Stein’s Method, Malliavin Calculus, Lévy White Noise Analysis, and their Applications in Financial Mathematics

by

Chu, Pin-Cheng

in the School of Mathematical Sciences

August 2015
Abstract

In this thesis we focus on Stein’s method, Malliavin calculus, Lévy white noise analysis, and their applications in financial mathematics. The principal novel contribution is the derivation and rigorous proof of the Stein bound for functionals of general Lévy processes. In particular, as far as we are aware, nowhere in the literature is there a Stein bound which unifies the Brownian and pure jump components in a natural way. A second contribution is to show that the Stein bounds in the literature for functionals of (i) pure jump Lévy processes, and (ii) the pure Brownian motion Lévy process, are both particular cases of the general bound in the thesis. To derive the Stein bound in the thesis, it is essential to adopt a unified approach to the Brownian part and pure jump part of the Lévy process. In the thesis, this was achieved by making use of the white noise approach developed by Lee and Shih. We give examples in which the Stein bound for Lévy processes are computed.

Finally, we use the closed form expression for an option price, whose underlying asset follows a normal distribution, to approximate an option price, whose underlying asset has stochastic volatility, by application of Stein’s method and Malliavin calculus.
Acknowledgements

I would like to thank my first supervisor Professor Andrew Wood, and my second supervisor Dr. Huiling Le, for all their help, patience, meetings and discussions, and excellent guidance throughout my research in Nottingham. I would never have been able to finish my thesis without the guidance of my supervisors, and support from my parents.

Many thanks to my parents for their support, not only during my PhD but throughout my whole life and education.

I would like to thank Dr. S. Utev for his time during the first three years.

Finally my infinite gratitude to the School of Mathematical Sciences for funding me, and all the administrative staff for all their time, help and support.
# Contents

Abstract i

Acknowledgements ii

1 Introduction 1
   1.1 Stein’s method and normal approximation of Lévy functionals ..... 1
   1.2 Financial option pricing approximation by Stein’s method and Malliavin calculus ........................................... 4

2 Review of relevant literature 7
   2.1 Introduction .............................................. 7
   2.2 Basic probability concepts .................................. 7
   2.3 Hida white noise analysis .................................. 17
   2.4 The Malliavin calculus .................................... 18
      2.4.1 The Wiener-Itô chaos expansion ...................... 19
         2.4.1.1 Iterated Itô integrals .......................... 19
         2.4.1.2 The Wiener-Itô chaos expansion ............... 21
      2.4.2 The Skorohod integral via chaos expansion ............ 22
      2.4.3 The Malliavin derivative as adjoint operator of the Skorohod integral 25
      2.4.4 The Ornstein-Uhlenbeck operator \( L \) .................................. 29

3 Stein’s method and normal approximation of Lévy functionals 30
   3.1 Introduction .............................................. 30
   3.2 Lévy white noise functionals analysis ...................... 31
      3.2.1 Segal-Bargmann transform (S-transform) ............... 39
      3.2.2 The spaces of test and generalised functions ............ 43
         3.2.2.1 Gelfand triple associated with \( L^2(\mathbb{R}, dt) \) .......... 43
         3.2.2.2 Gelfand triple associated with \( L^2(\mathbb{R}, \lambda) \) ............ 47
      3.2.3 Differential operator .................................. 49
      3.2.4 Adjoint operator ..................................... 51
      3.2.5 N operator ........................................... 58
<table>
<thead>
<tr>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.6 Derivative of analytic functionals and chain rule</td>
</tr>
<tr>
<td>3.2.6.1 Derivative of analytic functionals</td>
</tr>
<tr>
<td>3.2.6.2 Chain rule</td>
</tr>
<tr>
<td>3.3 Stein’s method and normal approximation of Lévy functionals</td>
</tr>
<tr>
<td>3.3.1 Normal approximation in the Wasserstein distance via Stein’s method</td>
</tr>
<tr>
<td>3.3.2 Upper bound of Wasserstein distance for Lévy functionals</td>
</tr>
<tr>
<td>3.4 The difference and connection between white noise analysis and Malliavin calculus</td>
</tr>
<tr>
<td>3.4.1 The white noise derivative is a directional derivative</td>
</tr>
<tr>
<td>3.4.2 The Malliavin derivative</td>
</tr>
<tr>
<td>3.4.3 The connection between the white noise derivative and the Malliavin derivative</td>
</tr>
<tr>
<td>3.4.4 Upper bound of Wasserstein distance</td>
</tr>
<tr>
<td>3.5 Summary</td>
</tr>
<tr>
<td>4 Option pricing approximation by Stein’s method and Malliavin calculus</td>
</tr>
<tr>
<td>4.1 Introduction</td>
</tr>
<tr>
<td>4.2 Financial option pricing approximation</td>
</tr>
<tr>
<td>4.2.1 Financial option pricing approximation by Stein’s method and Malliavin calculus</td>
</tr>
<tr>
<td>4.2.2 Financial spread option pricing for Gaussian and non-Gaussian processes</td>
</tr>
<tr>
<td>4.3 Difference between option prices (distance between probability measures)</td>
</tr>
<tr>
<td>4.3.1 Stochastic processes for spread between two assets</td>
</tr>
<tr>
<td>4.3.2 Wasserstein distance</td>
</tr>
<tr>
<td>4.4 Upper bound estimation by Stein’s method and Malliavin calculus</td>
</tr>
<tr>
<td>4.4.1 Calculate the first term in the bound</td>
</tr>
<tr>
<td>4.4.2 Calculate the second term in the bound</td>
</tr>
<tr>
<td>4.4.3 Upper bound estimation</td>
</tr>
<tr>
<td>4.5 Illustrative examples</td>
</tr>
<tr>
<td>4.6 Summary</td>
</tr>
<tr>
<td>5 Conclusions and further research</td>
</tr>
<tr>
<td>5.1 Conclusions</td>
</tr>
<tr>
<td>5.2 Further research</td>
</tr>
<tr>
<td>6 Questions and Actions Taken</td>
</tr>
<tr>
<td>6.1 First some overall comments:</td>
</tr>
<tr>
<td>6.2 Abstract</td>
</tr>
<tr>
<td>6.3 Chapter 1 (Introduction)</td>
</tr>
<tr>
<td>6.4 Chapter 2 (Review of relevant literature)</td>
</tr>
<tr>
<td>6.5 Chapter 3 (Stein’s methods and normal approximation of Lévy functionals)</td>
</tr>
<tr>
<td>6.6 Chapter 4</td>
</tr>
<tr>
<td>6.7 References</td>
</tr>
</tbody>
</table>
Bibliography
Chapter 1

Introduction

1.1 Stein’s method and normal approximation of Lévy functionals

The first question treated in this thesis is the development of Stein’s method and normal approximation of Lévy functionals via white noise analysis initiated by Hida.

According to Reinert [36], combining Malliavin calculus and Steins method has lead to a new framework for normal approximation. Nourdin and Peccati [29] obtain explicit bounds for the normal approximation of smooth functionals of Gaussian fields using Stein’s method and Malliavin calculus on a Gaussian space. Below, $d_W(F, X)$ denotes the Wasserstein distance defined in (3.73). If $F$ is functional of some centered isonormal Gaussian process, $X \sim \mathcal{N}(0, 1)$, $E[F] = 0$, the main result is

$$d_W(F, X) \leq E[|1 - <DF, -DL^{-1}F>|] \leq \sqrt{E[(1 - <DF, -DL^{-1}F>)^2]},$$

where $DF$ is the Malliavin derivative, and $L^{-1}$ is the inverse Ornstein-Uhlenbeck generator. For more details, refer to Theorem 4.4.1.

Furthermore, Peccati, Solé, Taqqu and Utzet [35] extend the analysis to the framework of normal approximation in the Wasserstein distance with a version of Malliavin calculus on the Poisson space.
We know that Brownian motions and Poisson processes are just special cases of Lévy processes. We want to find more general results when $F$ is a Lévy functional. Therefore, in this thesis, for normal approximation, we derive upper bounds for the Wasserstein distance based on generalized Lévy white noise analysis [24] initiated by Hida [12]. We can assess the distance between the law of a Gaussian random variable and the law of a Lévy random variable.

For more details and proof, refer to Theorem 3.3.1. Some examples have been computed, including Example 3.3.1, Example 3.3.2, and Example 3.3.3. We also derive the chain rule for Lévy white noise functionals in Theorem 3.2.10.

The principal novel contribution of Chapter 3:

1. The principal novel contribution is the derivation and rigorous proof of the Stein bound for functionals of general Lévy processes; see Theorem 3.3.1. In particular, as far as we are aware, nowhere in the literature is there a Stein bound which unifies the Brownian and pure jump components in a natural way.

2. A second contribution is to show that the Stein bounds in the literature for functionals of (i) pure jump Lévy processes, and (ii) the pure Brownian motion Lévy process, are both particular cases of the general bound in the thesis; see Corollary 3.3.1 and Corollary 3.3.2.

3. To derive the Stein bound in the thesis, it is essential to adopt a unified approach to the Brownian part and pure jump part of the Lévy process. In the thesis, this was achieved by making use of the the white noise approach developed by Lee and Shih (2004) [24]. We give Example 3.3.3 which computes the upper bound for a double integral and Example 3.3.1 which computes that for a single integral. We show that our approach unifies the Brownian and pure jump components in a natural way, and simplifies the computation.

4. However, whether this is the case remains an open question because, so far as we are aware, no one has yet developed a unified approach to functionals of general Lévy processes within the Malliavin framework. For example, in the book by Nunno, Øksendal and Proske (2009) [32], Brownian functionals and pure jump functionals are treated separately; the book by Nourdin and Peccati (2012) [30], only considers Brownian functionals; and the paper by Peccati, Solé, Taqqu and Utzet (2010) [35], gives a bound in the pure jump (which they refer to as the Poisson) case, and in particular they do not consider Lévy processes with both Brownian and pure jump components; see Remark 3.3.2.
5. In Peccati, Solé, Taqqu and Utzet [35], at page 447, in Remark 2.6, they review Lévy random measure because they want to say something about the product formula of multiple integrals, and in the end, they refer to another paper of Lee and Shih [25] for more details about the product formula of Lévy multiple integrals. However, Remark 2.6 in [35] is not concerned with the proof of Stein-Malliavin upper bound. Moreover, Lee and Shih have published a series of papers to build white noise analysis for Lévy functionals, mostly in Journal of Functional Analysis, including [22], [23], [24], [25], [26], [27], and [38]. They employ the Segal-Bargmann transform of a regular Lévy white noise functional, and define the generalised Lévy white noise functionals by means of their functional representations acting on test functionals. They provide the product formula of multiple Lévy-Itô integrals.

We shall review Stein’s method for normal approximation. According to [4] and [5] Stein’s method provides a way of accurately estimating of the approximation of one probability distribution by another via comparing expectations. An explicit upper bound is computed for the difference between the expectations of a large family of test functions under the two distributions. An associated metric is determined by the family of test functions. The test functions we consider in this thesis consist of all Lipschitz functions $h$ with constant bounded by 1. Then the associated metric is called Wasserstein distance.

In order to compute the explicit upper bound of Wasserstein distance, Nourdin and Peccati [29] [30] obtain the bounds for the normal approximation of smooth functionals of Gaussian fields using Stein’s method and Malliavin calculus on a Gaussian space. Furthermore, Peccati, Solé, Taqqu and Utzet [35] extend the analysis to the framework of normal approximation in the Wasserstein distance with a version of Malliavin calculus on the Poisson space.

Hida [12] has introduced the theory of generalized functions on infinite dimensional space in terms of Gaussian white noise. While we are studying such an analysis, we take a white noise, the time derivative $\dot{B}(t)$ of a Brownian motion, $B(t)$, to be the argument of the functional in question, instead of a Brownian motion itself. The advantage is that the $\{\dot{B}(t) : t \in \mathbb{R}\}$ forms a system of i.i.d. (independent identically distributed) random variables since $B(t)$ has independent increments. The i.i.d. property of the variables makes the analysis simpler. He considered $\{\dot{B}(t) : t \in \mathbb{R}\}$ as coordinate system. $\langle \dot{B}(t), \xi \rangle$ is well defined and should be understood as $\langle \dot{B}(t), \xi \rangle = -\langle B(t), \xi' \rangle$. Hida worked on generalized Brownian (or white noise) functionals $f(\dot{B}(t) : t \in \mathbb{R})$. White noise derivative is a directional derivative.
Moreover, the generalised function, the Dirac delta function, is one of the subjects in quantum mechanics [10]. The space of generalized white noise functionals is much bigger than the classical $L^2$-space of functionals of Brownian motion (Malliavin calculus is only defined in real Hilbert space). Therefore, we can carry on the calculus on a wider class of random complex systems (white noise calculus can work in complex space). Therefore, we can say Malliavin calculus is a special case of white noise calculus. Itô calculus is a special case of Malliavin calculus. If the stochastic process is adapted, Malliavin calculus coincides with the Itô calculus.

Y. Ito [18] constructed the Poisson counterpart of Hida’s theory. In the Poisson case, the differential operator is actually a difference operator. Lee and Shih [24] develop the Lévy white noise analysis for the general case of Lévy processes by means of the Segal-Bargmann transform (S-transform). They also introduce a differential, adjoint, N operators and derivatives of analytic functionals. In this thesis, we use these operators to derive our theorem for Stein’s method and the normal approximation of Lévy functionals.

1.2 Financial option pricing approximation by Stein’s method and Malliavin calculus

The second question treated in this thesis is to use Stein’s method and Malliavin calculus to construct approximations for financial spread option prices whose underlying asset has stochastic volatility.

Since awarding Harry Markowitz, William Shape, and Merton Miller the 1990 Nobel Prize in Economics, the theory of finance has become increasingly mathematical, to the point that problems in finance are now driving research in mathematics. Harry Markowitz’s 1952 Ph.D. thesis *Portfolio Selection* laid the groundwork for the mathematical theory of finance. He quantified the concept of diversification in a market using mean return and covariances for stocks. He argued that we should hold portfolios whose variance is minimal among all portfolios with a given mean return.

In 1969, Robert Merton, motivated by the desire to understand how prices are set in financial markets, introduced stochastic calculus into finance. At the same time as Merton’s work, Fischer Black and Myron Scholes were developing the celebrated option pricing formula, which won the 1997 Nobel Prize in Economics. The Black-Scholes option pricing formula provides solutions for both the pricing and hedging problems in
financial markets. In the period 1979-1983, Harrison, Kreps, and Pliska used the general theory of continuous-time stochastic processes to put the Black-Scholes formula on a solid basis. We will review applications of Malliavin calculus in finance later. Nowadays, besides stock options, both academic and industry apply financial mathematics to all asset types in financial markets, including interest rates derivatives, mortgage-backed security (MBS), collateralized debt obligation pricing (CDO), and credit default swap (CDS) [1] [15].

In financial mathematics, a financial European option price at time 0 could be calculated by risk-neutral pricing formula, which is the discounted expected payoff. $F$ is the underlying asset, and is a stochastic process. The option price at time 0 is

$$V(0) = E(e^{-rT}\Phi(F)) = E(h(F)),$$

where the payoff function, $\Phi(F)$, could be $(F - K)^+$ or $(K - F)^+$, and $r$ is a constant risk-free interest rate. $T$ is time to maturity. $K$ is the strike price.

The Malliavin calculus (also known as the stochastic calculus of variations) was first introduced by Paul Malliavin in 1978. Traditionally, it is an infinite-dimensional differential calculus on the Wiener space. The purpose of this calculus was to investigate regularity properties of the law of Wiener functionals such as solutions of stochastic differential equations driven by Brownian motion. Besides the study of the regularity of probability laws, other applications of the Malliavin calculus in financial mathematics have emerged. In 1991, Ocone and Karatzas [34] proved that the Clark-Ocone formula can be used to obtain an explicit formula for replicating portfolios of contingent claims in complete markets. The next pioneering paper came in 1999. Fournié et al. [7] presented an original probabilistic method for the numerical computations of Greeks (i.e. price sensitivities) in finance. In the recent years, many new and interesting applications of the Malliavin calculus have been discovered. The anticipative stochastic calculus allows us to discuss stochastic differential equations where the solution is not adapted to the Brownian filtration. A stochastic calculus with respect to fractional Brownian motion is also developed by using the Malliavin calculus. More applications of the Malliavin calculus in mathematical finance are pricing and hedging American options (an optimal stopping time problem), insider trading and partial information optimal control. The Malliavin calculus and its applications are discussed in detail by D. Nualart [31] and P. Malliavin and A. Thalmaier [28].
The Malliavin calculus is also extended from the original setting of Brownian motion to more general Lévy processes. There are some applications in finance where Lévy processes based models are used. The Malliavin calculus for Lévy processes is discussed by G. Di Nunno, B. Øksendal and F. Proske [32].

There are several ways to define the Malliavin derivative. One way to present the Malliavin derivative of a random variable $F = F(\omega)$, $\omega \in \Omega$, on a probability space $(\Omega, \mathcal{F}, P)$ is to regard it as a derivative with respect to the random parameter $\omega$. Traditionally, for the Brownian motion case, $\Omega$ is represented as the Wiener space $C_0([0, T])$ of continuous functions $\omega : [0, T] \to \mathbb{R}$ with $\omega(0) = 0$, equipped with the uniform topology. Alternatively, we can also use the Wiener-Itô chaos expansion [16] to introduce the Malliavin derivative. The advantage of the chaos expansion approach is that it gives us a relatively unified approach for both Brownian motion and Lévy processes/Poisson random measures.

In Chen and Shao [4], by comparing expectations, Stein’s method provides explicit estimates for normal approximation. The Wasserstein distance, $d_W(F, X)$, between the law of $X$ and the law of $F$ is defined as follows:

$$d_W(F, X) = \sup_{h \in \text{Lip}(1)} |Eh(F) - Eh(X)|,$$

where $\text{Lip}(1)$ denotes the class of real-valued Lipschitz functions, from $\mathbb{R}$ to $\mathbb{R}$, with Lipschitz constant less or equal to one. Therefore, we can use the Stein-Malliavin method to construct financial option price approximations. In Theorem 4.2.1, we have a closed form solution for financial options which has the underlying asset following a normal distribution. However, when the underlying asset follows a stochastic volatility process (4.15), we may not have a convenient closed form solution. By comparing these two different expectations, we can calculate the upper bound of Wasserstein distance to quantify the accuracy of the approximation. The main results are Theorem 4.4.2 and Corollary 4.4.1.
Chapter 2

Review of relevant literature

2.1 Introduction

In this chapter, first, we review some preliminary material in real analysis and probability. Then we review Hida’s work to pave the way for white noise calculus in the next chapter.

2.2 Basic probability concepts

According to [6], consider a set $\Omega$, called the set of scenarios, equipped with a $\sigma$-algebra $\mathcal{F}$. In a financial modelling context, $\Omega$ will represent the different scenarios which can occur in the market, each scenario $\omega \in \Omega$ being described in terms of the evolution of prices of different instruments. A probability measure on $(\Omega, \mathcal{F})$ is a positive finite measure $P$ with total mass 1. $(\Omega, \mathcal{F}, P)$ is then called a probability space. A measurable set $A \in \mathcal{F}$, called an event, is therefore a set of scenarios to which a probability can be assigned. A probability measure assigns a probability between 0 and 1 to each event.

Based on [39], we provide the definition of a random variable.

**Definition 2.2.1.** Let $(\Omega, \mathcal{F}, P)$ be a probability space. A random variable is a real-valued function $X$ defined on $\Omega$ with the property that for every Borel subset $B$ of $\mathbb{R}$, the
subset of $\Omega$ given by

$$\{X \in B\} = \{\omega \in \Omega; X(\omega) \in B\}$$  \hspace{1cm} (2.1)

is in the $\sigma$-algebra $\mathcal{F}$.

According to [37], we provide the definition of a measure space.

**Definition 2.2.2.** Let $(X, \mathcal{M}, \mu)$ be a measure space. The measure $\mu$ is called finite provided $\mu(X) < \infty$. It is called $\sigma$-finite provided $X$ is the union of a countable collection of measurable sets, each of which has finite measure. A measurable set $E$ is said to be of finite measure provided $\mu(E) < \infty$, and said to be $\sigma$-finite provided $E$ is the union of a countable collection of measurable sets, each of which has finite measure.

Based on [20], the definition of a stochastic process has been provided.

**Definition 2.2.3.** A stochastic process is a collection of random variables

$$\{X_t\}_{t \in [0,T]}$$  \hspace{1cm} (2.2)

defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assuming values in $\mathbb{R}^n$.

According to [20], we define a filtration below.

**Definition 2.2.4.** Let $\mathcal{F}$ be a $\sigma$-algebra of subsets of a given set $\Omega$. A family $\{\mathcal{F}_t, t \geq 0\}$ of sub-$\sigma$-algebras of $\mathcal{F}$ is called a filtration if

$$\mathcal{F}_s \subseteq \mathcal{F}_t \quad \text{whenever } s \leq t.$$  \hspace{1cm} (2.3)

Now let $X = \{X_t\}_{t \in [0,T]}$ be a stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We say that it is adapted to the filtration if $X_t$ is $\mathcal{F}_t$-measurable for each $t \geq 0$.

According to [6], it will be assumed that all random processes in the thesis satisfy the c\text{ad}c\text{ag} property defined below.
Definition 2.2.5. A function \( f : [0, T] \to \mathbb{R}^d \) is said to be cadlag if it is right-continuous with left limits: for each \( t \in [0, T] \) the limits

\[
    f(t-) = \lim_{s \to t, s < t} f(s) \quad f(t+) = \lim_{s \to t, s > t} f(s)
\]

exist and \( f(t) = f(t+) \).

Note that a right-continuous process is a cadlag process, which can have discontinuities. If \( t \) is a discontinuity point we denote by

\[
    \triangle f(t) = f(t) - f(t-),
\]

the jump of \( f \) at \( t \).

Based on [6], additive processes are obtained from Lévy processes by relaxing the condition of stationarity of increments.

Definition 2.2.6. A stochastic process \( \{X_t\}_{t \geq 0} \) on \( \mathbb{R}^d \) is called an additive process if it is cadlag, satisfies \( X_0 = 0 \) and has the following properties:

(i) Independent increments: for every increasing sequence of times \( t_0, \ldots, t_n \), the random variables \( X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}} \) are independent.

(ii) Stochastic continuity: \( \forall \varepsilon > 0, \mathbb{P}(|X_{t+h} - X_t| \geq \varepsilon) \to 0 \) as \( h \to 0 \).

Based on [6], we define a martingale and a Markov property.

Definition 2.2.7. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and \( X_t \) be a stochastic process adapted to a filtration \( \{\mathcal{F}_t\} \) and suppose \( E|X_t| < \infty \) for all \( t \in [0, T] \). Then \( X_t \) is called a martingale with respect to \( \{\mathcal{F}_t\} \) if for any \( s \leq t \in [0, T] \),

\[
    E[X_t | \mathcal{F}_s] = X_s, \quad \text{a.s. (almost surely)}.
\]
Definition 2.2.8. A stochastic process $X_t$, $a \leq t \leq b$, is said to satisfy the Markov property if for any $a \leq t_1 < t_2 < \cdots < t_n < t \leq b$, the equality

$$P(X_t \leq x|X_{t_1}, X_{t_2}, \ldots, X_{t_n}) = P(X_t \leq x|X_{t_n})$$  \hspace{1cm} \text{(2.7)}$$

holds for any $x \in \mathbb{R}$, or equivalently, the equality

$$P(X_t \leq x|X_{t_i} = y_i, i = 1, 2, \ldots, n) = P(X_t \leq x|X_{t_n} = y_n)$$  \hspace{1cm} \text{(2.8)}$$

holds for any $x, y_i \in \mathbb{R}$. By a Markov process, we mean a stochastic process $X_t$, $a \leq t \leq b$, satisfying the Markov property.

According to [6], a Wiener process is both a martingale and a Markov process.

Definition 2.2.9. A Wiener process is a stochastic process $\{W_t : t \geq 0\}$, on some $(\Omega, \mathcal{F}, P)$ with these properties:

(i) The process starts at 0: $W_0 = 0$.

(ii) For $0 = t_0 < t_1 < \cdots < t_m$, the increments

$$W_{t_i} = W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \ldots, W_{t_m} - W_{t_{m-1}}$$  \hspace{1cm} \text{(2.9)}$$

are independent.

(iii) Each of these increments is normally distributed with mean 0 and variance $t_{i+1} - t_i$.

According to [37], we provide the following definition.

Definition 2.2.10. Let $X$ be a nonempty set. A topology $T$ for $X$ is a collection of subsets of $X$, called open sets, possessing the following properties:

(i) The entire set $X$ and the empty-set $\emptyset$ are open;

(ii) The intersection of any finite collection of open sets is open;

(iii) The union of any collection of open sets is open.
A nonempty set $X$, together with a topology on $X$, is called a topological space. For a point $x$ in $X$, an open set that contains $x$ is called a neighborhood of $x$. We may denote a topological space by $(X,T)$.

According to [37], we provide the following definition.

**Definition 2.2.11.** A topological space $(X,T)$ is called Hausdorff, or a Hausdorff space, iff for every two distinct points $x$ and $y$ in $X$, there are open sets $U$ and $V$ with $x \in U$, $y \in V$, and $U \cap V = \emptyset$.

According to [37], we provide the following definition.

**Definition 2.2.12.** Let $E$ be a vector space. A seminorm on $E$ is a function $\| \cdot \|$ from $E$ into $[0,\infty)$ such that

(i) $\|cx\| = |c|\|x\|$ for all $c \in \mathbb{R}$ and $x \in E$, and

(ii) $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in E$.

A seminorm $\| \cdot \|$ is called a norm iff $\|x\| = 0$ only for $x = 0$.

Based on [33], the following definition has been provided.

**Definition 2.2.13.** A topological vector space $E$ is called locally convex if the topology of $E$ is Hausdorff and given by a family of seminorms $\{\| \cdot \|_\alpha\}_{\alpha \in A}$.

Based on [33], the following definition has been provided.

**Definition 2.2.14.** A seminorm $\| \cdot \|$ on a vector space $G$ over $\mathbb{R}$ is called Hilbertian if it is induced by some non-negative, symmetric bilinear form $(\cdot,\cdot)$ on $G \times G$, namely if $\|\xi\|^2 = (\xi,\xi)$ for all $\xi \in G$. Here it is not assumed that $(\xi,\xi) = 0$ implies $\xi = 0$.

According to [41], we provide the following definition.

**Definition 2.2.15.** Let $E, F$ be Hilbert spaces. A bounded linear operator $A : E \rightarrow F$ is said to be Hilbert-Schmidt if there exists a complete orthonormal sequence $(e_n)_{n=1}^\infty$ in $E$ such that

$$
\sum_{n=1}^\infty \| Ae_n \|^2 < \infty. \quad (2.10)
$$
Definition 2.2.16. If $A \subset B$, then the canonical map is the function

$$
\pi : A \to B, \pi(x) = x.
$$

(2.11)

Based on [32], we give the following definition of dual space.

Definition 2.2.17. Let $X$ be a Banach space, that is, complete, normed vector space over $\mathbb{R}$, and let $\|x\|$ denote the norm of the element $x \in X$. A linear functional on $X$ is a linear map

$$
T : X \to \mathbb{R}.
$$

Recall that $T$ is called linear if

$$
T(ax + y) = aT(x) + T(y)
$$

for all $a \in \mathbb{R}, x, y \in X$.

A linear functional $T$ is called bounded if

$$
|||T||| := \sup_{\|x\| \leq 1} |T(x)| < \infty.
$$

The set of all bounded linear functionals is called the dual of $X$ and is denoted by $X^*$. Equipped with the norm $|| \cdot ||$, the space $X^*$ is a Banach space.

The pairing of a functional $T$ in the dual space $X^*$ and an element $x$ of $X$ is denoted by a bracket:

$$
\langle T, x \rangle
$$

(2.12)

which is called "the action of $T$ on $x". The pairing defines a nondegenerate bilinear mapping

$$
\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{R}.
$$

(2.13)
Chapter 2. Review of relevant literature

**Definition 2.2.18.** A locally convex space $G$ equipped with defining Hilbertian semi-norms $\{\|\cdot\|_n, n \geq 0\}$ is called **nuclear** if for any $n$ there is $m$ with $n \leq m$ such that the canonical map $h_{n,m} : G_m \to G_n$ is of Hilbert-Schmidt type.

Based on [13] [14], the definition of characteristic functional is given.

**Definition 2.2.19.** A functional $C(\xi)$ on a nuclear space $G$, $C : G \to \mathbb{C}$, is called a **characteristic functional** if it satisfies the following conditions:

1. $C(0) = 1$;
2. $C(\xi)$ is positive definite, i.e., for any $n$, $z_j \in \mathbb{C}$, $\xi_j \in G$, $1 \leq j, k \leq n$,
   $$\sum_{j,k=1}^{n} z_j \bar{z}_k C(\xi_j - \xi_k) \geq 0.$$
3. $C(\xi)$ is continuous on $G$.

From [14], the following definition of Schwartz space $\mathcal{S}(\mathbb{R})$ is provided.

**Definition 2.2.20.** The **Schwartz space** $\mathcal{S}(\mathbb{R})$ is the linear space of all functions $f : \mathbb{R} \to \mathbb{C}$ which have derivatives of all orders and which satisfy the condition

$$p_{a,b}(f) \overset{\text{def}}{=} \sup_{x \in \mathbb{R}} |x^a f^{(b)}(x)| < \infty$$

for all $a, b \in \{0, 1, 2, \ldots\}$. The finiteness condition for $a \geq 1$ and $b \in \{0, 1, 2, \ldots\}$ implies that $x^a f^{(b)}(x)$ goes to 0 as $|x| \to \infty$, for all $a, b \in \{0, 1, 2, \ldots\}$, and so functions of this type are said to be rapidly decreasing.

Here, the space $G$ is the Schwartz space of test functions. Schwartz space is a typical example of a nuclear space. The dual space $G^*$ is the space of tempered distributions or the space of Schwartz distributions.

Let $G_{-n}$ be the dual space of $G_n$. Then, we have

$$\cdots \subset G_{n+1} \subset G_n \subset \cdots \subset L^2(\mathbb{R}) \subset \cdots \subset G_{-n} \subset G_{-n-1} \subset \cdots$$
Set $G = \bigcap G_n$. The space $G$ is called a nuclear space.
For examples, please refer to the section 3.2.2.1.

According to [13], and [14], we have the Bochner-Minlos Theorem.

**Theorem 2.2.1.** (Bochner-Minlos) Let $G$ denote a nuclear space with dual space $G^*$. Let a complex-valued function $\mathcal{C}(\xi)$ on $G$ be a characteristic functional. Then there exists a probability measure $\nu$ on the measurable space $(G^*, \mathcal{B})$ such that

$$\mathcal{C}(\xi) = \int_{G^*} e^{i\langle x, \xi \rangle} d\nu(x), \quad \xi \in G,$$

where $\langle x, \xi \rangle$ is the $G^* - G$ pairing, (2.12) and (2.13). Such a measure is unique.

The probability measure space $(G^*, \mathcal{B}, \nu)$ is called a generalized stochastic process with the characteristic functional $\mathcal{C}(\xi)$.

We will introduce some definitions which are useful in white noise analysis. For more details, refer to [12], [13], [14] and [33]

According to [32], we provide the following three definitions.

**Definition 2.2.21.** A real function is called symmetric if

$$g(s_{\sigma_1}, \cdots, s_{\sigma_n}) = g(s_1, \cdots, s_n) \quad (2.14)$$

for all permutations $\sigma = (\sigma_1, \cdots, \sigma_n)$ of $(1, 2, \cdots, n)$

**Definition 2.2.22.** The symmetrization $\tilde{f}$ is defined by

$$\tilde{f}(s_1, \cdots, s_n) = \frac{1}{n!} \sum_{\sigma} f(s_{\sigma_1}, \cdots, s_{\sigma_n}) \quad (2.15)$$

where the sum is taken over all permutation $\sigma$ of $(1, \cdots, n)$.

We note that $\tilde{f} = f$ if $f$ is symmetric.
**Definition 2.2.23.** The tensor product \( f \otimes g \) of two functions \( f, g \) is defined by

\[
(f \otimes g)(x_1, x_2) = f(x_1)g(x_2)
\]

and the symmetrized tensor product \( \hat{f} \otimes g \) is the symmetrization of \( f \otimes g \).

The \( L^p \)-spaces are useful and important examples of Banach spaces. From [37], \( L^p \) spaces are defined below. For definiteness, we consider real-valued functions. Analogous results apply to complex-valued functions.

**Definition 2.2.24.** Let \((T, \mathcal{A}, \mu)\) be a measure space and \(1 \leq p \leq \infty\). The space \( L^p(T) \) consists of equivalence classes of measurable functions \( f : T \to \mathbb{R} \) such that

\[
\int |f|^p d\mu < \infty,
\]

where two measurable functions are equivalent if they are equal \( \mu \)-a.e. The \( L^p \)-norm of \( f \in L^p(T) \) is defined by

\[
\|f\|_{L^p} = \left( \int |f|^p d\mu \right)^{1/p}.
\]

The notation \( L^p(T) \) assumes that the measure \( \mu \) on \( T \) is understood. We say that \( f_n \to f \) in \( L^p \) if \( \|f - f_n\|_{L^p} \to 0 \).

Due to [21], the definition of Poisson random measure is given below. In the definition, \( \mathcal{S} \), for instance, could be \( \mathcal{B}[0, \infty) \times \mathcal{B}(\mathbb{R} \setminus \{0\}) \).

**Definition 2.2.25.** We assume that \((S, \mathcal{S}, \eta)\) is an arbitrary \( \sigma \)-finite measure space. Let \( N : S \to \{0, 1, 2, \ldots \} \cup \{\infty\} \) in such a way that the family \( \{N(A) : A \in \mathcal{S}\} \) are random variables defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Then \( N \) is called a Poisson random measure on \((S, \mathcal{S}, \eta)\) (or sometimes a Poisson random measure on \( S \) with intensity \( \eta \)) if

(i) for mutually disjoint \( A_1, \ldots, A_n \) in \( \mathcal{S} \), the variables \( N(A_1), \ldots, N(A_n) \) are independent,

(ii) for each \( A \in \mathcal{S} \), \( N(A) \) is Poisson distributed with parameter \( \eta(A) \) (here we allow
Chapter 2. Review of relevant literature

According to [20], we now describe our class of functions for which the Itô integral will be defined:

Definition 2.2.26. Let \( \nu = \nu(S, T) \) be the class of functions

\[
f(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}
\]

such that

(i) \( (t, \omega) \rightarrow f(t, \omega) \) is \( B \times F \)-measurable, where \( B \) denotes the Borel \( \sigma \)-algebra on \([0, \infty)\).

(ii) \( f(t, \omega) \) is \( F_t \)-adapted.

(iii) \( E[\int_S^T f(t, \omega)^2 dt] < \infty \).

A function \( \phi \in \nu \) is called elementary if it has the form

\[
\phi(t, \omega) = \sum_j e_j(\omega) \cdot 1_{[t_j, t_{j+1})}(t).
\]

Note that since \( \phi \in \nu \) each function \( e_j \) must be \( F_{t_j} \)-measurable.

Definition 2.2.27. Let \( f \in \nu(S, T) \), and \( f(t, \omega) \) is \( F_t \)-adapted. \( B_t \) denotes standard Brownian motions. Then the Itô integral of \( f \) (from \( S \) to \( T \)) is defined by

\[
\int_S^T f(t, \omega) dB_t(\omega) = \lim_{n \rightarrow \infty} \int_S^T \phi_n(t, \omega) dB_t(\omega) \quad \text{(limit in } L^2(\mathbb{P}))
\]

where \( \{\phi_n\} \) is a sequence of elementary functions such that

\[
E\left[\int_S^T (f(t, \omega) - \phi_n(t, \omega))^2 dt\right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.
\]

Itô integral has the following properties:
Theorem 2.2.2. Let $f, g \in \nu(0, T)$ and let $0 \leq S < U < T$. Then

(i) we have
\[ \int_S^T f dB_t = \int_S^U f dB_t + \int_U^T f dB_t, \tag{2.20} \]

(ii) for any constant $c \in \mathbb{R}$,
\[ \int_S^T (cf + g) dB_t = c \int_S^T f dB_t + \int_S^T g dB_t, \tag{2.21} \]

(iii) we have
\[ E\left[ \int_S^T f dB_t \right] = 0, \tag{2.22} \]

(iv) the integral
\[ \int_S^T f dB_t \text{ is } \mathcal{F}_T \text{ measurable}, \tag{2.23} \]

(v) we have Itô isometry
\[ E \left[ \left( \int_S^T f(t, \omega) dB_t \right)^2 \right] = E \left[ \int_S^T f^2(t, \omega) \right] \text{ for all } f \in \nu(S, T). \tag{2.24} \]

2.3 Hida white noise analysis

T. Hida [12] introduced Gaussian white noise analysis in 1975 in his notes. The idea of white noise calculus is to switch from a functional of Brownian motion $g(B(t) : t \in \mathbb{R})$ to one of white noise $f(\dot{B}(t) : t \in \mathbb{R})$, where $\dot{B}(t)$ is a time derivative of a Brownian motion $B(t)$. This derivative does not exist in the usual sense, but can be defined via integration by parts (2.26). $\dot{B}(t)$ could be thought as a generalized stochastic process while $f$ is realized as a generalized white noise functional. We may consider \{\dot{B}(t) : t \in \mathbb{R}\} as a collection of infinitely many independent random variables and therefore a coordinate system in an infinite dimensional space.
Start with a Gelfand triple,
\[ G \subset L^2(\mathbb{R}) \subset G^*, \]  
(2.25)

where \( G \) is a nuclear space (Definition 2.2.18) dense in \( L^2(\mathbb{R}) \) and where \( G^* \) is the dual space (Definition 2.2.17) for \( G \). Usually we take the nuclear space \( G \) to be the Schwartz space (Definition 2.2.20).

The space \( G \) is viewed as a space of test functions on \( \mathbb{R} \) while the space \( G^* \) is considered as a space of generalised functions.

The canonical bilinear form connecting \( G \) and \( G^* \) could be denoted by
\[ \langle x, \xi \rangle, \]
where \( x \in G^*, \xi \in G \); for more details, refer to equation (2.12) and (2.13).

A sample path of \( \dot{B}(t) \) is a member of \( G^* \). Therefore, \( \langle \dot{B}(t), \xi \rangle \) is well defined and should be understood as
\[ \langle \dot{B}(t), \xi \rangle = -\langle B(t), \xi' \rangle, \]
that is
\[ \int \dot{B}(t)\xi dt = -\int B(t)\xi' dt, \]  
(2.26)

where \( \dot{B}(t) \) is a first time derivative of a Brownian motion \( B(t) \), and \( \xi' \) is a first time derivative of \( \xi \). The equation above is integration by parts.

