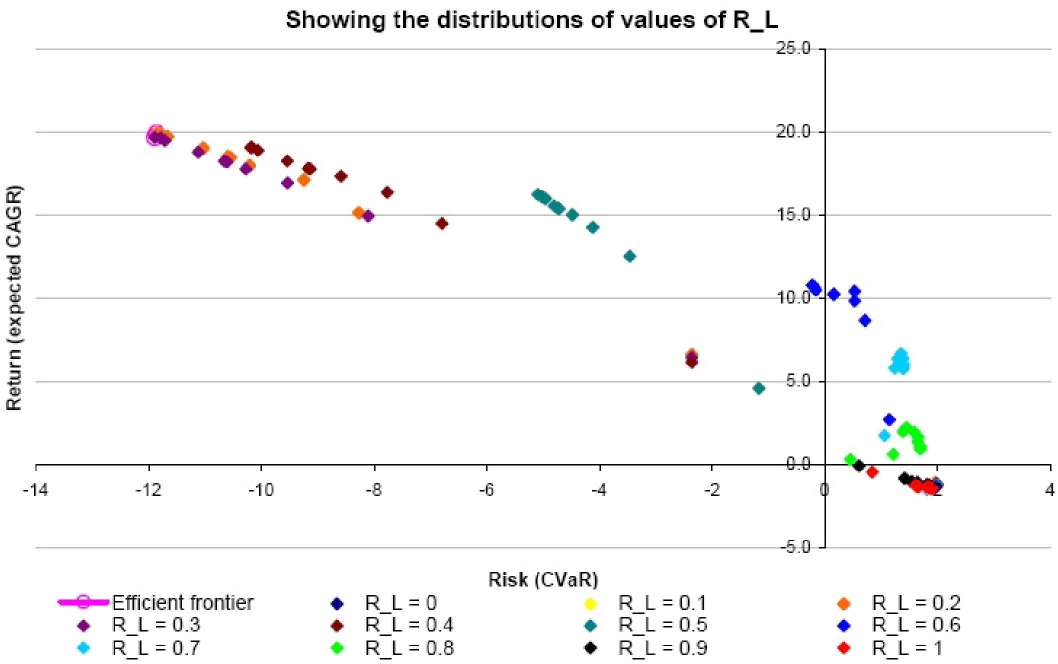

Figure C.5.1: The colour-coded version of Figure 8.7.4.1, illustrating the effects of varying R_L.

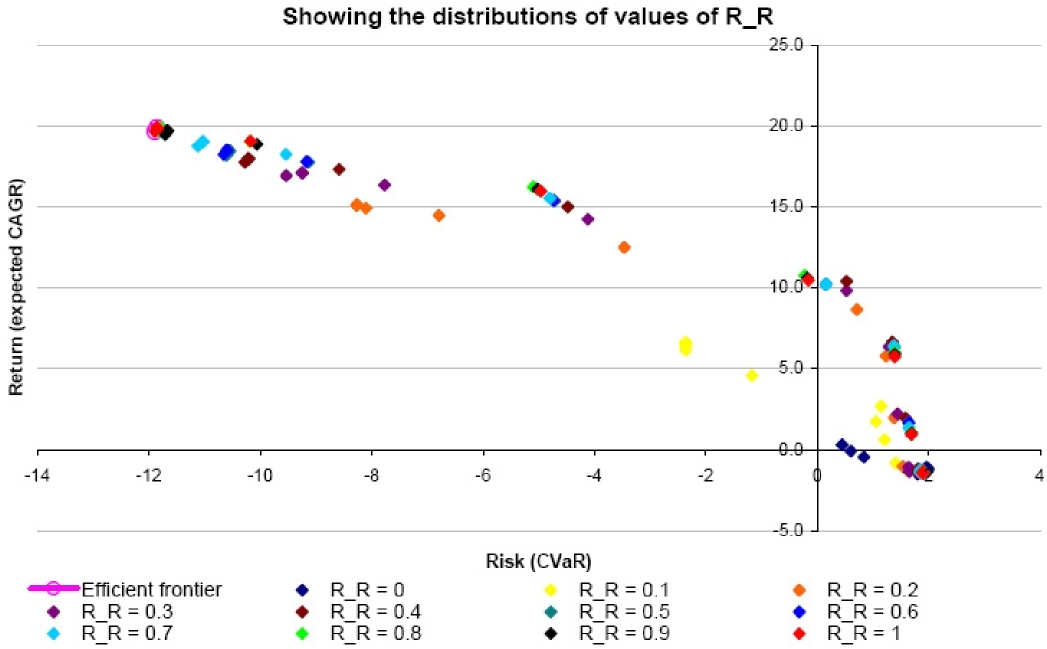

Figure C.5.2: The colour-coded version of Figure 8.7.4.1, illustrating the effects of varying R_R.

