Dyson type formula for pure jump Lévy processes and applications

Sixian Jin * Henry Schellhorn † Josep Vives ‡§

June 5, 2016

Abstract

In this paper we obtain a Dyson type formula for square integrable functionals of a pure jump Lévy process, that is, we obtain a representation of the conditional expectation of a certain functional F at a certain fixed time t as an exponential formula at a later time T in terms of the Malliavin derivatives evaluated along a frozen path. The series representation of this exponential formula turns out to be useful for different applications, and in particular in Quantitative Finance, as we show in several examples. We prove the result by a cylindrical approximation and the repeated use of a backward Taylor expansion.

Keywords: Additive processes, Lévy processes, Malliavin calculus, Dyson type formula.

Mathematical Subject Classification: 60H07, 91G80.

1 Introduction

In [11] a representation theorem for smooth Brownian martingales, also called Dyson type formula, was obtained. A similar representation was obtained in [12] for functionals of the fractional Brownian motion for $H > \frac{1}{2}$. In both case the representation involves the Malliavin derivative. Here we obtain an analogous formula for functionals of a pure jump Lévy process. Our work is based on Malliavin-Skorohod calculus techniques for Lévy processes. As general references for this calculus we refer the reader to books [9], [7] and [22].

Note that our result is essentially different from the previous Brownian results. In that case, the involved operator is the second order Malliavin derivative whereas in our case only the first order one is required. So, this shows that this type of results are intrinsically probabilistic and cannot be covered by algebraic methods based only on the Fock space structure of the space of square integrable functionals of a process with the chaotic representation property. See [17] or [27] for a discussion about the role of the Fock space structure in Malliavin-Skorohod calculus. Also, the method of proof is significantly different.

*Institute of Mathematical Sciences, Claremont Graduate University, Claremont CA 91711, USA. Email: Sixian. Jin@cgu.edu

†Institute of Mathematical Sciences, Claremont Graduate University, Claremont CA 91711, USA.E-mail: henry.schellhorn@cgu.edu

‡Faculty of Mathematics, Universitat de Barcelona, Gran Via 585, 08007 Barcelona (Catalunya), Spain. E-mail: josep.vives@ub.edu

§Supported by grant MEC MTM 2012 31192

The obtained formula is a new way to compute conditional expectations, based on the idea to *freeze* the path on the conditioning time instant. Taking into account that the price of a financial derivative is nothing more than the conditional expectation of the final payoff at the current time with respect the risk neutral probability, the obtained representation formula is potentially useful in pricing and hedging, as two examples in the paper shows. One of them, the quadratic Lévy model, is, as far as we know, an original result.

We prove our result by a cylindrical approximation and the repeated use of backward Taylor expansion.

The paper is organized as follows. Section 2 is devoted to preliminaries about Malliavin-Skorohod type calculus for Lévy processes and Poisson random measures. Section 3 is devoted to the main results: the backward Taylor expansion (which is of independent interest for numerical calculations, see [11]), the time evolution equation and the Dyson type formula. Since the proof of the Dyson type formula is long, we dedicate section 4 to sketching how two other more intuitive approaches can lead to the same result. We believe however that the "time-approach" taken in section 3 gives sharper results.Two examples are analyzed in detail in section 5. Proofs are left to the appendix.

2 Preliminaries and notation

2.1 Lévy processes, Poisson random measures and isonormal Lévy processes

Let $T > 0$. Consider a real Lévy process $X = \{X_t, t \in [0,T]\}$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote by E the expectation with respect to $\mathbb{P}$. Denote by $\{\mathcal{F}_t, t \geq 0\}$ the completed natural filtration of X and assume $\mathcal{F} = \mathcal{F}_T$. Recall that a Lévy process is a process with independent and stationary increments, null at the origin and with càdlàg trajectories. See for example [26] for the basic theory of Lévy processes.

The distribution of a Lévy process can be characterized by the triplet (γ, σ^2, ν), where $\gamma \in \mathbb{R}$, $\sigma^2 \geq 0$ and ν is a Lévy measure $\mathbb{R}$, that is, a $\sigma-$finite positive measure, null at the origin and such that $\int_{\mathbb{R}_0}(1 \wedge x^2)\nu(dx) < \infty$.