### 2.4 The Malliavin calculus

In this section, we review the Wiener-Itô chaos expansion and some Malliavin calculus operators, which are useful tools in this thesis. For a full introduction to the Malliavin calculus and its application in finance, we refer the reader to D. Nualart [31], G. Di Nunno, B. Øksendal and F. Proske [32], and P. Malliavin and A. Thalmaier [28].
2.4.1 The Wiener-Itô chaos expansion

In Stochastic analysis, the celebrated Wiener-Itô chaos expansion is fundamental and plays a crucial role in the Malliavin calculus as it is presented in this chapter. The representation of square integrable random variables in terms of an infinite orthogonal sum was proved in its first version by Wiener in 1938. In the Wiener space setting, Itô [16], in 1951, showed that the expansion could be expressed in terms of iterated Itô integrals. Based on the book by G. Di Nunno, B. Øksendal and F. Proske [32], before we introduce iterated Itô integrals and the Wiener-Itô chaos expansion, we also give some useful notation in this section.

2.4.1.1 Iterated Itô integrals

Let \( W = W(t) = W(\omega, t), \omega \in \Omega, t \in [0, T] (T > 0) \), be a one-dimensional Wiener process, on the complete probability space \((\Omega, \mathcal{F}, P)\) such that \( W(0) = 0 \) P-a.s. Let \( \mathcal{F}_t \) be the \( \sigma \)-algebra generated by \( W(s), 0 \leq s \leq t \), augmented by all the P-zero measure events. We denote the corresponding filtration by

\[
\mathbb{F} = \{ \mathcal{F}_t, \quad t \in [0, T] \}. \tag{2.27}
\]

According to [32], we summarize some definitions and notation.

Let \( \overline{L}^2([0, T]^n, \mathcal{Z}^n, \lambda^n) = L^2([0, T]^n) \) be the standard space of square integrable Borel real functions on \([0, T]^n\), where \( \mathcal{Z}^n \) is \( \sigma \)-algebra, and \( \lambda^n \) is Lebesgue measure, such that

\[
\|g\|_{\overline{L}^2([0, T]^n)}^2 := \int_{[0,T]^n} g^2(t_1, \ldots, t_n) \, dt_1 \cdots dt_n < \infty. \tag{2.28}
\]

Let \( \widetilde{L}^2([0, T]^n) \subset L^2([0, T]^n) \) be the space of symmetric square integrable Borel real functions on \([0, T]^n\). Let us consider the set

\[
S_n = \{(t_1, \ldots, t_n) \in [0, T]^n : \quad 0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq T \}. \tag{2.29}
\]

The set \( S_n \) occupies the fraction \( \frac{1}{n!} \) of the whole n-dimensional box \([0, T]^n\). Therefore, if \( g \in \widetilde{L}^2([0, T]^n) \), then
Chapter 2. Review of relevant literature

\[ \left\| g \right\|_{L^2([0,T]^n)}^2 = n! \int_{S^n} g^2(t_1, \ldots, t_n) dt_1 \cdots dt_n, = n! \left\| g \right\|_{L^2(S^n)}^2, \quad (2.30) \]

where \( \| \cdot \|_{L^2(S^n)} \) denotes the norm induced by \( L^2([0,T]^n) \) on \( L^2(S^n) \), the space of square integrable function on \( S^n \).

**Definition 2.4.1.** Let \( f \) be a deterministic function defined on \( S^n (n \geq 1) \) such that

\[ \left\| f \right\|_{L^2(S^n)}^2 := \int_{S^n} f^2(t_1, \ldots, t_n) dt_1 \cdots dt_n < \infty. \quad (2.31) \]

Then we define the \( n \)-fold iterated Itô integral as

\[ J_n(f) := \int_0^T \int_0^{t_n} \cdots \int_0^{t_1} f(t_1, \ldots, t_n) dW(t_1) dW(t_2) \cdots dW(t_{n-1}) dW(t_n). \quad (2.32) \]

In the equation (2.32), at each iteration \( i = 1, \ldots, n \), the corresponding Itô integral with respect to \( dW(t_i) \) is well defined. The integrand

\[ \int_0^{t_i} \cdots \int_0^{t_3} \int_0^{t_2} f(t_1, \ldots, t_n) dW(t_1) dW(t_2) \cdots dW(t_{i-1}) \]

is \( \mathcal{F} \)-adapted and square integrable with respect to \( dP \times dt_i \). \( t_i \in [0, t_{i+1}] \).

According to the construction of the Itô integral, we see that \( J_n(f) \) belongs to \( L^2(P) \), that is, the space of square integrable random variable. We denote the norm of \( X \in L^2(P) \) by

\[ \left\| X \right\|_{L^2(P)} := \left( E[X^2] \right)^{1/2} = \left( \int_{\Omega} X^2(\omega) P(d\omega) \right)^{1/2}. \]

Applying the Itô isometry iteratively, we summarize the results as follows. The proof is given in section 1.1 in G. Di Nunno, B. Øksendal and F. Proske [32].

**Proposition 2.4.1.** The following relation hold true:

\[ E[J_m(g)J_n(h)] = \begin{cases} 0, & n \neq m \\ (g, h)_{L^2(S^n)}, & n = m \end{cases}, \quad (m, n = 1, 2, \ldots) \]
where

\[(g, h)_{L^2(S_n)} := \int_{S_n} g(t_1, \ldots, t_n)h(t_1, \ldots, t_n)dt_1 \cdots dt_n\]

is the inner product of \(L^2(S_n)\).

We can define \(J_0(g) = g\), for \(n = 0\) or \(m = 0\), and \((g, h)_{L^2(S_0)} = gh\).

**Definition 2.4.2.** Let \(g \in \tilde{L}^2([0, T]^n)\). We define

\[I_n(g) := \int_{[0, T]^n} g(t_1, \ldots, t_n)dW(t_1) \cdots dW(t_n) \quad (2.33)\]

\[= n!J_n(g). \quad (2.34)\]

Line (2.33) is a special case of line (3.15).

Using Proposition 2.4.1 and Definition 2.4.2, we get

\[\|I_n(g)\|^2_{L^2(P)} = E[I_n^2(g)] = E[(n!)^2J_n^2(g)] = (n!)^2\|g\|^2_{L^2(S_n)} = n!\|g\|^2_{L^2([0, T]^n)}\]

for all \(g \in \tilde{L}^2([0, T]^n)\).

According to Proposition 2.4.1 and (2.34), the following result holds.

**Proposition 2.4.2.** if \(g \in \tilde{L}^2([0, T]^M)\) and \(h \in \tilde{L}^2([0, T]^n)\), we have

\[E[I_m(g)I_n(h)] = \begin{cases} 0, & n \neq m \\ n!(g, h)_{L^2([0, T]^n)}, & n = m \end{cases} \quad (m, n = 1, 2, \ldots)\]

where \((g, h)_{L^2([0, T]^n)} = n!(g, h)_{L^2(S_n)}\).

### 2.4.1.2 The Wiener-Itô chaos expansion

Based on [32], we give the Wiener-Itô chaos expansion below.
Theorem 2.4.1. (The Wiener-Itô chaos expansion) Let $\xi$ be an $\mathcal{F}_T$-measurable random variable in $L^2(P)$. Then there exists a unique sequence $\{f_n\}_{n=0}^{\infty}$ of functions $f_n \in \tilde{L}^2([0,T]^n)$ such that

$$\xi = \sum_{n=0}^{\infty} I_n(f_n)$$

(2.35)

where the convergence of the RHS of (2.35) to $\xi$ is in $L^2(P)$. We define $I_0(f_0) = f_0$.

Moreover, we have the isometry

$$\|\xi\|^2_{L^2(P)} = \sum_{n=0}^{\infty} n! \|f_n\|^2_{L^2([0,T]^n)}.$$  

(2.36)

We will represent Skorohod integral, Malliavin derivative, and Ornstein-Uhlenbeck generator $L$ in terms of the Wiener-Itô chaos expansion later. We can use chaos expansion to build Malliavin calculus.

### 2.4.2 The Skorohod integral via chaos expansion

The Wiener-Itô chaos expansion introduced in the last sub-section is a good starting point for studying Malliavin calculus. In this sub-section, we focus on the Skorohod integral. According to G. Di Nunno, B. Øksendal and F. Proske [32], the Skorohod integral, introduced for the first time by A. Skorohod in 1975, is an extension of the Itô integral to integrands that are not necessarily $\mathcal{F}$-adapted. Moreover, if the integrand $u$ is $\mathcal{F}$-adapted, then the two integrals coincide as elements of $L^2(P)$. The Skorohod integral is also the adjoint operator of the Malliavin derivative, which is introduced in the next sub-section.

According to [32], let $u = u(t,\omega)$, $t \in [0,T]$, $\omega \in \Omega$, be a measurable stochastic process such that, for all $t \in [0,T]$, $u(t)$ is a $\mathcal{F}_T$-measurable random variable and

$$E[u^2(t)] < \infty.$$  

For each $t \in [0,T]$, we can apply the Wiener-Itô chaos expansion to the random variable $u = u(t,\omega)$, $\omega \in \Omega$, and thus there exist symmetric functions $f_{n,t} = f_{n,t}(t_1,\ldots,t_n)$,
$(t_1, \ldots, t_n) \in [0, T]^n$, in $\tilde{L}^2([0, T]^n), n = 1, 2, \ldots$, such that

$$u(t) = \sum_{n=0}^{\infty} I_n(f_{n,t}).$$

The functions $f_{n,t}, n = 1, 2, \ldots$, depend on the parameter $t \in [0, T]$, and we can write

$$f_n(t_1, \ldots, t_n, t) := f_{n,t}(t_1, \ldots, t_n).$$

We can regard $f_n$ as a function of $n + 1$ variables. This function is symmetric with respect to its first $n$ variables, its symmetrization $\tilde{f}_n$ is given by

$$\tilde{f}_n(t_1, \ldots, t_n, t_{n+1}) = \frac{1}{n+1} \left[ f_n(t_1, \ldots, t_n, t_{n+1}) + f_n(t_2, \ldots, t_{n+1}, t_1) + \cdots + f_n(t_1, \ldots, t_{n+1}, t_n) \right]. \quad (2.37)$$

Based on [32], we give the definition of the Skorohod integral.

**Definition 2.4.3.** Let $u(t), t \in [0, T]$, be a measurable stochastic process such that for all $t \in [0, T]$ the random variable $u(t)$ is $\mathcal{F}_T$-measurable and $\mathbb{E}[\int_0^T u^2(t)dt] < \infty$. Let its Wiener-Itô chaos expansion be

$$u(t) = \sum_{n=0}^{\infty} I_n(f_{n,t}) = \sum_{n=0}^{\infty} I_n(f_{n}(\cdot, t)).$$

Then we define the Skorohod integral of $u$ by

$$\delta(u) := \int_0^T u(t) \delta W(t) := \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \quad (2.38)$$

when convergent in $L^2(P)$. Here $\tilde{f}_n, n = 1, 2, \ldots$, are the symmetric functions (2.37) derived from $f_{n}(\cdot, t), n = 1, 2, \ldots$. We say that $u$ is Skorohod integrable, and we write $u \in \text{Dom}(\delta)$ if the series in (2.38) converges in $L^2(P)$.
According to (2.36), a stochastic process $u$ belongs to $\text{Dom}(\delta)$ if and only if

$$E[\delta(u)^2] = \sum_{n=0}^{\infty} (n+1)! \|\tilde{f}_n\|_{L^2([0,T]^{n+1})}^2 < \infty.$$ 

**Remark 2.4.2.** Note that, the integrand, $u(t)$, is a $\mathcal{F}_T$-measurable random variable, so the function is not necessarily Itô integrable. For Itô calculus, the integrand has to be adapted, which is $\mathcal{F}_t$-measurable.

From [32], we provide some basic properties of the Skorohod integral.

**Proposition 2.4.3.** For any $u \in \text{Dom}(\delta)$, the Skorohod integral has zero expectation, that is,

$$E[\delta(u)] = 0.$$  

This is a simple consequence of the fact that Itô integrals and thus iterated Itô integrals have zero expectation.

The Skorohod integral is an extension of the Itô integral.

**Theorem 2.4.3.** Let $u = u(t), t \in [0,T]$, be a measurable $\mathcal{F}$-adapted stochastic process such that

$$E\left[ \int_0^T u^2(t)dt \right] < \infty.$$  

Then $u \in \text{Dom}(\delta)$ and its Skorohod integral coincides with the Itô integral

$$\int_0^T u(t)\delta W(t) = \int_0^T u(t)dW(t).$$

**Example 2.4.1.** Using (2.38), let us verify that

$$\int_0^T W(T)\delta W(t) = W^2(T) - T.$$
Proof. By the Itô lemma, we have

\[
\frac{1}{2}W^2(T) = f(W(T)) - F(W(0)) \\
= \int_0^T f'(W(t))dW(t) + \frac{1}{2}\int_0^T f''(W(t))dt \\
= \int_0^T W(t)dW(t) + \frac{1}{2}T \\
\Rightarrow \int_0^T W(t)dW(t) = \frac{1}{2}(W^2(T) - T) \tag{2.39}
\]

It is obvious that \(W(T)\) is not \(\mathcal{F}_T\)-measurable, so it is not Itô integrable, but

\[u(t) = W(T) = \int_0^T 1dW(t_1)\]

has a chaos expansion with \(f_{0,t} = 0, f_{1,t} = 1,\) and \(f_{n,t} = 0\) for \(n \geq 2\).

\[
\delta(u) = I_2(\tilde{f}_1) = I_2(1) = \int_0^T \int_0^T 1dW(t_1)dW(t_2) \\
= 2 \int_0^T \int_0^{t_2} 1dW(t_1)dW(t_2) \\
= 2 \int_0^T W(t_2)dW(t_2) \\
= W^2(T) - T. \text{ by (2.39)}
\]

We can see that when the integrand, \(W(T)\), is not adapted, we have to use Malliavin calculus to compute. The answer is different from that of adapted integrand, \(W(t)\).

2.4.3 The Malliavin derivative as adjoint operator of the Skorohod integral

There are different ways of introducing the Malliavin derivative. The original construction was given on the Wiener space \(\Omega = C_0([0,T])\) consisting of all continuous functions \(\omega : [0,T] \to \mathbb{R}\) with \(\omega(0) = 0\). This construction is outlined in G. Di Nunno, B. Øksendal and F. Proske [32].
In this chapter, according to [32], we mainly use the approach based on the Wiener-Itô chaos expansion (2.35). It is illuminating to show that the Malliavin derivative is the adjoint operator of the Skorohod integral.

From G. Di Nunno, B. Øksendal and F. Proske [32], we represent the Malliavin derivative in terms of the Wiener-Itô chaos expansion.

Definition 2.4.4. Let \( F \in L^2(P) \) be \( \mathcal{F}_T \)-measurable with chaos expansion

\[
F = \sum_{n=0}^{\infty} I_n(f_n),
\]

where the \( f_n \in \tilde{L}^2([0,T]^n), \ n = 1, 2, \ldots \)

(i) We say that \( F \in \mathbb{D}^{1,2} \) if

\[
\sum_{n=1}^{\infty} n^n! \| f_n \|^2_{L^2([0,T]^n)} < \infty \quad (2.40)
\]

(ii) If \( F \in \mathbb{D}^{1,2} \), we define the Malliavin derivative \( D_tF \) of \( F \) at time \( t \) as the expansion

\[
D_tF = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)), \quad t \in [0,T] \quad (2.41)
\]

We define

\[
I_0(f_1(t)) := f_1(t) \quad (2.42)
\]

where \( I_{n-1}(f_n(\cdot, t)) \) is the \( (n-1) \)-fold iterated integral of \( f_n(t_1, \ldots, t_{n-1}, t) \) with respect to the first \( n-1 \) variables \( t_1, \ldots, t_{n-1} \) and \( t_n = t \) left as parameter.

If (2.40) holds, then

\[
\| DF \|^2_{L^2(P \times \lambda)} = E \left[ \int_0^T (D_tF)^2 dt \right] = \sum_{n=1}^{\infty} \int_0^T n^2(n-1)! \| f_n(\cdot, t) \|^2_{L^2([0,T]^{n-1})} dt = \sum_{n=1}^{\infty} n n! \| f_n(\cdot, t) \|^2_{L^2([0,T]^n)} < \infty,
\]
so $D_t F, t \in [0, T]$ is well defined as an element of $L^2(P \times \lambda)$.

According to [32], we have the following chain rule.

**Theorem 2.4.4.** (Chain rule) Let $G \in D^{1,2}$ and $g \in C^1(\mathbb{R})$ with bounded derivative. Then $g(G) \in D^{1,2}$ and

$$Dg(G) = g'(G)DG.$$  

(2.43)

Here $g'(x) = \frac{d}{dx}g(x)$, and $DG$ denotes the Malliavin derivative of $G$.

**Example 2.4.2.** Set

$$S_T = S_0 \exp \left( \int_0^T (r - \frac{1}{2} \sigma) dt + \int_0^T \sigma dW_t \right) = g(G),$$

(2.44)

where $g(x) = ce^x$, and according to (2.35) $G = f_0 + I_1(f_1)$. $f_0 = \int_0^T (r - \frac{1}{2} \sigma) dt$ and $I_1(f_1) = \int_0^T \sigma dW_t$.

Due to the (2.41) and chain rule (2.43),

$$D_t S_T = Dg(G) = g'(G)DG = S_0 \exp \left( \int_0^T (r - \frac{1}{2} \sigma) dt + \int_0^T \sigma dW_t \right) \sigma = \sigma S_T.$$

**Example 2.4.3.** Set $F = \int_0^T S_t dt$, $\varphi(F) = \frac{1}{F}$

According to the chain rule (2.43),

$$D_t(\varphi(F)) = D_t \left( \frac{1}{\int_0^T S_t dt} \right) = - \frac{1}{(\int_0^T S_t dt)^2} D_t \int_0^T S_t dt.$$

Due to Corollary 3.13. in [32], we have the following corollary.
Chapter 2. Review of relevant literature

Corollary 2.4.1. Let \( u = u(s), \ s \in [0, T], \) be an \( \mathcal{F} \)-adapted stochastic process and assume that \( u(s) \in \mathbb{D}^{1,2} \) for all \( s \). Then

(1) \( D_t u(s), \ s \in [0, T], \) is \( \mathcal{F} \)-adapted for all \( t \);

(2) \( D_t u(s) = 0, \) for \( t > s \).

We cite Theorem 3.14. in [32] which says that the Malliavin derivative is the adjoint operator of the Skorohod integral.

Theorem 2.4.5. (Duality formula) Let \( F \in \mathbb{D}^{1,2} \) be \( \mathcal{F}_T \)-measurable and let \( u \) be a Skorohod integrable stochastic process. Then

\[
E \left[ F \int_0^T u(t) \delta W(t) \right] = E \left[ \int_0^T u(t) D_t F dt \right];
\]

or equivalently,

\[
E(F\delta(u)) = E(<DF,u>_{L^2([0,T])}); \quad (2.45)
\]

or equivalently,

\[
<F,\delta(u)>_{L^2(\Omega)}=<DF,u>_{L^2(\Omega \times [0,T])}. \]

The proof is provided for Theorem 3.14 in [32]. The duality formula is very useful in this thesis when we derive the main results.

From Theorem 3.15. in [32], we have the following proposition.

Proposition 2.4.4. Let \( u(t), \ t \in [0, T], \) be a Skorohod integrable stochastic process and \( F \in \mathbb{D}^{1,2} \) such that \( Fu(t), \ t \in [0, T], \) is Skorohod integrable. Then

\[
\delta(Fu) = F\delta(u) - <DF,u>_{L^2([0,T])}, \quad (2.46)
\]

or equivalently,

\[
\int_0^T Fu(t) \delta W(t) = F \int_0^T u(t) \delta W(t) - \int_0^T u(t) D_t F dt.
\]
2.4.4 The Ornstein-Uhlenbeck operator $L$

According to [29] and [35], we define $L$ and $L^{-1}$ operators.

**Definition 2.4.5.** The operator $L$, known as the Ornstein-Uhlenbeck operator, has the domain

$$\text{Dom}L = \{ F \in L^2(P), F = \sum_{n=0}^{\infty} I_n(f_n) : \sum_{n=1}^{\infty} n^2 n! \|f_n\|^2_{L^2([0,T]^n)} < \infty \},$$

and, for $F = \sum_{n=1}^{\infty} I_n(f_n) \in \text{Dom}L$, $L$ is defined by

$$LF = -\sum_{n=1}^{\infty} n I_n(f_n).$$

Note that $E(LF) = 0$ by definition.

**Definition 2.4.6.** With $E(F) = I_0(f_0) = 0$, the domain of $L^{-1}$, denoted by $L_0^2(P)$, is the space of centered random variables in $L^2(P)$. If $F \in L_0^2(P)$ and $F = \sum_{n=1}^{\infty} I_n(f_n)$, then we define the inverse operator $L^{-1}$ as

$$L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n).$$

(2.47)

and $LL^{-1}F = F$. 

Chapter 3

Stein’s method and normal approximation of Lévy functionals

3.1 Introduction

The principal novel contribution in this chapter is the derivation of the Stein bound for functionals of general Lévy processes; see Theorem 3.3.1. A second contribution is to show that the Stein bounds in the literature for functionals of (i) pure jump Lévy processes, and (ii) the pure Brownian motion Lévy process, are both particular cases of the general bound in the thesis; see Theorem 3.3.1 and Theorem 3.3.2. To derive the Stein bound in this chapter, it is essential to adopt a unified approach to the Brownian part and pure jump part of the Lévy process. In this chapter, this was achieved by making use of the the white noise approach developed by Lee and Shih (2004) [24]. We give Example 3.3.3 which computes the upper bound for a double integral and Example 3.3.1 which computes that for a single integral. We show that our approach unifies the Brownian and pure jump components in a natural way, and simplifies the computation. In new Example 3.3.2, we compute the upper bound for a functional which uses notation specific to white noise analysis.

In this chapter, first, we summarize the basic setting of the paper by Lee and Shih [24]. Then from section 3.2.1 to section 3.2.6, we summarize Segal-Bargmann transform (S-transform), differential, adjoint, N operators and derivative of analytic functionals created by Lee and Shih. In section 3.2.6.2, we derive the chain rule for Lévy white noise
Chapter 3. Stein’s method and normal approximation of Lévy functionals

functionals; see Theorem 3.2.10. Then, in section 3.3, we use these operators above to derive our theorem for Stein’s method and normal approximation of Lévy functionals; see Theorem 3.3.1. In section 3.4, we show the difference and connection between white noise analysis and Malliavin calculus by computing simple examples. Section 3.5 is summary.

3.2 Lévy white noise functionals analysis

After reviewing Hida’s work in the section 2.3, now we are ready to discuss Lévy white noise analysis.

Lee and Shih [24] develop the Lévy white noise analysis for the general case of Lévy processes by means of the Segal-Bargmann transform (S-transform).

In the setting of Lee and Shih [24], a Lévy process is an additive process \( X = \{X(t) : t \in \mathbb{R}\} \) on a probability space \((\Omega, \mathcal{M}, P)\). It is a right continuous process with left limits.

The characteristic function of \( X(t) - X(s) \) for \( s < t \) is given by

\[
E[\exp[ir(X(t) - X(s))]] = \exp[(t-s)\mu r - \frac{\sigma^2 r^2}{2} + \int_{|u|>0} (e^{iru} - 1 - iru - \frac{1}{1+u^2}) \frac{(1 + u^2)}{u^2} d\beta(u), \quad r \in \mathbb{R}]
\]

where \( \mu = \sqrt{-1}, \) \( \mu \) is a real constant, \( \beta \) is a positive finite measure on \( \mathbb{R} \) with

\[
\sigma^2 = \beta(\{0\}) = \int_{u=0} (1 + u^2) d\beta(u).
\]

The Lévy measure of \( X \) is denoted by \( \beta_0 \) which is a positive measure on \( \mathbb{R}_+ (= \mathbb{R} \setminus \{0\}) \) defined by

\[
\beta_0(E) = \int_E \frac{1 + u^2}{u^2} d\beta(u)
\]
for $E \in \mathcal{B}(\mathbb{R}_*)$

According to Lee and Shih [24], let $\mathcal{S}$ be the Schwartz space (Definition 2.2.20) with the dual $\mathcal{S}'$ (Definition 2.2.17) of tempered distributions on $\mathbb{R}$.

Under the condition \( \int_{-\infty}^{+\infty} |u| d\beta(u) < +\infty \), where $\beta$ is defined in (3.1), the Bochner-Minlos theorem (Theorem 2.2.1) guarantees that there exists an unique probability measure $\Lambda$ on $(\mathcal{S}', \mathcal{B}(\mathcal{S}'))$ such that the characteristic function $\mathcal{C}(\eta)$ is given by

\[
\int_{\mathcal{S}'} \exp\{i(y, \eta)\} \Lambda(dy) = \mathcal{C}(\eta) = \exp\{\int_{-\infty}^{+\infty} f_X(\eta(t)) dt\} \quad (\eta \in \mathcal{S}),
\]

where $(\cdot, \cdot)$ is the $\mathcal{S}' - \mathcal{S}$ pairing, which is defined in (2.12) and (2.13). Here we use $(\cdot, \cdot)$ instead of $\langle \cdot, \cdot \rangle$, so the notation is consistent with that in Lee and Shih [24]. $\Lambda$ is then called the Lévy white noise measure. $(\mathcal{S}', \mathcal{B}(\mathcal{S}'), \Lambda)$ will serve as the underlying probability space in this chapter.

Moreover, for each $\eta \in \mathcal{S}$, $(y, \eta)$ is a random variable on the probability space $(\mathcal{S}', \mathcal{B}(\mathcal{S}'), \Lambda)$. If $\beta$ satisfies the following absolute moment condition

\[
\int_{-\infty}^{+\infty} |u|^n d\beta(u) < \infty \quad \forall n \in \mathbb{N}.
\]  

(3.4)

Then the mean and variance of the random variables $(y, \eta)$ are given by the following lemma.

**Lemma 3.2.1.** Let $y \in \mathcal{S}'$ be a generalised function whose probability distribution has characteristic function $\mathcal{C}(\eta)$; see (3.3). Assume that the integral in (3.4) is finite for $n = 1, 2, 3$. Then for fixed $\eta \in \mathcal{S}$, we have

\[
E[(y, \eta)] = \tau_1 \int_{-\infty}^{+\infty} \eta(t) dt \quad \text{and} \quad Var[(y, \eta)] = \tau_2 \int_{-\infty}^{+\infty} \eta(t)^2 dt,
\]

(3.5)

where

\[
\tau_1 = \mu + \int_{-\infty}^{+\infty} u d\beta(u) \quad \text{and} \quad \tau_2 = \int_{-\infty}^{+\infty} (1 + u^2) d\beta(u).
\]

(3.6)

**Proof.** Lee and Shih [24] does not provide the proof for the lemma above, so we give it here.
As we do not know a priori that $C(\varepsilon \eta)$ has finite first and second $\varepsilon$-derivatives at $\varepsilon = 0$, we shall calculate these derivatives from the first principles. In other words, we shall calculate the RHS of

$$\frac{d}{d\varepsilon} C(\varepsilon \eta) \bigg|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{C(\varepsilon \eta) - C(0)}{\varepsilon}$$

and

$$\frac{d^2}{d\varepsilon^2} C(\varepsilon \eta) \bigg|_{\varepsilon=0} = \lim_{\varepsilon, \varepsilon_1, \varepsilon_2 \to 0} \frac{C((\varepsilon_1 + \varepsilon_2) \eta) + C(0) - C(\varepsilon_1 \eta) - C(\varepsilon_2 \eta)}{\varepsilon_1 \varepsilon_2} \tag{3.7}$$

directly, and show that the limits are given by (3.5) and (3.6).

Starting from (3.1),

$$f_X(r) = i\mu r - \frac{\sigma^2 r^2}{2} + \int_{|u|>0} \left( e^{iru} - 1 - \frac{iru}{1+u^2} \right) \frac{1+u^2}{u^2} d\beta(u), \quad r \in \mathbb{R}$$

$$= i\mu r - \frac{\sigma^2 r^2}{2} + \int_{|u|>0} \left( \frac{iru^3}{1+u^2} - \frac{(ru)^2}{2} \right) \left( 1 + \frac{u^2}{2} \right) d\beta(u) + R(r),$$

where

$$R(r) = \int_{|u|>0} \left( e^{iru} - 1 - iru - \frac{(iru)^2}{2} \right) \frac{1+u^2}{u^2} d\beta(u).$$

Consequently,

$$\int_{-\infty}^{+\infty} f_X(\varepsilon \eta(t)) dt = i\mu \varepsilon \int_{-\infty}^{+\infty} \eta(t) dt - \frac{\sigma^2 \varepsilon^2}{2} \int_{-\infty}^{+\infty} \eta(t)^2 dt$$

$$+ i\varepsilon \left( \int_{-\infty}^{+\infty} \eta(t) dt \right) \left( \int_{-\infty}^{+\infty} ud\beta(u) \right)$$

$$- \frac{\varepsilon^2}{2} \left( \int_{-\infty}^{+\infty} \eta(t)^2 dt \right) \left( \int_{|u|>0} (1+u^2) d\beta(u) \right)$$

$$+ \int_{-\infty}^{+\infty} R(\varepsilon \eta(t)) dt.$$
Using the fact (3.2),
\[ \sigma^2 = \beta(\{0\}) = \int_{u=0} (1 + u^2) d\beta(u), \]
and
\[ \tau_1 = \mu + \int_{-\infty}^{+\infty} u d\beta(u), \quad \tau_2 = \int_{-\infty}^{+\infty} (1 + u^2) d\beta(u), \]
we get
\[ \int_{-\infty}^{+\infty} f_X(\varepsilon \eta(t)) dt = i\varepsilon \tau_1 \int_{-\infty}^{+\infty} \eta(t) dt - \frac{\varepsilon^2}{2} \tau_2 \int_{-\infty}^{+\infty} \eta(t)^2 dt + \int_{-\infty}^{+\infty} R(\varepsilon \eta(t)) dt, \tag*{(3.8)} \]
where
\[ \left| \int_{-\infty}^{+\infty} R(\varepsilon \eta(t)) dt \right| \leq O(\varepsilon^3). \tag*{(3.9)} \]

The RHS of the stated order is for all fixed \( \eta \in \mathcal{S} \) because \( \mathcal{S} \) is a Schwartz space.

In deriving (3.9), we have used the inequality
\[ \left| e^{ix} - 1 - ix - \frac{(ix)^2}{2} \right| \leq C |x|^3, \]
where \( C \in (0, \infty) \) is a constant independent of \( x \).

Therefore, from (3.3) and (3.8),
\[
C(\varepsilon \eta) = \exp\left\{ \int_{-\infty}^{+\infty} f_X(\varepsilon \eta(t)) dt \right\} \\
= 1 + i\varepsilon \tau_1 \int_{-\infty}^{+\infty} \eta(t) dt - \frac{\varepsilon^2}{2} \tau_2 \int_{-\infty}^{+\infty} \eta(t)^2 dt \\
- \frac{\varepsilon^2}{2} \left[ \tau_1 \int_{-\infty}^{+\infty} \eta(t) dt \right]^2 + O(\varepsilon^3),
\]
as $\varepsilon \to 0$. Consequently,

$$\frac{d}{d\varepsilon} \mathcal{C}(\varepsilon \eta)\bigg|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{\mathcal{C}(\varepsilon \eta) - \mathcal{C}(0)}{\varepsilon} = i\tau_1 \int_{-\infty}^{+\infty} \eta(t)dt,$$

from which the first part of (3.5) follows. To show the second part of (3.5), first note that

$$\mathcal{C}((\varepsilon_1 + \varepsilon_2)\eta) + \mathcal{C}(0) - \mathcal{C}(\varepsilon_1 \eta) - \mathcal{C}(\varepsilon_2 \eta)$$

$$= -\varepsilon_1\varepsilon_2 \left[ \tau_1 \int_{-\infty}^{+\infty} \eta(t)dt \right]^2 + \tau_2 \int_{-\infty}^{+\infty} \eta(t)^2dt + O(|\varepsilon_1|^3 + |\varepsilon_2|^3),$$

(3.10)

as $\varepsilon_1, \varepsilon_2 \to 0$, from (3.7) and (3.10), it follows that

$$E[(y,\eta)^2] = \left[ \tau_1 \int_{-\infty}^{+\infty} \eta(t)dt \right]^2 + \tau_2 \int_{-\infty}^{+\infty} \eta(t)^2dt,$$

and therefore $\text{Var}[(y,\eta)]$ is given by (3.5) as required.

□

Now, for each $\rho \in L^1 \cap L^2(\mathbb{R}, dt)$ (Definition 2.2.24), choose a sequence $\{\eta_n\} \subset \mathcal{S}$ so that $\eta_n \to \rho$ in $L^1 \cap L^2(\mathbb{R}, dt)$ under the norm $|\cdot|_{L^1(\mathbb{R}, dt)} + |\cdot|_{L^2(\mathbb{R}, dt)}$. Then it follows from (3.5) that $\{\langle \cdot, \eta_n \rangle\}$ forms a Cauchy sequence in $L^2(\mathcal{S}', \Lambda)$. Denote by $\langle \cdot, \rho \rangle$ the $L^2$-limit of $\{\langle \cdot, \eta_n \rangle\}$.

When $\rho = 1_{(s,t)}$, the indicator of $(s,t]$, the characteristic function of $\langle \cdot, \rho \rangle$ is exactly the same as the one in (3.1). Therefore, the Lévy process $X$ on $(\mathcal{S}', \mathcal{B}(\mathcal{S}'), \Lambda)$ can be represented by

$$X(t; x) = \left\{ \begin{array}{ll} < x, 1_{[0,t]} > & \text{if } t \geq 0 \\ -< x, 1_{[t,0]} > & \text{if } t < 0, \end{array} \right. x \in \mathcal{S}'.$$

We get $\dot{X}(t; x) = x(t), x \in \mathcal{S}'$. Thus the elements of $\mathcal{S}'$ are regarded as the sample paths of Lévy white noise. The members of $L^2(\mathcal{S}', \Lambda)$ are called quadratic integrable Lévy white noise functionals.

For $\eta = \eta_1 + i\eta_2 \in L^1_c \cap L^2_c(\mathbb{R}, dt)$ with $\eta_1, \eta_2 \in L^1 \cap L^2(\mathbb{R}, dt)$, $\langle \cdot, \eta \rangle$ is defined to be $\langle \cdot, \eta_1 \rangle + i\langle \cdot, \eta_2 \rangle$, where $V_c$ denotes the complexification of a real locally convex
Chapter 3. *Stein’s method and normal approximation of Lévy functionals*

space $V$. Then the above identity in (3.5) also hold for complex-valued random variable $< \cdot, \eta >$.

The following Lévy white noise measures all satisfy condition (3.4)

- the Gaussian white noise measure $\Lambda_W$ with $\mu = 0$ and $\beta = \delta_0$ Then $\mathcal{C}(\eta) = \exp\{-\frac{1}{2} \int_{-\infty}^{+\infty} \eta(t)^2 dt\}$ for $\eta \in \mathcal{F}$, and $X$ is a Wiener process.

- the Poisson white noise measure $\Lambda_P$ with $\mu = 1/2$ and $\beta = (1/2)\delta_1$ Then $\mathcal{C}(\eta) = \exp\{\int_{-\infty}^{+\infty}[e^{i\eta(t)} - 1]dt\}$ for $\eta \in \mathcal{F}$, and $X$ is a Poisson process with intensity parameter 1.

According to Lee and Shih [24], let $B_{\delta}$ where $\delta_a$ is Dirac measure concentrated on the point $a$.

For $E \in \mathcal{B}_b(\mathbb{R}^2)$ be the class of all bounded Borel subsets $E$ of $\mathbb{R}^2 = \mathbb{R}^2 \setminus \{ (t,0) : t \in \mathbb{R} \}$. Define the product measure $d\nu(t,u) = d\beta_0(u)dt$ on $\mathcal{B}(\mathbb{R}^2)$.

For $E \in \mathcal{B}_b(\mathbb{R}^2)$, let $N(E; \cdot)$ be a random variable on $(\mathcal{F}', \mathcal{B}(\mathcal{F}'))$ defined by

$$N(E;x) = \# \{ (t,u) \in E : X(t;x) - X(t-;x) = u \}.$$

Then $N(E; \cdot)$ is Poisson distributed with intensity measure $\nu$. The system of $\{N(E;x) - \nu(E) : E \in \mathcal{B}_b(\mathbb{R}^2); x \in \mathcal{F}' \}$ forms an Poisson random measure (Definition 2.2.25) with zero mean. We denote $"N(E;x) - \nu(E)"$ by $N_0(E;x)$. Therefore, $E[N_0(E)N_0(F)] = \nu(E \cap F)$. Let $f = \sum_{i=1}^{m} a_i 1_{E_i}$, where $a_i \in \mathbb{C}$ and $E_i; s \in \mathcal{B}_b(\mathbb{R}^2)$ being disjoint, be a complex-valued simple function on $\mathbb{R}^2$. The stochastic integral $\int_{\mathbb{R}^2} f(s)dN_0(s)$ with respect to $N_0$ is defined by

$$\int_{\mathbb{R}^2} f(s)dN_0(s) := \sum_{i=1}^{m} a_i N_0(E_i).$$

Then we have the isometry

$$\left\| \int_{\mathbb{R}^2} f(s)dN_0(s) \right\|_{L^2(\mathcal{F}',\Lambda)}^2 = \int_{\mathbb{R}^2} |f(s)|^2 d\nu(s),$$

where $\int_{\mathbb{R}^2} f(s)dN_0(s)$ is then defined for any $f \in L^2(\mathbb{R}^2, \nu)$.