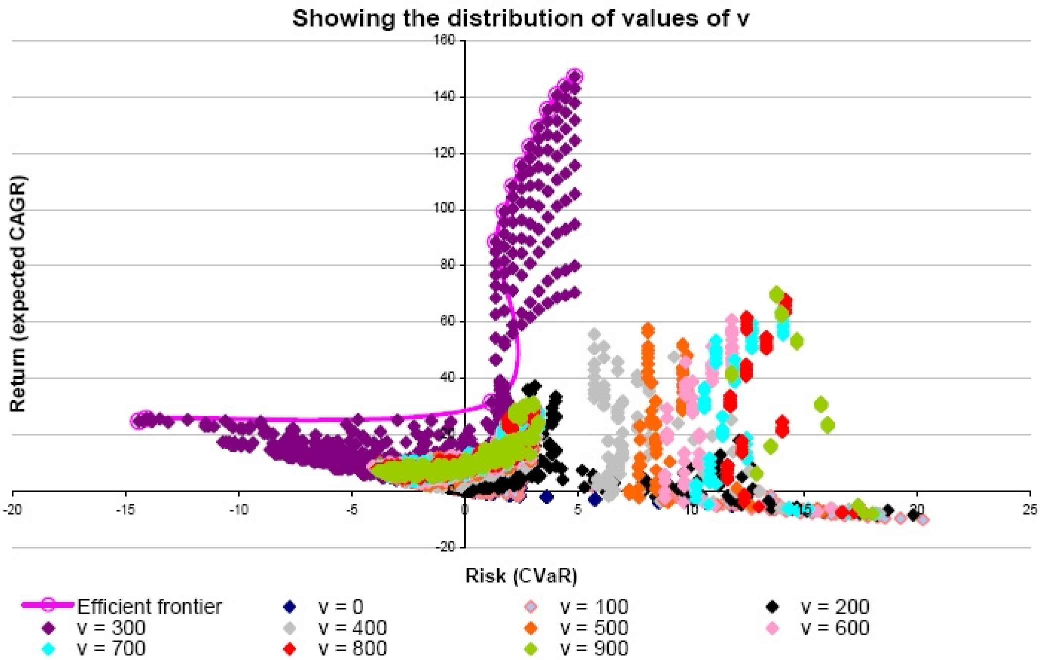

Figure C.5.3: The colour-coded version of Figure 8.7.4.2, illustrating the effects of varying v.

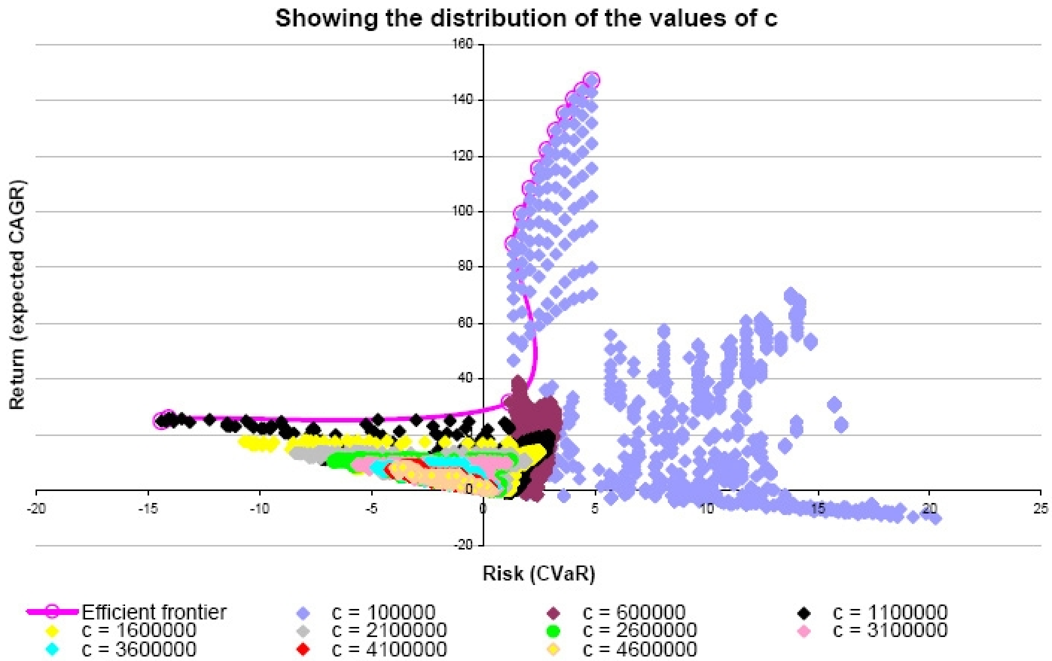

Figure C.5.4: The colour-coded version of Figure 8.7.4.2, illustrating the effects of varying c.