Let us denote by $\mathcal{B}_T$ and $\mathcal{B}$ the $\sigma-$algebras of Borel sets of $[0,T]$ and $\mathbb{R}$ respectively. Consider also $\mathbb{R}_0 := \mathbb{R} - \{0\}$ and denote by $\mathcal{B}_0$ its Borel $\sigma-$algebra. Consider the measure space $([0,T] \times \mathbb{R}_0, \mathcal{G}, \ell \otimes \nu)$ where $\mathcal{G} := \mathcal{B}_T \otimes \mathcal{B}_0$ and ℓ denotes the Lebesgue measure on $[0,T]$.

Given $G \in \mathcal{G}$ we introduce the random measure N associated to X, defined as

$$N(G) = \#\{t : (t, \Delta X_t) \in G\},$$

with $\Delta X_t = X_t - X_{t-}$.

Recall that N is a Poisson random measure on $\mathcal{G}$ with intensity $\ell \times \nu$. Let $\mathcal{G}^*$ be the family of Borel sets G such that $(\ell \otimes \nu)(G) < \infty$. Then, for any $G \in \mathcal{G}^*$, $N(G)$ is a Poisson random variable with $E[N(G)] = (\ell \otimes \nu)(G)$. Identifying, as usual, $\ell(dt) = dt$, we can consider the compensated random measure $\widetilde{N}(dt,dx) := N(dt,dx) - dt\nu(dx)$, that is a square integrable centered random measure such that for any G, G_1 and G_2, subsets of $\mathcal{G}^*$, we have

$$E[(\widetilde{N}(G))^2] = (\ell \otimes \nu)(G)$$

and

$$E[\widetilde{N}(G_1)\widetilde{N}(G_2)] = \nu(G_1 \cap G_2).$$

According to the Lévy-Itô decomposition, see for example [26], we can write:

$$X_t = \gamma t + \sigma W_t + L_t, \quad t \geq 0, \tag{2.1}$$

where W is the standard Brownian motion and L is a Lévy process with triplet $(0, 0, \nu)$ independent of W and defined by

$$L_t = \int_0^t \int_{|x|>1} xN(\mathrm{d}s, \mathrm{d}x) + \int_0^t \int_{|x|\leq 1} x\widetilde{N}(\mathrm{d}s, \mathrm{d}x) \tag{2.2}$$

where

$$\int_0^t \int_{|x|\leq 1} x\widetilde{N}(\mathrm{d}s, \mathrm{d}x) := \lim_{\epsilon \downarrow 0} \int_0^t \int_{\epsilon<|x|\leq 1} x\widetilde{N}(\mathrm{d}s, \mathrm{d}x)$$

and the convergence is $a.s.$ and uniform with respect to t on every bounded interval. Following the literature, we will call the process $L = \{L_t, t \geq 0\}$ a pure jump Lévy process.

Moreover, see [25], if $\{\mathcal{F}_t^W, t \geq 0\}$ and $\{\mathcal{F}_t^L, t \geq 0\}$ are, respectively, the completed natural filtrations of W and L, then, for every $t \geq 0$, we have $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^L$. Since W and L are independent, and since the Brownian case was treated in [11], we restrict the analysis in this paper to the case $\gamma = \sigma = 0$.

If we assume X is a square integrable process, that is equivalent to assume $\int_{\mathbb{R}} x^2 \nu(dx) < \infty$, we can write

$$X_t = t \int_{|x|>1} x\nu(\mathrm{d}x) + \int_0^t \int_{\mathbb{R}_0} x\widetilde{N}(\mathrm{d}s, \mathrm{d}x)$$

where the stochastic integral with respect to $\widetilde{N}$ is understood as the uniform limit defined above and the square integrability guarantees the existence of the deterministic integral.

2.2 Chaotic representations property

Following [25] and [27] we introduce the measures $M(\mathrm{d}t, \mathrm{d}x) := x\widetilde{N}(\mathrm{d}t, \mathrm{d}x)$ and $\mu(\mathrm{d}t, \mathrm{d}x) := x^2\nu(\mathrm{d}x)\mathrm{d}t$. For any G, G_1 and G_2, subsets of $\mathcal{G}^*$, we have $E[M(G)] = 0$, $E\big[(M(G))^2\big] = \mu(G)$ and

$$E[M(G_1)M(G_2)] = \mu(G_1 \cap G_2).$$

So, M is a square integrable centered and independent random measure defined on $([0, T] \times R_0, \mathcal{G})$.

Following [10], we can consider spaces

$$\mathbb{L}_n^2 := L^2\Big(([0, T] \times \mathbb{R}_0)^n, \mathcal{G}^{\otimes n}, \mu^{\otimes n}\Big)$$

and define the Itô multiple stochastic integrals $I_n(f)$ with respect M in the usual way. Then we have the so-called chaos representation property, that is, for any functional

$$F \in L^2(\mathbb{P}),$$

we have

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

for a certain unique family of symmetric kernels $f_n \in \mathbb{L}_n^2$.