In the setting of Lee and Shih [24], the Lévy-Itô decomposition ensures that for $b > a$,

$$X(b) - X(a) = \tau_1(b-a) + \sigma(B(b) - B(a)) + \int_{\mathbb{R}^2} u1_{(a,b) \times \mathbb{R}^2}(t,u)dN_0(t,u),$$

(3.11)
where $B = \{ B(t) : t \in \mathbb{R} \}$ is a 1-dimensional Wiener process, independent of the system of $\{ N(E) : E \in \mathcal{B}((\mathbb{R}^2)) \}$.
Throughout this thesis, $\lambda$ is a positive measure on $\mathcal{B}(\mathbb{R})^2$ defined by
\[ d\lambda(t,u) = (1 + u^2)d\beta(u)dt, \quad (3.12) \]
where $\beta(u)$ is the measure defined in (3.1).
Define a $L^2(S', \Lambda)$-valued function $M$ on $\{ E \in \mathcal{B}(\mathbb{R}^2) : \lambda(E) < +\infty \}$ by
\[ M(E) = \sigma \int_{-\infty}^{+\infty} 1_E(t,0)dB(t) + \int_{\mathbb{R}^2} u1_E(t,u)dN_0(t,u). \quad (3.13) \]
Then the system of $\{ M(E;x) : E \in \mathcal{B}(\mathbb{R}^2) \}$ forms an independent random measure with zero mean such that
\[ E[M(E)M(F)] = \lambda(E \cap F) \]
for all $E, F \in \mathcal{B}(\mathbb{R}^2)$.
In [17], Itô introduced the multiple stochastic integral with respect to $M$, which is described as follows. Let $V_n$ denote the class of all complex-valued symmetric simple function on $(\mathbb{R}^2)^n$ of the form
\[ \sum_{i=1}^n a_i \hat{1}_{E_1 \times \cdots \times E_n} \]
in which $\lambda(E_j^i) < +\infty$ for $j = 1, \ldots, n$, $E_1^i < \cdots < E_n^i$. $\hat{1}_{E_1^i \times \cdots \times E_n^i}$ is the symmetrization of $1_{E_1^i \times \cdots \times E_n^i}$. Here, "$A < B$" means that $t_1 < t_2$ for $(t_1, u_1) \in A$ and $(t_2, u_2) \in B$, which is partial ordering.
For any $g \in V_n$ of the form (3.14), the multiple stochastic integral $I_n(g)$ of $g$ with respect to $M$ is defined by $\sum_{i=1}^n a_i \prod_{j=1}^n M(E_j^i)$ which satisfies the isometry
\[ \|I_n(g)\|_{L^2(S', \Lambda)}^2 = n! \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} |g(s_1, \ldots, s_n)|^2d\lambda(s_1)\cdots d\lambda(s_n). \]
Since $V_n$ is dense in $\hat{L}^2_c((\mathbb{R}^2)^n, \lambda^{\otimes n})$, the closed subspace of $L^2_c((\mathbb{R}^2)^n, \lambda^{\otimes n})$ consisting
Chapter 3. Stein’s method and normal approximation of Lévy functionals

of all symmetric complex-valued functions in $L^2_c((\mathbb{R}^2)^n, \lambda^{\otimes n})$, $I_n(g)$ can be extended by
continuity to all $g \in \hat{L}^2_c((\mathbb{R}^2)^n, \lambda^{\otimes n})$. $I_n(g)$ is the multiple Lévy-Itô integral of order $n$
with respect to $M$. That is

$$I_n(g) = \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} g(s_1, \ldots, s_n) dM(s_1) \cdots dM(s_n). \tag{3.15}$$

According to the definition of multiple Lévy-Itô integrals, it is clear that as $n \neq m$, $I_n(g)$ and $I_m(h)$ are orthogonal to each other in $L^2_c(\mathcal{S}', \Lambda)$ for $g \in \hat{L}^2_c((\mathbb{R}^2)^n, \lambda^{\otimes n})$ and $h \in \hat{L}^2_c((\mathbb{R}^2)^m, \lambda^{\otimes m})$. For any $g \in L^2_c(\mathbb{R}^2, \lambda)$

$$I_1(g) = \sigma \int_{-\infty}^{+\infty} g(t,0) dB(t) + \int_{\mathbb{R}^2} u g(t,u) dN_0(t,u). \tag{3.16}$$

According to Lee and Shih [24], the centered Lévy process $\tilde{X} = \{X(t) - \tau_1 t : t \in \mathbb{R}\}$ is
a martingale and we can consider the $n$-multiple stochastic integral with respect to $\tilde{X}$

In Lee and Shih [24], due to Itô [17], the following theorem is an orthogonal decom-
position theorem for the quadratic integrable Lévy white noise functionals in terms of
multiple Lévy-Itô integrals.

**Theorem 3.2.1.** (Lévy-Itô decomposition). Let $\varphi$ be given in $L^2(\mathcal{S}', \Lambda)$. Then there
exist uniquely a series of kernel functions $\phi_n \in \hat{L}^2_c((\mathbb{R}^2)^n, \lambda^{\otimes n})$, $n \in \mathbb{N} \cup \{0\}$, such that
$\varphi$ is equal to the orthogonal direct sum.

$$\sum_{n=0}^{\infty} \oplus I_n(\phi_n). \tag{3.17}$$

In notation, we may write

$$\varphi \sim (\phi_n).$$

We also have

$$\|\varphi\|^2_{L^2(\mathcal{S}', \Lambda)} = \sum_{n=0}^{\infty} n! \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} |\phi_n(x_1, \ldots, x_n)|^2 d\lambda(x_1) \cdots d\lambda(x_n) < \infty. \tag{3.18}$$
Moreover, let \( g \in \hat{L}_2^2((\mathbb{R}^2)^n, \lambda^\otimes n) \)

\[
\|I_n(g)\|_{L_2(\mathcal{F}', \Lambda)}^2 = n! \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} |g(x_1, \ldots, x_n)|^2 d\lambda(x_1) \cdots d\lambda(x_n).
\]

For \( n \neq m \), \( g \in \hat{L}_2^2((\mathbb{R}^2)^n, \lambda^\otimes n) \) and \( h \in \hat{L}_2^2((\mathbb{R}^2)^m, \lambda^\otimes m) \).

Then

\[
E(I_n(g)I_m(h)) = \int_{\mathcal{F}'} I_n(g)(y) \cdot \overline{I_m(h)(y)} \Lambda(dy) = 0,
\]

where \( \overline{I_m(h)(y)} \) denotes complex conjugate of \( I_m(h)(y) \).

### 3.2.1 Segal-Bargmann transform (S-transform)

Following Lee and Shih [23] and Lee and Shih [24], the S-transform is defined in this section.

For \( \varphi \in L^2(\mathcal{F}', \Lambda) \), the well-known Itô theorem [17] asserts that \( \varphi \) can be written in an unique manner as an orthogonal direct sum of multiple Lévy-Itô integrals \( I_n(\phi_n) \) with kernel functions \( \phi_n \in \hat{L}_2^2((\mathbb{R}^2)^n, \lambda^\otimes n) \), where \( \lambda \) is the product measure \( dt(1 + u^2)d\beta(u) \).

Define the S-transform of \( \varphi \) on \( L^2(\mathbb{R}^2, \Lambda) \) by

**Definition 3.2.1.** Let \( \varphi \in L^2(\mathcal{F}', \Lambda) \) and \( \varphi \sim (\phi_n) \). Then the S-transform of \( \varphi \) is defined on \( L^2(\mathbb{R}^2, \lambda) \) by

\[
S\varphi(g) = \sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \phi_n(x_1, \ldots, x_n) g(x_1) \cdots g(x_2) d\lambda(x_1) \cdots d\lambda(x_n)
= \sum_{n=0}^{\infty} (\phi_n, g^\otimes n).
\]

We define the exponential vector functional \( \text{Exp}(g) \) associated with \( g \) by

\[
\text{Exp}(g) = \sum_{m=0}^{\infty} \oplus I_m(g^\otimes m)/m!.
\]
Theorem 3.2.2. Let $H = L^2(\mathbb{R}^2, \lambda)$. For $g \in H$, $\varphi \in L^2(\mathcal{H}', \Lambda)$, and $\varphi \sim (\phi_n)$, we have

$$S\varphi(g) = \sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \phi_n(x_1, \ldots, x_n) g(x_1) \cdots g(x_2) d\lambda(x_1) \cdots d\lambda(x_n)$$

$$= \int_{\mathcal{H}'} \varphi(y) \text{Exp}(g)(y) \Lambda(dy).$$

Proof. First, we prove that

$$\int_{\mathcal{H}'} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} I_n(\phi_n) I_m(g^{\otimes m}) / m! \Lambda(dy) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \int_{\mathcal{H}'} I_n(\phi_n) I_m(g^{\otimes m}) / m! \Lambda(dy).$$

We know that, for all $a_n \geq 0$,

$$\sum_{n=0}^{\infty} a_n < \infty \Rightarrow \lim_{N \to \infty} \sum_{n>N} a_n \to 0.$$ 

According to the formula (3.18), we get

$$\lim_{N \to \infty} \int_{\mathcal{H}'} \sum_{n=N}^{\infty} |I_n(\phi_n)|^2 \Lambda(dy) = 0$$

and

$$\lim_{M \to \infty} \int_{\mathcal{H}'} \sum_{m=M}^{\infty} \frac{|I_m(g^{\otimes m})|^2}{(m!)^2} \Lambda(dy) = 0.$$ 

Moreover, we decompose

$$\int_{\mathcal{H}'} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} I_n(\phi_n) I_m(g^{\otimes m}) / m! \Lambda(dy)$$
into 4 pieces:

\[
\left| \int_{\mathcal{F}} \sum_{n=0}^{\infty} \sum_{m=0}^{M} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy) - \sum_{n=0}^{N} \sum_{m=0}^{M} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy) \right|
\]

\[
= \left| \int_{\mathcal{F}} \sum_{n=N+1}^{\infty} \sum_{m=0}^{M} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy) + \int_{\mathcal{F}} \sum_{n=0}^{N} \sum_{m=M+1}^{\infty} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy) \right|
\]

\[
\leq \left| \int_{\mathcal{F}} \sum_{n=N+1}^{\infty} I_n(\phi_n) \left( \sum_{m=0}^{M} I_m(g^{\otimes m})/m! \right) \Lambda(dy) + \int_{\mathcal{F}} \sum_{n=0}^{N} \sum_{m=M+1}^{\infty} I_n(\phi_n) \left( \sum_{m=0}^{M} I_m(g^{\otimes m})/m! \right) \Lambda(dy) \right|
\]

As \( N, M \to \infty \), the upper bound is zero. Therefore, we can say

\[
\int_{\mathcal{F}} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy).
\]
Substitute $\text{Exp}(g) = \sum_{m=0}^{\infty} \oplus I_m(g^{\otimes m})/m!$ into the equation below:

\[
E(\varphi \text{Exp}(g)) = \int_{\mathbb{R}} \varphi(y) \text{Exp}(g)(y) \Lambda(dy)
\]

\[
= \int_{\mathbb{R}} \sum_{n=0}^{\infty} I_n(\phi_n) \sum_{m=0}^{\infty} I_m(g^{\otimes m})/m! \Lambda(dy)
\]

\[
= \int_{\mathbb{R}} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy)
\]

\[
= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \int_{\mathbb{R}} I_n(\phi_n) I_m(g^{\otimes m})/m! \Lambda(dy)
\]

\[
= \sum_{n=0}^{\infty} 1/n! \int_{\mathbb{R}} I_n(\phi_n) I_n(g^{\otimes n}) \Lambda(dy)
\]

\[
= \sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \phi_n(x_1, \ldots, x_n) g(x_1) \cdots g(x_n) d\lambda(x_1) \cdots d\lambda(x_n)
\]

\[
= S\varphi(g).
\]

\[\square\]

According to Lee and Shih [24], an example is given below.

**Example 3.2.1.** Let $h, g \in L^2(\mathbb{R}^2, \lambda)$. Then

\[
S(\text{Exp}(h))(g) = \exp \left\{ \int_{\mathbb{R}^2} g(x) h(x) d\lambda(x) \right\}.
\]
To see this, substitute \( \text{Exp}(h) \) and \( \text{Exp}(g) \) into the equation below:

\[
S(\text{Exp}(h))(g) = \int_{\mathcal{S}} \text{Exp}(h)(y)\text{Exp}(g)(y)\Lambda(dy)
\]

\[
= \int_{\mathcal{S}} \sum_{n=0}^{\infty} I_n(h^{\otimes n})/n! \sum_{m=0}^{\infty} I_m(g^{\otimes m})/m! \Lambda(dy)
\]

\[
= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{S}} I_n(h^{\otimes n})I_n(g^{\otimes n})\Lambda(dy)
\]

\[
= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} h(x_1) \cdots h(x_n)g(x_1) \cdots g(x_n)d\lambda(x_1) \cdots d\lambda(x_n)
\]

\[
= \sum_{n=0}^{\infty} \frac{1}{n!} \left( \int_{\mathbb{R}^2} g(x)h(x)d\lambda(x) \right)^n
\]

\[
= \exp\{ \int_{\mathbb{R}^2} g(x)h(x)d\lambda(x) \}.
\]

### 3.2.2 The spaces of test and generalised functions

#### 3.2.2.1 Gelfand triple associated with \( L^2(\mathbb{R}, dt) \)

According to Lee and Shih [24], consider the Hilbert space \( L^2(\mathbb{R}, dt) \) with the norm \( |\cdot|_0 \). Let \( H_n \) be the Hermite polynomial of degree \( n \), \( n \geq 0 \), defined by

\[
H_n(t) = (-1)^n e^{t^2} D_t^n e^{-t^2}.
\]

Let \( e_n \) be the Hermite function of order \( n \), \( n \geq 0 \), defined by

\[
e_n(t) = (\sqrt{\pi}^n n!)^{-1/2} H_n(t) e^{-t^2/2}.
\]

This family is a complete orthonormal basis for \( L^2(\mathbb{R}, dt) \). Let \( A \) be a densely defined self-adjoint operator on \( L^2(\mathbb{R}, dt) \)

\[
A = -\frac{d^2}{dt^2} + t^2 + 1.
\]

It is a fact that

\[
Ae_n = (2n + 2)e_n, \quad n \geq 0,
\]
where $2n + 2$ are eigenvalues.

For any $p \in \mathbb{R}$, and $f \in L^2(\mathbb{R}, dt)$, define

$$|f|_p := |A^p f|_{L^2(\mathbb{R}, dt)} = |A^p f|_0,$$

and let $\mathcal{S}_p$ be the completion of the class \{ $f \in L^2(\mathbb{R}, dt)$ : $|f|_p < +\infty$ \} with respect to $|\cdot|_p$-norm. Then $\mathcal{S}_p$ is a real separable Hilbert space and we have the continuous inclusions:

$$\mathcal{S} = \lim_{\rightarrow} \mathcal{S}_p \subset \mathcal{S}_q \subset L^2(\mathbb{R}, dt) \subset \mathcal{S}_{-q} \subset \mathcal{S}_{-p} \subset \lim_{\rightarrow} \mathcal{S}_{-q} \subset \mathcal{S}' \quad p > q \geq 0.$$

$f$ has the expansion in terms of the Hermite functions below.

$$f = \sum_{n=0}^{\infty} (f, e_n) e_n.$$

Then,

$$|f|_p^2 = \sum_{n=0}^{\infty} (2n + 2)^{2p} (f, e_n)^2.$$

where $(\cdot, \cdot)$ is the inner product of $L^2(\mathbb{R}, dt)$.

Now we show that the operator $A^{-p}$ is a Hilbert-Schmidt operator of $L^2(\mathbb{R}, dt)$ if and only if $p > \frac{1}{2}$.

$$||A^{-p}||_{HS}^2 = \sum_{n=0}^{\infty} |A^{-p} e_n|^2 = \sum_{n=0}^{\infty} |(2n + 2)^{-p} e_n|^2 = \sum_{n=0}^{\infty} (2n + 2)^{-2p}. \quad (3.22)$$

According to Definition 2.2.15, and line (3.22), we see if $p > \frac{1}{2}$, the operator $A^{-p}$ is a Hilbert-Schmidt operator of $L^2(\mathbb{R}, dt)$. 
In applications, a generalised function can be defined as the limit of a sequence of test functions. We provide a proof below to show its property.

**Theorem 3.2.3.** For $p > 0$, $f \in \mathcal{S}_p$, $\delta_t^\epsilon \in \mathcal{S}_p$ and $\delta_t \in \mathcal{S}_{-p}$. Choosing

$$\delta_t^\epsilon(u) = \frac{1}{\sqrt{2\pi\epsilon}} \exp\left(-\frac{1}{2\epsilon}(u-t)^2\right).$$

Then

$$(\delta_t, f) = \lim_{\epsilon \to 0^+} (\delta_t^\epsilon, f) = f(t).$$

**Proof.** Since $f \in \mathcal{S}_p$ is a test function, (i) $f(t)$ is continuous for all $t \in \mathbb{R}$ and (ii) $\int_{-\infty}^{+\infty} |f(u)| du < \infty$. Properties (i) and (ii) are used below.

We have

$$\int_{-\infty}^{+\infty} \delta_t^\epsilon(u) du = 1.$$

Therefore,

$$\lim_{\epsilon \to 0^+} \left| f(t) - \int_{-\infty}^{+\infty} \delta_t^\epsilon(u) f(u) du \right|$$

$$= \lim_{\epsilon \to 0^+} \left| \int_{-\infty}^{+\infty} \delta_t^\epsilon(u)(f(t) - f(u)) du \right|$$

$$\leq \lim_{\epsilon \to 0^+} \left| \int_{A_{\epsilon,c}} \delta_t^\epsilon(u)(f(t) - f(u)) du \right| + \lim_{\epsilon \to 0^+} \left| \int_{u \not\in A_{\epsilon,c}} \delta_t^\epsilon(u)(f(t) - f(u)) du \right|$$

$$\leq \lim_{\epsilon \to 0^+} \left( \sup_{u \in A_{\epsilon,c}} |f(t) - f(u)| + \sup_{u \not\in A_{\epsilon,c}} \delta_t^\epsilon(u) \int_{u \not\in A_{\epsilon,c}} |f(u)| du + |f(t)| \int_{u \not\in A_{\epsilon,c}} \delta_t^\epsilon(u) du \right)$$

$$= 0,$$

where

$$A_{\epsilon,c} = \{ u : |u - t| < \epsilon^{(1/2-c)} \},$$
and \(c\) is a constant, such that

\[
0 < c < 1/2.
\]

\(\square\)

The example below shows how we choose the generalised function space.

**Example 3.2.2.** Consider the delta function \(\delta_t\) in \(S'\). Its expansion in terms of the Hermite functions is given by

\[
\delta_t = \sum_{n=0}^{\infty} (\delta_t, e_n) e_n = \sum_{n=0}^{\infty} e_n(t) e_n.
\]

Using Theorem 3.2.3 above, \(\int_{-\infty}^{+\infty} e_n(t') \delta_t(t') dt' = e_n(t)\).

Therefore, for \(p > 0\), we have

\[
|\delta_t|_p^2 = |A^{-p}\delta_t|_0^2
\]

\[
= \int_{-\infty}^{+\infty} A^{-p}\delta_t(t') A^{-p}\delta_t(t') dt'
\]

\[
= \int_{-\infty}^{+\infty} \sum_{n=0}^{\infty} (2n + 2)^{-p} (\delta_t, e_n) e_n(t') \sum_{m=0}^{\infty} (2m + 2)^{-p} (\delta_t, e_m) e_m(t') dt'
\]

\[
= \sum_{n=0}^{\infty} (2n + 2)^{-2p} \int_{-\infty}^{+\infty} (\delta_t, e_n)^2 e_n(t') e_n(t') dt'
\]

\[
= \sum_{n=0}^{\infty} (2n + 2)^{-2p} (\delta_t, e_n)^2 \int_{-\infty}^{+\infty} e_n(t') e_n(t') dt'
\]

\[
= \sum_{n=0}^{\infty} (2n + 2)^{-2p} (\delta_t, e_n)^2
\]

\[
= \sum_{n=0}^{\infty} \frac{1}{(2n + 2)^{2p}} e_n(t)^2.
\]

In the above equations, line (3.23) follows from the definition of \(|\cdot|_p\)-norm (3.21); line (3.24) follows from the substitution of \(L^2\) expansion for \(A^{-p}\delta_t = \sum_{n=0}^{\infty} (2n + \)
2) \(-p(\delta_t, e_n)e_n;\) line (3.25) follows from the properties of the orthonormal basis; line (3.26) follows because \((\delta_t, e_n)\) does not depend on \(t';\) lines (3.27) and (3.28) follow because \(\delta_t\) is a Dirac delta function, i.e. \(\int_{-\infty}^{+\infty} \delta_t(t')dt' = 1\) and \(\int_{-\infty}^{+\infty} e_n(t')\delta_t(t')dt' = e_n(t).\) The Dirac delta is not a function in the traditional sense as no function defined on the real numbers has these properties.

For the sup-norm of the Hermite functions we have

\[
\|e_n\|_{\infty} = O(n^{-1/12}).
\]

Hence \(|\delta_t|_{-p} < \infty\) if \(2p + \frac{1}{6} > 1,\) i.e., \(p > \frac{5}{12}\). Thus \(\delta_t \in S_{-p}\) for any \(p > \frac{5}{12}\).

Now we introduce an example for the generalised functionals. The multiple Wiener integral can be expressed in terms of white noise. \(I_n(f_n)(x) = (: x^{\otimes n} :, f_n),\) where \(x^{\otimes n} :\) is Wick product.

If \(f_n = \delta_t^{\otimes n},\) then \(\hat{B}(t)^n := I_n(\delta_t^{\otimes n}) = (: \cdot^{\otimes n} :, \delta_t^{\otimes n}) \in \mathcal{S}'.\)

A generalised function can be defined as the limit of a sequence of test functions. We can find \(\delta_{t_\epsilon} \in \mathcal{S}_p,\) be a smooth approximation of \(\delta_t \in \mathcal{S}_{-p}.\) For example,

\[
\delta_{t_\epsilon}(u) = \frac{1}{\sqrt{2\pi\epsilon}} \exp(-\frac{1}{2\epsilon}(u - t)^2).
\]

Then, \((\delta_{t_\epsilon})^{\otimes n}\) converges to \((\delta_t)^{\otimes n}\) as \(\epsilon\) tends to zero. According to Hida, Kuo, Potthoff, and Streit [14], for heuristic interpretations and also for calculations, as \(\epsilon\) tends to zero, the limit of \(W_{t,\epsilon}^{(n)} = (: x^{\otimes n} :, (\delta_{t_\epsilon})^{\otimes n})\) converges to

\[
W_t^{(n)} = (: x^{\otimes n} :, \delta_t^{\otimes n}).
\]

### 3.2.2.2 Gelfand triple associated with \(L^2(\mathbb{R}, \lambda)\)

Now we discuss the space for pure jump Lévy processes, and then we combine the results with the space in section 3.2.2.1 to get the space for general Lévy processes. Consider the real Hilbert space \(L^2(\mathbb{R}, \gamma),\) where \(d\gamma(u) = (1 + u^2)d\beta(u)\) is finite measure on \((\mathbb{R}, \mathcal{B}(\mathbb{R})),\) which is related to the pure jump part of Lévy processes. Define a linear operator \(A_\beta\) on \(L^2(\mathbb{R}, \gamma)\) by \(A_\beta \zeta_n = r_n \zeta_n\) for \(n = 0, 1, \ldots,\) where the eigenvalues \(r_n\)'s
Chapter 3. *Stein’s method and normal approximation of Lévy functionals*

satisfy the assumption that $1 < r_0 \leq r_1 \leq \cdots$ and the subscript $\beta$ means the operator $A_\beta$ depends on what pure jump Lévy processes we have. According to Definition 2.2.15, $\|A^{-\alpha}_\beta\|_{HS} = \Sigma_n 1/r_n^{2\alpha} < +\infty$ for some $\alpha > 0$. In this thesis, set

$$\alpha_\beta = \inf\{\alpha > 0: \|A^{-\alpha}_\beta\|_{HS} < +\infty\},$$

where $\alpha_\beta$ means $\alpha$ depends on the measure $\beta$, and different Lévy processes have different $\beta$. For each $p \geq 0$, let $\mathcal{E}_p$ be the set of all $\zeta \in L^2(\mathbb{R}, \gamma)$ with $|A^{p}_\beta \zeta|_{L^2(\mathbb{R}, \gamma)} < +\infty$. Then $\mathcal{E}_p$ is a real separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{p,\beta}$ given by

$$\langle \zeta, \psi \rangle_{p,\beta} = \int_{-\infty}^{+\infty} A^{-p}_\beta \zeta(u) A^{-p}_\beta \psi(u) d\gamma(u) \quad \text{for all } \zeta, \psi \in \mathcal{E}_p$$

and induced norm by $|\cdot|_{p,\beta} := \sqrt{\langle \cdot, \cdot \rangle_{p,\beta}}$.

$\mathcal{E}_{-p}$ denotes the dual of $\mathcal{E}_p$. Then $\mathcal{E}_{-p}$ is isometrically isomorphic to the completion of $L^2(\mathbb{R}, \gamma)$ with the inner product $\langle \cdot, \cdot \rangle_{-p,\beta}$ and $|\cdot|_{-p,\beta}$ norm below

$$\langle \zeta, \psi \rangle_{-p,\beta} = \int_{-\infty}^{+\infty} A^{-p}_\beta \zeta(u) A^{-p}_\beta \psi(u) d\gamma(u) \quad \text{for all } \zeta, \psi \in L^2(\mathbb{R}, \gamma),$$

where we identify $x \in \mathcal{E}_{-p}$ with $\Sigma_n (x, \zeta_n) \zeta_n$. Let $\mathcal{E} = \lim_{p \to 0} \mathcal{E}_p$. Then $\mathcal{E}$ is a nuclear space with the dual $\mathcal{E}' = \lim_{p \to 0} \mathcal{E}_{-p}$. $\mathcal{E} \subset L^2(\mathbb{R}, \gamma) \subset \mathcal{E}'$ forms a Gel’fand triple.

Now we combine the results with the space in section 3.2.2.1 to get the space for general Lévy processes. For $p \in \mathbb{R}$, $\mathcal{N}_p$ denote the Hilbert space tensor product $\mathcal{S} \otimes \mathcal{E}_p$ with $|||\cdot|||_p$ norm defined by $|||e_i \otimes \zeta_j|||_p = |e_i|_p |\zeta_j|_{p,\beta}$. So, $\mathcal{N}_0 = L^2_c(\mathbb{R}^2, \lambda)$, and $d\lambda = dt \otimes d\gamma(u) = dt \otimes (1 + u^2) d\beta(u)$. Let $\mathcal{N} = \mathcal{S} \otimes \mathcal{E}$ with the dual $\mathcal{N}' = \mathcal{S}' \otimes \mathcal{E}'$.

$\mathcal{N} = \lim_{p \to 0} \mathcal{N}_p$ is a nuclear space induced by the family $\{\mathcal{N}_p, |||\cdot|||_p : p \geq 0\}$.

The dual $\mathcal{N}’_p$ of $\mathcal{N}_p$, $p \in \mathbb{R}$, is isometrically isomorphic to $\mathcal{N}_{-p}$ by identifying $x \in \mathcal{N}_{p}'$ with $\Sigma_{n,m}(x, e_n \otimes \zeta_m) e_n \otimes \zeta_m$.

$\mathcal{N}' = \lim_{p \to 0} \mathcal{N}_{-p}$ and $\mathcal{N} \subset L^2_c(\mathbb{R}^2, \lambda) \subset \mathcal{N}'$ forms a Gel’fand triple. Moreover, the inclusions below are all continuous.

$$\mathcal{N} \subset \mathcal{N}_p \subset \mathcal{N}_q \subset L^2(\mathbb{R}^2, \lambda) \subset \mathcal{N}_{-q} \subset \mathcal{N}_{-p} \subset \mathcal{N}' \quad p > q \geq 0.$$  \hspace{1cm} (3.29)
Then $\mathcal{N}_p$ is a real separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_p$ given by

$$<\zeta, \psi>_p = \int_{-\infty}^{+\infty} (A^p \otimes A^p_{\beta}) \zeta(u)(A^p \otimes A^p_{\beta}) \psi(u)d\lambda(u) \quad \text{for all } \zeta, \psi \in \mathcal{N}_p,$$

and induced norm by $|||\cdot|||_p := \sqrt{<\cdot, \cdot>_p}$, where $d\lambda = dt \otimes d\gamma(u) = dt \otimes (1 + u^2)d\beta(u)$.

Then $\mathcal{N}_{-p}$ is isometrically isomorphic to the completion of $L^2(\mathbb{R}, \lambda)$ with the inner product $<\cdot, \cdot>_{-p}$ and $|||\cdot|||_{-p}$ norm below

$$<\zeta, \psi>_{-p} = \int_{-\infty}^{+\infty} (A^{-p} \otimes A^{-p}_{\beta}) \zeta(u)(A^{-p} \otimes A^{-p}_{\beta}) \psi(u)d\lambda(u) \quad \text{for all } \zeta, \psi \in L^2(\mathbb{R}, \lambda),$$

where $d\lambda = dt \otimes d\gamma(u) = dt \otimes (1 + u^2)d\beta(u)$.

### 3.2.2.3 Test and generalised functionals

For $p \in \mathbb{R}$, define

$$|||\varphi|||_p^2 = \sum_{n=0}^{\infty} \frac{1}{n!} |||D^nS\varphi(0)|||^2_{\mathcal{N}_{-p}},$$

where $\varphi \in L^2(\mathcal{S}, \Lambda)$, and let $\mathcal{L}_p$ be the completion of the class $\{\varphi \in L^2(\mathcal{S}, \Lambda) : |||\varphi|||_p < +\infty\}$ with respect to $|||\cdot|||_p$ norm. Then $\mathcal{L}_p$, $p \in \mathbb{R}$, is a Hilbert space with the inner product $\ll \cdot, \cdot \gg_p$ induced by $|||\cdot|||_p$ norm. Let $\mathcal{L} = \lim_{\leftarrow p>0} \mathcal{L}_p$. Then $\mathcal{L}$ is a nuclear space. $\mathcal{L}$ will serve as the space of test functions and the dual space $\mathcal{L}'$ of $\mathcal{L}$ the space of generalised functions. In this thesis, members of $\mathcal{L}'$ are called generalised Lévy white noise functionals. Therefore, we obtain a Gel’fand triple $\mathcal{L} \subset L^2(\mathcal{S}, \Lambda) \subset \mathcal{L}'$ and have the continuous inclusion:

$$\mathcal{L} \subset \mathcal{L}_p \subset \mathcal{L}_q \subset L^2(\mathcal{S}, \Lambda) \subset \mathcal{L}_q' \subset \mathcal{L}_p' \subset \lim_{\rightarrow} \mathcal{L}_p' = \mathcal{L}' \quad p > q \geq 0,$$

where $\mathcal{L}_p'$ is isometrically isomorphic to $\mathcal{L}_{-p}$. Two isometrically isomorphic normed vector spaces share the same structure, so they are usually identified with each other.
Moreover, let $F_n \in \mathcal{N}_{-p,c}^{\otimes n}$

\[
\ll F, \varphi \gg = \sum_{n=0}^{\infty} n! (F_n, \phi_n),
\]

(3.32)

and

\[
\|F\|_p^2 = \sum_{n=0}^{\infty} n! \|F_n\|_p^2,
\]

(3.33)

where $(\cdot, \cdot)$ is the $\mathcal{N}_{-p,c}^{\otimes n}-\mathcal{N}_{p,c}^{\otimes n}$ pairing, where subscript $c$ denotes the complexification.

Now we define the S-transform for all generalised Lévy white noise functionals as follows.

**Definition 3.2.2.** The S-transform $SF$ of a generalised Lévy white noise functional $F \in \mathcal{L}_{-p}$, $p > 0$, is defined to be a complex-valued function on $\mathcal{N}_{p,c}$ such that

\[
SF(g) = \ll F, \text{Exp}(g) \gg = \sum_{n=0}^{\infty} (F_n, g^{\otimes n})
\]

(3.34)

where $\text{Exp}(g) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(g^{\otimes n}) \in \mathcal{L}_p$, $g \in \mathcal{N}_c$, and $(\cdot, \cdot)$ is the $\mathcal{N}_c^{\otimes n}-\mathcal{N}_c^{\otimes n}$ pairing.

From the equation (3.34), we can verify the following results.

**Proposition 3.2.1.** Let $F \sim (F_n)$ be in $\mathcal{L}_{-p}$, $p \in \mathbb{R}$. Then $SF$ is analytic on $\mathcal{N}_{p,c}$, satisfying the exponential growth condition:

\[
|SF(g)| \leq \|F\|_p e^{(1/2)\|g\|_p^2} \quad \text{for any } g \in \mathcal{N}_{p,c}.
\]

(3.35)

**Proof.** Lee and Shih [24] do not provide the proof, so we provide it here.

Using first a continuous version of the Cauchy-Schwarz inequality for each $n$, and then a discrete form of the Cauchy-Schwarz inequality, we have the general inequality.

\[
\left| \sum_{n=0}^{\infty} \int f_n g_n d\mu_n \right| \leq \sum_{n=0}^{\infty} \left( \int |f_n|^2 d\mu_n \right)^{1/2} \left( \int |g_n|^2 d\mu_n \right)^{1/2} \leq \left( \sum_{n=0}^{\infty} \int |f_n|^2 d\mu_n \right)^{1/2} \left( \sum_{n=0}^{\infty} \int |g_n|^2 d\mu_n \right)^{1/2}.
\]

(3.36)
Using the equation (3.34),

$$|SF(g)| = \left| \sum_{n=0}^{\infty} (F_n, g^{\otimes n}) \right|$$

$$= \sum_{n=0}^{\infty} \left| (A^{\otimes n} \otimes A_\beta^{\otimes n})^{-p} F_n, (A^{\otimes n} \otimes A_\beta^{\otimes n})^p g^{\otimes n} \right|$$

$$= \sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} (A^{\otimes n} \otimes A_\beta^{\otimes n})^{-p} F_n (A^{\otimes n} \otimes A_\beta^{\otimes n})^p g^{\otimes n} d\lambda(x_1) \cdots d\lambda(x_n)$$

$$\leq \sqrt{\sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \left| (A^{\otimes n} \otimes A_\beta^{\otimes n})^{-p} F_n (x_1, \ldots, x_n) \right|^2 d\lambda(x_1) \cdots d\lambda(x_n)}$$

$$\leq \sqrt{\sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \left| (A \otimes A_\beta)^p g(x_1) \cdots (A \otimes A_\beta)^p g(x_n) \right|^2 d\lambda(x_1) \cdots d\lambda(x_n)}.$$}

We use (3.36) and (3.37). Then,

$$|SF(g)| \leq \left[ \sum_{n=0}^{\infty} n! \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \left| (A^{\otimes n} \otimes A_\beta^{\otimes n})^{-p} F_n (x_1, \ldots, x_n) \right|^2 d\lambda(x_1) \cdots d\lambda(x_n) \right]^{1/2}$$

$$\leq \left[ \sum_{n=0}^{\infty} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \frac{\left| (A \otimes A_\beta)^p g(x_1) \cdots (A \otimes A_\beta)^p g(x_n) \right|^2}{n!} d\lambda(x_1) \cdots d\lambda(x_n) \right]^{1/2}$$

$$\leq \|F\|_{-p} \left[ \sum_{n=0}^{\infty} \frac{\left( \int_{\mathbb{R}^2} \left| (A^p \otimes A_\beta^p) g(x) \right|^2 d\lambda(x) \right)^n}{n!} \right]^{1/2}$$

$$= \|F\|_{-p} e^{1/2\|g\|_p^2}.$$}

Using the definitions of \(\|F\|_{-p}\) and \(\|g\|_p\) given in (3.33) and (3.30) respectively. □

### 3.2.3 Differential operator

We use S-transform to derive the differential operator, and then according to the duality formula, we will find that \(\partial_\xi^*\) is the adjoint operator of the differential operator in the next section. The purpose here is to deliver the intuitive ideas about main operators in analysis of generalized Lévy white noise functionals. Detailed proofs can be found in Lee and Shih [24].
Definition 3.2.3. For $F \in \mathcal{L}_p$, $p \in \mathbb{R}$, let $\xi \in \mathcal{N}_{-p,c}$ be fixed. Then we define the partial derivative $\partial_\xi$ acting on $\mathcal{L}_p$ in the direction $\xi$ by

$$
\partial_\xi F = S^{-1} \left( \frac{d}{dz} \bigg|_{z=0} SF(\cdot + z\xi) \right),
$$

where $\partial_\xi$ is a continuous operator from $\mathcal{L}_p$ into $\mathcal{L}_{p-l}$, and $l > 1$.

That is

$$
\| \partial_\xi F \|_{p-l} \leq \rho \| \xi \|_{-p} \| F \|_p,
$$

where $\rho$ is a constant. For white noise calculus, we choose $\xi = \delta_{(t,u)}$.

In particular, $\partial_\xi$ is continuous from $\mathcal{L}$ into itself. Accordingly, we define its adjoint operator $\partial_\xi^*$ from $\mathcal{L}'$ into $\mathcal{L}'$ by the following duality formula.

Definition 3.2.4. For $F \in \mathcal{L}'$ and $\varphi \in \mathcal{L}$, we have the duality formula.

$$
\langle \partial_\xi^* F, \varphi \rangle := \langle F, \partial_\xi \varphi \rangle.
$$

(3.38)

Let $\xi = \delta_{(t,u)}$, we get

$$
\langle \partial_{\delta_{(t,u)}}^* F, \varphi \rangle := \langle F, \partial_{\delta_{(t,u)}} \varphi \rangle,
$$

that is,

$$
E(\partial_{\delta_{(t,u)}}^* F \varphi) = E(F \partial_{\delta_{(t,u)}} \varphi).
$$

(3.39)

Lemma 3.2.2. For $\phi \in \mathcal{N}_{p,c}^\otimes n$ (3.29), subscript $c$ denotes complexification. $n \geq 1$, and $\xi \in \mathcal{N}_{-p,c}$ with $p \in \mathbb{R}$, define $K_{\xi,\phi} \in \mathcal{N}_{-p,c}^\otimes (n-1)$ (3.29) by

$$
K_{\xi,\phi}(\psi) = (\xi \otimes \psi, \phi), \quad \psi \in \mathcal{N}_{c}^\otimes (n-1),
$$

(3.40)

where $(\cdot, \cdot)$ is the $\mathcal{N}_{-p,c}^\otimes - \mathcal{N}_{p,c}^\otimes$ pairing. Then, for $p \geq q$,

$$
||| K_{\xi,\phi} |||_q \leq \theta_0^{(n-1)(q-p)} ||| \xi |||_{-p} ||| \phi |||_p.
$$

(3.41)
Chapter 3. Stein's method and normal approximation of Lévy functionals

**Proof.** Let \( \rho_0 = \inf \text{Spec}(A) > 1 \), \( r_0 = \inf \text{Spec}(A_\beta) > 1 \), and

\[
\theta_0 = \rho_0 \times r_0,
\tag{3.42}
\]

where \( \inf \text{Spec}(\cdot) \) denotes the smallest eigenvalue. Set \( p \geq q \), and \( \xi \otimes \psi \) is symmetric. 

| \cdot | is the absolute value of a complex number.