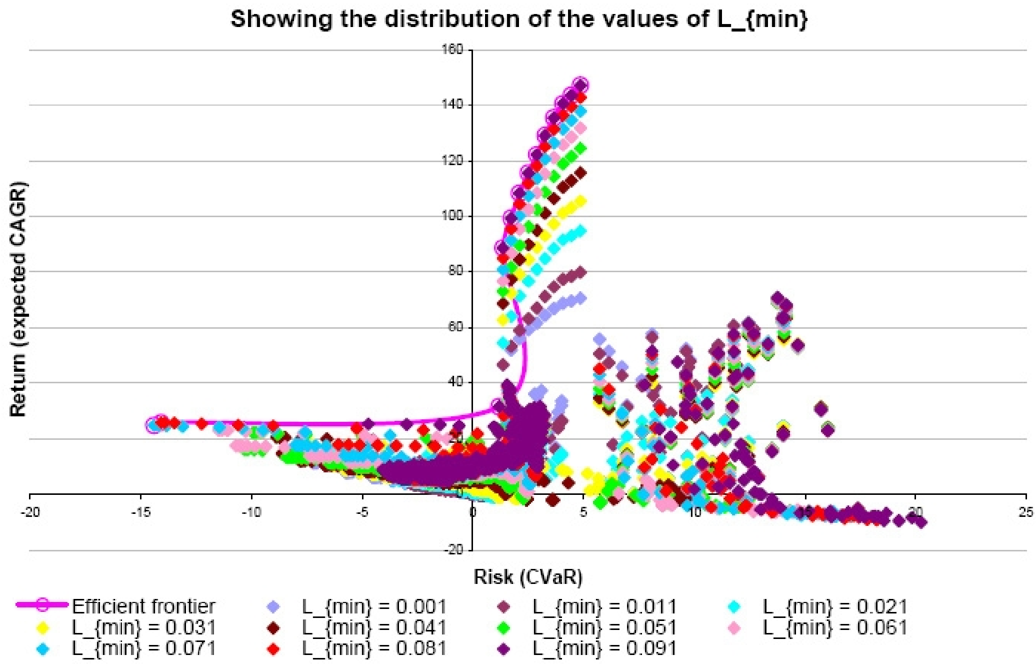

Figure C.5.5: The colour-coded version of Figure 8.7.4.2, illustrating the effects of varying L_{min}.

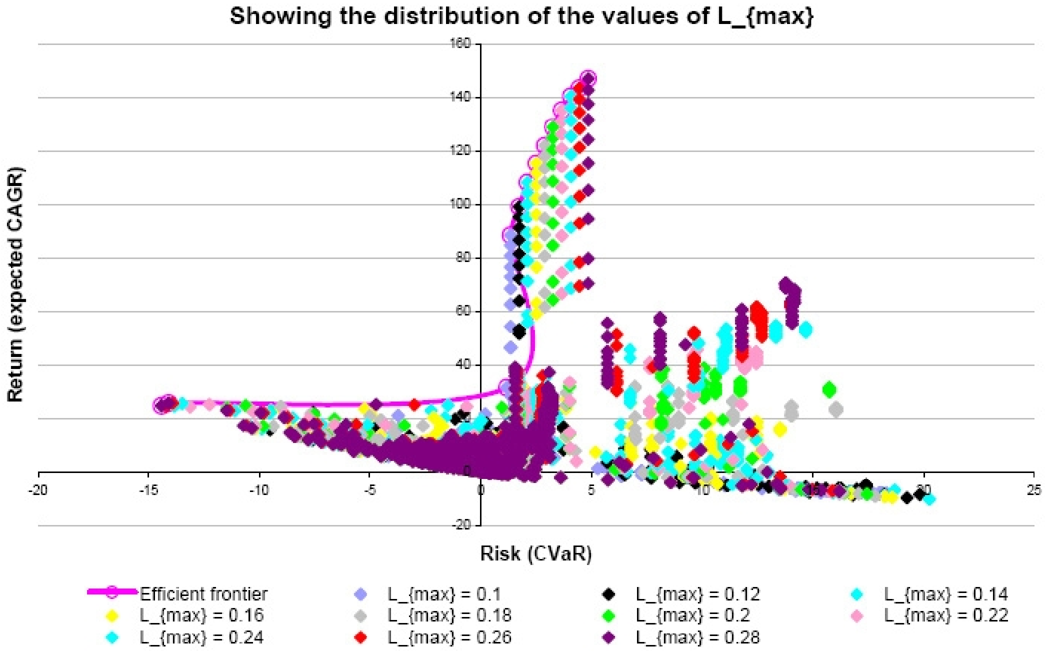

Figure C.5.6: The colour-coded version of Figure 8.7.4.2, illustrating the effects of varying $L_{\max}$.