We have moreover the following result.

Lemma 2.1 *Consider the functionals $e^{I_1(h)}$ where $h \in \mathbb{L}_1^2$ and $xh(s,x)$ is bounded. Then*

1. $$E\left[e^{I_1(h\chi_G)}\right] = \exp\left(\int_0^T \int_{\mathbb{R}_0} \chi_G(s,x)(e^{h(s,x)x} - 1 - h(s,x)x)\nu(\mathrm{d}x)\mathrm{d}s\right)$$

   for any $G \in \mathcal{G}$.

2. *The family is total in $L^2(\Omega)$.*

Proof. See [6] for the first part and [28] for the second. ■

2.3 Malliavin-Skorohod calculus

The chaos representation property of $L^2(\mathbb{P})$ shows that this space has a Fock space structure. Thus it is possible to apply all the machinery related to the annihilation operator (Malliavin derivative) and the creation operator (Skorohod integral) as it is exposed, for example, in [17] and [27].

Consider $F = \sum_{n=0}^{\infty} I_n(f_n)$, with f_n symmetric and such that $\sum_{n=1}^{\infty} n\, n! \|f_n\|^2_{\mathbb{L}_n^2} < \infty$. The Malliavin derivative of F is an object of $L^2([0,T] \times \mathbb{R}_0 \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mu \otimes \mathbb{P})$, defined as

$$D_{t,x}F = \sum_{n=1}^{\infty} nI_{n-1}\Big(f_n\big(t,x,\cdot\big)\Big), \quad t \in [0,T], x \in \mathbb{R}_0. \tag{2.3}$$

We will denote by $\mathbb{D}$ the domain of this operator.

Of course the Malliavin derivative can be iterated in the usual way. We will write $D^n_{t_1,x_1,\ldots,t_n,x_n}$. The domain of the n-th iterated operator will be denoted by $\mathbb{D}^n$.

On other hand, let $u \in L^2([0,T] \times \mathbb{R}_0 \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mu \otimes \mathbb{P})$. For every $t \in [0,T]$ and $x \in \mathbb{R}_0$, we have the chaos decomposition

$$u_{t,x} = \sum_{n=0}^{\infty} I_n(f_n(t,x,\cdot))$$

where $f_n \in \mathbb{L}^2_{n+1}$ is symmetric in the last n pairs of variables. Let $\tilde{f}_n$ be the symmetrization in all $n+1$ pairs of variables. Then we define the Skorohod integral of u by

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n), \tag{2.4}$$

in $L^2(\mathbb{P})$, provided $u \in \operatorname{Dom}\delta$ that means

$$\sum_{n=0}^{\infty}(n+1)!\, \|\tilde{f}_n\|^2_{\mathbb{L}^2_{n+1}} < \infty.$$

Moreover according to Proposition 5.4 in [28], we have the following integration by parts formula:

Lemma 2.2 *If $u \in \operatorname{Dom}\delta$ and $F \in \mathbb{D}$ such that $E\int_{[0,T]\times\mathbb{R}_0} u_{t,x}^2 (F^2 + (D_{t,x}F)^2)\,\mu(dt,dx) < \infty$. Then the following relation holds:*

$$\delta(Fu) = F\delta(u) - \int_{[0,T]\times\mathbb{R}_0} u_{t,x}\, D_{t,x}F\, \mu(\mathrm{d}t, \mathrm{d}x) - \delta((DF)\,u) \tag{2.5}$$

provided that one of the two sides of the equality exists.

2.4 Clark-Haussmann-Ocone formula

In this setting we can establish an abstract Clark-Haussmann-Ocone (CHO) formula. Given $A \in \mathcal{G}$ we can consider the $\sigma-$algebra $\mathcal{F}_A$ generated by $\{\widetilde{N}(A') : A' \in \mathcal{G}^*, A' \subseteq A\}$. We have, see [17], that F is $\mathcal{F}_A-$measurable if for any $n \geq 1$, $f_n(t_1, x_1, \dots, t_n, x_n) = 0$, $\mu^{\otimes n} - a.e.$ unless $(t_i, x_i) \in A \quad \forall\ i = 1, \dots n$.