Applying Cauchy Schwartz inequality,

\[
||| K_{\xi,\phi} |||_q
= \sup_{||\psi||_0 = 1} \left| \left( \xi \hat{\otimes} (A^{(n-1)} \otimes A_\beta^{(n-1)})^q \psi, \phi \right) \right|
= \sup_{||\psi||_0 = 1} \left| \left( (A^{(n)} \otimes A_\beta^{(n)})^{-p} \xi \hat{\otimes} (A^{(n-1)} \otimes A_\beta^{(n-1)})^q \psi, (A^{(n)} \otimes A_\beta^{(n)})^p \phi \right) \right|
\leq \sup_{||\psi||_0 = 1} \left| \left( (A^{(n)} \otimes A_\beta^{(n)})^{-p} \xi \hat{\otimes} (A^{(n-1)} \otimes A_\beta^{(n-1)})^q \psi \right) \right|_0 \left| \left( (A^{(n)} \otimes A_\beta^{(n)})^p \phi \right) \right|_0
= \sup_{||\psi||_0 = 1} \left| \left( (A^{(n-1)} \otimes A_\beta^{(n-1)})^{-p} (A^{(n-1)} \otimes A_\beta^{(n-1)})^q \psi \right) \right|_0 \left| \left( (A \otimes A_\beta)^{-p} \xi \right) \right|_0 \left| \left( (A^{(n)} \otimes A_\beta^{(n)})^p \phi \right) \right|_0
= \sup_{||\psi||_0 = 1} \left| \left( (A^{(n-1)} \otimes A_\beta^{(n-1)})^{-p} \xi \right) \right|_0 \left| \left( (A^{(n-1)} \otimes A_\beta^{(n-1)})^q \psi \right) \right|_0 \left| \left( (A \otimes A_\beta)^{-p} \xi \right) \right|_0 \left| \left( (A^{(n)} \otimes A_\beta^{(n)})^p \phi \right) \right|_0
\leq \theta_0^{(n-1)(q-p)} ||| \xi |||_{-p} ||| \phi |||_p.
\tag{3.43}
\]

Line (3.43) follows because \( \theta_0 \) is the product of the two smallest eigenvalues and \( q - p \leq 0 \). Therefore, \( \theta_0^{(n-1)(q-p)} \) is the largest.

\[\square\]

In this thesis and Lee and Shih [24], for white noise calculus, we choose \( \xi = \delta(t,u) \), and \( \psi = \delta(t_2,u_2) \hat{\otimes} \cdots \hat{\otimes} \delta(t_n,u_n) \) for the following operator.

\[K_{\xi,\phi}(\psi) = (\xi \hat{\otimes} \psi, \phi).\]

Then Lee and Shih [24] provide the following Lemma.
Lemma 3.2.3. We choose \( \xi = \delta_{(t,u)} \), and \( \psi = \delta_{(t_2,u_2)} \circ \cdots \circ \delta_{(t_n,u_n)} \) in the equation 3.40, and we get

\[
K_{\delta_{(t,u)},\phi}(\psi) = \phi(s, s_2, \cdots, s_n) 
= \phi((t, u), (t_2, u_2), \ldots, (t_n, u_n)),
\]

where \( s = (t, u), s_2 = (t_2, u_2), \ldots, s_n = (t_n, u_n) \).

Proof. Lee and Shih [24] does not provide the proof for this lemma, so we provide the proof step by step.

Let \( s = (t, u), s_2 = (t_2, u_2), \ldots, s_n = (t_n, u_n) \).

Substitute \( \xi = \delta_{(t,u)} \) into (3.40) and using the definitions of symmetric function and symmetrization. (2.14), (2.15) and (2.16)

\[
K_{\delta_{(t,u)},\phi}(\psi) = \left( \delta_{(t,u)} \circ \psi, \phi \right) 
= \int_{\mathbb{R}^2} \left( \delta_{(t,u)} \circ \psi, \phi((t, u)) \right) d\lambda((t, u)) 
= \int_{\mathbb{R}^2} \left( \delta_s \circ \psi, \phi((s, s)) \right) d\lambda(s) 
= \frac{1}{n} \left( \int_{\mathbb{R}^2} \left( \delta_s \circ \psi, \phi((s, s)) \right) d\lambda(s) + \int_{\mathbb{R}^2} \left( \delta_s \circ \psi, \phi((s_2, s_2)) \right) d\lambda(s_2) + \cdots 
+ \int_{\mathbb{R}^2} \left( \delta_s \circ \psi, \phi((s_n, s_n)) \right) d\lambda(s_n) \right) 
= \int_{\mathbb{R}^2} \left( \delta_s \circ \psi, \phi((s, s)) \right) d\lambda(s) 
= (\psi, \phi(s, s_2, \cdots, s_n)).
\]

Using the fact that the symmetrized tensor product \( \delta_s \circ \psi \) is the symmetrization of \( \delta_s \circ \psi \).

Set \( \psi = \delta_{(t_2,u_2)} \circ \cdots \circ \delta_{(t_n,u_n)} = \delta_{(t_2,u_2)} \circ \psi', \) and \( \psi' = \delta_{(t_3,u_3)} \circ \cdots \circ \delta_{(t_n,u_n)} \). Then

\[
(\psi, \phi(s, s_2, \cdots, s_n)) = (\delta_{(t_2,u_2)} \circ \psi', \phi(s, s_2, \cdots, s_n)).
\]
Repeat the process above \( n - 1 \) more times.

In the end, we get

\[
K_{\delta(t,u),\phi}(\psi) = (\delta(t,u) \otimes \psi, \phi)
\]

\[
= (\delta(t_2,u_2) \otimes \psi', \phi(s, s_2, \ldots, s_n))
\]

\[
= \cdots \text{ repeat the process in (3.46) for } n-1 \text{ more times}
\]

\[
= \phi(s, s_2, \ldots, s_n)
\]

\[
= \phi((t, u), (t_2, u_2), \ldots, (t_n, u_n)).
\]

All integrals and Dirac delta functions cancel each other out.

Therefore, \( K_{\delta(t,u),\phi}(\psi) = \phi((t, u), (t_2, u_2), \ldots, (t_n, u_n)) \)

Now go back to the equation (3.47) and substitute.

We have

\[
K_{\delta(t,u),\phi}(\psi)
\]

\[
= (\delta(t,u) \otimes \psi, \phi)
\]

\[
= (\psi, K_{\delta(t,u),\phi}).
\]

Moreover, according to the permutation, we can also get

\[
\delta(t,u) \otimes \delta(t_2,u_2) \otimes \cdots \otimes \delta(t_n,u_n) = \frac{1}{n!}\delta(t,u) \otimes \delta(t_2,u_2) \otimes \cdots \otimes \delta(t_n,u_n)
\]

\[
= \delta(t,u) \otimes \delta(t_2,u_2) \otimes \cdots \otimes \delta(t_n,u_n).
\]

Therefore,

\[
\delta(t,u) \otimes \delta(t_2,u_2) \otimes \cdots \otimes \delta(t_n,u_n) \text{ is the symmetrization of } \delta(t,u) \otimes \delta(t_2,u_2) \otimes \cdots \otimes \delta(t_n,u_n).
\]

Now we are ready to deliver the differential operator of Lévy functionals in terms of the Lévy-Itô decomposition of \( F \), which has been provided in Lee and Shih [24].
Theorem 3.2.4. (Differential operator)

Let $F \sim (F_n) \in \mathcal{L}_p$ given in (3.31), and $\delta_{(t,u)} \in \mathcal{N}_{-p,c}$ given in (3.29) for $p \in \mathbb{R}$. Then, for $q < p$,

$$\partial \delta_{(t,u)} F = (\delta_{(t,u)}, F_1) + \sum_{n=2}^\infty n I_{n-1}(K \delta_{(t,u)}, F_n)$$

$$= \sum_{n=1}^\infty n I_{n-1}(K \delta_{(t,u)}, F_n) \in \mathcal{L}_q.$$  \hspace{1cm} (3.48)

Moreover,

$$\|\partial \delta_{(t,u)} F\|_q \leq \omega_{p-q} \theta_0^{p-q} \||\delta_{(t,u)}\||_{-p} \|F\|_p,$$

where $\theta_0$ is given in (3.42), and $\omega_r^2 = \sup\{x \theta_0^{-2xr} : x \geq 0\}$.

Proof. We provide the proof here for the self-contained purpose.

According to the definition of $S$-transform (3.34), we get

$$SF(g) = \sum_{n=0}^\infty (F_n, g^\otimes n) \equiv \sum_{n=0}^\infty \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} F_n(x_1 \cdots, x_n) g(x_1) \cdots g(x_n) d\lambda(x_1) \cdots d\lambda(x_n).$$  \hspace{1cm} (3.49)

When $F = I_n(F_n)$, we have

$$S(I_n(F_n))(g) = (F_n, g^\otimes n) \equiv \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} F_n(x_1 \cdots, x_n) g(x_1) \cdots g(x_n) d\lambda(x_1) \cdots d\lambda(x_n).$$  \hspace{1cm} (3.50)
\[ \partial_{\delta(t,u)} F = S^{-1} \left( \frac{d}{dz} \bigg|_{z=0} SF(\cdot + z\delta(t,u)) \right) \]

\[ \Rightarrow S(\partial_{\delta(t,u)} F)(g) = \frac{d}{dz} \bigg|_{z=0} SF(g + z\delta(t,u)) \]

\[ = \frac{d}{dz} \sum_{n=0}^{\infty} (F_n, (g + z\delta(t,u))^\otimes n) \quad \text{(according to (3.49))} \]

\[ = \sum_{n=1}^{\infty} \left( F_n, \frac{d}{dz} \bigg|_{z=0} (g + z\delta(t,u))^\otimes n \right) \]

\[ = \sum_{n=1}^{\infty} n \left( F_n, g^{\otimes (n-1)} \otimes \delta(t,u) \right) \]

\[ = \sum_{n=1}^{\infty} n \left( K_{\delta(t,u)}, F_n, g^{\otimes (n-1)} \right). \]  

Step (3.51) follows because \( G(z) = SF(g + z\delta(t,u)) = \sum_{n=0}^{\infty} (F_n, (g + z\delta(t,u))^\otimes n) \) is an entire function of \( z \), and according to (3.35), \( G(z) \) is bounded. Hence its derivative with respect to \( z \) is obtained by differentiating term by term.  

Taking \( S^{-1} \) on both sides,

\[ \partial_{\delta(t,u)} F = \sum_{n=1}^{\infty} nS^{-1} \left( K_{\delta(t,u)}, F_n, g^{\otimes (n-1)} \right) \]

\[ = \sum_{n=1}^{\infty} nI_{n-1}(K_{\delta(t,u)}, F_n). \quad \text{(according to (3.50))} \]
Moreover, setting $\xi = \delta_{(t,u)}$,

$$\|\partial_\xi F\|_q^2 = \left\| \sum_{n=1}^{\infty} nI_{n-1}(K_{\xi,F_n}) \right\|_q^2$$

Using (3.48)

$$= \sum_{n=1}^{\infty} n^2(n-1)!||K_{\xi,F_n}||_q^2$$

Using (3.33)

$$\leq \sum_{n=1}^{\infty} n^2(n-1)!\theta_0^{2(n-1)(q-p)}||\xi||_{-p}^2||F_n||_p^2$$

Using (3.41)

$$= ||\xi||^2_p\sum_{n=1}^{\infty} nn!\theta_0^{-2n(p-q)}\theta_0^{2(p-q)}||F_n||_p^2$$

$$= \omega_r^{2(p-q)}||\xi||_{-p}^2\|F\|_p^2$$

Using $||F||_p^2 = \sum_{n=0}^{\infty} nn!||F_n||_p^2$,

where

$$\omega_r^2 = \sup\{x\theta_0^{-2x} : x \geq 0\}. \quad (3.52)$$

We cite Example 5.12 from Lee and Shih [24] below to show the derivative in the generalised function space.

**Example 3.2.3.** Let $F \in \mathcal{N}_{-p,c}$ for $p > 0$. Then

$$\partial_{(t,u)}I_1(F) = F(t,u).$$

### 3.2.4 Adjoint operator

We get the differential operator now, and the following result shows that the operator $\partial_{(t,u)}^*$ is the adjoint operator of the differential operator. We express $\partial_{(t,u)}^* F$ in terms of the Lévy-Itô decomposition of $F$ below. For the complete proof, please refer to Lee and Shih [24].
Theorem 3.2.5. (Adjoint operator) Let $F \sim (F_n) \in \mathcal{L}_p$ with $p \in \mathbb{R}$, and $\delta_{(t,u)} \in \mathcal{N}'_c$. Then

$$\partial^*_{\delta_{(t,u)}} F = \sum_{n=0}^{\infty} I_{n+1}(\delta_{(t,u)} \hat{\otimes} F_n) \text{ in } \mathcal{L}'^\prime.$$ 

Moreover, for $\delta_{(t,u)} \in \mathcal{N}_-q$ with $q > p$,

$$\|\partial^*_{\delta_{(t,u)}} F\|_{-q} \leq \omega_q^{-p} \theta_0^{q-p} \|\delta_{(t,u)}\| \|F\|_{-p},$$

where $\theta_0$ is given in (3.42), and $\omega^2_r = \sup\{x^{\theta_0-2x} : x \geq 0\}$ given in (3.52).

**Proof.** We provide some intuition of the proof here, and detailed proofs can be found in Lee and Shih [24].

According to the duality formula (3.39)

$$\langle \partial^*_{\delta_{(t,u)}} F, \varphi \rangle = \langle F, \partial_{\delta_{(t,u)}} \varphi \rangle,$$

that is,

$$E(\partial^*_{\delta_{(t,u)}} F \varphi) = E(F \partial_{\delta_{(t,u)}} \varphi).$$

Insert $\partial^*_{\delta_{(t,u)}} F = \sum_{n=0}^{\infty} I_{n+1}(\delta_{(t,u)} \hat{\otimes} F_n)$ into the duality formula.

First,

$$\langle \partial^*_{\delta_{(t,u)}} F, \varphi \rangle = E\left(\sum_{n=0}^{\infty} I_{n+1}(\delta_{(t,u)} \hat{\otimes} F_n) \sum_{m=0}^{\infty} I_m(\phi_m)\right)$$

$$= E\left(\sum_{n=0}^{\infty} I_{n+1}(\delta_{(t,u)} \hat{\otimes} F_n) I_{n+1}(\phi_{n+1})\right)$$

$$= \sum_{n=0}^{\infty} E(I_{n+1}(\delta_{(t,u)} \hat{\otimes} F_n) I_{n+1}(\phi_{n+1}))$$

$$= \sum_{n=0}^{\infty} (n+1)! (\delta_{(t,u)} \hat{\otimes} F_n, \phi_{n+1}). \quad (3.53)$$
Let \( s = (t, u), s_2 = (t_2, u_2), \ldots, s_n = (t_n, u_n) \). We show that

\[
\sum_{n=0}^{\infty} (n+1)! (\delta_{(t,u)} \otimes F_n, \phi_{n+1}) = \sum_{n=0}^{\infty} (n+1)! \left( \frac{1}{n+1} \left( \int_{\mathbb{R}^2} (\delta_s \otimes F_n, \phi_{n+1}(\cdot, s)) d\lambda(s) \right) \right) + \sum_{n=0}^{\infty} (n+1)! \left( \int_{\mathbb{R}^2} (\delta_{s_1} \otimes F_n, \phi_{n+1}(\cdot, s_1)) d\lambda(s_1) + \cdots \right) + \sum_{n=0}^{\infty} (n+1)! \left( \int_{\mathbb{R}^2} (\delta_{s_n} \otimes F_n, \phi_{n+1}(\cdot, s_n)) d\lambda(s_n) \right)
\]

\[
= \sum_{n=0}^{\infty} (n+1)! \left( \frac{1}{n+1} \left( \int_{\mathbb{R}^2} (\delta_s \otimes F_n, \phi_{n+1}(\cdot, s)) d\lambda(s) \right) \right) + \sum_{n=0}^{\infty} (n+1)! \int_{\mathbb{R}^2} (\delta_s \otimes F_n, \phi_{n+1}(\cdot, s)) d\lambda(s)
\]

\[
= \sum_{n=0}^{\infty} (n+1)! \int_{\mathbb{R}^2} (\delta_s \otimes F_n, \phi_{n+1}(\cdot, s)) d\lambda(s)
\]
\[ 
\sum_{n=0}^{\infty} (n+1)! \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \delta_s \otimes F_n(s_1, s_2, \ldots, s_n) \\
\phi_{n+1}(s, s_1, s_2, \ldots, s_n) d\lambda(s) d\lambda(s_1) d\lambda(s_2) \cdots d\lambda(s_n) \\
= \sum_{n=0}^{\infty} (n+1)! \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} \delta_s^\ast \otimes F_n(s_1, s_2, \ldots, s_n) \\
\phi_{n+1}(s, s_1, s_2, \ldots, s_n) d\lambda(s) d\lambda(s_1) d\lambda(s_2) \cdots d\lambda(s_n) \\
= \sum_{n=0}^{\infty} (n+1)! (\delta(t,u) \otimes F_n, \phi_{n+1}) . 
\]  

(3.57) 

(3.58) 

(3.59) 

Line (3.56) follows because we apply \( K \) operator. Line (3.57) follows because the Dirac delta function, \( \delta_s \), and the integral \( \int_{\mathbb{R}^2} d\lambda(s) \) cancel each other out; line (3.58) follows because the symmetrized tensor product \( f \otimes g \) is the symmetrization of \( f \otimes g \), (3.54) and (3.55). 

We can see that (3.53) is equal to (3.59), so 

\[ \ll \partial^*_{\delta(t,u)} F, \varphi \gg = \ll F, \partial_{\delta(t,u)} \varphi \gg . \] 

Therefore, \( \partial^*_{\delta(t,u)} F = \sum_{n=0}^{\infty} I_n(\delta(t,u) \otimes F_n) \) is the adjoint operator of the differential operator. 

\[ \square \] 

### 3.2.5 N operator 

According to Lee and Shih [24], set \( F = \sum_{n=0}^{\infty} I_n(F_n) \in L^2(\mathcal{F}', \Lambda) \), and \( E(F) = I_0(F_0) = F_0 = 0 \), we have the following \( N \) operator.
Theorem 3.2.6. \((N \text{ operator})\) \(N\) is a self-adjoint bounded operator from \(L^2(\mathscr{S}', \Lambda)\) into \(L_{-q}\) for some sufficiently large \(q > 0\). We have

\[
NF = N(F) = \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial^*_t(t,u) \partial_t(t,u) F d\lambda(t,u)
\]

\[
= \sum_{n=1}^{\infty} n I_n(F_n),
\]

and the operator norm of \(N\) is bounded by

\[
\int_{\mathbb{R}^2} \|\partial^*_t(t,u) \partial_t(t,u)\|_{-q} d\lambda(t,u) \leq \omega_q/2 (1 + \omega_q^2 \theta_0^{q/2} \|F\|_0 \left( \int_{\mathbb{R}^2} \|\delta(t,u)\|^2_{-q} d\lambda(t,u) \right)^{1/2},
\]

where \(\theta_0\) is given in (3.42), and \(\omega_r^2 = \sup \{x \theta_0^{-2r} : x \geq 0\}\) given in (3.52).

Now we define \(M\) operator.

Definition 3.2.5. Set \(E(F) = I_0(F_0) = F_0 = 0\), and define the operator \(M\) from \(L^2(\mathscr{S}', \Lambda)\) to \(L^2(\mathscr{S}', \Lambda)\) by

\[
MF = \sum_{n=1}^{\infty} \frac{1}{n} I_n(F_n),
\]

where \(F \sim (F_n)\) and \(F \in L^2(\mathscr{S}', \Lambda)\).

Therefore, when \(E(F) = I_0(F_0) = F_0 = 0\), we have the following equation.

\[
NMF = F.
\]

3.2.6 Derivative of analytic functionals and chain rule

3.2.6.1 Derivative of analytic functionals

Lee and Shih [27] [24] provide the celebrated Le\'vy white noise derivative of analytic functionals below.
We define $DF(x)\delta_t$ as a directional derivative below:

$$DF(x)\delta_t = \lim_{h \to 0} \frac{F(x + h\delta_t) - F(x)}{h}.$$

For a Gâteaux differentiable function $F$ on $\mathcal{S}'$, define $T_F : \mathbb{R}^2 \times \mathcal{S}' \to \mathbb{R}$ by

$$T_F(t, u; x) = \frac{F(x + u\delta_t) - F(x)}{u} \cdot 1_{\mathbb{R}\backslash\{0\}}(u) + (DF(x)\delta_t) \cdot 1_{\{0\}}(u), \quad (3.62)$$

where $\delta_t$ is the Dirac measure concentrated on the point $t$. Let $C(\mathcal{S}', \mathcal{S})$ be the class consisting all smooth cylinder functionals $F$ on $\mathcal{S}'$ such that $F$ can be expressed as

$$F = g((x, \eta_1), \ldots, (x, \eta_n)) \quad (3.63)$$

for some $g \in C^1(\mathbb{R}^n)$ and $\eta_1, \ldots, \eta_n \in \mathcal{S}$.

The following result is obtained in Lee and Shih [27].

**Theorem 3.2.7.** (Lee and Shih’s derivative of analytic functionals) For $(t, u) \in \mathbb{R}^2$, the derivative of a smooth random variable $F$ of the form (3.63) is given by

$$\partial_{(t,u)} F(x) = T_F(t, u; x) \quad \text{in } L^2(\mathcal{S}', \Lambda). \quad (3.64)$$

Or equivalently,

$$\partial_{(t,u)} g((x, \eta_1), \ldots, (x, \eta_n)) = T_{g((x, \eta_1), \ldots, (x, \eta_n))}(t, u; x) \quad \text{in } L^2(\mathcal{S}', \Lambda),$$

where $\partial_{(t,u)} F(x)$ is defined in Definition 3.2.3, and $T_F(t, u; x)$ is defined in the equation (3.62).

**Remark 3.2.8.** In the equation (3.62), the first term in RHS is a difference operator with respect to the pure jump part of Lévy white noise. Generally speaking, $F(x + u\delta_t)$ means that we add a jump of size of $u$ at time $t$ in the trajectory $x$.

The second term, $DF(x)\delta_t$ is the directional derivative with respect to the Lévy white noise diffusion part at the point $x$ in the direction $\delta_t$. We will compute an example later.
Based on (3.62) and (3.64), we define our notations $D^p F$, where $D^p$ denotes a difference operator with respect to the pure jump part of Lévy white noise.

**Definition 3.2.6.** For $F \in L^2(\mathcal{F}', \Lambda)$, let $D^p$, where the superscript $p$ denotes Poisson, denote the difference operator defined by

$$D^p_{(t,u)} F(x) = \frac{F(x + u\delta_t) - F(x)}{u}. \quad (3.65)$$

Then,

$$F(x + u\delta_t) - F(x) = uD^p F,$$

and from (3.62) and (3.64), we have

$$\partial_{(t,u)} F(x) = \left( \frac{F(x + u\delta_t) - F(x)}{u} \right) 1_{\{u:|u| > 0\}}(u) + (DF(x)\delta_t) 1_{\{0\}}(u) \quad (3.66)$$

$$= D^p F 1_{\{u:|u| > 0\}}(u) + (DF(x)\delta_t) 1_{\{0\}}(u). \quad (3.67)$$

There are two terms in the derivative of analytic Lévy functional (3.62) and (3.64). The first term is for the pure jump part while the second term is for the diffusion part. Now we derive the chain rule for the diffusion part of a Lévy functional.

**Theorem 3.2.9.** (Chain rule for the the diffusion part of a Lévy functional). If $f$ and $F$ are smooth cylinder functionals of the form given in (3.63), then

$$Df(F(x))\delta_t = f'(F)DF(x)\delta_t. \quad (3.68)$$

**Proof.** By the definition of directional derivative,

$$DF(x)\delta_t = \lim_{h \to 0} \frac{F(x + h\delta_t) - F(x)}{h}. \quad (3.69)$$

Suppose that $f$ is differentiable at the point $u = F(x)$, and $F$ is differentiable at $x$. Let the function $M(k)$ be defined by

$M(0) = 0$
\[ M(k) = \frac{f(u+k) - f(u)}{k} - f'(u), \text{ if } k \neq 0 \]

By definition,

\[
\lim_{k \to 0} M(k) = f'(u) - f'(u) = 0 = M(0)
\]

and we have

\[
f(u + k) - f(u) = (f'(u) + M(k))k.
\]

Now we put \( u = F(x) \), and \( k = F(x + h\delta_t) - F(x) \)

so that \( u + k = F(x + h\delta_t) \), and obtain

\[
f(F(x + h\delta_t)) - f(F(x)) = (f'(F(x)) + M(k))(F(x + h\delta_t) - F(x)).
\]

Hence

\[
Df(F(x))\delta_t = \lim_{h \to 0} \frac{f(F(x + h\delta_t)) - f(F(x))}{h} = \lim_{h \to 0} \left( f'(F(x)) + M(k) \right) \frac{F(x + h\delta_t) - F(x)}{h} = \lim_{h \to 0} \left( f'(F(x)) + 0 \right) \frac{F(x + h\delta_t) - F(x)}{h} = f'(F(x))DF(x)\delta_t,
\]

where \( \lim_{h \to 0} M(k) = \lim_{k \to 0} M(k) = M(0) = 0 \), and

\[
DF(x)\delta_t = \lim_{h \to 0} \frac{F(x + h\delta_t) - F(x)}{h}
\]

which was to be proved.

According to Definition 3.2.6, for \( F \in L^2(\mathcal{S}', \Lambda) \), and \( f(F) \in L^2(\mathcal{S}', \Lambda) \), we define our notation \( D^p f(F) \) below. \( D^p \) is a difference operator for the pure jump part of Lévy
functional.

\[ D_{(t,u)}^p f(F)(x) = f(F(x + u\delta_t)) - f(F(x)) \frac{u}{u}. \]  \hspace{1cm} (3.70)

According to (3.70), we derive the following useful Lemma. We will use it to derive our chain rule for Lévy functional later.

**Lemma 3.2.4.** For \((t, u) \in \mathbb{R}^2\), \(F\) is a smooth random variable of the form (3.63). The chain rule for the difference operator (3.70) is given by

\[ D_{(t,u)}^p f(F)(x) = f'(F(x))D^p F + \frac{f''(s)}{2!}u(D^p F)^2, \]  \hspace{1cm} (3.71)

where \(s\) is some number between \(F(x + u\delta_t)\) and \(F(x)\).

**Proof.** According to (3.70), we get

\[ D^p f(F) = \frac{f(F(x + u\delta_t)) - f(F(x))}{u} \]

\[ \implies uD^p f(F) = f(F(x + u\delta_t)) - f(F(x)) \] \hspace{1cm} (by Taylor’s formula)

\[ = f'(F(x))(F(x + u\delta_t) - F(x)) + \frac{f''(s)}{2!}((F(x + u\delta_t) - F(x))^2 \]

\[ = f'(F(x))uD^p F + \frac{f''(s)}{2!}(uD^p F)^2 \]

\[ \implies uD^p f(F) = f'(F(x))uD^p F + \frac{f''(s)}{2!}(uD^p F)^2 \]

\[ \implies D^p f(F) = f'(F(x))D^p F + \frac{f''(s)}{2!}u(D^p F)^2, \]

where \(s\) is some number between \(F(x + u\delta_t)\) and \(F(x)\).

### 3.2.6.2 Chain rule

We derive the chain rule for Lévy white noise functionals. We will use this chain rule in the next section.

**Theorem 3.2.10.** (Chain rule for Lévy white noise functionals) Let \(f : \mathbb{R} \to \mathbb{R}\) and \(f \in \mathcal{C}^2\). \(F\) is a smooth random variable of the form (3.63). The derivative of \(f(F)\) is
Chapter 3. Stein’s method and normal approximation of Lévy functionals

The purpose of this section is to provide the analysis in the framework of the normal approximation (in the Wasserstein distance) of regular functionals of Lévy measures. This research is initiated by Nourdin and Peccati [29].
3.3.1 Normal approximation in the Wasserstein distance via Stein’s method

Now we need to review Stein’s method for normal approximation in the Wasserstein distance. According to [4] and [35], given two real-valued random variables $F$ and $X$, the Wasserstein distance between the laws of $F$ and $X$, written as $d_W(F, X)$, is defined as follows.

**Definition 3.3.1.** For any two real-valued random variables $F$ and $X$,

$$d_W(F, X) = \sup_{h \in \text{Lip}(1)} |Eh(F) - Eh(X)|,$$

(3.73)

where $\text{Lip}(1)$ denotes the class of real-valued Lipschitz functions, from $\mathbb{R}$ to $\mathbb{R}$, with Lipschitz constant less or equal to one, that is, $\|h'\|_\infty \leq 1$. For a function $g : \mathbb{R} \to \mathbb{R}$, $\|g\|_\infty$ denotes $\sup_{x \in \mathbb{R}} |g(x)|$.

Based on [4] and [35], now consider a real-valued function $h$ such that the expectation $E[h(X)]$ is well defined. Set $X \sim \mathcal{N}(0, 1)$. The Stein equation associated with $h$ and $X$ is given by

$$h(F) - E[h(X)] = f'(F) - Ff(F) \quad F \in \mathbb{R}.$$  

(3.74)

A solution to (3.74) is a function $f$ depending on $h$.

From [4] and [35], Stein’s lemma is given below.

**Lemma 3.3.1.** Let $F$ be a random variable. Then $F \overset{\text{Law}}{=} X \sim \mathcal{N}(0, 1)$ if and only if,

$$Ef'(F) = E[Ff(F)]$$

(3.75)

for every continuous and piecewise continuously differentiable function $f$ such that $E|f'(X)| < \infty$.

If $h$ is absolutely continuous with bounded derivative, then (3.74) has a solution $f_h$ which is twice differentiable and such that $\|f_h'\|_\infty \leq \|h'\|_\infty$ and $\|f_h''\|_\infty \leq 2\|h'\|_\infty$. 

According to [35], $\mathcal{F}_W$ denote the class of twice differentiable functions, whose first derivative is bounded by 1 and whose second derivative is bounded by 2. If $h \in \text{Lip}(1)$ and $\|h'\|_\infty \leq 1$, then $f_h$ satisfies $\|f_h'\|_\infty \leq 1$ and $\|f_h''\|_\infty \leq 2$. Therefore, $f_h \in \mathcal{F}_W$.

Now let $F$ be a random variable such that $E(F^2) < \infty$, and Gaussian random variable $X \sim \mathcal{N}(0, 1)$. By integrating both sides of (3.74) with respect to the law of $F$ and by using (3.73) and Stein’s lemma 3.3.1, we can get the following lemma.

**Lemma 3.3.2.** Suppose $X \sim \mathcal{N}(0, 1)$. Then

$$d_W(F, X) = \sup_{h \in \text{Lip}(1)} |Eh(F) - Eh(X)|$$

$$= \sup_{h \in \text{Lip}(1)} |Ef_h'(F) - Ff_h(F)|$$

$$\leq \sup_{f_h \in \mathcal{F}_W} |Ef_h'(F) - Ff_h(F)|.$$  

(3.76) (3.77) (3.78)

In the subsequent section, we will show that we can bound the quantity appearing on the right-hand side of (3.78) by means of the Lévy white noise operators introduced from the section 3.2.1 to the section 3.2.6.

### 3.3.2 Upper bound of Wasserstein distance for Lévy functionals

After we review Stein’s method for normal approximation in the Wasserstein distance, we are now ready to derive the upper bound of Wasserstein distance for Lévy functionals. First, we compute $E[If(F)]$ in the lemma below which will be used later.

**Lemma 3.3.3.** Let $F \in L^2(\mathcal{S}', \Lambda)$ be smooth cylinder functionals of the form given in (3.63). Using duality formula (3.39), we get

$$E[If(F)] = E[\int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t, u)}MF\partial_{(t, u)}f(F)d\lambda(t, u)]$$

$$= E[(\partial_{(t, u)}MF, \partial_{(t, u)}f(F))].$$  

(3.79)
Proof. According to duality formula (3.39),

\[
E[F f(F)] = E[NMF f(F)] \\
= E\left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial^*_t \partial_u MF d\lambda(t, u) f(F) \right] \\
= \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E[\partial^*_t \partial_u MF] f(F) d\lambda(t, u) \\
= \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E[\partial_t MF \partial_u f(F)] d\lambda(t, u) \quad \text{(By the duality (3.39))} \\
= E\left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_t MF \partial_u f(F) d\lambda(t, u) \right] \\
= E\left[ \left( \partial_t MF, \partial_u f(F) \right) \right].
\]

\[
\square
\]

The following theorem is the main result in this chapter. Since \( F \) is Lévy functionals, we provide Stein’s Method and Normal Approximation of Lévy functionals, which is the most general results. Gaussian and poisson functionals are the only special case of Lévy functionals. We also use white noise operators instead of Malliavin calculus operators, which shows how we can apply analysis of generalized Lévy white noise functionals. The result gives an upper bound for the Wasserstein distance in terms of the sum of two quantities.

**Theorem 3.3.1.** (Upper bound for Lévy functionals) Assume \( F \in L^2(\mathcal{F}', \Lambda) \), and \( F \) are smooth cylinder functionals given in (3.63). \( E(F) = 0 \), and \( X \sim \mathcal{N}(0, 1) \). Then

\[
d_W(F, X) \leq E\left[ \left| 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_t MF \partial_u f(F) d\lambda(t, u) \right| \right] + \tag{3.80}
\]

\[
\int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E\left[ \left| \partial_t MF \partial_u f(F) \right| \right] d\lambda(t, u) \tag{3.81}
\]

\[
\leq \sqrt{E\left[ \left( 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_t MF \partial_u f(F) d\lambda(t, u) \right)^2 \right]} + \tag{3.82}
\]

\[
\int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E\left[ \left| \partial_t MF \partial_u f(F) \right| \right] d\lambda(t, u), \tag{3.83}
\]

where \( \partial_{(t,u)} F \) is white noise derivative (3.48), and we use \( M \) (3.60) operator, and \( X \sim \mathcal{N}(0, 1) \).
Proof. According to the Stein-type bound (3.78), it is sufficient to prove that, for every function \( f \) such that \( \|f'\|_\infty \leq 1 \) and \( \|f''\|_\infty \leq 2 \) (that is, for every \( f \in \mathcal{F}_W \)), the quantity \( |E[f'(F) - Ff(F)]| \) is smaller than (3.82) + (3.83).

First, by the linearity of expectations and (3.79)

\[
|E[f'(F) - Ff(F)]| = |E f'(F) - E(Ff(F))| = |E f'(F) - E \left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} f(F) d\lambda(t,u) \right].
\]

(by (3.79))

Using linearity of expectations

\[
= \left| E f'(F) - E \left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF f'(F) \partial_{(t,u)} F d\lambda(t,u) + \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF f''(s) \frac{u(D^p F)^2}{2^1} 1_{\{u:|u|>0\}}(u) d\lambda(t,u) \right] \right|.
\]
Taking out $f'(F)$, and apply triangle inequality.

\[
= \left| E \left[ f'(F)(1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} F d\lambda(t,u)) \right] + E \left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} -\partial_{(t,u)} MF \frac{f''(s)}{2!} u(D^p F)^2 1_{|u| > 0} (u) d\lambda(t,u) \right] \right|
\leq \left| E \left[ f'(F)(1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} F d\lambda(t,u)) \right] + E \left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \frac{f''(s)}{2!} u(D^p F)^2 1_{|u| > 0} (u) d\lambda(t,u) \right] \right|
\leq A + B. \tag{3.84}
\]

By the fact that $\|f'\|_\infty \leq 1$ and by Cauchy-Schwarz,

\[
A = \left| E \left[ f'(F)(1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} F d\lambda(t,u)) \right] \right|
\leq E \left[ \left| 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} F d\lambda(t,u) \right| \right]
\leq \sqrt{E \left[ \left( 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} F d\lambda(t,u) \right)^2 \right]. \tag{3.85}
\]

Due to the fact that $\|f''\|_\infty \leq 2$,

\[
B = \left| E \left[ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \frac{f''(s)}{2!} u(D^p F)^2 1_{|u| > 0} (u) d\lambda(t,u) \right] \right|
\leq E \left[ \left| \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \frac{f''(s)}{2!} u(D^p F)^2 1_{|u| > 0} (u) d\lambda(t,u) \right| \right]
\leq \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ \left| \frac{f''(s)}{2!} \right| \partial_{(t,u)} MF(D^p F)^2 1_{|u| > 0} (u) \right| d\lambda(t,u)
\leq \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ \left| \partial_{(t,u)} MF \left( D^p F 1_{|u| > 0} (u) \right)^2 \right| u \right] d\lambda(t,u)
\leq \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ \left| \partial_{(t,u)} MF \left( D^p F 1_{|u| > 0} (u) \right)^2 \right| u \right] d\lambda(t,u). \tag{3.86}
\]

Substitute (3.85) and (3.86) into (3.84), (3.83) + (3.82) is obtained. \square
Remark 3.3.2. 1. Peccati, Solé, Taqqu and Utzet [35] only provide Stein normal approximation for Poisson functionals as they say in the paper’s title. In Peccati, Solé, Taqqu and Utzet [35], at page 445, the equation (2.4), Lévy-Khinchine exponent we cite below only has a pure jump part, so the Malliavin calculus they use to derive their main result is only for pure jump processes, not for stochastic processes which have both a diffusion part and a pure jump part. The Lévy-Khinchine exponent for pure jumps in Peccati, Solé, Taqqu and Utzet [35] is

\[
\psi(h, \lambda) = \log \mathbb{E}[e^{i\lambda \hat{N}(h)}] = \int_Z \{e^{i\lambda h(z)} - 1 - i\lambda h(z)\} \mu(dz).
\]

On the contrary, in this thesis, we provide a Stein normal approximation for Lévy functionals which have both a diffusion part and a pure jump part. The Lévy-Khinchine exponent we use below has both a diffusion part and a pure jump part:

\[
\psi(\mu, \sigma^2, \eta) = \log \mathbb{E}[e^{i(y, \eta)}] = \int_{-\infty}^{+\infty} f(\mu, \sigma^2)(\eta(t)) dt \quad (\eta \in \mathcal{F}),
\]

where \((y, \eta) = \tau_1 \int_{-\infty}^{+\infty} \eta(t) dt + I_1(\eta)(y), \quad \tau_1 = \mu + \int_{-\infty}^{+\infty} ud\beta(u),\) and

\[
f(\mu, \sigma^2) = i\mu r - \frac{\sigma^2 r^2}{2} + \int_{|u| > 0} (e^{iru} - 1 - \frac{iru}{1 + u^2}) \frac{1 + u^2}{u^2} d\beta(u), \quad r \in \mathbb{R}
\]

\[
= i\mu r - \frac{\sigma^2 r^2}{2} + \int_{-\infty}^{+\infty} \left( e^{iru} - 1 - iru \mathbf{1}_{|u| \leq 1}(u) \right) d\beta_0(u), \quad r \in \mathbb{R}, \quad (3.87)
\]

where \(i = \sqrt{-1}, \mu \) is a real constant, \(\beta \) is a positive finite measure on \(\mathbb{R}\) with

\[
\sigma^2 = \beta(\{0\}).
\]

The Lévy measure is denoted by \(\beta_0\) which is a positive measure on \(\mathbb{R}_* (= \mathbb{R} \setminus \{0\})\) defined by

\[
\beta_0(E) = \int_E \frac{1 + u^2}{u^2} d\beta(u)
\]
for \( E \in \mathcal{B}(\mathbb{R}_*) \)

2. We know that Malliavin derivative operator for pure jump processes is a difference operator. Therefore, in Peccati, Solé, Taqqu and Utzet [35], at page 450, the equation (2.20), Malliavin derivative operator they define has only the difference operator, and does not have the directional derivative for the diffusion part in a Lévy process.