C.6 Risk-return analysis in a rising market

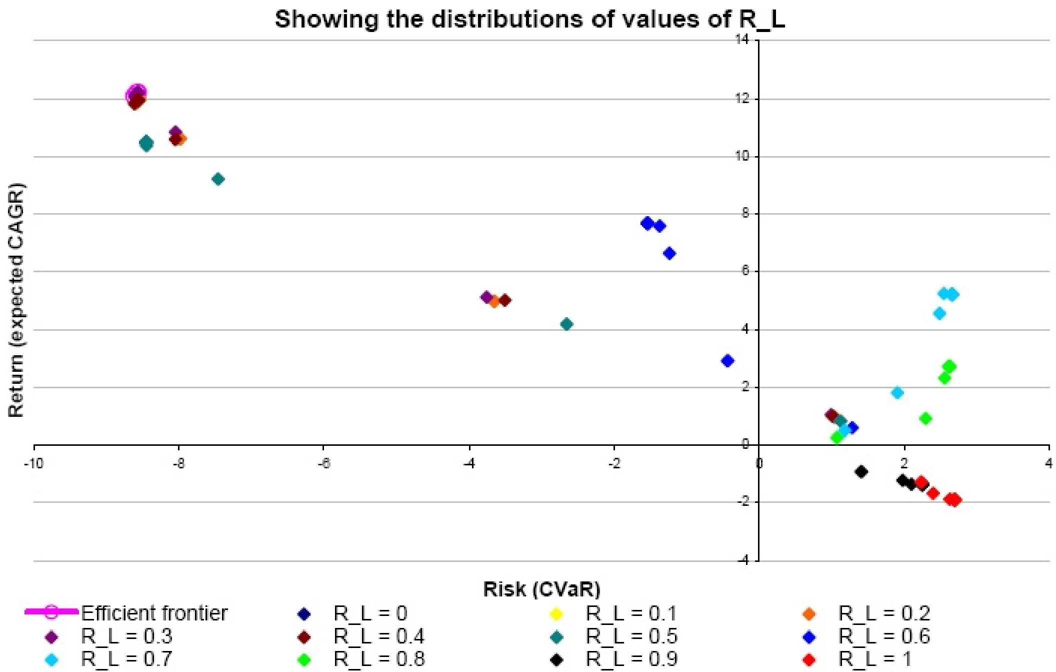

Figure C.6.1: The colour-coded version of Figure 8.7.5.1, illustrating the effects of varying R_L.

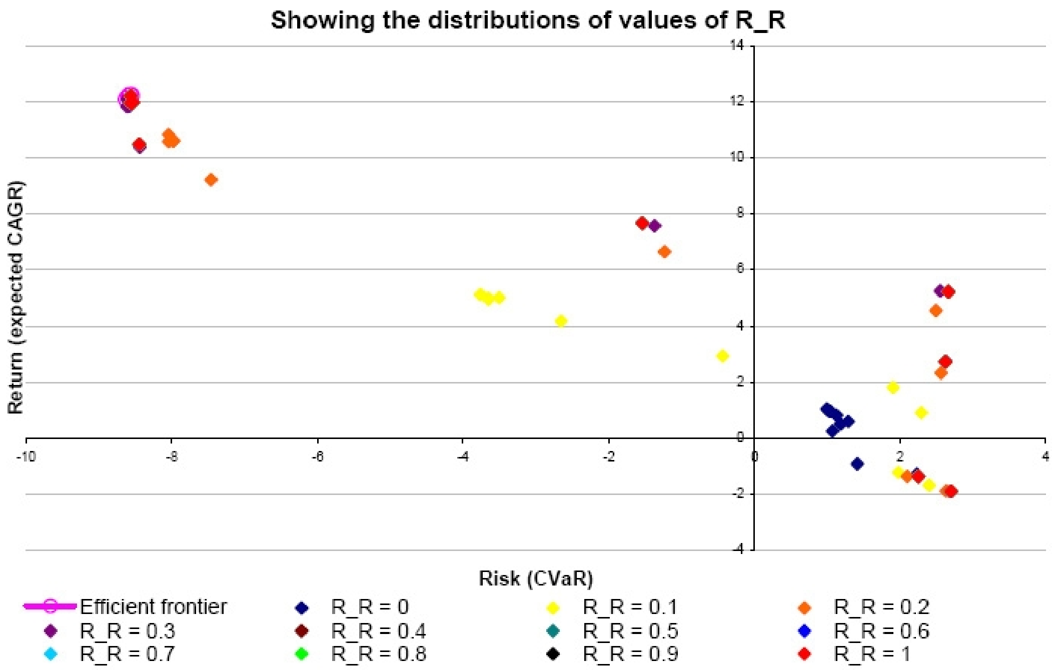

Figure C.6.2: The colour-coded version of Figure 8.7.5.1, illustrating the effects of varying R_R.