On the contrary, in this thesis, the white noise derivative we use defined in Lee and Shih [24] is for Lévy processes which have both a diffusion part and a pure jump part. We have the equation 3.64

\[
\partial_{(t,u)} F(x) = T_F(t,u;x) = \frac{F(x + u\delta_t) - F(x)}{u} \cdot 1_{\mathbb{R}\setminus\{0\}}(u) + (DF(x)\delta_t) \cdot 1_{\{0\}}(u)
\]

The first term \( D^p_{(t,u)} F(x) \) is a difference operator for the pure jump part. The second term \( DF(x)\delta_t \) is a directional derivative for the diffusion part.

According to the equation (3.16), and \( \mathbb{R}_2^* = \mathbb{R}_2 \setminus \{(t,0) : t \in \mathbb{R}\} \), we define

\[
I_1(f) = I_{1G}(f_G) + I_{1J}(f_J),
\]

where \( f_G \) is the kernel function for the Gaussian component, and \( f_J \) is the kernel function for the pure jump component, and

\[
I_{1G}(f_G) = \int_{-\infty}^{+\infty} f(t,0)\sigma dB(t), \tag{3.88}
\]

and

\[
I_{1J}(f_J) = \int_{\mathbb{R}_2^*} uf(t,u)dN_0(t,u). \tag{3.89}
\]

\( B = \{B(t) : t \in \mathbb{R}\} \) is a 1-dimensional Wiener process, independent of the system of \( \{N(E) : E \in \mathcal{B}_b(\mathbb{R}_2^*)\} \). Let \( \mathcal{B}_b(\mathbb{R}_2^*) \) be the class of all bounded Borel subsets \( E \) of
$\mathbb{R}_*^2 = \mathbb{R}^2 \setminus \{(t, 0) : t \in \mathbb{R}\}$. Then $N(E; \cdot)$ is Poisson distributed with intensity measure $\nu$. The system of $\{N(E; x) - \nu(E) : E \in \mathcal{B}(\mathbb{R}_*^2) ; x \in \mathcal{S} \}$ forms an independent random measure (Definition 2.2.25) with zero mean. We denote $"N(E; x) - \nu(E)"$ by $N_0(E; x)$. Therefore, $E[N_0(E)N_0(F)] = \nu(E \cap F)$.

In Malliavin calculus for Brownian motion, they define
\[ I_{1G}(f_G) = \int_{-\infty}^{+\infty} g(t, 0) dB(t), \quad (3.90) \]
and in Malliavin calculus for Poisson functional, they define
\[ I_{1J}(f_J) = \int_{\mathbb{R}_*^2} g(t, u) dN_0(t, u). \quad (3.91) \]

From (3.88), we have
\[ E(I_{1G}(f_G)I_{1G}(f_{2G})) = \int_{-\infty}^{+\infty} f_1(t, 0)f_2(t, 0)\sigma^2 dt, \]
where
\[ \sigma^2 = \beta(\{0\}) = \int_{u=0} (1 + u^2) d\beta(u). \]

From (3.89), we have
\[ E(I_{1J}(f_{1J})I_{1J}(f_{2J})) = \int_{\mathbb{R}_*^2} f_1(t, u)f_2(t, u)(1 + u^2) d\beta dt, \]
where $d\lambda = (1 + u^2)d\beta dt$.

Therefore,
\[ E(I_1(f_1)I_1(f_2)) = \int_{\mathbb{R}_*^2} f_1(t, u)f_2(t, u) d\lambda(t, u) \]
\[ = \int_{\mathbb{R}} \int_{u=0} f_1(t, 0)f_2(t, 0)(1 + u^2) d\beta dt \]
\[ + \int_{\mathbb{R}_*^2} f_1(t, u)f_2(t, u)(1 + u^2) d\beta dt \]
\[ = \int_{\mathbb{R}} f_1(t, 0)f_2(t, 0)\sigma^2 dt + \int_{\mathbb{R}_*^2} f_1(t, u)f_2(t, u)(1 + u^2) d\beta dt. \]
For Malliavin calculus for Brownian functional, from the equation (3.90), we have
\[ E(I_1G(f_1G)I_1G(f_2G)) = \int_{-\infty}^{+\infty} f_1(t,0)f_2(t,0)dt. \]

For Malliavin calculus for Poisson functional, from the equation (3.91), we have
\[ E(I_1J(f_1J)I_1J(f_2J)) = \int_{\mathbb{R}^2} f_1(t,u)f_2(t,u)d\beta dt, \]
where \( \beta \) is Lévy measure, \( d\nu = d\beta dt \) and recall that
\[ d\beta(u) = \frac{1 + u^2}{u^2}d\beta(u). \]

Now we show that Stein bounds for functionals of pure jump Lévy processes are particular cases of the general bound in Theorem 3.3.1.

**Corollary 3.3.1.** Let \( F \in L^2(\mathcal{F},\Lambda), \ E(F) = 0, \) and \( X \sim \mathcal{N}(0,1). \) \( F \) has the only pure jump part. Then
\[ d_W(F,X) \leq E \left[ \left| 1 - \int_{t\in\mathbb{R}}\int_{u\in\mathbb{R}\setminus\{0\}} (\partial_{(t,u)}F)(\partial_{(t,u)}MF)u^2 d\nu(t,u) \right| + \right. \]
\[ \left. \int_{t\in\mathbb{R}}\int_{u\in\mathbb{R}} E \left[ \left| \partial_{(t,u)}F1_{\{u>0\}}(u) \right|^2 \left| \partial_{(t,u)}MF \right| \right] |u|u^2 d\nu(t,u), \right. \]
where \( d\nu = d\beta_0(u)dt. \)

**Proof.** In Theorem 3.3.1, our main result is
\[ d_W(F,X) \leq E \left[ \left| 1 - \int_{t\in\mathbb{R}}\int_{u\in\mathbb{R}} \partial_{(t,u)}MF\partial_{(t,u)}Fd\lambda(t,u) \right| + \right. \]
\[ \left. \int_{t\in\mathbb{R}}\int_{u\in\mathbb{R}} E \left[ \left| \partial_{(t,u)}MF \left( \partial_{(t,u)}F1_{\{u>0\}}(u) \right) \right|^2 \right] |u|d\lambda(t,u), \right. \]
where
\[ d\lambda(t, u) = dt(1 + u^2)d\beta(u). \]
and

$$\sigma^2 = \beta(\{0\}) = \int_{u=0} (1 + u^2) d\beta(u).$$

The Lévy measure is denoted by $\beta_0$ which is a positive measure on $\mathbb{R}_*(= \mathbb{R} \setminus \{0\})$ defined by

$$\beta_0(E) = \int_E \frac{1 + u^2}{u^2} d\beta(u)$$

for $E \in \mathcal{B}(\mathbb{R}_*)$.

If $F$ has the only pure jump part, and we use $d\lambda(t, u) = dt(1+u^2)d\beta(u)$, then from (3.92), we can get

$$d_W(F, X) \leq E \left[ 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R} \setminus \{0\}} (\partial_{(t,u)} F)(\partial_{(t,u)} MF)(1 + u^2)d\beta(u)dt \right]$$

$$+ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ |\partial_{(t,u)} F 1_{\{u:|u|>0\}}(u)|^2 |\partial_{(t,u)} MF| |u| (1 + u^2) d\beta(u)dt \right].$$

Recall that $d\beta_0(u) = \frac{1 + u^2}{u^2} d\beta(u)$, we get

$$d_W(F, X) \leq E \left[ 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R} \setminus \{0\}} (\partial_{(t,u)} F)(\partial_{(t,u)} MF)u^2 d\beta_0(u)dt \right]$$

$$+ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ |\partial_{(t,u)} F 1_{\{u:|u|>0\}}(u)|^2 |\partial_{(t,u)} MF| |u| u^2 d\beta_0(u)dt \right]$$

$$= E \left[ 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R} \setminus \{0\}} (\partial_{(t,u)} F)(\partial_{(t,u)} MF)u^2 d\nu(t, u) \right]$$

$$+ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ |\partial_{(t,u)} F 1_{\{u:|u|>0\}}(u)|^2 |\partial_{(t,u)} MF| |u| u^2 d\nu(t, u),$$

where $d\nu = d\beta_0(u)dt$. \(\square\)
Assume \( F \) is a pure jump Lévy process, we define
\[
D_{(t,u)} F(x) = F(x + u\delta_t) - F(x),
\]
so
\[
\partial_{(t,u)} F(x) = \frac{D_{(t,u)} F(x)}{u}.
\]
Set \( \mu(dz) = d\nu(t, u) \), and \( Z = (t, u) \in \mathbb{R}^2 \). Hence we get the main result in Peccati, Solé, Taqqu and Utzet [35] below when \( F \) is a pure jump Lévy process.

\[
d_W(F, X) \leq E[|1 - \int_Z (D_z F)(-D_z L^{-1} F)\mu(dz)|] + \int_Z E[|D_z F|^2|D_z L^{-1} F|]\mu(dz).
\]

There is an extra negative sign before \( D_z L^{-1} F \). This is because they define \( L^{-1} F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n) \) while we define \( M F = \sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n) \). \( L \) operator is for Poisson functional while \( L \) operator is for Brownian functional.

Moreover, we have jump size \( u \) in our main result because we define our random measure below differently.

\[
M(E) = \sigma \int_{-\infty}^{+\infty} 1_E(t, 0)dB(t) + \int_{\mathbb{R}^2} u 1_E(t, u)dN_0(t, u).
\]

In Peccati, Solé, Taqqu and Utzet [35], they define their random measure as

\[
\int_{\mathbb{R}^2} 1_E(t, u)dN_0(t, u).
\]

There is no \( u \) in it, so it is like \( u = 1 \). Here we can ignore the first term \( \sigma \int_{-\infty}^{+\infty} 1_E(t, 0)dB(t) \) because we assume \( F \) has the only pure jump part.

Now we show that Stein bounds for functionals of the pure Brownian motion Lévy process are particular cases of the general bound in Theorem 3.3.1.

**Corollary 3.3.2.** Let \( F \in L^2(\mathcal{F}', \Lambda), E(F) = 0, \) and \( X \sim \mathcal{N}(0, 1) \). \( F \) has the only pure Brownian motion part. Then

\[
d_W(F, X) \leq E[|1 - \int_{t \in \mathbb{R}} (\partial_{(t,0)} F)(\partial_{(t,0)} M F)\sigma^2 dt|].
\]
**Proof.** In Theorem 3.3.1, our main result is

\[
    d_W(F, X) \leq E \left[ \left| 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)} MF \partial_{(t,u)} F \, d\lambda(t,u) \right| \right] + \\
    \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ \left| \partial_{(t,u)} MF \left( \partial_{(t,u)} F \mathbf{1}_{\{u:|u|>0\}}(u) \right)^2 \right| \right] \, d\lambda(t,u),
\]

where

\[
    d\lambda(t, u) = dt(1 + u^2)d\beta(u)
\]

and

\[
    \sigma^2 = \beta(\{0\}) = \int_{u=0} (1 + u^2)d\beta(u).
\]

The Lévy measure is denoted by \(\beta_0\) which is a positive measure on \(\mathbb{R}_*(= \mathbb{R} \setminus \{0\})\) defined by

\[
    \beta_0(E) = \int_E \frac{1 + u^2}{u^2}d\beta(u)
\]

for \(E \in \mathcal{B}(\mathbb{R}_*)\).

Assume \(F\) has only the diffusion part, and using

\[
    d\lambda(t, u) = dt(1 + u^2)d\beta(u),
\]

and

\[
    \sigma^2 = \beta(\{0\}) = \int_{u=0} (1 + u^2)d\beta(u).
\]
Then we get
\[
d_W(F, X) \leq E \left[ 1 - \int_{t \in \mathbb{R}} \int_{u=0} \partial_{(t,u)} MF \partial_{(t,u)} F (1 + u^2) d\beta(u) dt \right]
\]
\[
= E \left[ 1 - \int_{t \in \mathbb{R}} (\partial_{(t,0)} F)(\partial_{(t,0)} MF) \sigma^2 dt \right].
\]

\[
\square
\]

Nourdin and Peccati [29] obtain the explicit bound for a Brownian motion functional given below:
\[
d_W(F, X) \leq E \left[ |1 - \int_{K} (DF)(-DL^{-1}F) dt| \right],
\]
(3.93)
where $K \subset \mathbb{R}$.
Compared with the equation 3.93, we have an extra $\sigma^2$. This is because we define the random measure below differently.
\[
M(E) = \sigma \int_{-\infty}^{+\infty} 1_E(t,0) dB(t) + \int_{\mathbb{R}^2} u1_E(t,u) dN_0(t,u).
\]
In Nourdin and Peccati [29], they define their random measure as
\[
\int_{K} 1_E(t,0) dB(t).
\]
Moreover, there is an extra negative sign before $DL^{-1}F$. This is because they define $L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n)$ while we define $MF = \sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n)$.

Let $F$ be given in $L^2(\mathcal{Z}', \Lambda)$. Then $F$ has a unique Wiener-Itô chaos expansion (3.17).
For the purpose of exposition, set $F = I_1(F_1)$, which means $F_0 = 0$, and $F_n = 0$, for $n \geq 2$, and then we compute Wasserstein distance, $d_W(F, X)$ using (3.80) and (3.81) in the following example.

Based on (3.13), (3.16) $F$ is given below.
Set
\[
F_1(t) = F_1(t, u) = \frac{1}{\sigma \sqrt{T}} 1_{[0,T]}(t) 1_{\{0\}}(u) + \varepsilon \frac{1}{\sqrt{T}} 1_{[0,T]}(t) 1_{\{u|u|>0\}}(u).
\]
Chapter 3. Stein’s method and normal approximation of Lévy functionals

\[ F = I_1(F_1) = \int_0^T \frac{1}{\sigma \sqrt{T}} \sigma dB(t) + \int_0^T \int_{\mathbb{R}\setminus\{0\}} \varepsilon \sqrt{T} \mu dN_0(t, u) \]
\[ = \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} F_1(t, u) dM(t, u) \]
\[ = \int_{-\infty}^{+\infty} F_1(t) d\tilde{X}(t) \]
\[ = \int_{-\infty}^{+\infty} F_1(t)(\tilde{X}(t) - \tau_1) dt. \quad (3.94) \]

Here we choose Lévy process \( \tilde{X}(t) \) as a jump-diffusion process without drift term, and pure jump part is a Poisson process with intensity parameter 1.

**Example 3.3.1.** \( F = I_1(F_1) \) given in (3.94), and \( X \sim \mathcal{N}(0,1) \)

\[ d_W(F, X) \leq \varepsilon^2 + \varepsilon^3 \frac{1}{\sqrt{T}}. \quad (3.95) \]

As \( \varepsilon \to 0 \), the upper bound is zero.

To see this, we apply upper bound Theorem 3.3.1. According to (3.48) and (3.60), we get

\[ \partial_{(t,u)} F = F_1(t, u). \quad (3.97) \]

\[ MF = I_1(F_1). \]

\[ \partial_{(t,u)} MF = F_1(t, u). \quad (3.98) \]
Moreover, due to the (3.12) and (2.16)

\[ d\lambda(t, u) = dt \otimes (1 + u^2)d\beta(u) \]

and from (3.2)

\[ \sigma^2 = \beta(\{0\}) = \int_{u \in \mathbb{R}} 1_{\{0\}}(u)(1 + u^2)d\beta(u). \]  \hspace{1cm} (3.99)

For centered Poisson,

\[ \beta = \frac{1}{2}\delta_1, \]

where \( \delta_a \) is the Dirac measure concentrated on the point \( a \).

Substitute (3.97) and (3.98) into \( A \), we have

\[
A = E \left[ \left. 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} F_1(t, u)^2d\lambda(t, u) \right|_{t \in [0,T]} \right]
\]

\[
= \left| 1 - \left\{ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \frac{1}{\sigma^2T}1_{\{0,T\}}(t)1_{\{0\}}(u)d\lambda(t, u) + \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \varepsilon^2 \frac{1}{T}1_{\{0,T\}}(t)1_{\{u:|u|>0\}}(u)d\lambda(t, u) \right\} \right|
\]

\[
= \left| 1 - \left\{ \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \frac{1}{\sigma^2T}1_{\{0,T\}}(t)1_{\{0\}}(u)dt \otimes (1 + u^2)d\beta(u) + \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \varepsilon^2 \frac{1}{T}1_{\{0,T\}}(t)1_{\{u:|u|>0\}}(u)dt \otimes (1 + u^2)d\beta(u) \right\} \right|
\]

Using (3.99),

\[
= \left| 1 - \left\{ \int_0^T \frac{1}{\sigma^2T}dt \int_{u \in \mathbb{R}} 1_{\{0\}}(u)d\beta(u) + \int_0^T \varepsilon^2 \frac{1}{T}dt \int_{u \in \mathbb{R}} 1_{\{u:|u|>0\}}(u)(1 + u^2)\frac{1}{2}d\delta_1(u) \right\} \right|
\]

\[
= \left| 1 - \left\{ \int_0^T \frac{1}{T}dt + \int_0^T \varepsilon^2 \frac{1}{T}dt \right\} \right|
\]

\[= \varepsilon^2. \]  \hspace{1cm} (3.100)
Substitute (3.97) and (3.98) into $B$, we have

$$B = \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ \left| \frac{\partial}{\partial (t,u)} MF \left( \frac{\partial}{\partial (t,u)} F 1_{\{|u|>0\}}(u) \right) \right| \right] d\lambda(t,u)$$

$$= \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ F_1(t,u)^3 1_{\{|u|>0\}}(u) \right] d\lambda(t,u)$$

$$= \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} E \left[ F_1(t,u)^3 1_{\{|u|>0\}}(u) \right] d\lambda(t,u)$$

$$= \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \varepsilon^3 \left( \frac{1}{\sqrt{T}} \right)^3 1_{[0,T]}(t) 1_{\{|u|>0\}}(u) \right| dt \otimes (1 + u^2) d\beta(u)$$

$$= \int_{t \in \mathbb{R}} \varepsilon^3 \left( \frac{1}{\sqrt{T}} \right)^3 1_{[0,T]}(t) \right| dt \int_{u \in \mathbb{R}} 1_{\{|u|>0\}}(u) \left| (1 + u^2) \frac{1}{2} d\delta(u) \right|$$

$$= \int_0^T \varepsilon^3 \left( \frac{1}{\sqrt{T}} \right)^3 dt$$

$$= \varepsilon^3 \frac{1}{\sqrt{T}}. \quad (3.101)$$

Substitute (3.100) and (3.101) into (3.96), we get (3.95).

Now we consider Brownian motion white noise functional, and we introduce an example using $A^{-p}\delta_s$ as a kernel function.

**Example 3.3.2.** Let

$$F = \frac{1}{|\delta_s|^{-p}} I_1(A^{-p}\delta_s),$$

where $|\delta_s|^{-p} = \int_{-\infty}^{+\infty} A^{-p}\delta_s(t)A^{-p}\delta_s(t)dt$, and $X \sim \mathcal{N}(0,1)$.

We have the upper bound

$$d_W(F, X) \leq 0.$$

To see this, we consider Brownian motion white noise functional.

According to Example 3.2.2, Theorem 3.2.3, and the section 3.2.2.1, we know that

$$\delta_t = \sum_{n=0}^{\infty} (\delta_t, e_n) e_n.$$
Chapter 3. Stein’s method and normal approximation of Lévy functionals

\[ A^{-p}\delta_s = \sum_{m=0}^{\infty} (2m + 2)^{-p}(\delta_s, e_m)e_m. \]

Using the fact,

\[
(\delta_t, A^{-p}\delta_s) = \int_{-\infty}^{+\infty} \delta_{t'}(t')A^{-p}\delta_s(t')dt'
\]

\[
= \int_{-\infty}^{+\infty} \sum_{n=0}^{\infty} (\delta_t, e_n)\delta_{t'}(t) \sum_{m=0}^{\infty} (2m + 2)^{-p}(\delta_s, e_m)e_m(t')dt'
\]

\[
= \sum_{n=0}^{\infty} (2n + 2)^{-p} \int_{-\infty}^{+\infty} (\delta_t, e_n)\delta_{t'}(t) e_n(t')dt'
\]

\[
= \sum_{n=0}^{\infty} (2n + 2)^{-p} e_n(s)e_n(t)
\]

\[
= \sum_{n=0}^{\infty} (2n + 2)^{-p} e_n(t)
\]

\[
= A^{-p}\delta_s(t).
\]

For the sup-norm of the Hermite functions we have

\[
\|e_n\|_\infty = O(n^{-1/12}).
\]

Hence \((\delta_t, A^{-p}\delta_s) < \infty\) if \(p + \frac{1}{6} > 1\), i.e., \(p > \frac{5}{6}\).

Then,

\[
\partial_t F = \frac{1}{|\delta_s|_p} \partial_t I_1(A^{-p}\delta_s) = \frac{1}{|\delta_s|_p} (\delta_t, A^{-p}\delta_s) = \frac{1}{|\delta_s|_p} A^{-p}\delta_s(t).
\]

\[
MF = \frac{1}{|\delta_s|_p} M I_1(A^{-p}\delta_s) = \frac{1}{|\delta_s|_p} I_1(A^{-p}\delta_s).
\]

\[
\partial_t MF = \frac{1}{|\delta_s|_p} \partial_t M I_1(A^{-p}\delta_s) = \frac{1}{|\delta_s|_p} (\delta_t, A^{-p}\delta_s) = \frac{1}{|\delta_s|_p} A^{-p}\delta_s(t).
\]
Set $\sigma = 1$, then,

\[
d_{W}(F, X) \leq E \left[ 1 - \int_{-\infty}^{+\infty} \partial_t MF\partial_tF dt \right]
\]
\[
= \left| 1 - \frac{1}{|\delta_s|^{-p}} \int_{-\infty}^{+\infty} A^{-p} \delta_s(t) A^{-p} \delta_s(t) dt \right|
\]
\[
= \left| 1 - \frac{1}{|\delta_s|^2_{-p}} |A^{-p}\delta_s|_0^2 \right|
\]
\[
= \left| 1 - \frac{1}{|\delta_s|^2_{-p}} |\delta_s|^2_{-p} \right|
\]
\[
= 0.
\]

This lemma will be used later.

**Lemma 3.3.4.** Let $I_{1J}(f_{iJ}) \in L^2(\mathcal{X}^i, \Lambda)$, and

\[
I_{1J}(f_{iJ}) = \int_{\mathbb{R}^2} u f_{iJ}(t,u)dN_0(t,u).
\]

We have

\[
E[I_{1J}(f_{1J})I_{1J}(f_{2J})I_{1J}(f_{3J})I_{1J}(f_{4J})]
\]
\[
= (f_{1J}, f_{2J})_{\lambda J}(f_{3J}, f_{4J})_{\lambda J}
\]
\[
+ (f_{1J}, f_{3J})_{\lambda J}(f_{2J}, f_{4J})_{\lambda J}
\]
\[
+ (f_{1J}, f_{4J})_{\lambda J}(f_{2J}, f_{3J})_{\lambda J}
\]
\[
+ (f_{1J}, f_{2J}, f_{3J}, f_{4J})_{\lambda J},
\]

where

\[
(f_{iJ}, f_{jJ})_{\lambda J} = \int_{\mathbb{R}^2} f_{iJ}(t,u)f_{jJ}(t,u)u^2 d\beta_0(u) dt, \quad 1 \leq i, j \leq 4
\]
\[
= (uf_{iJ}, uf_{jJ})_\nu,
\]
and \( d\nu = d\beta_0 dt \). \( \beta_0 \) is Lévy measure. \( d\lambda = u^2 d\beta_0(u) dt \), and recall that
\[
d\beta_0(u) = \frac{1 + u^2}{u^2} d\beta(u).
\]

**Proof.** We define
\[
I_{1J}(f_{iJ}) = \int_{\mathbb{R}_2^*} u f_i(t,u)dN_0(t,u).
\]

\( B = \{ B(t) : t \in \mathbb{R} \} \) is a 1-dimensional Wiener process, independent of the system of \( \{ N(E) : E \in \mathcal{B}_b(\mathbb{R}_2^*) \} \). Let \( \mathcal{B}_b(\mathbb{R}_2^*) \) be the class of all bounded Borel subsets \( E \) of \( \mathbb{R}_2^* = \mathbb{R}_2 \setminus \{(t,0) : t \in \mathbb{R}\} \). Then \( N(E;\cdot) \) is Poisson distributed with intensity measure \( \nu \). The system of \( \{ N(E;x) - \nu(E) : E \in \mathcal{B}_b(\mathbb{R}_2^*); x \in \mathcal{S}' \} \) forms an independent random measure (Definition 2.2.25) with zero mean. We denote "\( N(E;x) - \nu(E) \)" by \( N_0(E;x) \). Therefore, \( E[N_0(E)N_0(F)] = \nu(E \cap F) \).

Also, define
\[
(f_{iJ}, f_{jJ})_{\lambda_J} = \int_{\mathbb{R}_2^*} f_{iJ}(t,u)f_{jJ}(t,u)u^2 d\beta_0(u) dt, \quad 1 \leq i, j \leq 4,
\]
where \( d\nu = d\beta_0 dt \) and \( \beta_0 \) is Lévy measure.

Let us first calculate \( E[I_{1J}(f_{iJ})I_{1J}(f_{2J})I_{1J}(f_{3J})I_{1J}(f_{4J})] \) for simple functions multiplied by \( 1/u \).

\[
f_{iJ}^{(n)}(t,u) = \frac{1}{u} \sum_{j=1}^{n} a_{ij} 1_{A_j},
\]
where \( a_{ij} \in \mathbb{R} \) are constants, and
\[
I_{1J}(f_{iJ}^{(n)}) = \sum_{j=1}^{n} a_{ij} N_0(A_j).
\]
Then

\[
E[I_{1,J}(f_{1,J}^{(n)})I_{1,J}(f_{2,J}^{(n)})I_{1,J}(f_{3,J}^{(n)})I_{1,J}(f_{4,J}^{(n)})]
= E\left[ \sum_{i=1}^{n} a_{1i}N_0(A_i) \sum_{j=1}^{n} a_{2j}N_0(A_j) \sum_{k=1}^{n} a_{3k}N_0(A_k) \sum_{l=1}^{n} a_{4l}N_0(A_l) \right].
\] (3.102)

Since Poisson random measure is a martingale measure, we can decompose the formula (3.102) into four parts.

\[
\sum_{i=j \neq k=l} + \sum_{i=k \neq j=l} + \sum_{i=l \neq k=j} + \sum_{i=j=k=l}.
\]

All these sums must be understood as being defined over indices \(\{i, j, k, l\} \in \{1, \ldots, n\}^4\).

Therefore,

\[
E[I_{1,J}(f_{1,J}^{(n)})I_{1,J}(f_{2,J}^{(n)})I_{1,J}(f_{3,J}^{(n)})I_{1,J}(f_{4,J}^{(n)})]
= (u_{f_{1,J}^{(n)}}, u_{f_{2,J}^{(n)}})_\nu (u_{f_{3,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu
+ (u_{f_{1,J}^{(n)}}, u_{f_{3,J}^{(n)}})_\nu (u_{f_{2,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu
+ (u_{f_{1,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu (u_{f_{2,J}^{(n)}}, u_{f_{3,J}^{(n)}})_\nu
- 3 \sum_{i=1}^{n} a_{1i}a_{2i}a_{3i}a_{4i} \left( E[N_0(A_i)^2] \right)^2
+ \sum_{i=1}^{n} a_{1i}a_{2i}a_{3i}a_{4i} E[N_0(A_i)^4]
= (u_{f_{1,J}^{(n)}}, u_{f_{2,J}^{(n)}})_\nu (u_{f_{3,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu
+ (u_{f_{1,J}^{(n)}}, u_{f_{3,J}^{(n)}})_\nu (u_{f_{2,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu
+ (u_{f_{1,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu (u_{f_{2,J}^{(n)}}, u_{f_{3,J}^{(n)}})_\nu
+ (u_{f_{1,J}^{(n)}}, u_{f_{2,J}^{(n)}}^2, u_{f_{3,J}^{(n)}}, u_{f_{4,J}^{(n)}})_\nu.
\] (3.103)

Now let \(n \to \infty\), and choose the \(a_{ij}\) in such a way that \(\sum_{j=1}^{n} a_{ij}1_{A_j} \to u_{f_{i,J}(t,u)}\) pointwise almost everywhere with respect to \(\nu\), and the lemma is proved. Such a choice for the \(a_{ij}\) is possible because the limit function is measurable, and by Theorem 5.2 in Kingman and Taylor [19], all measurable functions can be expressed as the pointwise
limit of simple functions.

Line (3.103) follows because we use the two facts:

(i) For a random variable \( X \) with zero mean,

\[
E[X^4] = K_4(X) + 3(K_2(X))^2,
\]

where \( K_n(X) \) is the nth cumulant of \( X \). For the normal distribution, \( K_n = 0, n \geq 3 \).

(ii) The cumulants of a Poisson random variable are all equal to the mean.

From (i) and (ii),

\[
E[N_0(A_i)^4] = \nu(A_i) + 3\nu(A_i)^2,
\]

where \( \nu(A_i) = E[N_0(A_i)^2] \). Therefore,

\[
\sum_{i=1}^{n} a_1 a_2 a_3 a_4 E[N_0(A_i)]^3 - 3 \sum_{i=1}^{n} a_1 a_2 a_3 a_4 (E[N_0(A_i)^2])^2 = \sum_{i=1}^{n} a_1 a_2 a_3 a_4 \nu(A_i) = (u^2 f_{1,1}^{(n)} f_{2,1}^{(n)}, u^2 f_{3,1}^{(n)} f_{4,1}^{(n)}) \nu.
\]

□

This lemma will be used later.

**Lemma 3.3.5.** Let \( I_1(f_i) = I_{1G}(f_i) + I_{1J}(f_i) \), where \( I_{1G}(f_i) = \int_{-\infty}^{+\infty} f_i(t,0) dB(t) \), and \( I_{1J}(f_i) = \int_{\mathbb{R}^2} u f_i(t,u) dN_0(t,u) \).

We have

\[
E[I_1(f_1) I_1(f_2) I_1(f_3) I_1(f_4)] = (f_1, f_2)(f_3, f_4) + (f_1, f_3)(f_2, f_4) + (f_1, f_4)(f_2, f_3) + (f_1, f_2, f_3, f_4) \lambda,
\]

(3.104)
where

\[(f_i, f_j) = \int_{\mathbb{R}^2} f_i(t, u)f_j(t, u)d\lambda(t, u) = \int_{\mathbb{R}} f_i(t, 0)f_j(t, 0)\sigma^2dt + \int_{\mathbb{R}^2} f_i(t, u)f_j(t, u)(1 + u^2)d\beta dt,\]

and

\[(f_{1J}f_{2J}, f_{3J}f_{4J})_{\lambda_J} = \int_{\mathbb{R}^2} f_{1J}f_{2J}f_{3J}f_{4J}(1 + u^2)d\beta dt,\]

where \(\mathbb{R}_2^* = \mathbb{R}^2 \setminus \{(t, 0) : t \in \mathbb{R}\}\), \(d\lambda = (1+u^2)d\beta dt\), and \(\sigma^2 = \beta(\{0\}) = \int_{u=0}(1+u^2)d\beta(u)\).

**Proof.** According to the equation (3.16), and \(\mathbb{R}_2^* = \mathbb{R}^2 \setminus \{(t, 0) : t \in \mathbb{R}\}\), we define

\[I_1(f_i) = I_{1G}(f_{iG}) + I_{1J}(f_{iJ}),\]

where

\[I_{1G}(f_{iG}) = \int_{-\infty}^{+\infty} f_i(t, 0)\sigma dB(t),\]

and

\[I_{1J}(f_{iJ}) = \int_{\mathbb{R}^2} u f_i(t, u)dN_0(t, u).\]

\(B = \{B(t) : t \in \mathbb{R}\}\) is a 1-dimensional Wiener process, independent of the system of \(\{N(E) : E \in \mathcal{B}_b(\mathbb{R}_2^*)\}\). Let \(\mathcal{B}_b(\mathbb{R}_2^*)\) be the class of all bounded Borel subsets \(E\) of \(\mathbb{R}_2^* = \mathbb{R}^2 \setminus \{(t, 0) : t \in \mathbb{R}\}\). Then \(N(E; \cdot)\) is Poisson distributed with intensity measure \(\nu\). The system of \(\{N(E; x) - \nu(E) : E \in \mathcal{B}_b(\mathbb{R}_2^*); x \in \mathcal{I}\}\) forms an independent random measure (Definition 2.2.25) with zero mean. We denote "\(N(E; x) - \nu(E)\)" by \(N_0(E; x)\). Therefore, \(E[N_0(E)N_0(F)] = \nu(E \cap F)\).
Define

$$\sigma_{ij} = \int_{-\infty}^{+\infty} f_iG(t)f_jG(t)\sigma^2 dt, \quad 1 \leq i, j \leq 4,$$

where

$$\sigma^2 = \beta(\{0\}) = \int_{u=0}^\infty (1 + u^2)d\beta(u).$$

Then $I_{1G}(f_1G), I_{1G}(f_2G), I_{1G}(f_3G), \text{ and } I_{1G}(f_4G)$ are jointly Gaussian with zero means and covariances given by

$$\text{COV}(I_{1G}(f_iG), I_{1G}(f_jG)) = \sigma_{ij}, \quad 1 \leq i, j \leq 4.$$

Also, define

$$(f_iJ, f_jJ)_{\lambda_J} = \int_{R^2} f_iJ(t, u)f_jJ(t, u)(1 + u^2)d\beta dt, \quad 1 \leq i, j \leq 4,$$

where $d\lambda = (1 + u^2)d\beta dt$ and $\beta_0$ is Lévy measure, and recall that

$$d\beta_0(u) = \frac{1 + u^2}{u^2}d\beta(u).$$
We note that

\[
E[I_1(f_1)I_1(f_2)I_1(f_3)I_1(f_4)] \\
= E[I_{1G}(f_1G)I_{1G}(f_2G)I_{1G}(f_3G)I_{1G}(f_4G)] \\
+ E[I_{1G}(f_1G)I_{1G}(f_2G)I_{1J}(f_3J)I_{1J}(f_4J)] \\
+ E[I_{1G}(f_1G)I_{1G}(f_3G)I_{1J}(f_2J)I_{1J}(f_4J)] \\
+ E[I_{1G}(f_1G)I_{1G}(f_4G)I_{1J}(f_2J)I_{1J}(f_3J)] \\
+ E[I_{1G}(f_2G)I_{1G}(f_3G)I_{1J}(f_1J)I_{1J}(f_4J)] \\
+ E[I_{1G}(f_3G)I_{1G}(f_4G)I_{1J}(f_1J)I_{1J}(f_2J)] \\
+ E[I_{1J}(f_1J)I_{1J}(f_2J)I_{1J}(f_3J)I_{1J}(f_4J)].
\]

Using Lemma 3.3.4 and Itô isometry, then

\[
E[I_1(f_1)I_1(f_2)I_1(f_3)I_1(f_4)] \\
= \sigma_{12}\sigma_{34} + \sigma_{13}\sigma_{24} + \sigma_{14}\sigma_{23} \\
+ \sigma_{12}(f_{3J},f_{4J})\lambda_J + \sigma_{13}(f_{2J},f_{4J})\lambda_J \\
+ \sigma_{14}(f_{2J},f_{3J})\lambda_J + \sigma_{23}(f_{1J},f_{4J})\lambda_J \\
+ \sigma_{24}(f_{1J},f_{3J})\lambda_J + \sigma_{34}(f_{1J},f_{2J})\lambda_J \\
+ (f_{1J},f_{2J})\lambda_J(f_{3J},f_{4J})\lambda_J \\
+ (f_{1J},f_{3J})\lambda_J(f_{2J},f_{4J})\lambda_J \\
+ (f_{1J},f_{4J})\lambda_J(f_{2J},f_{3J})\lambda_J \\
+ (f_{1J},f_{2J},f_{3J},f_{4J})\lambda_J.
\]
Finally, we have

\[
E[I_1(f_1)I_1(f_2)I_1(f_3)I_1(f_4)] \\
= (\sigma_{12} + (f_1J, f_2J)_{\lambda_J})(\sigma_{34} + (f_3J, f_4J)_{\lambda_J}) \\
+ (\sigma_{13} + (f_1J, f_3J)_{\lambda_J})(\sigma_{24} + (f_2J, f_4J)_{\lambda_J}) \\
+ (\sigma_{14} + (f_1J, f_4J)_{\lambda_J})(\sigma_{23} + (f_2J, f_3J)_{\lambda_J}) \\
+ (f_1J, f_2J, f_3J, f_4J)_{\lambda_J},
\]

Example 3.3.3. Let \( F = I_2(f_2) \in L^2(\mathcal{F}', \Lambda) \), and \( X \sim \mathcal{N}(0, 1) \)

\[
d_W(F, X) \leq A + V,
\]

where \( A \) and \( B \) are given by (3.107) and (3.109).

To see this, to compute the upper bound of \( F \), first, using (3.2.4) and (3.2.5), we calculate

\[
\partial_{(t,u)} F = 2I_1(K_{\delta(t,u), f_2}),
\]

and

\[
MF = \frac{1}{2} I_2(f_2),
\]

and

\[
\partial_{(t,u)} MF = I_1(K_{\delta(t,u), f_2}).
\]
According to (3.45), set $\psi = \delta_{(t_2, u_2)}$, we have

\[
K_{\delta_{(t,u)}, f_2}(\psi) = \left(\delta_{(t,u)} \natural \delta_{(t_2, u_2)}, f_2\right) = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \delta_{(t,u)}((t', u')) \otimes \delta_{(t_2, u_2)}((t_2', u_2')) f_2((t', u'), (t_2', u_2')) d\lambda((t', u')) d\lambda((t_2', u_2')) = f_2((t, u), (t_2, u_2)).
\]  

(3.105)

According to (3.82) and (3.83),

\[
d_W(F, X) \leq \sqrt{\mathbb{E}\left[\left(1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t,u)}MF \partial_{(t,u)}F d\lambda(t, u)\right)^2\right]} + \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \mathbb{E}\left[\left|\partial_{(t,u)}MF \partial_{(t,u)}F 1_{\{u: |u| > 0\}}(u)\right|^2 \right] |u| d\lambda(t, u) = A + B.
\]  

(3.106)

\[
A = \sqrt{\mathbb{E}\left[\left(1 - 2 \int_{\mathbb{R}^2} I_1(K_{\delta_{(t,u)}, f_2}) I_1(K_{\delta_{(t,u)}, f_2}) d\lambda(t, u)\right)^2\right]} = \sqrt{1 - 4\mathbb{E}[C_1] + 4\mathbb{E}[C_2]},
\]  

(3.107)

where

\[
C_1 = \int_{\mathbb{R}^2} I_1(K_{\delta_{(t,u)}, f_2}) I_1(K_{\delta_{(t,u)}, f_2}) d\lambda(t, u),
\]

and using (3.105) and Itô isometry, we have

\[
\mathbb{E}[C_1] = \int_{\mathbb{R}^2} \mathbb{E}[I_1(K_{\delta_{(t,u)}, f_2}) I_1(K_{\delta_{(t,u)}, f_2})] d\lambda(t, u) = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f_2((t, u), (t_2, u_2)) f_2((t, u), (t_2, u_2)) d\lambda(t_2, u_2) d\lambda(t, u),
\]

and

\[
C_2 = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} I_1(K_{\delta_{(t,u)}, f_2}) I_1(K_{\delta_{(t,u)}, f_2}) I_1(K_{\delta_{(s,v)}, f_2}) I_1(K_{\delta_{(s,v)}, f_2}) d\lambda(t, u) d\lambda(s, v),
\]
Chapter 3. Stein’s method and normal approximation of Lévy functionals

and using (3.104) and (3.105), we have

\[
E[C_2] = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} E[I_1(K_{\delta(t,u),f_2})I_1(K_{\delta(t,u),f_2})I_1(K_{\delta(s,v),f_2})I_1(K_{\delta(s,v),f_2})]d\lambda(t,u)d\lambda(s,v)
\]

\[
= 3 \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f_2((t,u), (t_2, u_2))f_2((t,u), (t_2, u_2)) f_2((s,v), (t_2, u_2))f_2((s,v), (t_2, u_2))d\lambda(t_2, u_2)d\lambda(t, u)d\lambda(s,v)
\]

\[
+ \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f_2((t,u), (t_2, u_2))f_2((t,u), (t_2, u_2)) f_2((s,v), (t_2, u_2))f_2((s,v), (t_2, u_2))d\lambda(t_2, u_2)d\lambda(t, u)d\lambda(s,v),
\]

where \(\mathbb{R}^2_0 = \mathbb{R}^2 \setminus \{(t,0) : t \in \mathbb{R}\}\).

Moreover,

\[
B = 4 \int_{\mathbb{R}^2} E\left[ I_1(K_{\delta(t,u),f_2})\left(I_1(K_{\delta(t,u),f_2})1_{\{u:|u|>0\}}(u)\right)^2 |u|\right]d\lambda(t,u)
\]

\[
\leq 4 \int_{\mathbb{R}^2} E\left[ I_1(K_{\delta(t,u),f_2})^3 |u|d\lambda(t,u)
\right.
\]

\[
\leq 4 \int_{\mathbb{R}^2} \left[ E(I_1(K_{\delta(t,u),f_2})^4\right]^{3/4} |u|d\lambda(t,u)
\]

(3.108)

\[
\leq 4 \int_{\mathbb{R}^2} \left[ 4 \int_{\mathbb{R}^2} f_2((t,u), (t_2, u_2))f_2((t,u), (t_2, u_2)) f_2((s,v), (t_2, u_2))f_2((s,v), (t_2, u_2))d\lambda(t_2, u_2)d\lambda(t, u)
\]

\[
\right)^{3/4} |u|d\lambda(t,u)
\]

(3.109)

\[
= V.
\]

Line (3.108) follows from Hölder’s inequality.

Line (3.109) follows from (3.104) and (3.105). Here we computes the upper bound for a double Lévy integral using our main theorem, which shows that our approach unifies the Brownian and pure jump components in a natural way, and simplifies the computation.
3.4 The difference and connection between white noise analysis and Malliavin calculus

Instead of Malliavin calculus, we choose to use white noise analysis provided by Lee and Shih [24] to derive the upper bound of Wasserstein distance for Lévy functionals because Lee and Shih [24] provide the celebrated white noise derivative formula (3.64), based on the concept of a directional derivative. In their paper, they rigorously derive the white noise derivative for Lévy functionals as a whole from the beginning to the end while Solé, Utzet and Vives [40] derive the Malliavin derivative for the diffusion part and the pure jump part separately, and they borrow the increment quotient operator from Lee and Shih [24]. Therefore, we believe that Lee and Shih [24] provide a more unified and rigorous framework to work with. By means of a simple example, we show below the difference and connection between white noise analysis and Malliavin calculus.

3.4.1 The white noise derivative is a directional derivative

By (3.17), let \( F = \sum_{n=0}^{\infty} \oplus I_n(F_n) \).
So far we have two white noise derivative formulas. First, based on equation (3.48), we represent white noise derivative in terms of Lévy-Itô decomposition (3.17):

\[
\partial \delta_{(t,u)} F = (\delta_{(t,u)}, F_1) + \sum_{n=2}^{\infty} nI_{n-1}(K\delta_{(t,u)}, F_n)
\]

\[
= \sum_{n=1}^{\infty} nI_{n-1}(K\delta_{(t,u)}, F_n), \tag{3.111}
\]

Second, using equation (3.64), we can express the white noise derivative as a directional derivative:

\[
\partial_{(t,u)} F(x) = \left( \frac{F(x + u\delta_t) - F(x)}{u} \right) 1_{\{u:|u|>0\}}(u) + (DF(x)\delta_t) 1_{\{0\}}(u), \tag{3.112}
\]

where \( \partial_{(t,u)} \) is a shorthand of \( \partial\delta_{(t,u)} \), and in the second term in RHS, \( \delta_t \) is the Dirac measure concentrated on the point \( t \).

To derive Stein bound for Lévy processes, we need to use both versions of Malliavin derivative or white noise derivative, Lévy-Itô decomposition and directional derivative. In Geiss and Laukkarin [8], they show the domain for Malliavin derivative in terms of
Lévy-Itô decomposition is $D^{1,2} \in L^2(\Omega, \mathcal{F}, P)$, and the domain for Malliavin derivative in terms of a directional derivative is smooth functions, $\mathcal{S}$. The intersection of two domains is $MD = L^2(\Omega, P) \cap \mathcal{S}$; on the other hand, according to Theorem 3.3.1, the intersection of two domains for white noise derivatives is $WD = L^2(\mathcal{S}', \Lambda) \cap \mathcal{S}$. We believe that $MD \subset WD$.

Moreover, in Lee and Shih [24], at page 39, they clarify that, for convenience, they denote $\partial_t^\delta, \partial^{\ast}_t, \partial_{t,u}^\delta$, and $\partial^{\ast}_{t,u}$ respectively by $\partial_t, \partial_t^\ast, \partial_{t,u}$ and $\partial^{\ast}_{t,u}$ for $t \in \mathbb{R}$ and $(t, u) \in \mathbb{R}^2$. We use the same abbreviations in this thesis.

Now we will check that (3.111) and (3.112) are equal in a simple example.

**Example 3.4.1.** Based on (3.13), and (3.16), $F$ is given below.

\[
F = I_1(F_1) = \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} F_1(t, u) dM(t, u) \\
= \int_{-\infty}^{+\infty} F_1(t) d\tilde{X}(t) \\
\overset{d}{=} \int_{-\infty}^{+\infty} F_1(t)(\dot{X}(t) - \tau_1) dt,
\]

where $\overset{d}{=} \text{ in the final equation means equality in distribution.}$

According to (3.111), we know that every time we differentiate $F$, we reduce one integral. Since in this simple example $F$ has the only one integral, we get one deterministic function after we differentiate $F$:

\[
\partial_{t,u} F = F_1(t, u).
\]

On the other hand, by (3.112), and (3.113), we get

\[
\partial_{t,u} F(x) = A + B,
\]
where $A$ and $B$ are calculated below.

First, for the pure jump part, where $x$ is the Lévy white noise $\dot{X}(t)$, we get

\[
A = \left( \frac{F(x + u\delta_t) - F(x)}{u} \right) 1_{\{|u| > 0\}}(u) \quad \text{using (3.113)}
\]

\[
= \left( \int_{-\infty}^{+\infty} F_1(t) (\dot{X}(t) - \tau_1 + u\delta_t) dt - \int_{-\infty}^{+\infty} F_1(t) (\dot{X}(t) - \tau_1) dt \right) 1_{\{|u| > 0\}}(u)
\]

\[
= \int_{-\infty}^{+\infty} F_1(t) u\delta_t dt u 1_{\{|u| > 0\}}(u) \quad \text{if and } \int_{-\infty}^{+\infty} dt \text{ cancel each other out}
\]

\[
= F_1(t) 1_{\{|u| > 0\}}(u). \quad (3.116)
\]

Then, for the diffusion part, we compute the directional derivative and get

\[
B = (DF(x)\delta_t) 1_{\{0\}}(u) \quad \delta_t \text{ is the direction}
\]

\[
= \frac{d}{d\varepsilon} [F(x + \varepsilon\delta_t)]_{\varepsilon=0} 1_{\{0\}}(u) \quad \text{apply directional derivative formula}
\]

\[
= \frac{d}{d\varepsilon} \int_{-\infty}^{+\infty} F_1(t) \left( \dot{X}(t) - \tau_1 + \varepsilon\delta_t \right) dt |_{\varepsilon=0} 1_{\{0\}}(u) \quad \text{using (3.113)}
\]

\[
= \int_{-\infty}^{+\infty} F_1(t) \delta_t dt 1_{\{0\}}(u)
\]

\[
= F_1(t) 1_{\{0\}}(u). \quad (3.117)
\]

Therefore, from (3.115), (3.116) and (3.117), we have

\[
\partial_{(t,u)} F(x) = A + B = F_1(t) = F_1(t, u). \quad (3.118)
\]

From (3.114) and (3.118), we get the same result from the two different white noise derivative formulas.

### 3.4.2 The Malliavin derivative

On the given probability space $(\Omega, \mathcal{F}, P)$, the Malliavin derivative of a given random variable $F = F(\omega)$, $\omega \in \Omega$, can be regarded as a derivative with respect to the random parameter $\omega$. This is the original approach used by Malliavin. For the Brownian motion case, $\Omega$ is represented as the Wiener space $C_0([0, T])$ of continuous functions $\omega : [0, T] \rightarrow \mathbb{R}$.
R with \( \omega = 0 \), equipped with the uniform topology. In this Wiener space, we can regard each path

\[
t \rightarrow W(t, \omega)
\]

of the Wiener process starting at 0 as an element \( \omega \) of \( C_0([0, T]) \). Thus we may identify \( W(t, \omega) \) with the value \( \omega(t) \) at time \( t \) of an element \( \omega \in C_0([0, T]) \):

\[
W(t, \omega) = \omega(t).
\]

This measurable space is equipped with probability measure \( P \), which is given by the probability law of the Wiener process. The measure \( P \) is called the Wiener measure on \( \Omega \).

Following G. Di Nunno, B. Øksendal and F. Proske [32], we give the following definition.

**Definition 3.4.1.** Let \( F : \Omega \rightarrow \mathbb{R} \) be a random variable, choose \( g \in L^2([0, T], \mathbb{Z}, \lambda) = L^2([0, T]) \). and consider

\[
u(t) = \int_0^t g(s)ds \in \Omega.
\]

Then we define the directional derivative of \( F \) at the point \( \omega \in \Omega \) in the direction \( u \in \Omega \) by

\[
D^u F(\omega) = \frac{d}{d\varepsilon}[F(\omega + \varepsilon u)]_{\varepsilon=0},
\]

if the derivative exists in some sense. The superscript \( u \) denotes direction.

The set of \( u \in \Omega \), which can be written on the form (3.119), for some \( g \in L^2([0, T]) \), is called the Cameron-Martin space and it is denoted by \( H \).

Based on G. Di Nunno, B. Øksendal and F. Proske [32], we have
Definition 3.4.2. Assume that $F: Ω → ℝ$ has a directional derivative in all directions $u$ of the form $u ∈ H$, that is,

$$D^u F(ω) := \lim_{ε \to 0} \frac{F(ω + εu) - F(ω)}{ε}$$

exists in $L^2(P)$. Assume in addition that there exists $ψ(ω, t) ∈ L^2(P \times λ)$ such that

$$D^u F(ω) := \int_0^T ψ(t, ω)g(t)dt, \quad \text{for all} \quad u ∈ H. \quad (3.120)$$

Then we say that $F$ is differentiable and we set

$$D_t F(ω) := ψ(t, ω). \quad (3.121)$$

Remark 3.4.1. From (3.120) and (3.121), we can see that $D^u F(ω)$ is a directional derivative while the Malliavin derivative $D_t F(ω)$ is just part of it. However, from (3.112),

$$∂_{(t,u)} F(x) = \left(\frac{F(x + uδ_t) - F(x)}{u}\right)1_{\{u;|u|>0\}}(u) + (DF(x)δ_t)1_{\{0\}}(u),$$

the white nose derivative, is a directional derivative itself.

3.4.3 The connection between the white noise derivative and the Malliavin derivative

According to Definition (3.4.1) and Definition (3.4.2), we compute the Malliavin derivative.

Example 3.4.2. Assume $F = \int_0^T σdB(s) = \int_0^T σdω(s)$, where $σ ∈ L^2([0, T])$. Then for $u ∈ H$, we have

$$F(ω + εu) = \int_0^T σ(dω(s) + εdu(s))$$

$$= \int_0^T σdω(s) + ε \int_0^T σg(s)ds,$$
and hence
\[
\frac{F(\omega + \varepsilon u) - F(\omega)}{\varepsilon} = \int_0^T \sigma g(s)ds,
\]
for all \( \varepsilon > 0 \). Comparing (3.120) with (3.121), we get
\[
D_t F(\omega) = \sigma, \quad t \in [0, T], \quad \omega \in \Omega.
\]  

(3.122)

Now we compute the white noise derivative for the same \( F \). In [24], according to (3.3), (3.11), and (3.64), when \( \Lambda \) is the Gaussian white noise measure \( \Lambda_W \),
\[
\partial_t F(x) = DF(x)\delta_t,
\]  
when \( \Lambda \) is the Poisson white noise measure \( \Lambda_P \), and then
\[
\partial_t F(x) = F(x + \delta_t) - F(x).
\]

**Example 3.4.3.** Suppose
\[
F = \int_0^T \sigma dB(s) = \int_0^T \sigma \dot{B}(s)dt = \int_0^T \dot{X}(s)dt.
\]  

(3.124)

In (3.124), the Lévy white noise process has only a diffusion part, and there is neither a drift term, nor a pure jump part.
From (3.123) and (3.124),

\[ \partial_t F(x) = DF(x) \delta_t \]

\[ = \frac{d}{d\varepsilon} [F(x + \varepsilon \delta_t)]_{\varepsilon=0} \quad \text{apply directional derivative formula} \]

\[ = \frac{d}{d\varepsilon} \left[ \int_{-\infty}^{+\infty} (\dot{X}(t') + \varepsilon \delta_t(t')) \, dt' \right]_{\varepsilon=0} \]

\[ = \frac{d}{d\varepsilon} \left[ \int_{-\infty}^{+\infty} \dot{X}(t') \, dt' + \varepsilon \int_{-\infty}^{+\infty} \delta_t(t') \, dt' \right]_{\varepsilon=0} \]

\[ = \int_{-\infty}^{+\infty} \delta_t(t') \, dt' \]

\[ = 1, \quad (3.126) \]

as \( \delta_t \) is the Dirac delta function.

Remark 3.4.2. According to (3.122) and (3.126), without formal proof, for the diffusion part only, we may propose that

\[ \partial_t F = DF \delta_t = \frac{1}{\sigma} D_t F. \]

The white noise derivative coincides with the Malliavin derivative, except it is divided by \( \sigma \), due to the fact that in \( \partial_t F = DF \delta_t \), we are differentiating with respect to \( \sigma \dot{B} \).

3.4.4 Upper bound of Wasserstein distance

In section 3.3.2, we derived the upper bound of Wasserstein distance for Lévy functionals using the S-transform and white noise analysis based on Lee and Shih [24]. Therefore, the explicit bound Nourdin and Peccati [29] obtained for Gaussian functionals using Malliavin calculus should be a special case of our general result (3.83) because Brownian motion is a special case of Lévy processes. Now we compute an example using both theorems to get the same upper bound for a diffusion process.
Based on (3.13), (3.16), $F$ is given below:

\[
F = I_1(F_1) = \int_0^T \frac{1}{\sigma \sqrt{T}} \sigma dB(t) = \int_{-\infty}^{+\infty} F_1(t) d\tilde{X}(t) = \int_{-\infty}^{+\infty} F_1(t)(\tilde{X}(t) - \tau_1) dt, \tag{3.127}
\]

where

\[
F_1(t) = F_1(t, u) = \frac{1}{\sigma \sqrt{T}} \mathbf{1}_{[0, T]}(t) \mathbf{1}_{(0)}(u).
\]

and $\overset{d}{=} \text{ means equality in distribution.}

Here we choose Lévy process $\tilde{X}(t)$ as a diffusion process without a drift term, i.e. $\tau_1 = 0$.

**Example 3.4.4.** If $F = I_1(F_1)$ given in (3.127), and $X \sim \mathcal{N}(0, 1)$, then

\[
d_W(F, X) = 0. \tag{3.128}
\]

**Proof.** We apply upper bound Theorem 3.3.1.

According to (3.80) and (3.81), we get

\[
d_W(F, X) \leq E \left[ 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \partial_{(t, u)} MF \partial_{(t, u)} F d\lambda(t, u) \right]. \tag{3.129}
\]

In the equation above, we have only one term because $F$ has only a diffusion part. According to (3.48) and (3.60), we get

\[
\partial_{(t, u)} F = F_1(t, u) = \frac{1}{\sigma \sqrt{T}} \mathbf{1}_{[0, T]}(t) \mathbf{1}_{(0)}(u), \tag{3.130}
\]

\[
MF = I_1(F_1)
\]

and

\[
\partial_{(t, u)} F = F_1(t, u) = \frac{1}{\sigma \sqrt{T}} \mathbf{1}_{[0, T]}(t) \mathbf{1}_{(0)}(u). \tag{3.131}
\]
Moreover, due to (3.12) and (2.16),
\[ d\lambda(t,u) = dt \otimes (1 + u^2) d\beta(u), \]  
and from (3.2),
\[ \sigma^2 = \beta(\{0\}) = \int_{u \in \mathbb{R}} 1_{\{0\}}(u) d\beta(u). \]  
Substituting (3.130), (3.131), (3.132), and (3.133) into (3.129), we have
\[ d_W(F, X) \leq E \left| 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} F_1(t, u)^2 d\lambda(t, u) \right| \]
\[ = \left| 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \frac{1}{\sigma^2 T} 1_{\{0\}}(t) 1_{\{0\}}(u) d\lambda(t, u) \right| \]
\[ = \left| 1 - \int_{t \in \mathbb{R}} \int_{u \in \mathbb{R}} \frac{1}{\sigma^2 T} 1_{[0,T]}(t) 1_{\{0\}}(u) dt \otimes (1 + u^2) d\beta(u) \right| \]
\[ = \left| 1 - \int_{0}^{T} \frac{1}{\sigma^2 T} dt \int_{u \in \mathbb{R}} 1_{\{0\}}(u) d\beta(u) \right| \]
\[ = \left| 1 - \int_{0}^{T} \frac{1}{T} dt \right| \]
\[ = 0. \]
and we know that the metric, \( d_W(F, X) \geq 0 \). Therefore,
\[ d_W(F, X) = 0. \]

Now we compute the same example using Malliavin calculus and Stein’s method. According to Theorem 2.4.1, Definition 2.4.4 and Definition 2.4.6, we have the following
Malliavin operators:

\[ F = \sum_{n=0}^{\infty} I_n(f_n), \]
\[ D_tF = \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot, t)) \]

and

\[ L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n). \]

According to Nourdin and Peccati [29], we have the following theorem. We will discuss the Malliavin calculus and this theorem in more details in the next chapter. Here the whole theorem is given for the convenience.

**Theorem 3.4.3.** Let \( N \sim \mathcal{N}(0, 1) \), and \( F \in \mathbb{D}^{1,2} \) (Definition 2.4.4) be such that \( E(F) = 0 \). We have the following bound for Wasserstein distance between the law of \( N \) and the law of \( F \):

\[ d_W(F, N) \leq E|x| \leq \sqrt{E x^2}, \quad (3.134) \]

where

\[ x = 1 - <DF, -DL^{-1}F>_{L^2([0,T])}, \quad (3.135) \]

and \( DF \) and \( L^{-1}F \) are defined in Definition 2.4.4, and Definition 2.4.6. In the rest of this thesis, we use \( <\cdot, \cdot> \) instead of \( <\cdot, \cdot>_{L^2([0,T])} \).

Assume

\[ F = I_t(F_1) = \int_0^T \frac{1}{\sigma \sqrt{T}} \sigma dB(t) = \int_0^T \frac{1}{\sqrt{T}} dB(t). \]
Example 3.4.5. If $F = I_1(F_1)$ as shown above, and $X \sim \mathcal{N}(0,1)$, then

$$d_W(F,X) = 0.$$ (3.136)

Proof. According to Theorem 2.4.1, Definition 2.4.4, and Definition 2.4.6, we get

$$D_tF = \frac{1}{\sqrt{T}}$$ (3.137)

$$L^{-1}F = -\int_0^T \frac{1}{\sqrt{T}} dB(t)$$

$$-DL^{-1}F = \frac{1}{\sqrt{T}}.$$ (3.138)

Substitute (3.137) and (3.138) into (3.135), we have

$$x = 1 - \langle DF, -DL^{-1}F \rangle_{L^2([0,T])}$$

$$= 1 - \int_0^T \frac{1}{\sqrt{T}} \frac{1}{\sqrt{T}} dt$$

$$= 1 - 1$$

$$= 0.$$ 

Therefore, due to (3.134), we get

$$d_W(F,N) \leq E|x|$$

$$= 0.$$ 

Moreover, we know that $d_W(F,N) \geq 0$. Therefore,

$$d_W(F,N) = 0.$$ 

$\square$
Remark 3.4.4. We compute the upper bound for the following function (3.127).

\[
F = I_1(F_1) = \int_0^T \frac{1}{\sigma\sqrt{T}} \sigma dB(t) \\
= \int_{-\infty}^{+\infty} F_1(t) d\tilde{X}(t) \\
= \int_{-\infty}^{+\infty} F_1(t)(\dot{\tilde{X}}(t) - \tau_1) dt.
\]

According to (3.128) and (3.136), we get the same upper bound using our theorem with white noise operators (3.80) (3.81) and the theorem from Nourdin and Peccati [29] with Malliavin operators. This is what we expect.

3.5 Summary

In this chapter, we summarize Segal-Bargmann transform (S-transform), differential, adjoint, N operators and derivative of analytic functionals in the celebrated paper by Lee and Shih [24]. We derive the chain rule for Lévy white noise functionals, (Theorem 3.2.10). Then, we compute the upper bound of Wasserstein distance normal approximation for Lévy functionals via Stein’s method (Theorem 3.3.1). Instead of Malliavin calculus Nourdin and Peccati(2009) [29] use, we use white noise analysis [24] initiated by Hida. Another contribution is to show that the Stein bounds in the literature for functionals of (i) pure jump Lévy processes, and (ii) the pure Brownian motion Lévy process, are both particular cases of the general bound in the thesis; see Corollary 3.3.1 and Corollary 3.3.2. Therefore, up to date, we have the most general result, which paves the way for further research and applications. In the end, we give some examples to show the difference and connection between white noise analysis and Malliavin calculus. We will introduce a new method for option pricing approximation via Stein’s method and Malliavin calculus in the next chapter.
Chapter 4

Option pricing approximation by Stein’s method and Malliavin calculus

4.1 Introduction

We introduce financial option pricing approximation by Stein’s method and Malliavin calculus. We show that the difference between two option prices can be expressed as the distance between two probability measures. The explicit upper bound of Wasserstein distance has been derived by Stein’s method and Malliavin calculus. Numerical examples are provided. This chapter is another application of Stein’s method although it does not use the outcomes of the previous chapter.

This chapter is organized in the following manner: Section 4.2 describes the theoretical framework and intuition in this chapter. Section 4.3 defines the difference between option prices. We compute upper bound for Wasserstein distance in Section 4.4. The numerical results are also provided in Section 4.5.
4.2 Financial option pricing approximation

4.2.1 Financial option pricing approximation by Stein’s method and Malliavin calculus

In financial mathematics, according to [2], [6], and [9], suppose the state of a arbitrage-free market can be represented by the filtered probability space, \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})\), where \(\mathbb{P}\) is a risk-neutral measure. Let \(\{S_t\}_{t\in[0,T]}\) be a stochastic price process on this space. The price of all instruments, \(g(0, S_0)\), may be computed as discounted expectations of their terminal payoff with respect to some equivalent martingale measure \(\mathbb{P}\). We have the following pricing formula:

\[
g(0, S_0) = E(e^{-rT}g(T, S_T)) = E(h(S_T)), \tag{4.1}
\]

where \(r\), a constant, is a risk-free rate, \(T\) is time to maturity. Here we use bank account, \(BA\) as numeraire. We know \(BA = 1\) at time zero, and \(BA = e^{rT}\) at time \(T\). Moreover,

\[
h(S_T) = e^{-rT}g(T, S_T) \tag{4.2}
\]

and, for instance,

\[
g(T, S_T) = (S_T - K)^+,
\]

where \(K\) is the strike price.

We can write \(g(x) = (x - K)^+\), where \(x\) is the dummy variable.

The function \(g\) is not differentiable at \(x = K\). However, we can approximate \(g\) by \(C^1\) function \(g_n\).

Based on Shreve [39], we define a sequence of functions \(\{g_n\}_{n=1}^{\infty}\) by the formula

\[
g_n(x) = \begin{cases} 0 & \text{if } x \leq K - \frac{1}{2n}, \\ \frac{n}{2}(x - K)^2 + \frac{1}{2}(x - K) + \frac{1}{8n} & \text{if } K - \frac{1}{2n} \leq x \leq K + \frac{1}{2n}, \\ x - K & \text{if } x \geq K + \frac{1}{2n}. \end{cases} \tag{4.3}
\]
We can see that
\[ g_n'(x) = \begin{cases} 
0 & \text{if } x \leq K - \frac{1}{2n}, \\
 n(x-K) + \frac{1}{2} & \text{if } K - \frac{1}{2n} \leq x \leq K + \frac{1}{2n}, \\
 1 & \text{if } x \geq K + \frac{1}{2n}. 
\end{cases} \tag{4.4} \]

Then it is straightforward to check the following: \( g_n \in C^1 \), \( \lim_{n \to \infty} g_n(x) = g(x) \) and \( 0 \leq \lim_{n \to \infty} g_n'(x) \leq 1 \), for all \( x \in \mathbb{R} \).

By (4.2), \( h(x) = e^{-rT}g(x) \), and \( e^{-rT} \leq 1 \). Therefore, we can approximate \( h(x) \) by
\[ h_n(x) = e^{-rT}g_n(x) \tag{4.5} \]
and \( h_n(x) \in \text{Lip}(1) := \{ h : \mathbb{R} \to \mathbb{R} ; \| h' \|_{\infty} \leq 1 \} \), where \( \text{Lip}(1) \) denotes the class of real-valued Lipschitz functions, from \( \mathbb{R} \) to \( \mathbb{R} \), with Lipschitz constant less or equal to one.

Stein’s method is a way of deriving explicit estimates of the accuracy of the approximation of one probability distribution by another. According to Lemma 3.3.2 and Theorem 4.4.1, we have the following inequality. For \( F \in \mathbb{D}^{1,2} \),

\[ d_W(F,N) = \sup_{h \in \text{Lip}(1)} |E(h(F)) - E(h(N))| \leq \sup_{h \in \text{Lip}(1)} |E[f'_h(F) - Ff_h(F)]| \leq \sup_{f_h \in \mathcal{F}_W} |E[f'_h(F) - Ff_h(F)]| \leq E|X| \leq \sqrt{E(X^2)}, \tag{4.8} \]

where \( f_h \) is defined in the equation (3.74), \( N \sim \mathcal{N}(0,1) \), and \( d_W(F,N) \) is Wasserstein distance between the law of \( N \) and the law of \( F \), \( d \) denotes metric, \( X = 1 - < DF, -DL^{-1}F > \), and \( D \) and \( L^{-1} \) are Malliavin operators defined in the last section. \( \mathcal{F}_W \) denote the class of twice differentiable functions, whose first derivative is bounded by 1 and whose second derivative is bounded by 2. If \( h \in \text{Lip}(1) \) and \( \| h' \|_{\infty} \leq 1 \), then \( f_h \) satisfies \( \| f'_h \|_{\infty} \leq 1 \) and \( \| f''_h \|_{\infty} \leq 2 \). Therefore, \( f_h \in \mathcal{F}_W \).
Usually the supremum of (4.7) which has only the random variable $F$ is simpler to bound than that of (4.6).

Nourdin and Peccati(2009) [29] give explicit bounds from (4.7) to (4.8) using Malliavin operators. Therefore, combining (4.8) and (4.7) with (4.1) we can calculate the upper bound of the difference between financial option prices when the underlying follows different probability distribution. Here, the law of $N$ and the law of $F$.

Note that the upper bound (4.8) we calculate can be applied to a class of function $(h \in \text{Lip}(1))$ in (4.6), not just one specific function. The result is very general. For example, in (4.2) and (4.5), we show that we can even apply this to a function which is not differentiable at one point.

When we know the closed form solution of one option price, we may use it to approximate another option price which does not have the closed form solution by using the Stein-Malliavin bound.

### 4.2.2 Financial spread option pricing for Gaussian and non-Gaussian processes

According to [3], in financial mathematics, the pricing and hedging of spread options are very popular research topics. These options are ubiquitous in all financial markets, such as equity, fixed income, foreign exchange, commodities, or energy markets. A spread option is an option written on the spread, which is defined as the difference of two underlying assets.

In most applications in financial mathematics, the underlying assets are stock prices that are modeled by means of log-normal distributions as prescribed by the geometric Brownian motion model. One of the advantages of this model is that it produces underlying prices that are inherently positive. However, the positivity restriction does not apply to the spreads themselves, since the latter can be negative as difference of two positive quantities.

When we plot histograms of historical spread values, we find that the marginal distribution of a spread at a given time extends on both sides of zero; moreover, the normal distribution can give a reasonable fit [3]. This suggests us to use arithmetic Brownian motion (4.9)(as opposed to the geometric Brownian motion leading to the lognormal
distribution) for the dynamics of spreads. In so doing, prices of spread options can be derived by computing Gaussian integrals leading to simple closed form formulae (4.13).

According to [3], we assume that the risk-neutral dynamics of the spread $S_t$ is given by a stochastic differential equation which has a Gaussian process as a solution:

$$dS_t = \mu S_t dt + \sigma dW_t$$  \hspace{1cm} (4.9)

for some standard Brownian motion $\{W_t\}_{t \geq 0}$ and some positive $\sigma$. Here $\mu$ stands for the short interest rate $r$. Following (4.9), we get the solution:

$$S_T = S_0 e^{\mu T} + \sigma \int_0^T e^{\mu (T-u)} dW_u$$ \hspace{1cm} (4.10)

with

$$\begin{align*}
\mu_{S_T} &= E(S_T) = S_0 e^{\mu T}, \hspace{1cm} (4.11) \\
\sigma_{S_T} &= \sqrt{\frac{\sigma^2}{2\mu}(e^{2\mu T} - 1)}, \hspace{1cm} (4.12)
\end{align*}$$

where $S_T$ is a Gaussian process with $\mu_{S_T}$ and $\sigma_{S_T}$. $S_T \sim N(\mu_{S_T}, \sigma_{S_T}^2)$. We can just use Theorem 4.2.1 to get call spread option price because $S_T$ follows a normal distribution (4.10).

**Theorem 4.2.1.** Let $P = E[e^{-rT}(S_T - K)^+]$ denote the call spread option price formula, where $r$, $T$, and $K$ are constants, and $S_T = \mu_{S_T} - \sigma_{S_T}X$, where $\mu_{S_T}$ is given in (4.11), $\sigma_{S_T}$ is given in (4.12), and $X \sim N(0,1)$. Then

$$P = e^{-rT}(\mu_{S_T} - K)N(d) + e^{-rT}\sigma_{S_T}n(d),$$ \hspace{1cm} (4.13)

where $d = \frac{\mu_{S_T} - K}{\sigma_{S_T}}$, and $N(\cdot)$ stands for the standard Gaussian cumulative distribution function while $n(\cdot)$ stands for the standard Gaussian density function.

**Proof.** We now derive the call spread option price formula for a Gaussian process:
First, we compute $d$:

\[
S_T - K > 0 \\
\Rightarrow \mu_{S_T} - \sigma_{S_T}X - K > 0 \\
\Rightarrow X < \frac{\mu_{S_T} - K}{\sigma_{S_T}} = d.
\]

Then

\[
P = E[e^{-rT}(S_T - K)^+] \\
= e^{-rT}E[(S_T - K)I_{\{S_T > K\}}] \\
= e^{-rT}\left( \int_{-\infty}^{+\infty} (\mu_{S_T} - K)I_{\{S_T > K\}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\
- \int_{-\infty}^{+\infty} \sigma_{S_T}xI_{\{S_T > K\}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \right) \\
= e^{-rT}\left( \int_{-\infty}^{d} (\mu_{S_T} - K) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\
- \int_{-\infty}^{d} \sigma_{S_T}x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \right) \\
= e^{-rT}(\mu_{S_T} - K)N(d) + e^{-rT}\sigma_{S_T}n(d).
\]

However, when we have stochastic volatility SDE, and $h_t$ is a deterministic function.

\[
d\tilde{S}_t = \mu S_t dt + \sigma(h_t + \varepsilon W_t) dW_t \tag{4.14}
\]

and

\[
\tilde{S}_T = S_0e^{\mu T} + \sigma \int_0^T e^{\mu(T-u)}(h_u + \varepsilon W_u) dW_u, \tag{4.15}
\]

where $\tilde{S}_T$ does not follow a normal distribution any more.
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus

The model (4.9) has been used in financial modelling based on [3]. The simple perturbed version of (4.9) given in (4.14) has been introduced for illustrative purposes, where its volatility growing with time is not very realistic. For further research, we may consider a more complicated stochastic volatility model, such as Heston stochastic volatility model [11]. The popular Heston model is a commonly used SV model, in which the randomness of the variance process varies as the square root of variance. In this case, the stochastic differential equation for variance takes the form:

\[ dV_t = \kappa(\theta - V_t)dt + \xi \sqrt{V_t} dB_t, \]

where \( \theta \) is the long run variance, \( \kappa \) is the mean reversion rate, \( \xi \) is the volatility of variance, \( B_t \) is standard Brownian movements, and \( dW_t \) and \( dB_t \) are correlated with the constant correlation value, \( \rho \). Therefore, the Heston SV model assumes that the variance is a random process that

1. exhibits a tendency to revert towards a long-term mean \( \theta \) at a rate \( \kappa \),
2. exhibits a variance proportional to the square root of its level,
3. and whose source of randomness is correlated with the randomness of the spread processes.

We wonder if we can still use the closed form solution; that is, under what conditions can we use

\[ P = E[e^{-rT}(S_T - K)^+] \] (4.10)
\[ = e^{-rT}(\mu S_T - K)N(d) + e^{-rT} \sigma S_T n(d) \]

in Theorem 4.2.1 to approximate

\[ \tilde{P} = E[e^{-rT}(\tilde{S}_T - K)^+]. \] (4.16)

We can compute the Wasserstein distance, \( d_W(\tilde{S}_T, S_T) \) between the law of \( S_T \) and the law of \( \tilde{S}_T \). If \( d_W(\tilde{S}_T, S_T) \) is small, we could ask whether

\[ P = E[e^{-rT}(S_T - K)^+] \] (4.10)
\[ = e^{-rT}(\mu S_T - K)N(d) + e^{-rT} \sigma S_T n(d) \]
is a good normal approximation for

\[ \tilde{P} = E[e^{-rT}(\tilde{S}_T - K)^+] \]  

(\(\tilde{S}_T\) follows 4.15))

in that situation. This is what we discuss and derive in this chapter.

4.3 Difference between option prices (distance between probability measures)

4.3.1 Stochastic processes for spread between two assets

We consider two stochastic processes given by

\[ dS_t = \mu S_t dt + \sigma W_t, \]  

(4.17)

and

\[ dS_t = \mu S_t dt + \sigma (h_t + \varepsilon W_t) dW_t, \]  

(4.18)

for some standard Brownian motion \( \{W_t\}_{t \geq 0} \).

The stochastic differential equation (4.17) is often used for financial spread option pricing in equity, energy and fixed-income markets [3]. The spread \( S_t \) follows a normal distribution, and we have closed form solutions for European spread call options.

In (4.18), the volatility is not a constant \( \sigma \), but a stochastic process \( \sigma (h_t + \varepsilon W_t) \). \( h_t \) is a deterministic function, and \( \int_0^t h_u^2 du < \infty \), for all \( t \in (0, \infty) \). \( \varepsilon \) is a constant.

This is a one-factor stochastic volatility process, and \( S_t \) does not follow normal distribution. From (4.17) and (4.18), we get (4.19) and (4.20).