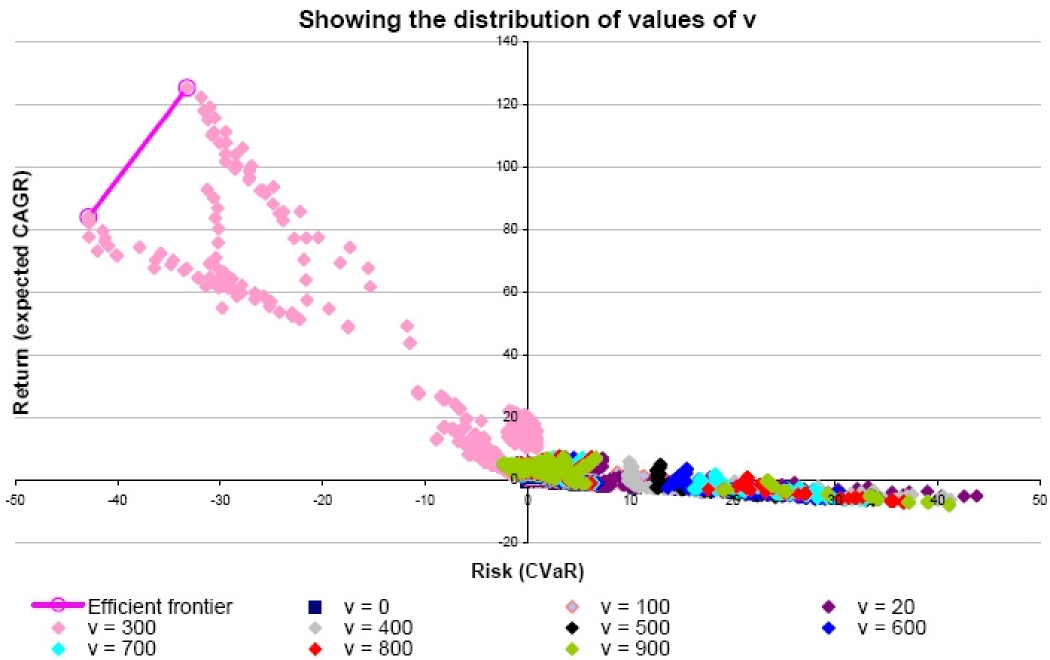

Figure C.6.3: The colour-coded version of Figure 8.7.5.2, illustrating the effects of varying v.

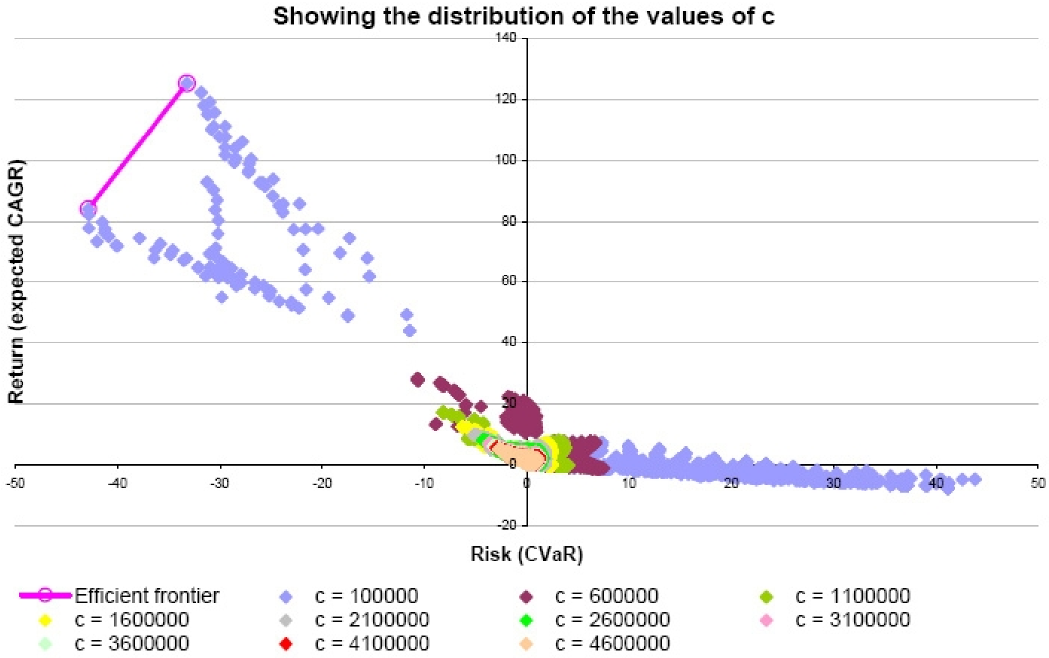

Figure C.6.4: The colour-coded version of Figure 8.7.5.2, illustrating the effects of varying c.