**Lemma 4.3.1.** Assume \( \mu = r \), which is a risk-free interest rate, and \( T \) is time to maturity. \( \varepsilon \) is a constant, and \( \int_0^t h_u^2 du < \infty \).

\[ S_T^N = S_0 e^{\mu T} + \sigma \int_0^T e^{\mu (T-u)} dW_u, \]  

(4.19)
and
\[
S_T^F = S_0 e^{\mu T} + \sigma \int_0^T e^{\mu(T-u)}(h_u + \varepsilon W_u) dW_u,
\]
where \(S_T^N\) means \(S_T\) follows normal distribution while \(S_T^F\) is for our stochastic volatility process. Let \(a = S_0 e^{\mu T}\) and
\[
\begin{align*}
Y &= \sigma \int_0^T e^{\mu(T-u)} dW_u \\
V &= \sigma \int_0^T e^{\mu(T-u)}(h_u + \varepsilon W_u) dW_u,
\end{align*}
\]
and
\[
\frac{Y}{M} = N,
\]
where \(N \sim \mathcal{N}(0, 1)\),
\[
\frac{V}{M} = F,
\]
and
\[
M = \sigma \sqrt{\frac{e^{2\mu T} - 1}{2\mu}}.
\]
Then,
\[
S_T^N = a + MN,
\]
and
\[
S_T^F = a + MF.
\]

**Proof.** Set \(Y_t = S_t e^{-\mu t}\).

According to Itô Lemma, \(df(t, X_t) = f_t(t, X_t) dt + f_x(t, X_t) dX_t + \frac{1}{2} f_{xx}(t, X_t) dX_t dX_t\)
\[
dY_t = S_t e^{-\mu t}(-\mu) dt + e^{-\mu t} dS_t.
\]
If $dS_t = \mu S_t dt + \sigma (h_t + \varepsilon W_t) dW_t$ (4.18) then

$$dY_t = S_t e^{-\mu t} (-\mu)dt + e^{-\mu t}(\mu S_t dt + \sigma (h_t + \varepsilon W_t) dW_t)$$

$$= e^{-\mu t} \sigma (h_t + \varepsilon W_t) dW_t. \quad (4.24)$$

Integrating on both side,

$$Y_t - Y_0 = \sigma \int_0^t e^{-\mu u} (h_u + \varepsilon W_u) dW_u$$

$$S_t e^{-\mu t} = S_0 + \sigma \int_0^t e^{-\mu u} (h_u + \varepsilon W_u) dW_u$$

$$S_t = S_0 e^{\mu t} + \sigma \int_0^t e^{\mu (t-u)} (h_u + \varepsilon W_u) dW_u.$$

In a similar way, if $dS_t = \mu S_t dt + \sigma dW_t$ (4.17), then

$$S_t = S_0 e^{\mu t} + \sigma \int_0^t e^{\mu (t-u)} dW_u.$$

Therefore, we can get

$$S^N_T = S_0 e^{\mu T} + \sigma \int_0^T e^{\mu (T-u)} dW_u,$$

and

$$S^F_T = S_0 e^{\mu T} + \sigma \int_0^T e^{\mu (T-u)} (h_u + \varepsilon W_u) dW_u.$$

Let $a = S_0 e^{\mu T}$,

$$Y = \sigma \int_0^T e^{\mu (T-u)} dW_u$$

$$V = \sigma \int_0^T e^{\mu (T-u)} (h_u + \varepsilon W_u) dW_u.$$

Then

$$S^N_T = a + Y.$$
Chapter 4. *Option pricing approximation by Stein’s method and Malliavin calculus* 117

and

\[ S_T^F = a + V. \]

Set \( \sqrt{\text{Var}(Y)} = M \), and we know \( \text{Var}(Y) = \frac{\sigma^2}{2\mu} [e^{2\mu T} - 1] \).

Then \( M = \sigma \sqrt{\frac{e^{2\mu T} - 1}{2\mu}} \).

Let \( \frac{Y}{M} = N \), where \( N \sim N(0,1) \),

\[ \frac{V}{M} = F. \]

Therefore,

\[ S_T^N = a + MN, \]

and

\[ S_T^F = a + MF. \]

In (4.24), for Itô’s lemma to hold, we need to prove that

\[ S_t = S_0 + \int_0^t \mu S_u du + \int_0^t \sigma (h_u + \varepsilon W_u) dW_u \quad (4.25) \]

is a Itô process, that is, to prove that, for all \( t \geq 0 \),

\[ \int_0^t |\mu S_u| du < \infty \]

with probability 1, and

\[ \int_0^t (\sigma (h_u + \varepsilon W_u))^2 du < \infty \]

with probability 1, see Øksendal [20](2003, p.44).
First, we know that $h_t$ is a deterministic function, and $\int_0^t h_u^2 \, du < \infty$ for all $t \in (0, \infty)$. Recall that $\varepsilon$ is a constant. Then

\[
\int_0^t \sigma(h_u + \varepsilon W_u) \, dW_u = \sigma \int_0^t h_u \, dW_u + \sigma \varepsilon \int_0^t W_u \, dW_u
\]

\[
= \sigma X_t + \frac{\sigma \varepsilon}{2}(W_t^2 - t),
\]

where $X_t = \int_0^t h_u \, dW_u$. Since $W_t$ is a Brownian motion and $X_t$ is a time-changed Brownian motion, it follows that, for every $T \in (0, \infty)$,

\[
\sup_{0 \leq t \leq T} |X_t| < \infty \text{ with probability 1, \quad (4.26)}
\]

and

\[
\sup_{0 \leq t \leq T} |W_t| < \infty \text{ with probability 1. \quad (4.27)}
\]

Therefore, using (4.26) and (4.27), for every $T \in (0, \infty)$,

\[
a_T = \sup_{0 \leq t \leq T} \left| \int_0^t \sigma(h_u + \varepsilon W_u) \, dW_u \right|
\]

\[
= \sup_{0 \leq t \leq T} \left| \sigma X_t + \frac{\sigma \varepsilon}{2}(W_t^2 - t) \right|
\]

\[
\leq \sigma \sup_{0 \leq t \leq T} |X_t| + \frac{\sigma \varepsilon}{2} \sup_{0 \leq t \leq T} |W_t^2| + \frac{\sigma \varepsilon T}{2}
\]

\[
< \infty \text{ with probability 1.}
\]

Consequently, from (4.25),

\[
|S_t| \leq |S_0| + \left| \int_0^t \mu S_u \, du \right| + \left| \int_0^t \sigma(h_u + \varepsilon W_u) \, dW_u \right|
\]

\[
\leq |S_0| + |\mu| \int_0^t |S_u| \, du + a_T.
\]
Using Gronwall inequality, we may conclude immediately that

\[ |S_t| \leq (|S_0| + a_T)e^{\mu t} \quad \text{for } 0 \leq t \leq T. \]

Therefore, for each \( t \geq 0 \),

\[ \int_0^t |\mu S_u| du \leq (|S_0| + a_T)(e^{\mu t} - 1) < \infty, \]

with probability 1.

Second, using that \( \int_0^t h_u^2 du < \infty \) for all \( t \in (0, \infty) \). \( E(W_u^2) = u \) and \( \varepsilon \) is a constant.

\[ \left| \int_0^t h_u W_u du \right| \leq \left( \int_0^t h_u^2 du \int_0^t W_u^2 du \right)^{1/2} < \infty \quad \text{with probability 1.} \]

Therefore,

\[ \int_0^t (\sigma(h_u + \varepsilon W_u))^2 du \]
\[ = \sigma^2 \int_0^t (h_u^2 + 2\varepsilon h_u W_u + \varepsilon^2 W_u^2) du \]
\[ < \infty \quad \text{with probability 1.} \]

Consequently, according to [20], \( S_t \) in (4.25) is a one-dimensional Itô process and therefore we can apply the one-dimensional Itô’s lemma.

\[ \square \]
4.3.2 Wasserstein distance

In the following statement, $S_T^N$ defined in (4.19) is the solution of the SDE (4.17) while $S_T^F$ in (4.20) is the solution of the SDE (4.18). Using Lemma 4.3.1, we find that

$$d_W(S_T^F, S_T^N) = d_W(S_T^F - a, S_T^N - a) = d_W(V, Y) = d_W(MF, MN) = |M|d_W(F, N), \quad (4.28)$$

where $d_W(F, N)$ is defined in Theorem 4.4.1 and in (4.6).

From Lemma 4.3.1, we get the Wiener-Itô chaos expansion for $F$

**Lemma 4.3.2.** $F \in \mathbb{D}^{1,2}$. Then we have

$$F = \frac{V}{M} = K I_1(f_1) + K_1 I_2(f_2), \quad (4.29)$$

where

$$I_1(f_1) = \int_0^T e^{-\mu u} h_u dW_u,$$

$$I_2(f_2) = 2! \int_0^T \int_0^{t_2} \frac{1}{2!} \left( e^{-\mu t_2} \mathbb{1}_{\{t_1 < t_2\}} + e^{-\mu t_1} \mathbb{1}_{\{t_2 < t_1\}} \right) dW_{t_1} dW_{t_2}$$

$$= 2! J_2(f_2), \quad \text{(by (2.34))}$$

$$K = \frac{\sigma}{M} e^{\mu T}, \quad (4.30)$$

$$K_1 = K \varepsilon. \quad (4.31)$$
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus

**Proof.** From (4.21), (4.22), and (4.23)

\[
F = \frac{V}{M} = \frac{\sigma}{M} \int_0^T e^{\mu(T-u)} (h_u + \varepsilon W_u) dW_u
\]

\[
= \frac{\sigma}{M} e^{\mu T} \left( \int_0^T e^{-\mu u} h_u dW_u + \varepsilon \int_0^T e^{-\mu u} W_u dW_u \right)
\]

\[
= K \left( \int_0^T e^{-\mu u} h_u dW_u + \varepsilon \int_0^T e^{-\mu u} W_u dW_u \right)
\]

\[
= K \int_0^T e^{-\mu u} h_u dW_u + K_1 \int_0^T e^{-\mu u} W_u dW_u \quad (K = \frac{\sigma}{M} e^{\mu T} \text{ and } K_1 = K \varepsilon)
\]

\[
= K I_1(f_1) + K_1 I_2(f_2),
\]

where

\[
I_1(f_1) = \int_0^T f_1(u) dW_u,
\]

and \( f_1 = e^{-\mu u} h_u \),

\[
I_2(f_2) = \int_0^T e^{-\mu u} W_u dW_u = \int_0^T e^{-\mu t_2} W_{t_2} dW_{t_2}
\]

\[
= \int_0^T \int_0^{t_2} e^{-\mu t_2} dW_{t_1} dW_{t_2} = \int_0^T \int_0^{t_2} e^{-\mu t_2} I_{(t_1 < t_2)} dW_{t_1} dW_{t_2}
\]

\[
= 2! \int_0^T \int_0^{t_2} \frac{1}{2!} \left( e^{-\mu t_2} I_{(t_1 < t_2)} + e^{-\mu t_1} I_{(t_2 < t_1)} \right) dW_{t_1} dW_{t_2}
\]

\[
= 2! J_2(f_2), \quad \text{by (2.34)}
\]

and \( f_2 = \frac{1}{2!} \left( e^{-\mu t_2} I_{(t_1 < t_2)} + e^{-\mu t_1} I_{(t_2 < t_1)} \right) \) is a symmetric function. \( \square \)

### 4.4 Upper bound estimation by Stein’s method and Malliavin calculus

According to Nourdin and Peccati(2009) [29], and Theorem 2.4.1, Definition 2.4.4, and Definition 2.4.6, we have the following theorem.
Theorem 4.4.1. Let $N \sim \mathcal{N}(0,1)$, and $F \in \mathbb{D}^{1,2}$ be such that $E(F) = 0$. We have the following bound for Wasserstein distance between the law of $N$ and the law of $F$:

$$d_W(F, N) \leq E|x| \leq \sqrt{E|x^2|},$$

and

$$x = 1 - <DF, -DL^{-1}F>_{L^2([0,T])},$$

where $DF$ and $L^{-1}F$ are defined in Definition 2.4.4, and 2.4.6.

In the rest of this thesis, we use $<\cdot, \cdot>$ instead of $<\cdot, \cdot>_{L^2([0,T])}$.

Now, according to Lemma 4.3.2, set $F = KI_1(f_1) + K_1I_2(f_2)$, and

$$I_1(f_1) = \int_0^T e^{-\mu t} h_t dW_t$$

$$f_1 = e^{-\mu t} h_t$$

$$I_2(f_2) = 2! \int_0^T \int_0^t \frac{1}{2!}(e^{-\mu t_2}I_{t_1 < t_2} + e^{-\mu t_1}I_{t_2 < t_1}) dW_t_1 dW_t_2$$

$$f_2 = \frac{1}{2!}(e^{-\mu t_2}I_{t_1 < t_2} + e^{-\mu t_1}I_{t_2 < t_1}), \quad (f_2 \text{ is symmetric})$$

We need to compute $DF$, and $-DL^{-1}F$ first, and finally get $\sqrt{E|x^2|}$.

From (4.29), and using (2.41) and (2.47), we obtain the following expressions. $F \in \mathbb{D}^{1,2}$ and $E(F) = 0$. Then we have

$$D_t F = Kf_1(t) + 2K_1I_1(f_2(t_1, t))$$

$$L^{-1}F = -KI_1(f_1) - \frac{1}{2}K_1I_2(f_2)$$

$$-D_t L^{-1}F = Kf_1(t) + K_1I_1(f_2(t_1, t))$$

where by (2.41) and (2.42),

$$I_0(f_1(t)) = f_1(t)$$
and

\[ f_1 = e^{-\mu t} h_t, \]  

(4.33)

\[ I_1(f_2(t_1, t)) = \int_0^T \frac{1}{2!} (e^{-\mu t} \mathbb{1}_{t_1 < t} + e^{-\mu t} \mathbb{1}_{t < t_1}) dW_{t_1}, \]  

(4.34)

with

\[ f_2(t_1, t) = \frac{1}{2!} (e^{-\mu t} \mathbb{1}_{t_1 < t} + e^{-\mu t} \mathbb{1}_{t < t_1}). \]  

(4.35)

Therefore, by (4.32),

\[
<DF, -DL^{-1}F> =<Kf_1 + 2K_1I_1(f_2), Kf_1 + K_1I_1(f_2)>
\]

\[ = K^2 <f_1, f_1> + KK_1 <f_1, I_1(f_2)> + 2KK_1 <I_1(f_2), f_1>
\]

\[ + 2K_1^2 <I_1(f_2), I_1(f_2)>
\]

\[ = K^2 <f_1, f_1> + 3KK_1 <f_1, I_1(f_2)>
\]

\[ + 2K_1^2 <I_1(f_2), I_1(f_2)>. \]  

(4.36)

4.4.1 Calculate the first term in the bound

In Theorem 4.4.1

\[ d_W(F, N) \leq E|x| \leq \sqrt{Ex^2}. \]

Now

\[ x = 1 - <DF, -DL^{-1}F>, \]

\[ x^2 = (1 - <DF, -DL^{-1}F>)^2 \]

\[ = 1 - 2 <DF, -DL^{-1}F> + (<DF, -DL^{-1}F>)^2. \]
Taking the expectation,

\[ E(x^2) = 1 - 2E(DF, -DL^{-1}F) + E[(DF, -DL^{-1}F)^2]. \] (4.37)

Note that we need to calculate the first term \( E(DF, -DL^{-1}F) \) and the second term \( E[(DF, -DL^{-1}F)^2] \) in (4.37), and finally get \( \sqrt{E x^2} \), which is a constant.

Here we need to calculate the first expectation in (4.37). We prove that the first expectation is the sum of two constants.

**Lemma 4.4.1.** \( F \in \mathbb{D}^{1,2} \) and \( E(F) = 0 \). Then we have

\[ E(DF, -DL^{-1}F) = \Delta_1 + \Delta_3, \]

where

\[ \Delta_1(h_t) = \Delta_1 = K^2 \int_0^T e^{-2\mu t} h_t^2 dt. \] (4.38)

\[ \Delta_3 = \frac{1}{2} K^2 \left( \frac{2\mu - 4\mu T - 1}{4\mu^2} e^{-2\mu T} + \frac{1 - 2\mu}{4\mu^2} \right). \] (4.39)

**Proof.** From (4.36) and (4.37), we have

\[
E(DF, -DL^{-1}F) = K^2 E(f_1, f_1) + 3K K_1 E(f_1, I_1(f_2)) \\
+ 2K^2 E(I_1(f_2), I_1(f_2)) \\
= \Delta_1 + \Delta_2 + \Delta_3. \] (4.40)

The first term in the above formula is a constant.

\[
\Delta_1 = K^2 E(f_1, f_1) = K^2 \int_0^T e^{-\mu t} h_t e^{-\mu t} h_t dt \\
= K^2 \int_0^T e^{-2\mu t} h_t^2 dt. \] (4.41)
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus 125

The second term is zero because

\[ E(\langle f_1, I_1(f_2) \rangle) = E\left( \int_0^T e^{-\mu t} h_t I_1(f_2(\cdot, t)) \, dt \right) \]
\[ = E\left( \int_0^T e^{-\mu t_1} \int_0^T \frac{1}{2!} (e^{-\mu t_1 \mathbb{1}_{t_1 < t}} + e^{-\mu t_1 \mathbb{1}_{t < t_1}}) \, dW_{t_1} \, dt \right) \]
\[ = \int_0^T E\left( \int_0^T \frac{1}{2!} e^{-\mu t_1} (e^{-\mu t_1 \mathbb{1}_{t_1 < t}} + e^{-\mu t_1 \mathbb{1}_{t < t_1}}) \, dW_{t_1} \right) \, dt \]
\[ = 0 \quad \text{since} \quad E\left( \int_0^T (\cdot) \, dW_{t_1} \right) = 0, \]

so

\[ \Delta_2 = 0. \tag{4.42} \]

Starting from (4.40), we get the third term, which is a constant too.

\[ \Delta_3 = 2K_1^2 E(\langle I_1(f_2), I_1(f_2) \rangle) \]
\[ = 2K_1^2 \int_0^T E(I_1(f_2(t_1, t))^2) \, dt \]
\[ = 2K_1^2 \int_0^T E(I_1(f_2(t_1, t))^2) \, dt \]
\[ = 2K_1^2 \int_0^T E \left[ \int_0^T \frac{1}{2!} (e^{-\mu t_1 \mathbb{1}_{t_1 < t}} + e^{-\mu t_1 \mathbb{1}_{t < t_1}}) \, dW_{t_1} \right] \, dt. \]
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus

\[ = 2K^2 \int_0^T \int_0^T \left( \frac{1}{2} \left( e^{-\mu t_1} I_{t_1 < t} + e^{-\mu t_1} I_{t < t_1} \right) \right)^2 dt_1 dt \]

\[ = \frac{1}{2} K^2 \int_0^T \int_0^T (e^{-2\mu t_1} I_{t_1 < t} + e^{-2\mu t_1} I_{t < t_1}) dt_1 dt \]

\[ = \frac{1}{2} K^2 \left( \int_0^T \int_0^T e^{-2\mu t_1} dt_1 dt + \int_0^T \int_0^t e^{-2\mu t_1} dt_1 dt \right) \]

\[ = \frac{1}{2} K^2 \left( \int_0^T e^{-2\mu t} dt + \int_0^T \frac{-1}{2\mu} e^{-2\mu t} \right) \quad (4.43) \]

\[ = \frac{1}{2} K^2 \left( \frac{2\mu - 4\mu T - 1}{4\mu^2} e^{-2\mu T} + \frac{1 - 2\mu}{4\mu^2} \right), \quad (4.44) \]

where the integrals in (4.43) are elementary to compute.

The proof is complete. \[ \square \]

4.4.2 Calculate the second term in the bound

We know that \((a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc\)

Therefore, starting from (4.36)

\[
\langle DF, -DL^{-1}F \rangle^2 = (K^2 \langle f_1, f_1 \rangle + 3KK_1 \langle f_1, I_1(f_2) \rangle + 2K^2 \langle I_1(f_2), I_1(f_2) \rangle)^2 \\
= (K^2 ||f_1||^2)^2 + (3KK_1 \langle f_1, I_1(f_2) \rangle)^2 + (2K^2 \langle I_1(f_2), I_1(f_2) \rangle)^2 \\
+ 6K^3K_1||f_1||^2 \langle f_1, I_1(f_2) \rangle + 4K^2K_1^2||f_1||^2 \langle I_1(f_2), I_1(f_2) \rangle.
\]

Now we need to calculate the second expectation in (4.37). The second expectation in (4.37) is given by

**Lemma 4.4.2.** \(F \in \mathbb{D}^{1,2}\) and \(E(F) = 0\). Then we have

\[ E((\langle DF, -DL^{-1}F \rangle)^2) = \Delta_4 + \Delta_5 + \Delta_6 + \Delta_8, \]

where \(\Delta_4 = (4.46)\), \(\Delta_5 = (4.48)\), \(\Delta_6 = (4.50)\), \(\Delta_8 = (4.51)\).
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus 127

**Proof.** Starting from (4.37),

\[
E((< DF, -DL^{-1} F >)^2) = E((K^2 ||f_1||^2)^2) + E((3KK_1 < f_1, I_1(f_2)>)^2) \\
+ E((2K_1^2 < I_1(f_2), I_1(f_2)>)^2) \\
+ 6K^3 K_1 E(||f_1||^2 < f_1, I_1(f_2)>) \\
+ 4K^2 K_1^2 E(||f_1||^2 < I_1(f_2), I_1(f_2)>) \\
+ 12K K_1^2 E(< f_1, I_1(f_2) >< I_1(f_2), I_1(f_2)>) \\
= \Delta_4 + \Delta_5 + \Delta_6 + \Delta_7 + \Delta_8 + \Delta_9 \quad (4.45)
\]

We know that \( \Delta_7 = \Delta_9 = 0 \) because \( E(I_1(f_2)) = 0 \) and \( E(I_1(f_2)I_1(f_2)I_1(f_2)) = 0 \)

Now we need to calculate \( \Delta_4, \Delta_5, \Delta_6 \) and \( \Delta_8 \). According to (4.33) (4.34) and (4.35), we get

\[
f_1(t) = e^{-\mu t} h_t \\
I_1(f_2) = I_1(f_2(t_1, t)) = \int_0^T \frac{1}{2}(e^{-\mu t} \mathbb{1}_{t_1 < t} + e^{-\mu t_1} \mathbb{1}_{t < t_1}) dW_t \\
f_2(t_1, t) = \frac{1}{2}(e^{-\mu t} \mathbb{1}_{t_1 < t} + e^{-\mu t_1} \mathbb{1}_{t < t_1}).
\]

Since \( f_1(t) = e^{-\mu t} h_t \) is deterministic,

\[
\Delta_4(h_t) = \Delta_4 = E((K^2 ||f_1||^2)^2) = K^4 \left( \int_0^T e^{-2\mu h_t^2 dt} \right)^2. \quad (4.46)
\]
Therefore,

\[
\Delta_5 = E\left( (3K_1 < f_1(t), I_1(f_2) >)^2 \right) \\
= 9K^2K_1^2E\left( \left( \int_0^T f_1(t)I_1(f_2(t_1, t))dt \right)^2 \right) \\
= 9K^2K_1^2E\left( \left( \int_0^T f_1(t)I_1(f_2(t_1, t))dt \right) \left( \int_0^T f_1(s)I_1(f_2(s_1, s))ds \right) \right) \\
= 9K^2K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s)I_1(f_2(t_1, t))I_1(f_2(s_1, s))dt ds \right) \\
= 9K^2K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) \left( I_1(f_2(t_1, t))I_1(f_2(s_1, s)) \right) dt ds \right) \\
= 9K^2K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) \int_0^T f_2^2(t_1, t)dt_1 dt ds \right), \quad \text{by Itô isometry (4.47)}
\]

where in the final step, the Itô isometry has been used. Using the fact that \( f_2(t_1, t) = \frac{1}{2T} \left( e^{-\mu t}1_{t_1 < t} + e^{-\mu t}1_{t < t_1} \right) \), and continuing from (4.47), we obtain

\[
\Delta_5 = \frac{9}{4}K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) \int_0^T \left( e^{-\mu t}1_{t_1 < t} + e^{-\mu t}1_{t < t_1} \right)^2 dt_1 dt ds \right) \\
= \frac{9}{4}K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) \int_0^T \left( e^{-2\mu t}1_{t_1 < t} + 0 + e^{-2\mu t}1_{t < t_1} \right) dt_1 dt ds \right) \\
= \frac{9}{4}K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) \left( \int_0^t e^{-2\mu t} dt_1 + \int_t^T e^{-2\mu t} dt_1 \right) dt ds \right) \\
= \frac{9}{4}K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) e^{-2\mu t} dt_1 ds + \int_0^T \int_0^T f_1(t)f_1(s) e^{-2\mu t} dt_1 ds \right) \\
= \frac{9}{4}K_1^2\left( \int_0^T \int_0^T f_1(t)f_1(s) e^{-2\mu t} dt ds + \int_0^T \int_0^T f_1(t)f_1(s) \frac{-1}{2\mu} \left( e^{-2\mu t} - e^{-2\mu t} \right) dt ds \right).
\]

(4.48)
Moreover,

\[
\Delta_6 = E\left( (2K_1^2 < I_1(f_2), I_1(f_2) > )^2 \right)
\]

\[
= 4K_1^4 E\left( \int_0^T I_1(f_1)(f_1)dt \right)^2
\]

\[
= 4K_1^4 E\left( \int_0^T \int_0^T I_1(f_2(u_1, t))I_1(f_2(u_2, t))I_1(f_2(u_3, s))I_1(f_2(u_4, s))dt ds \right)
\]

\[
= 4K_1^4 \int_0^T \int_0^T E\left( I_1(f_2(u_1, t))I_1(f_2(u_2, t))I_1(f_2(u_3, s))I_1(f_2(u_4, s)) \right) dt ds
\]

\[
= 4K_1^4 \int_0^T \int_0^T \left( < f_2(t), f_2(t) > < f_2(s), f_2(s) > + < f_2(t), f_2(s) > < f_2(t), f_2(s) > \right) dt ds
\]

Defining

\[
F(t, s) = < f_2(t), f_2(t) > < f_2(s), f_2(s) > + < f_2(t), f_2(s) > < f_2(t), f_2(s) > + < f_2(t), f_2(s) > < f_2(t), f_2(s) >,
\]

the proof continues as follows:

\[
\Delta_6 = 4K_1^4 \int_0^T \int_0^T F(t, s)dt ds
\]

\[
= 8K_1^4 \int_0^T \int_0^T F(t, s) dt ds \quad \text{since} \quad F(t, s) \text{ is symmetric}
\]

\[
= 8K_1^4 \int_0^T \int_0^s \left( < f_2(t), f_2(t) > < f_2(s), f_2(s) > + < f_2(t), f_2(s) > < f_2(t), f_2(s) > \right) dt ds
\]

\[
= 8K_1^4 \left( \int_0^T \int_0^s < f_2(t), f_2(t) > < f_2(s), f_2(s) > dt ds + \int_0^T \int_0^s < f_2(t), f_2(s) > < f_2(t), f_2(s) > dt ds + \int_0^T \int_0^s < f_2(t), f_2(s) > < f_2(t), f_2(s) > dt ds \right)
\]
Moreover,

\[
\Delta_6 = K^4_1 \left\{ A(\mu, T) + \int_0^T \int_0^s e^{-2\mu(t+s)} t^2 dtds \\
+ \int_0^T \int_0^s \left[ \frac{1}{\mu^2} (e^{-2\mu s} - e^{-\mu(s+t)})^2 + \left( \frac{1}{4\mu^2} \right)(e^{-2\mu T} - e^{-2\mu s})^2 \\
+ 2 \frac{1}{2\mu^2} (e^{-2\mu s} - e^{-\mu(s+t)})(e^{-2\mu T} - e^{-2\mu s}) \right] dtds \\
+ \int_0^T \int_0^s \left[ 2e^{-\mu(s+t)} t \left( \frac{1}{\mu} \right)(e^{-2\mu s} - e^{-\mu(s+t)}) \\
+ 2e^{-\mu(s+t)} t \left( \frac{1}{2\mu} \right)(e^{-2\mu T} - e^{-2\mu s}) \right] dtds \right\}
\]

\[
= K^4_1 \left\{ A(\mu, T) + B(\mu, T) + C(\mu, T) + D(\mu, T) \right\}, \tag{4.49}
\]

where \(A(\mu, T), B(\mu, T), C(\mu, T), \) and \(D(\mu, T)\) are calculated below

in Lemmas 4.4.3, 4.4.4, 4.4.5, 4.4.6. Then substituting the expressions obtained in these
Lemmas into (4.49) and rearranging, we find that
\[
\Delta_6 = K_1^3 \left\{ \frac{32\mu^2 T^2 - 24\mu T - 94}{64\mu^4} e^{-4\mu T} + \frac{8}{3\mu^4} e^{-3\mu T} + \frac{-2\mu T - 11}{8\mu^4} e^{-2\mu T} + \frac{17}{96\mu^4} \right\}
\]
\[
= K_1^3 \frac{1}{192\mu^4} \left\{ (96\mu^2 T^2 - 72\mu T - 282) e^{-4\mu T} + 512e^{-3\mu T} - (48\mu T + 264) e^{-2\mu T} + 34 \right\}.
\]
(4.50)

Finally,
\[
\Delta_8 = 4K^2 K_1^2 \| f_1 \|^2 E( < I_1(f_2), I_1(f_2) > )
\]
\[
= 4K^2 K_1^2 \left( \int_0^T f_1^2 dt \right) E( < I_1(f_2), I_1(f_2) > )
\]
\[
= K^2 K_1^2 \left( \int_0^T e^{-2\mu t} h_1^2 dt \right) \left( \frac{2\mu - 4\mu T - 1}{4\mu^2} e^{-2\mu T} + \frac{1}{4\mu^2} \right).
\]
(4.51)

Substitute (4.46) (4.48) (4.50) (4.51) into (4.45), and \( \Delta_7 = 0, \Delta_9 = 0 \). The proof is complete. \( \square \)

To compute (4.49), we provide the following four lemmas. Calculate \( A(\mu, T) \) in (4.49).

**Lemma 4.4.3.** Suppose \( \mu \) and \( T \) are constants, and define
\[
A(\mu, T) = \frac{1}{2} \int_0^T \int_0^s \left( e^{-2\mu t} - \frac{1}{2\mu} (e^{-2\mu T} - e^{-2t}) \right) \left( e^{-2\mu s} - \frac{1}{2\mu} (e^{-2\mu T} - e^{-2s}) \right) dt ds.
\]

Then we have
\[
A(\mu, T) = \frac{4\mu^2 T^2 + 4\mu T + 1}{16\mu^4} e^{-4\mu T} + \frac{-2\mu T - 1}{8\mu^4} e^{-2\mu T} + \frac{1}{16\mu^4}.
\]
(4.53)
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus

**Proof.** Starting from (4.49)

\[
A(\mu, T) = \frac{1}{2} \int_0^T \int_0^s \left( e^{-2\mu t} - \frac{1}{2\mu} e^{-2\mu T} - e^{-2\mu t} \right) \left( e^{-2\mu s} - \frac{1}{2\mu} e^{-2\mu T} - e^{-2\mu s} \right) dtds
\]

\[
= \frac{1}{2} \int_0^T \int_0^s q(t)q(s) dtds \quad (Q(t, s) = q(t)q(s) \text{ is symmetric})
\]

\[
= \frac{1}{4} \int_0^T \int_0^T q(t)q(s) dtds
\]

\[
\leq \frac{1}{4} \left[ \frac{(-2\mu T - 1)}{2\mu^2} e^{-2\mu T} + \frac{1}{2\mu^2} \right]^2
\]

\[
= \frac{1}{4} \left[ \frac{4\mu^2 T^2 + 4\mu T + 1}{4\mu^4} e^{-4\mu T} + \frac{-2\mu T - 1}{2\mu^4} e^{-2\mu T} + \frac{1}{4\mu^4} \right]
\]

\[
= \frac{4\mu^2 T^2 + 4\mu T + 1}{16\mu^4} e^{-4\mu T} + \frac{-2\mu T - 1}{8\mu^4} e^{-2\mu T} + \frac{1}{16\mu^4}.
\]

\[\square\]

\(B(\mu, T)\) in (4.49) is given by

**Lemma 4.4.4.** Suppose \(\mu\), and \(T\) are constants, and define

\[
B(\mu, T) = \int_0^T \int_0^s e^{-2\mu(t+s)} t^2 dtds = \int_0^T e^{-2\mu s} \int_0^s e^{-2\mu t} t^2 dtds.
\]

Then we have

\[
B(\mu, T) = \frac{1}{64\mu^4} \left( 1 - 8e^{-2\mu T} + e^{-4\mu T} (7 + 12\mu T + 8\mu^2 T^2) \right). \tag{4.54}
\]

**Proof.** From (4.49)

\[
B(\mu, T) = \int_0^T \int_0^s e^{-2\mu(t+s)} t^2 dtds = \int_0^T e^{-2\mu s} \int_0^s e^{-2\mu t} t^2 dtds,
\]

and we know that

\[
\int_0^s e^{-2\mu t} t^2 dt = -\frac{1}{2\mu} s^2 e^{-2\mu s} + \frac{1}{2\mu^2} [ -se^{-2\mu s} - \frac{1}{2\mu} (e^{-2\mu s} - 1)].
\]
Therefore,

\[ B(\mu, T) = \int_0^T \left( \frac{1}{4\mu^3} e^{-2\mu s} - \frac{1}{4\mu^3} e^{-4\mu s} - \frac{1}{2\mu} s^2 e^{-4\mu s} - \frac{1}{2\mu^2} s e^{-4\mu s} \right) ds \]

\[ = b_1 - b_2 - b_3 - b_4, \]

where

\[ b_1 = \frac{1}{8\mu^4} (1 - e^{-2\mu T}), \]

\[ b_2 = \frac{1}{16\mu^4} (1 - e^{-4\mu T}), \]

\[ b_3 = \frac{1}{64\mu^4} \left[ 1 - e^{-4\mu T} (1 + 8\mu^2 T^2 + 4\mu T) \right], \]

and

\[ b_4 = \frac{1}{32\mu^4} (1 - e^{-4\mu T} - 4\mu T e^{-4\mu T}). \]

Therefore,

\[ B(\mu, T) = \frac{1}{64\mu^4} \left( 1 - 8e^{-2\mu T} + e^{-4\mu T} \left( 7 + 12\mu T + 8\mu^2 T^2 \right) \right). \]

\[ C(\mu, T) \text{ in (4.49) is given by} \]

**Lemma 4.4.5.** Suppose \( \mu \), and \( T \) are constants, and define

\[ C(\mu, T) = \int_0^T \int_0^s \left[ \frac{1}{\mu^2} (e^{-2\mu s} - e^{-\mu(s+t)})^2 + \frac{1}{4\mu^2} (e^{-2\mu T} - e^{-2\mu s})^2 \right. \]

\[ + \left. 2 \frac{1}{2\mu^2} (e^{-2\mu s} - e^{-\mu(s+t)})(e^{-2\mu T} - e^{-2\mu s}) \right] dt ds. \]
Then we have
\[
C(\mu, T) = \frac{8\mu^2T^2 - 20\mu T - 49}{64\mu^4} e^{-4\mu T} + \frac{4}{3\mu^4} e^{-3\mu T} - \frac{5}{8\mu^4} e^{-2\mu T} + \frac{11}{192\mu^4}. \tag{4.55}
\]

**Proof.** From (4.49)

\[
C(\mu, T) = \int_0^T \int_0^s \left[ \frac{1}{\mu^2} (e^{-2\mu s} - e^{-\mu(s+t)})^2 + \left( \frac{1}{4\mu^2} \right) (e^{-2\mu T} - e^{-2\mu s})^2 \right] dtds
\]
\[
+ \frac{1}{2\mu^2} (e^{-2\mu s} - e^{-\mu(s+t)})(e^{-2\mu T} - e^{-2\mu s}) \right] dtds
\]
\[
= \int_0^T \int_0^s \left[ \frac{1}{4\mu^2} e^{-4\mu s} - \frac{1}{\mu^2} e^{-3\mu s - \mu t} + \frac{1}{\mu^2} e^{-2\mu s - 2\mu t} + \frac{1}{4\mu^2} e^{-4\mu T} \right] dtds
\]
\[
+ \frac{1}{2\mu^2} e^{-2\mu T - 2\mu s} - \frac{1}{\mu^2} e^{-\mu s - \mu t - 2\mu T} \right] dtds
\]
\[
= \left[ \frac{-4\mu T - 7}{64\mu^4} e^{-4\mu T} + \frac{1}{64\mu^4} \right] + \left[ \frac{-1}{4\mu^4} e^{-4\mu T} + \frac{1}{3\mu^4} e^{-3\mu T} - \frac{1}{12\mu^4} \right]
\]
\[
+ \left[ \frac{-2\mu T - 7}{8\mu^4} e^{-4\mu T} + \frac{1}{8\mu^4} e^{-2\mu T} \right] + \left[ \frac{-1}{2\mu^4} e^{-4\mu T} + \frac{1}{\mu^4} e^{-3\mu T} - \frac{1}{2\mu^4} e^{-2\mu T} \right]
\]
\[
= \frac{8\mu^2T^2 - 20\mu T - 49}{64\mu^4} e^{-4\mu T} + \frac{4}{3\mu^4} e^{-3\mu T} - \frac{5}{8\mu^4} e^{-2\mu T} + \frac{11}{192\mu^4}.
\]

Compute \(D(\mu, T)\) in (4.49).

**Lemma 4.4.6.** Suppose \(\mu\), and \(T\) are constants, and define

\[
D(\mu, T) = \int_0^T \int_0^s \left[ 2e^{-\mu(s+t)} \left( \frac{1}{\mu} \right) (e^{-2\mu s} - e^{-\mu(s+t)}) \right. \\
\left. + 2e^{-\mu(s+t)} \left( \frac{1}{2\mu} \right) (e^{-2\mu T} - e^{-2\mu s}) \right] dtds.
\]

Then we have
\[
D(\mu, T) = \frac{-4\mu T - 7}{8\mu^4} e^{-4\mu T} + \frac{4}{3\mu^4} e^{-3\mu T} - \frac{1}{2\mu^4} e^{-2\mu T} + \frac{1}{24\mu^4}. \tag{4.56}
\]
Proof. From (4.49),

\[
D(\mu, T) = \int_0^T \int_0^s \left[ 2e^{-\mu(s+t)}t(\frac{-1}{2\mu})(e^{-2\mu s} - e^{-\mu(s+t)})
+ 2e^{-\mu(s+t)}t(\frac{-1}{2\mu})(e^{-2\mu T} - e^{-2\mu s}) \right] dt ds
\]

\[
= \int_0^T \int_0^s \left[ \frac{-1}{\mu}e^{-3\mu s-\mu t} + \frac{2}{\mu}e^{-2\mu s-2\mu t} - \frac{1}{\mu}e^{-\mu s-\mu t-2\mu T} \right] dt ds
\]

\[
= \left[ \frac{-4\mu T - 5}{16\mu^4} e^{-4\mu T} + \frac{1}{3\mu^4} e^{-3\mu T} - \frac{1}{48\mu^4} \right]
+ \left[ \frac{4\mu T + 3}{16\mu^4} e^{-4\mu T} - \frac{1}{4\mu^4} e^{-2\mu T} + \frac{1}{16\mu^4} \right]
+ \left[ \frac{-2\mu T - 3}{4\mu^4} e^{-4\mu T} + \frac{1}{3\mu^4} e^{-3\mu T} - \frac{1}{4\mu^4} e^{-2\mu T} \right]
\]

\[
= \frac{-4\mu T - 7}{8\mu^4} e^{-4\mu T} + \frac{4}{3\mu^4} e^{-3\mu T} - \frac{1}{2\mu^4} e^{-2\mu T} + \frac{1}{24\mu^4}.
\]

\[ \square \]

4.4.3 Upper bound estimation

According to Theorem 4.4.1, Lemma 4.4.1, and Lemma 4.4.2, we get the following main theorem.

Theorem 4.4.2. Let \( S^N_T \) and \( S^F_T \) be defined by (4.19) and (4.20) respectively. The Wasserstein distance between the probability laws of \( S^N_T \) and \( S^F_T \) is bounded by

\[
d_W(S^F_T, S^N_T) \leq |M|\sqrt{1 - 2(\Delta_1 + \Delta_3)} + \Delta_4 + \Delta_5 + \Delta_6 + \Delta_8,
\]  

(4.57)

where \( M, \Delta_1, \Delta_3, \Delta_4, \Delta_5, \Delta_6, \) and \( \Delta_8 \), are given by (4.23), (4.38), (4.39), (4.46), (4.48), (4.50), (4.51) respectively; moreover \( K = \frac{\sigma}{M} e^{\mu T} \), and \( K_1 = K\epsilon \) are given by (4.30), (4.31).