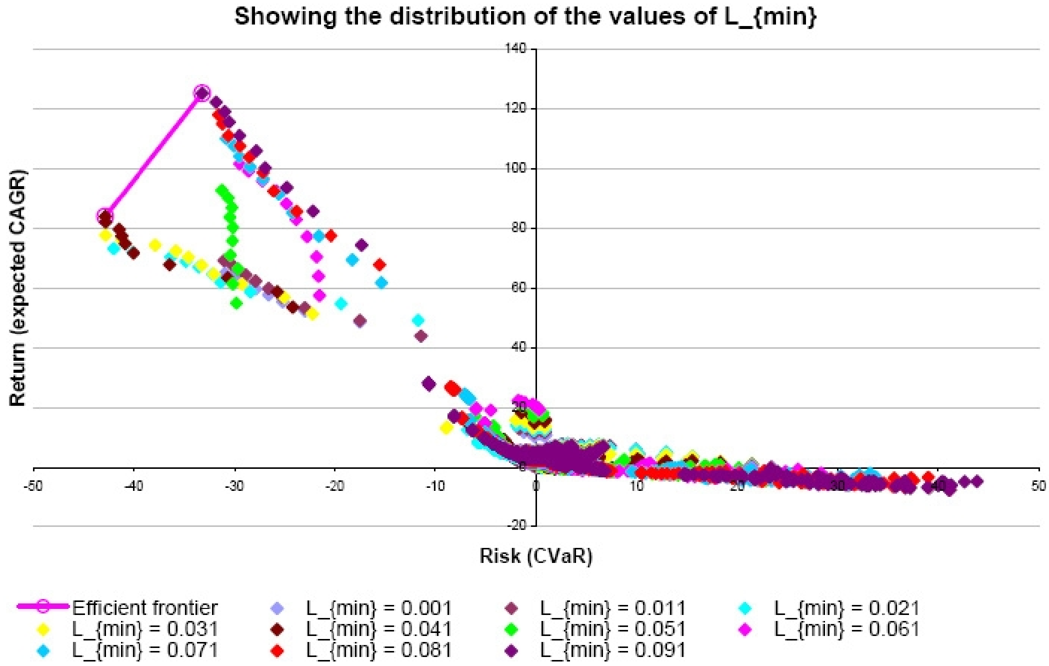

Figure C.6.5: The colour-coded version of Figure 8.7.5.2, illustrating the effects of varying $L_{\min}$.

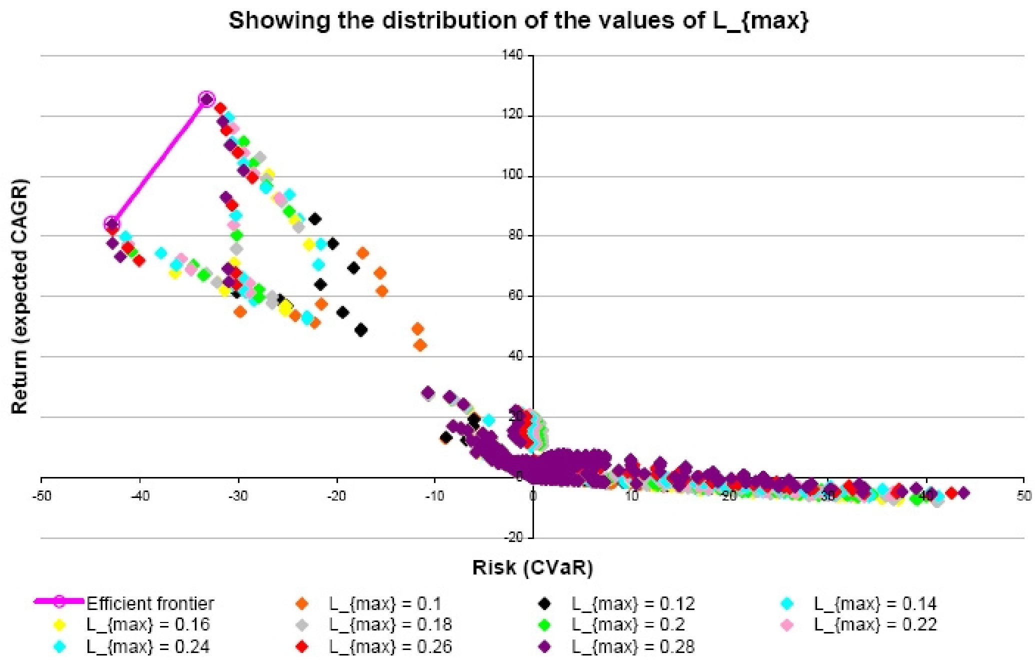

Figure C.6.6: The colour-coded version of Figure 8.7.5.2, illustrating the effects of varying $L_{\max}$.

References

ABRAMOWITZ, M., & STEGUN, I. A. 1964. *Handbook of mathematical functions.* Applied Mathematics, vol. 55. Washington: National Bureau of Standards.

AÏT-SAHALIA, Y., WANG, Y., & YARED, F. 2001. Do option markets correctly price the probabilities of movement of the underlying asset? *Journal of Econometrics*, **102**, 67–110.

APPLEBAUM, D. 2004. *Lévy processes and stochastic calculus.* Cambridge University Press.

BARNDORFF-NIELSEN, O., & SHEPHARD, N. 2002. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, **17**, 457–477.

BARNDORFF-NIELSEN, O., & SHEPHARD, N. 2006. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, **4**(1), 1–30.

BENTH, F., NUNNO, G. DI, LØKKA, A., ØKSENDAL, B., & PROSKE, F. 2003. Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. *Mathematical Finance*, **13**(1), 55–72.

BERTOIN, J. 1996. *Lévy processes.* Cambridge University Press, Cambridge.

BINGHAM, N., & KIESEL, R. 2001. *Risk-neutral valuation, pricing and hedging of financial derivatives.* Springer-Verlag.

BLACK, F., & JONES, R. 1987. Simplifying portfolio insurance. *The Journal of Portfolio Management*, 48–51.

BLASKOWITZ, O. 2001. Trading on deviations of implied and historical distributions. *Diploma Thesis, Humboldt-University of Berlin, School of Business and Economics*, October.