Moreover, define

\[
\|h_t - 1\|_{\infty} = \sup_{0 \leq t \leq T} |h_t - 1|.
\]
Then as $\varepsilon \to 0$ and $\|h_t - 1\|_\infty \to 0$, we have

$M$ does not depend on $\varepsilon$ or $\|h_t - 1\|_\infty$;

$\Delta_1 (h_t) = \Delta_1 (1) + O(\|h_t - 1\|_\infty)$;

$\Delta_3 = O(\varepsilon^2)$;

$\Delta_4 (h_t) = \Delta_4 (1) + O(\|h_t - 1\|_\infty)$;

$\Delta_5 = O(\varepsilon^2) [C_5 + O(\|h_t - 1\|_\infty)]$, where $C_5$ is a constant;

$\Delta_6 = O(\varepsilon^4)$;

$\Delta_8 = O(\varepsilon^2) [C_8 + O(\|h_t - 1\|_\infty)]$, where $C_8$ is a constant;

**Proof.** Using the equation (4.28) and Theorem 4.4.1, we have

$$d_W(S_T^F, S_T^N) = |M|d_W(F, N) \leq |M|\sqrt{Ex^2}$$

Substitute (4.37) into (4.58), and using Lemma 4.4.1, and Lemma 4.4.2, we get

$$d_W(S_T^F, S_T^N) \leq |M|\sqrt{1 - 2(\Delta_1 + \Delta_3) + \Delta_4 + \Delta_5 + \Delta_6 + \Delta_7 + \Delta_8},$$

where $\Delta_2 = \Delta_7 = \Delta_0 = 0$

Therefore,

$$d_W(S_T^F, S_T^N) \leq |M|\sqrt{1 - 2(\Delta_1 + \Delta_3) + \Delta_4 + \Delta_5 + \Delta_6 + \Delta_8}.$$

Moreover, we know that

$$\Delta_1 (h_t) = K^2 \int_0^T e^{-2\mu t} h_t^2 dt.$$
Then,

\[ |\Delta_1(h_t) - \Delta_1(1)| = \left| K^2 \int_0^T e^{-2\mu t}(h_t^2 - 1)dt \right| \]
\[ = \left| K^2 \int_0^T e^{-2\mu t}(h_t - 1)(h_t + 1)dt \right| \]
\[ \leq K^2\|h_t - 1\|_\infty \int_0^T e^{-2\mu t}|h_t + 1|dt. \]

Therefore,

\[ \Delta_1(h_t) = \Delta_1(1) + O(\|h_t - 1\|_\infty). \]

According to (4.44),

\[ \Delta_3 = O(\varepsilon^2). \]

Based on (4.46),

\[ \Delta_4(h_t) = K^4\left( \int_0^T e^{-2\mu t}h_t^2dt \right)^2. \]

Then,

\[ |\Delta_4(h_t) - \Delta_4(1)| = \left| K^4\left( \int_0^T e^{-2\mu t}h_t^2dt \right)^2 - K^4\left( \int_0^T e^{-2\mu t}dt \right)^2 \right| \]
\[ = K^4\left| \left( \int_0^T e^{-2\mu t}h_t^2dt \right)^2 - \left( \int_0^T e^{-2\mu t}dt \right)^2 \right| \]
\[ = K^4\left| \int_0^T e^{-2\mu t}(h_t^2 + 1)dt \int_0^T e^{-2\mu t}(h_t^2 - 1)dt \right| \]
\[ = K^4\left| \int_0^T e^{-2\mu t}(h_t^2 + 1)dt \int_0^T e^{-2\mu t}(h_t - 1)(h_t + 1)dt \right| \]
\[ \leq K^4\|h_t - 1\|_\infty \left| \int_0^T e^{-2\mu t}(h_t^2 + 1)dt \int_0^T e^{-2\mu t}(h_t + 1)dt \right|. \]

Therefore,

\[ \Delta_4(h_t) = \Delta_4(1) + O(\|h_t - 1\|_\infty). \]
According to (4.48), we have

\[ |\Delta_5| = \left| \frac{9}{4} K^2 K_1^2 \left( \int_0^T \int_0^T f_1(t)f_1(s)e^{-2\mu t}dt ds + \int_0^T \int_0^T f_1(t)f_1(s)\frac{-1}{2\mu}(e^{-2\mu T} - e^{-2\mu t})dt ds \right) \right| \]

\[ = \frac{9}{4} K^2 K_1^2 \left| \int_0^T \int_0^T f_1(t)f_1(s)G_0(t, s)dt ds \right| \]

\[ = \frac{9}{4} K^2 K_1^2 \left| \int_0^T \int_0^T e^{-\mu t}e^{-\mu s}h_tG_0(t, s)dt ds \right| \]

\[ = \frac{9}{4} K^2 K_1^2 \left| \int_0^T \int_0^T h_tG_1(t, s)dt ds \right| \]

\[ \leq \frac{9}{4} K^2 K_1^2 \left( \left| \int_0^T \int_0^T (h_t - 1)(h_s - 1)G_1(t, s)dt ds \right| + \left| \int_0^T \int_0^T (h_t - 1)G_1(t, s)dt ds \right| \right. \]

\[ + \left. \left| \int_0^T \int_0^T (h_s - 1)G_1(t, s)dt ds \right| + \left| \int_0^T \int_0^T G_1(t, s)dt ds \right| \right). \]

Therefore,

\[ \Delta_5 = O(\varepsilon^2)[C_5 + O(\|h_t - 1\|_{\infty})], \]

where \( C_5 \) is a constant.

Based on (4.50),

\[ \Delta_6 = K_1^2 \frac{1}{192\mu^4} \left\{ (96\mu^2 T^2 - 72\mu T - 282)e^{-4\mu T} + 512e^{-3\mu T} - (48\mu T + 264)e^{-2\mu T} + 34 \right\}. \]

We have \( \Delta_6 = O(\varepsilon^4). \)

According to (4.52),

\[ |\Delta_8| = K^2 K_1^2 \left| \left( \int_0^T e^{-2\mu t}h_t^2 dt \right) \left( \frac{2\mu - 4\mu T - 1}{4\mu^2} e^{-2\mu T} + \frac{1 - 2\mu}{4\mu^2} \right) \right| \]

\[ = K^2 K_1^2 \left| \int_0^T h_t^2 K(t)dt \right| \]

\[ = K^2 K_1^2 \left| \left( \int_0^T (h_t - 1 + 1)(h_t - 1 + 1)K(t)dt \right) \right|. \]
Therefore, \( \Delta_8 = O(\varepsilon^2)[C_8 + O(||h_t - 1||_\infty)] \), where \( C_8 \) is a constant.

The proof is complete. \( \Box \)

From Theorem 4.4.2, we get the following corollary.

**Corollary 4.4.1.** When \( h_t = 1 \) in Theorem 4.4.2 , we have

\[
d_W(S_T^F, S_T^N) \leq \varepsilon \sigma \sqrt{b_5 + \varepsilon^2 b_6}, \tag{4.59}
\]

where \( b_5 = (4.71) \) and \( b_6 = (4.72) \).

**Proof.** According to (4.41), if \( h_t = 1 \),

\[
\Delta_1 = K^2 \left( \frac{-1}{2\mu} e^{-2\mu T} + \frac{1}{2\mu} \right) = \frac{2\mu e^{2\mu T}}{e^{2\mu T} - 1} (1 - e^{-2\mu T}) \frac{1}{2\mu} = 1. \tag{4.60}
\]

Due to (4.46), if \( h_t = 1 \),

\[
\Delta_4 = \Delta_1^2 = 1. \tag{4.61}
\]

From (4.51), if \( h_t = 1 \),

\[
\Delta_8 = 4K^2 K_1^2 \left( \int_0^T f_1^2 \, dt \right) E( < I_1(f_2), I_1(f_2) > ) = 4K^2 K_1^2 \left( \int_0^T e^{-2\mu t} \, dt \right) \frac{1}{4} \left( \frac{2\mu - 4\mu T - 1}{4\mu^2} e^{-2\mu T} + \frac{1 - 2\mu}{4\mu^2} \right) = K^2 K_1^2 \left( \frac{1 + 4\mu T - 2\mu}{8\mu^3} e^{-2\mu T} + \frac{2\mu - 2\mu T - 1}{4\mu^3} e^{-2\mu T} + \frac{1 - 2\mu}{8\mu^3} \right),
\]

and

\[
\Delta_8 = 2\Delta_3. \tag{4.62}
\]
Substituting (4.60) (4.61) (4.62) into (4.57), we get

\[ d_W(S_T^F, S_T^N) \leq M \sqrt{\Delta_5 + \Delta_6}, \]  

where, based on (4.23),

\[ M = \sigma \sqrt{e^{2\mu T} - \frac{1}{2\mu}} > 0. \]  

From (4.30) and (4.31), we have

\[ K = \frac{\sigma}{M} e^{\mu T} = \frac{e^{\mu T}}{\sqrt{e^{2\mu T} - \frac{1}{2\mu}}}, \]

\[ K_1 = K\varepsilon = \frac{\varepsilon e^{\mu T}}{\sqrt{e^{2\mu T} - \frac{1}{2\mu}}}. \]

Now substitute \( K \) and \( K_1 \) into the equations below, we compute \( \Delta_5 \) and \( \Delta_6 \).

According to (4.48), if \( h_t = 1 \),

\[ \Delta_5 = \frac{9}{4} K^2 K_1^2 \left( \int_0^T \int_0^T t e^{-\mu s} e^{-3\mu t} dt ds + \int_0^T \int_0^T e^{-\mu t} e^{-\mu s} - \frac{1}{2\mu} (e^{-2\mu t} - e^{-2\mu s}) dt ds \right) \]

\[ = \frac{9}{4} K^2 K_1^2 \left( \frac{1}{9\mu^3} (1 - (3\mu T + 1) e^{-3\mu T})(1 - e^{-\mu T}) \right. \]

\[ + \frac{1}{6\mu^3} (1 + 2 e^{-3\mu T} - 3 e^{-2\mu T})(1 - e^{-\mu T}) \]  

\[ = \varepsilon^2 \frac{4\mu^2 e^{4\mu T}}{(e^{2\mu T} - 1)^2} \frac{9}{4} \left( \frac{1}{9\mu^3} (1 - (3\mu T + 1) e^{-3\mu T})(1 - e^{-\mu T}) \right. \]

\[ + \frac{1}{6\mu^3} (1 + 2 e^{-3\mu T} - 3 e^{-2\mu T})(1 - e^{-\mu T}) \]  

\[ = \varepsilon^2 \frac{4\mu^2 e^{4\mu T}}{(e^{2\mu T} - 1)^2} \frac{9}{4} \lambda_1 \]  

\[ = \varepsilon^2 a_5. \]
Moreover, from (4.50)

$$
\Delta_6 = K_1 \frac{1}{192 \mu^4} \left\{ (96 \mu^2 T^2 - 72 \mu T - 282) e^{-4 \mu T} + 512 e^{-3 \mu T} - (48 \mu T + 264) e^{-2 \mu T} + 34 \right\}
$$

$$
= \varepsilon^4 \frac{4 \mu^2 e^{4 \mu T}}{(e^{3 \mu T} - 1)^2} \frac{1}{192 \mu^4} \left\{ (96 \mu^2 T^2 - 72 \mu T - 282) e^{-4 \mu T} + 512 e^{-3 \mu T} - (48 \mu T + 264) e^{-2 \mu T} + 34 \right\}
$$

$$
= \varepsilon^4 \frac{e^{4 \mu T}}{48 \mu^4 (e^{3 \mu T} - 1)^2} \left\{ (96 \mu^4 T^2 - 72 \mu^3 T - 282 \mu^2) + 512 e^{-3 \mu T} - (48 \mu^3 T + 264 \mu^2) e^{2 \mu T} + 34 \mu^2 e^{4 \mu T} \right\}
$$

$$
= \varepsilon^4 \frac{e^{4 \mu T}}{48 \mu^4 (e^{3 \mu T} - 1)^2} \lambda^2
$$

(4.67)

$$
= \varepsilon^4 a_6.
$$

(4.68)

Substitute (4.66) and (4.68) into (4.63), we get

$$
d_W(S_F^T, S_N^T) \leq \varepsilon M \sqrt{a_5} + \varepsilon^2 a_6.
$$

(4.69)

Substitute $M$ (4.64) into (4.69), we get

$$
d_W(S_F^T, S_N^T) \leq \varepsilon \sigma \sqrt{b_5} + \varepsilon^2 b_6,
$$

(4.70)

where $b_5$ = (4.71) and $b_6$ = (4.72):

$$
b_5(\mu, T) = \frac{9}{2} \frac{\mu e^{4 \mu T}}{(e^{3 \mu T} - 1)} \left( \frac{1}{9 \mu^3} (1 - (3 \mu T + 1) e^{-3 \mu T}) (1 - e^{-\mu T}) + \frac{1}{6 \mu^4} (1 + 2 e^{-3 \mu T} - 3 e^{-2 \mu T}) (1 - e^{-\mu T}) \right).
$$

(4.71)
Chapter 4. Option pricing approximation by Stein’s method and Malliavin calculus

\[ b_0(\mu, T) = \frac{1}{96\mu^5(e^{2\mu T} - 1)} \left\{ \left( 96\mu^4 T^2 - 72\mu^3 T - 282\mu^2 \right) + 512\mu^2 e^{\mu T} - \left( 48\mu^3 T + 264\mu^2 \right) e^{2\mu T} + 34\mu^2 e^{4\mu T} \right\}. \]

(4.72)

According to (4.65) and (4.67),

\[ \lambda_1 = \left( \frac{1}{9\mu^8} (1 - (3\mu T + 1) e^{-3\mu T}) (1 - e^{-\mu T}) + \frac{1}{6\mu^3} (1 + 2 e^{-3\mu T} - 3 e^{-2\mu T}) (1 - e^{-\mu T}) \right), \]

and

\[ \lambda_2 = \left\{ \left( 96\mu^4 T^2 - 72\mu^3 T - 282\mu^2 \right) + 512\mu^2 e^{\mu T} - \left( 48\mu^3 T + 264\mu^2 \right) e^{2\mu T} + 34\mu^2 e^{4\mu T} \right\}. \]

The proof is complete. \( \Box \)

In the very beginning, (4.17) and (4.18) are shorthand for (4.19) and (4.20).

Now set \( h_t = 1 \) and \( \varepsilon = 0 \), the following two stochastic processes are the same

\[ dS_t = \mu S_t dt + \sigma dW_t, \]

\[ dS_t = \mu S_t dt + \sigma (h_t + \varepsilon W_t) dW_t. \]

Therefore, in Corollary 4.4.1 we get \( dW(S_T^F, S_T^N) = |M| dW(F, N) \leq 0 \). The upper bound is zero, which is what we expect.

4.5 Illustrative examples

In section 4.2.1, we state that the upper bound (4.8) we calculate can be applied to a class of functions \( h \in Lip(1) \) in (4.6), not just one specific function. The result is very general. In (4.2) and (4.5), we show that we can even apply the upper bound to a function which is not differentiable at one point. In practice, although the volatility is stochastic based on observations, traders have to quote a constant volatility to their
clients. Therefore, traders can use our results in this chapter to see if their constant volatility quote is close to the more realistic stochastic volatility model. We provide numerical results in this section.

For convenience, we recall the definition of $S_T^N$ in (4.19):

$$S_T^N = S_0 e^{\mu T} + \sigma \int_0^T e^{\mu (T-u)} dW_u. \quad (4.73)$$

Note that, as defined, $S_T^N$ is normally distributed.

Due to the (4.2) and (4.5), we can apply Corollary 4.4.1 to the function $(x-K)^+$. According to Theorem 4.2.1 and (4.10) (4.11) (4.12), when the spread call option is at the money (that is $S_0 = K$), for example, if $S_0 = 60$ and $K = 60$, $\sigma = 0.3$, $\mu = r = 0.02$, $T = 1$ are given, we get

$$P = E[e^{-rT}(S_T^N - K)^+]$$
$$= e^{-rT}(\mu S_T^N - K)N(d) + e^{-rT}\sigma S_T^N n(d)$$
$$= 1.18808,$$

where numerical accuracy is presented up to 5 decimal places.

In the money (that is $S_0 > K$), $S_0 = 65$ and $K = 60$, keeping the parameters $\sigma, \mu = r, T$ the same as above, we have

$$P = e^{-rT}(\mu S_T^N - K)N(d) + e^{-rT}\sigma S_T^N n(d)$$
$$= 6.18808.$$

Out of the money (that is $S_0 < K$), $S_0 = 55$ and $K = 60$, keeping the parameters $\sigma, \mu = r, T$ the same as above, and then

$$P = e^{-rT}(\mu S_T^N - K)N(d) + e^{-rT}\sigma S_T^N n(d)$$
$$= 1.21452e^{-39}$$
$$\approx 0.$$
Example 4.5.1. In Corollary 4.4.1 set $h_t = 1$, and use the same values for the parameters, $\sigma = 0.3$, $\mu = r = 0.02$, $\varepsilon = 0.01$, and $T = 1$. Then we get

$$d_W(S^F_T, S^N_T) \leq 0.00453,$$

which is about 0.381% of the at-the-money option price, and only 0.073% of in-the-money option value.

Therefore, as we discuss in the section 4.2.2, we may consider

$$P = E[e^{-rT}(S_T - K)^+] = e^{-rT}(\mu S_T - K)N(d) + e^{-rT}\sigma S_T n(d)$$

as an acceptable normal approximation for

$$\tilde{P} = E[e^{-rT}(\tilde{S}_T - K)^+],$$

where $\tilde{S}_T$ follows (4.15) with $S^N_T = S_T$ and $S^F_T = \tilde{S}_T$.

Note that $S_0$ only appears in the drift term in (4.73), and it has been eliminated during the proof in Theorem 4.28, so it has no influence.

Example 4.5.2. Using Corollary 4.4.1, we now choose $\varepsilon = 1$ but with the other quantities unchanged, i.e. $h_t = 1$, $\sigma = 0.3$, $\mu = r = 0.02$, and $T = 1$. Then we get

$$d_W(S^F_T, S^N_T) \leq 0.52267$$

which is a big 43.993% of the at-the-money option price, and also 8.446% of in-the-money option value.

From the two examples above, $h_t$ being constant to 1, we see that when $\varepsilon$ is small,

$$P = E[e^{-rT}(S_T - K)^+] = e^{-rT}(\mu S_T - K)N(d) + e^{-rT}\sigma S_T n(d)$$

could be a good normal approximation for $\tilde{P} = E[e^{-rT}(\tilde{S}_T - K)^+]$.

However, when $\varepsilon$ is large, normal approximation is not a good choice.

Example 4.5.3. Set $h_t = 1$, $\sigma = 0.3$.

In the equation (4.59), $\sigma$ only appears once as a multiplier, so its effect upon upper bound is clear. $\varepsilon$ appears twice, but $b_5(\mu, T)$ and $b_6(\mu, T)$ are complicated functions of $\mu$ and $T$. Therefore, under different values of $\varepsilon$, that is 0.01, 0.5, and 1, in Figure 4.1, 4.2, and 4.3, we show upper bound surface for $d_W(S^F_T, S^N_T)$ in Corollary 4.4.1 in the domain of $\mu \times T$.

The ranges, 0.5% $\sim$ 20% for $\mu$, and 0.1 $\sim$ 5 years for $T$ are chosen. We see that when we shift $\varepsilon$, we change the magnitude of upper bound surface significantly.
Figure 4.1: When $\varepsilon = 0.01$, the values of upper bound surface of combination of $\mu$ and $T$ are all less than 0.035.

Figure 4.2: When $\varepsilon = 0.5$, the values of upper bound surface of combination of $\mu$ and $T$ are all less than 2.
4.6 Summary

We use the closed form solution of one option price whose underlying asset follows normal distribution to approximate the option price whose underlying asset has stochastic volatility by Stein’s method and Malliavin calculus since financial option prices could be calculated by expected discounted payoffs, and an upper bound of a metric, Wasserstein distance, is computed for the difference between two expectations. The main results are Theorem 4.4.2 and Corollary 4.4.1. Some numerical examples are given.
Chapter 5

Conclusions and further research

5.1 Conclusions

By comparing expectations, Stein’s method provides a way accurately approximating one probability distribution by another. An upper bound is computed for the difference between the expectations of a large family of test functions under the two distributions. An associated metric is determined by the family of test functions. In this thesis, the test functions we consider consist of all Lipschitz functions $h$ with constant bounded by 1. Then the associated metric is called Wasserstein distance. In order to compute the explicit upper bound of Wasserstein distance, Nourdin and Peccati [29] obtain the bounds for the normal approximation of smooth functionals of Gaussian fields using Stein’s method and Malliavin calculus on a Gaussian space. Furthermore, Peccati, Solé, Taqqu and Utzet [35] extend the analysis to the framework of normal approximation in the Wasserstein distance with a version of Malliavin calculus on the Poisson space.

In Chapter 3, we derive upper bounds for the Wasserstein distance between a Lévy functional and a normal random variable using generalized Lévy white noise analysis [24] initiated by Hida [12]. We get upper bound formula for Lévy functionals using Lévy white noise analysis operators. We also derive the chain rule for Lévy white noise functionals; see Theorem 3.2.10. The principal novel contribution is the rigorous proof of the Stein bound for functionals of general Lévy processes; see Theorem 3.3.1. A second contribution is to show that the Stein bounds in the literature for functionals of (i) pure jump Lévy processes, and (ii) the pure Brownian motion Lévy process, are both
particular cases of the general bound in the thesis; see Corollary 3.3.1 and Corollary 3.3.2. To derive the Stein bound in the thesis, it is essential to adopt a unified approach to the Brownian part and pure jump part of the Lévy process. In the thesis, this was achieved by making use of the white noise approach developed by Lee and Shih (2004). We give a new Example 3.3.3 which computes the upper bound for the double integral. We show that our approach unifies the Brownian and pure jump components in a natural way, and simplifies the computation. It appears that white noise theory and Malliavin calculus are not mathematically equivalent theories; e.g. the Malliavin calculus has no role for generalized functions. Nevertheless, it might be the case that, when specialized to deriving Stein bounds, the two theories are mathematically equivalent, in the sense that each theory produces the same general Stein bound. In new Example 3.3.2, we compute the upper bound for a functional which uses notation specific to white noise analysis.

In financial mathematics, financial option prices could be calculated by a risk-neutral pricing formula, which is the expected discounted payoff. Assume the payoff is a Lipschitz function. Stein’s method provides a way of constructing an accurate approximation of one probability distribution by another. An upper bound of the Wasserstein distance is computed for the difference between the expectations of a large family of Lipschitz functions under the two distributions. Therefore, we can use Stein’s method and Malliavin calculus to do financial option price approximation. In Chapter 4, we use the closed form expression of one option price whose underlying asset follows a normal distribution to approximate an option price whose underlying asset has stochastic volatility. The main results are Theorem 4.4.2 and Corollary 4.4.1.

5.2 Further research

As we see in the equation (4.29), when dealing with the stochastic volatility processes (4.18), one needs explicit upper bounds for random variables of the type $F = I_1(F_1) + I_2(F_2)$, that is, random variables that are the sum of a single and a double integral. For Lévy functionals, in Example 3.3.3, we compute the explicit upper bound of Wasserstein distance of a double integral. Therefore, one further research direction is to compute explicit upper bounds of Wasserstein distance, $d_W(F, X)$, where $F = I_1(F_1) + I_2(F_2) + I_3(F_3)$, and $F = I_1(F_1) + I_2(F_2) + I_3(F_3) + I_4(F_4)$ using our general Theorem 3.3.1 for Lévy functionals.
In Chapter 4, we calculate an explicit upper bound for the Wasserstein distance between two stochastic differential equations, which are

\[ dS_t = \mu S_t dt + \sigma dW_t \]
\[ dS_t = \mu S_t dt + \sigma (h_t + \varepsilon W_t) dW_t, \]

(5.1)

where \( \{W_t\}_{t \geq 0} \) is standard Brownian motion. For further research, we can replace the stochastic differential equation (5.1) above with a more complicated Lévy process. Then, using our general Theorem 3.3.1 for Lévy functionals, we can compute an explicit upper bound for the Wasserstein distance.
Chapter 6

Questions and Actions Taken

6.1 First some overall comments:

I have searched carefully throughout the theses for instances of Point 1-6 below and have corrected all the instances I have found.

1. Fairly often the period at the end of a sentence is missing, in particular when the sentence ends with a mathematical formula.
2. No sentence should start with a mathematical symbol.
3. Only those formulas should be labelled to which you refer to later in the text.
4. Examples should not have proofs either the example is a corollary, or you just write ”To see this, ...” or similar.
5. The type face should be unified; do not suddenly start using bold face sometimes but only from page 113 onwards.
6. In a lot of formulas, numbers are assigned to each line. Only those formulas and lines should be numbered which are then used in the text.
6.2 Abstract

1. The Abstract does not follow the usual requirements as it is not self-contained and refers to particular theorems and results from the thesis.
   A: It has been restructured.

2. Also, it is not the case that each theory produces the same general Stein bound; bounds are derived based on the skills of the researcher. Rather it would be prudent to say that the two approaches can be applied so that they give the same bound for those test functions for which both approaches apply.
   A: It has been deleted.

6.3 Chapter 1 (Introduction)

1. The formulas on page 2 are not well explained and are not helpful for a first reading of the material. It may be better to omit them.
   A: P.1, P.2, we delete them.

2. Page 3 point 4: see the discussion for the Abstract.
   A: P.2, it has been deleted.

3. Page 4 3rd paragraph: what do you mean by efficient in this context? Delete if not essential.
   A: P.3, it has been deleted.

4. Page 7 last paragraph: rephrase; the part of the sentence which starts with "if the test functions ..." is awkward. Perhaps better: The Wasserstein distance is defined as follows: and then give a proper definition of the Wasserstein distance as a formula.
   A: P.6, it has been reworded.

6.4 Chapter 2 (Review of relevant literature)

1. Give references for the definitions.
   A: P.7, P.8, P.9, P.10, P.11, they have been provided.
2. **Page 9:** what is a filtered probability space? Should there be the filtration in the notation?
A: P.8, it has been provided.

3. **Definition 2.2.14** The sentence about the Banach space does not make sense as the Banach space requires a norm; delete.
A: P.11, it has been deleted.

4. **Definition 2.2.18** Define first what a nuclear space is.
A: P.13, it has been restructured.

5. **Page 14 line -4:** "The" probability space, as you refer to the notation in the Definition above. Also "stochastic" is not spelled correctly.
A: P.14, it has been reworded.

6. **Definition 2.2.25:** you refer to this definition as independent random measure use the name Poisson random measure throughout the dissertation as it is more precise.
A: P.36, it has been reworded.

7. **Definition 2.2.27** what is B, and what is the filtration?
A: P.16, it has been provided.

8. **Page 17** Define or explain "elementary functions"
A: P.16, it has been provided.

9. **(2.26)** you use many different notations for derivatives (dot, prime, \( D \))...; make clear what you mean by the dot and by the prime.
A: P.18, it has been provided.

10. **Malliavin calculus** should be introduced in this chapter rather than in Chapter 4, as you already use it in Chapter 3.
A: P.18, it has been restructured.

### 6.5 Chapter 3 (Stein’s methods and normal approximation of Lévy functionals)

1. **Page 19:** see the discussion from the Abstract about producing bounds.
A: P. 30, it has been deleted.
Chapter 6. Questions and Actions Taken

2. **Page 20** Re-phrase ”After reviewing Hidas work and notation” so that you refer to a section where this material has been presented in Chapter 2 (currently, Section 2.3).
   A: P. 31, we provide the reference.

3. **Page 20**, say that (3.1) is the Lévy-Khintchine formula.
   A: P. 31, it has been reworded.

4. **Lemma 3.2.1** should need a condition on the integral of $|\eta|^3$.
   A: P. 34, we add the description of the formula (3.9).

5. (3.8) and (3.9) are misleading as stated as exchanging derivative and expectation has not yet been justified; delete.
   A: P. 33, it has been deleted.

6. (3.23) Here $C$ should depend on $\eta$.
   A: P. 34, $C$ does not depend on $\eta$. $C \in (0, \infty)$ is a constant.

7. **Page 25** last paragraph: here is an example where you write independent random measure but you seem to mean Poisson random measure.
   A: P. 36, it has been reworded.

8. **Page 27**: label the definition of the integral as you refer to it later.
   A: P. 38, it has been done.

9. **Page 28** line -2: the notation for the product measure is different to (3.33) where you do not write the direct product; stick to one formulation.
   A: P. 39, it has been reworded.

10. **Theorem 3.2.2**. The definition of the exponential vector should not be part of the theorem.
    A: P. 39, it has been restructured.

11. Explain why (3.46) and (3.47) hold.
    A: P. 40, we provide the explanation.

12. **Page 31** first line: there is a word missing after upper.
    A: P. 41, it has been fixed.

13. **Define in Chapter 2** what a Gelfand triple is.
    A: P. 18, it has been defined.
14. Example 3.2.2 is not an example but the content of this subsubsection, and the paragraph before it should be mostly deleted to avoid redundancy. A: P.43, P. 44, it has been restructured.

15. Page 32 first paragraph: conclude what the calculation of the Hilbert Schmidt norm shows. A: P. 45, we provide the conclusion.

16. Page 32 second paragraph: you have by now used the delta function quite often and there is no need for this paragraph in this section. If you want to keep it then move it to Chapter 2. A: P. 45, we delete this paragraph.

17. Theorem 3.2.3: the assumptions include an assumption on $p$; make that clearer; also state what $\delta_t^\epsilon$ is in the statement of the theorem. A: P. 45, it has been restructured.

18. Page 35 4th line: this is an inequality, not an equality. A: P. 45, it has been changed.

19. (3.54) What is $c$? A: P. 46, it has been provided.

20. P.36 delete the last two sentences about the delta function; they do not belong here. A: P. 47, it has been deleted.

21. P. 36 line -2: explain the $\cdot$ notation A: P. 47, it has been defined.

22. Section 3.2.2.2 rephrase from [19] A: P. 47, P. 48, it is done.

23. Page 38 last paragraph define the p-norm A: P. 49, it has been defined.

24. (3.73) why is it allowed to take the inner product with the Exponential? A: P. 50, because $Exp(g) \in L_p$, 

25. Definition 3.2.3: the last part is not part of the definition. Also we do not need to choose delta in a certain way, we just choose it in that way. (See
also first line on p.43.)
A: P. 52, it has been fixed.

26. Definition 3.2.4: the last part is not part of the definition.
A: P. 52, it has been fixed.

27. (3.88): explain why this inequality holds.
A: P. 53, it is done.

28. Theorem 3.2.4: do not just write the equation numbers after the expression but write something like ”given in”. What are $\omega_r$ and $\theta_0$? State it in the theorem and not only in the proof.
A: P. 56, it is done.

29. Remark 3.2.5 belongs in the proof.
A: P. 57, it is done.

30. Theorem 3.2.6: state what $\omega_r$ and $\theta_0$ are, or refer to their definition if you give it earlier.
A: P. 59, it is done.

31. Pages 48 and 49 and elsewhere: do not leave so much empty space at the end of pages; either use an appropriate LaTeX package to take care of splitting long formulas or split them manually to avoid empty spaces.
A: P. 59, P. 60, it is done.

32. (3.123) Explain why this works.
A: P. 61, it is done.

33. (3.133) $DF$ should be defined before this equation and not just in the proof of Theorem 3.2.10.
A: P. 63, it is done.

34. Definition 3.2.6: what are the conditions on $F$ for $D^p$ to exist?
A: P. 64, it has been provided.

35. Theorem 3.2.10: what are the conditions on $F$ and on $f$?
A: P. 64, it has been provided.

36. Definition 3.2.7: this is not a definition but should follow from Definition 3.2.6 under appropriate conditions on $F$ and $f$.
A: P. 65, P. 66, it has been fixed.
37. **Lemma 3.3.3.** what are the conditions on $F$?
   A: P. 69, it has been provided.

38. (3.154) should not be part of the statement of the theorem.
   A: P. 70, it has been deleted.

39. Explain (3.178), which derivative is this?
   A: P. 74, we delete this formula. We have already provided explanation and the reference, and readers can refer to the original paper.

40. (3.179) what are $f_G$ and $f_I$?
   A: P. 74, we provide the explanation.

41. **Theorem 3.3.3 is a corollary, and so is Theorem 3.3.4.**
   A: P. 76, P. 78, it has been fixed.

42. (3.205) What is $Z$?
   A: P. 78, it has been defined.

43. Explain the difference between $L$ and $\mathcal{L}$; and how they relate to the Poisson and to the Brownian integral; use different notations for these integrals.
   A: P. 78, $\mathcal{L}$ operator is for Poisson functional while $L$ operator is for Brownian functional.

44. (3.214) What is $K$?
   A: P. 80, it has been provided.

45. Page 70: the definition of $F_1$ should appear before (3.215) and not afterwards.
   A: P. 80, it has been fixed.

46. Example 3.3.1 specify ”above”
   A: P. 81, it has been fixed.

47. **Lemma 3.3.4 What are the conditions on the functions?**
   A: P. 85, it has been provided.

48. (3.272) Are there conditions on the coefficients?
   A: P. 86, P.87, they have been provided.

49. (3.276) re-phrase what you mean; this is too short.
   A: P. 87, it is done.
50. **Page 84:** B should be a different symbol as it is the upper bound in (3.345) which you denote by B in the statement of the result.
A: P. 94, it is done.

51. **Page 84:** The inequality used is Hölders inequality.
A: P. 94, it has been fixed.

52. **Comment on the example; is there something you would like to draw attention to?**
A: P. 94, it has been provided.

53. **Page 85:** make clear how the set of functions for which Malliavin derivatives exist relates to the set of functions where white noise analysis works.
A: P. 95, P. 96, it has been provided.

54. **End of Page 85:** do you also use these abbreviations?
A: P. 96, we use the same abbreviations in this thesis.

55. **Page 86:** correct the wrong citing of formulas on the first line of Example 3.4.1
A: P. 96, it has been fixed.

56. **Page 89:** what is the stochastic derivative? Define it.
A: P. 99, we delete this paragraph.

57. **Example 3.4.4 specify ”above”.**
A: P. 102, it has been provided.

58. **Page 95 last paragraph:** you give some examples but you should not claim that you have completely discussed the differences and connections between Malliavin calculus and white noise analysis.
A: P. 106, it has been reworded.

### 6.6 Chapter 4

1. **Page 96 and perhaps in Chapter 1:** it should be made more clear that Chapter 4 does not use the outcomes of Chapter 3 but its a rather separate material.
A: P. 107, it is done.
Chapter 6. Questions and Actions Taken

2. Page 97 first paragraph: the next chapter is already the conclusion!
   A: P. 18, it has been restructured.

3. Page 97 line -3: explain the notation.
   A: P. 19, it has been provided.

4. (4.10) compare to the integral obtained through white noise analysis.
   A: P. 21, it is done.

5. P 100, Theorem 4.2.1 re-phrase where the convergence so that its clear convergence of what is meant here.
   A: P. 22, it is done.

6. P 101, (4.16), the left-hand side does not have $t$ while the middle and right-hand side have. Though its not difficult to guess the meaning, it requires to add either an explanation or change of the notation.
   A: P. 23, it has been reworded.

7. Example 4.2.1 Make it clearer what you are trying to say you are not calculating the same functionals with both approaches. What do you conclude?
   A: P. 25, it has been provided.

8. (4.32) What is $DG$?
   A: P. 27, it has been provided.

9. Page 107: the operator is called the Ornstein-Uhlenbeck operator in analogy with the Ornstein-Uhlenbeck generator.
   A: P. 29, it has been fixed.

10. Definition 4.2.5 is poorly worded and its end is not a part of the definition.
    A: P. 29, it has been reworded.

11. The title 4.3 should be more informative.
    A: P. 108, it has been reworded.

12. Page 107, the start of Section 4.3.1: write a bit more, add references, dont be abrupt.
    A: P. 108, it has been reworded.
13. **P. 108** and in what follows: two different letters are used for payoff; use just one, e.g. $g(x)$.
   A: P. 108, it has been fixed.

14. **End of Page 108- beginning of page 109**: presentation is poor, sentences are incomplete and abrupt.
   A: P. 109, it has been restructured.

15. **(4.56)** what are the conditions on $F$ for this to hold?
   A: P. 109, it has been provided.

16. **Theorem 4.3.1** has to be reformulated. In particular, appearance of $\mu_{S_T}$ is very confusing and $\sigma_{S_T}$ is not explained.
   A: P. 111, it has been fixed.

17. **Page 112**: correct the phrase about SDE solution.
   A: P. 111, it has been fixed.

18. **P. 113**: the form of the stochastic vol is rather unusual as variance of this vol grows with time which contradicts to common stochastic vol models. Some explanation/justification is required, preferably with references and discussion of commonly used stochastic vol models.
   A: P. 112, P. 113, it has been provided.

19. **Page 113**: is there a motivation for looking at this particular functional?
   A: P. 112, P. 113, the simple perturbed version of (4.9) given in (4.14) has been introduced for illustrative purposes.

20. **The Stein-Malliavin method** for normal approximation is much broader than how you define it on page 113.
   A: P. 114, it has been deleted.

21. **Remark 4.4.1**: belongs in the proof.
   A: P. 117, it has been restructured.

22. **Page 121**: you do not need to write down the Malliavin operators again.
   A: P. 121, it has been deleted.

23. **Lemma 4.5.1** has no proof and should be an immediate consequence.
   A: P. 122, it has been restructured.
24. (4.115) does not make sense.
A: P. 124, it has been deleted.

25. Theorem 4.5.2 and 4.5.3: should these be lemmas instead?
A: P. 124, P. 126, it has been fixed.

26. Calculate the expression (4.134) explicitly.
A: P. 131, it has been provided.

27. Theorem 4.5.4 ”it is useful to know” is not a good statement in a theorem.
The order estimates are interesting but not proven; provide a proof.
A: P. 135, P. 136, P. 137, P. 138, P. 139, it has been provided.

28. Page 140: delete the last paragraph.
A: P. 142, it has been deleted.

29. Section 4.6: Explain at an appropriate place what the practical usefulness
of the derived bounds (this was discussed during the viva).
A: P. 142, P. 143, it has been provided.

30. Page 143: your conclusions hold only in the case that $h_t$ is constant to 1,
make that clear.
A: P. 144, it has been provided.

6.7 References

1. The references should be unified; either abbreviate all journals or spell
them all out; either abbreviate all first names or spell them all out. Give the
place of publication for books.
A: They have been provided.