BLASKOWITZ, O., & SCHMIDT, P. 2002. *Trading on deviations of implied and historical density.* in W. Härdle, T. Kleinow, G. Stahl: Applied Quantitative Finance, Springer Verlag, Heidelberg.

BLASKOWITZ, O., HÄRDLE, W., & SCHMIDT, P. 2004. Skewness and kurtosis trades. *Center for Applied Statistics and Economics*, 1–17.

BOWMAN, A., & AZZALINI, A. 1997. *Applied smoothing techniques for data analysis.* Oxford University Press.

Carr, P., & Madan, D. 1999. Option valuation using the Fast Fourier Transform. *Journal of Computational Finance*, **3**, 463–520.

Carr, P., Geman, H., & Madan, D. 2001. Pricing and hedging in incomplete markets. *Journal of Financial Economics*, **62**, 131–167.

Carr, P., Geman, H., Madan, D., & Yor, M. 2003. Stochastic volatility for Lévy processes. *Mathematical Finance*, **13**(3), 345–382.

Cherny, A., & Shiryaev, A. 2002. Vector stochastic integrals and the fundamental theorems of asset pricing. *Proceedings of the steklov institute of mathematics*, **237**, 6ąV49.

Chourdakis, K. 2005. Option pricing using the fractional FFT. *Journal of Computational Finance*, **8**(2), 1–18.

Cont, R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, **1**, 223–236.

Cont, R., & Tankov, P. 2003. *Financial modelling with jump processes*. CRC Press.

Cont, R., Tankov, P., & Voltchkova, E. 2005. Hedging with options in models with jumps. *Abel Symposium 2005 on Stochastic Analysis and Applications*.

Corcuera, J. M., Nualart, D., & Schoutens, W. 2005. Completion of a Lévy market by power-jump assets. *Finance and Stochastics*, **9**, 109–127.

Corcuera, J. M., J. Guerra, D. Nualart, & Schoutens, W. 2006. Optimal investment in a Lévy market. *Applied Mathematics and Optimization*, **53**(3), 279–309.

Davis, M., & Johansson, M. 2006. Malliavin Monte Carloo Greeks for jump diffusions. *Stochastic processes and their applications*, **116**, 101–129.

Delbaen, F., & Schachermayer, W. 1994. A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, **300**, 463–520.

Delbaen, F., & Schachermayer, W. 1998. The fundamental theorem of asset pricing for unbounded stochastic processes. *Annals of mathematics*, **312**, 215ąV250.

Demeterfi, K., Derman, E., Kamal, M., & Zou, J. 1999. A guide to volatility and variance swaps. *The Journal of Derivatives*, 9–32.

Derman, E., Ergener, D., & Kani, I. 1995. Static options replication. *Journal of Derivatives*, **2**(4).

Dritschel, M., & Protter, P. 1999. Complete markets with discontinuous security price. *Finance and Stochastics*, **3**, 203–214.

EDDAHBI, M., SOLÉ, J.L., & VIVES, J. 2005. A Stroock formula for a certain class of Lévy processes and applications to finance. *Journal of Applied Mathematics and Stochastic Analysis*, **2005**(3), 211–235.

GRADSHTEYN, I., & RYZHIK, I. 1965. *Table of integrals, series, and products. Corrected and enlarged edition.* 6th edn. Academic Press, INC.

HARRISON, J., & KREPS, D. 1979. Martingales and arbitrage in multiperiod security markets. *Journal of economic theory*, **2**, 381ąV408.

HARRISON, J., & PLISKA, S. 1981. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, **11**, 215ąV260.

HE, C., KENNEDY, J., COLEMAN, T., FORSYTH, P., LI, Y., & VETZAL, K. 2005. Calibration and hedging under jump diffusion. *Working paper.*

HIGHAM, D. J. 2001. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, **43**(3), 525–546.

HIGHAM, D. J., & KLOEDEN, P.E. 2002. *MAPLE and MATLAB for stochastic differential equations in finance.* Programming Languages and Systems in Computational Economics and Finance (S.S. Nielsen, ed.), Kluwer Academic Publishers, Amsterdam.

HUEHNE, F. 2005. Malliavin calculus for the computation of greeks in markets driven by pure-jump Lévy processes. *Social science research network*, December.

HULL, J. 2003. *Options, futures, and other derivatives.* 5th edn. Prentice Hall Finance Series.

IKEDA, N., & WATANABE, S. 1989. *Stochastic differential equations and diffusion processes.* Amsterdam: North-Holland.

ITÔ, K. 1951. Multiple Wiener integral. *Journal of the Mathematical Society of Japan*, **3**, 157–169.

ITÔ, K. 1956. Spectral type of the shift transformation of differential processes with stationary increments. *Transactions of the American Mathematical Society*, **81**, 253–263.

JACOD, J., & SHIRYAEV, A. 2002. *Limit theorems for stochastic processes.* Springer: Berlin.

JAMSHIDIAN, F. 2005. Chaotic expansion of powers and martingale representation. *Working paper.* http://ideas.repec.org.

JEFFREYS, H., & JEFFREYS, B. 1988. *Taylor's Theorem, methods of Mathematical Physics.* 3rd edn. Cambridge, England: Cambridge University Press. 1.133.

KANNAN, D., & LAKSHMIKANTHAM, V. 2001. *Handbook of stochastic analysis and applications.* Marcel Dekker.

KHAN, I., & OHBA, R. 2003. Taylor series based finite difference approximations of higher-degree derivatives. *Journal of Computational and Applied Mathematics*, **154**, 115–124.

Kijima, M. 2002. *Stochastic processes with applications to finance*. Chapman and Hall.

Klebaner, F. 2005. *Stochastic calculus with applications*. Imperial College Press.

Kloeden, P. E. 2002. The systematic derivation of higher order numerical schemes for stochastic differential equations. *Milan Journal of Mathematics*, **70**, 187–207.

Kloeden, P. E., & Platen, E. 1999. *Numerical solution of stochastic differential equations*. Springer.

Koekeok, R., & Swarttouw, R. F. 1998. The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue. *Report 98-17, Delft University of Technology*, 39.

Leland, H. 1979. Who should buy portfolio insurance? *The Journal of Finance*, **35**(2), 581–594.

Léon, J., Solé, J., Utzet, F., & Vives, J. 2002. On Lévy processes, Malliavin Calculus and market models with jumps. *Finance and Stochastics*, **6**, 197–225.

Løkka, A. 2004. Martingale representation of functionals of Lévy processes. *Stochastic Analysis and Applications*, **22**(4), 867–892.

Madan, D., Carr, P., & Chang, E. 1998. The Variance Gamma process and option pricing. *European Finance Review*, **2**, 79–105.

Markowitz, H. 1952. *Portfolio selection*. Vol. 7. Journal of Finance.

Martin, R. 2004. *Credit portfolio modeling handbook*. Credit Suisse First Boston.

Miyahara, Y. 2005. Martingale measures for the geometric Lévy process models. *Discussion papers in Economics*, 1–14.

Monat, P., & Stricker, C. 1995. Follmer-Schweizer decomposition and mean-variance hedging for general claims. *The Annals of Probability*, **23**(2), 605–628.

Nualart, D. 1995. *The Malliavin Calculus and related topics*. Probability and Its Applications. Springer-Verlag New York INC.

Nualart, D., & Schoutens, W. 2000. Chaotic and predictable representations for Lévy processes. *Stochastic Processes and their Applications*, **90**, 109–122.

Nualart, D., & Schoutens, W. 2001. BSDE's and Feynman-Kac formuola for Lévy processes with applications in finance. *Bernoulli*, **7**, 761–776.

Nualart, D., & Vives, J. 1990. Anticipative calculus for the Poisson process based on the Fock space. *Séminaire de Probabilités XXIV*, 154–165.

Perold, A. 1986. *Constant portfolio insurance*. Harvard Business School. Unpublished manuscript.

Perold, A., & Sharpe, W. 1988. Dynamic strategies for asset allocation. *Financial Analyst Journal*, January-February, 16–27.

Protter, P. 2004. *Stochastic integration and differential equations*. 2nd edn. Springer.

Protter, P., & Talay, D. 1997. The Euler scheme for Lévy driven stochastic differential equations. *The Annals of Probability*, **25**, 393–423.

Revuz, D., & Yor, M. 1999. *Continuous martingales and brownian motion*. Springer: Berlin.

Sato, K. 1999. *Lévy processes and infinitely divisible distribution*. Vol. 68. Cambridge University Studies in Advanced Mathematics, Cambridge University Press, Cambridge.

Schachermayer, W. 2002. No arbitrage: on the work of david kreps, positivity. *Positivity*, **6**, 359ąV368.

Schoutens, W. 2000. *Stochastic processes and orthogonal polynomials*. Springer.

Schoutens, W. 2003. *Lévy processes in finance: pricing financial derivatives,*. Chichester, New York, N.Y. : J. Wiley.

Schoutens, W. 2005. Moment swaps. *Quantitative Finance*, **5**(6), 525–530.

Solé, J., Utzet, F., & Vives, J. 2006. Chaos expansions and Malliavin calculus for Lévy processes. *Preprint*.

Solé, J., Utzet, F., & Vives, J. 2007. Canonical lévy process and Malliavin calculus. *Stochastic processes and their applications*, **117**, 165–187.

Weisstein, E. 1999a. Incomplete Gamma function. *From MathWorld - A Wolfram Web Resource. http://mathworld.wolfram.com/incompletegammafunction.html*.

Weisstein, E. 1999b. Laguerre polynomial. *From MathWorld - A Wolfram Web Resource. http://mathworld.wolfram.com/laguerrepolynomial.html*.

Windcliff, H., Forsyth, P., & Vetzal, K. 2006. Pricing methods and hedging strategies for volatility derivatives. *Journal of Banking and Finance*, **30**(2), 409–431.

www.euronext.com. 2007. *Ftse 100 index option (european-style)*.