In particular, we are interested in the cases $A := [0,t) \times \mathbb{R}_0$ that correspond with $\sigma-$algebra $\mathcal{F}_t$. Obviously, if $F \in \operatorname{Dom} D$ and it is $\mathcal{F}_t-$measurable then $D_{s,x}F = 0$ for a.e. $s \geq t$ and any $x \in \mathbb{R}_0$. Recall also that if $u \in \operatorname{Dom}\delta$ is actually predictable with respect to $\mathcal{F}_t$, then the Skorohod integral coincides with the (non anticipating) Itô integral in the L^2-setting with respect to M.

From the chaos representation property we can see that for $F \in L^2(\mathbb{P})$,

$$E[F|\mathcal{F}_t] = \sum_{n=0}^{\infty} I_n(f_n(t_1, x_1, \dots, t_n, x_n) \prod_{i=1}^{n} \chi_{[0,t)}(t_i)),$$

(see e.g. [7]). Then, for $F \in \mathbb{D}$ we have

$$D_{s,x}E[F|\mathcal{F}_t] = E[D_{s,x}F|\mathcal{F}_t]\chi_{[0,t)}(s),\ (s,x) \in [0,T] \times \mathbb{R}_0.$$

Using these facts we can prove, see for example [7], the CHO formula says that if $F \in \mathbb{D}$ we have

$$F = E[F] + \delta(E[D_{t,x}F|\mathcal{F}_t]).$$

Being the integrand of a predictable process, the Skorohod integral δ here above is actually an Itô integral. Then, the CHO formula can be rewritten as

$$F = E[F] + \int_{[0,T]\times\mathbb{R}_0} E[D_{s,x}F|\mathcal{F}_s] x\tilde{N}(\mathrm{d}s, \mathrm{d}x). \tag{2.6}$$

See [4].

2.5 The Malliavin type derivative on the canonical space

We are interested in the probabilistic interpretation of the operator D defined before in the canonical space of the Poisson random measure. It is well known, see for example [25], that the canonical space of a Poisson random measure can be seen as the set of finite or infinite sequences of pairs (t_i, x_i) such that for any $\epsilon > 0$, only a finite number of them are in $[0,T] \times S_\epsilon$

where $S_\epsilon := \{x \in \mathbb{R} : |x| > \epsilon\} \subseteq \mathbb{R}_0$. In this setting it is well known that operator $D_{t,x}$ is essentially equivalent to the operator

$$\Psi_{t,x}F = \frac{F(\omega_{t,x}) - F(\omega)}{x}$$

where $\omega_{t,x}$ means the addition of the pair (t,x) to the sequence ω. This operator is linear, closed and well defined from $L^0(\Omega)$ to $L^0(\Omega \times [0,T] \times \mathbb{R}_0)$ and for any $F, G \in L^0(\Omega)$ satisfies the well known property

$$\Psi(FG) = G \cdot \Psi F + F \cdot \Psi G + x\Psi F \cdot \Psi G.$$

This is a probabilistic interpretation of operator D in the sense that $F \in \mathbb{D}$ if and only if $\Psi F \in L^2(\Omega \times [0,T] \times \mathbb{R}_0)$ and in this case

$$DF = \Psi F,\ \ell \times \nu \times \mathbb{P} - a.e.$$

and thus it is not hard to obtain such a chain rule: for any t and x,

$$xD_{t,x}F(G) = F(G + xD_{t,x}G) - F(G).$$

In relation with operator δ we have that it is an extension of the Itô type (and pathwise) integral in the sense that for a predictable process $u \in L^2([0,T] \times \mathbb{R}_0)$ and assuming X is an square integrable process, we have

$$\delta(u) = \int_0^T \int_{\mathbb{R}_0} u(s,x)\widetilde{N}(\mathrm{d}s, \mathrm{d}x).$$

2.6 A particular case: the compound Poisson process

As a particular case of the pure jump Lévy process, we denote the standard Poisson process with intensity λ by N and the compensated Poisson process by:

$$\tilde{N}_t := N_t - \lambda t.$$

The iterated stochastic integral of f_n is denoted by

$$I_n(f_n) = n! \int_{0 \leq s_1 \leq \ldots \leq s_n \leq T} f_n(s_1, ..., s_n)\mathrm{d}\tilde{N}_{s_1} ... \mathrm{d}\tilde{N}_{s_n}.$$

The Malliavin derivative of Poisson process, denoted by D_t, satisfies the following chain rule like the general pure jump process:

$$\begin{aligned} D_t(FG) &= D_tF + D_tG + D_tFD_tG; && (2.7) \\ D_tF(G) &= F(G + D_tG) - F(G). && (2.8) \end{aligned}$$

Then according to [22] page 114, we have the Clark-Ocone formula for Poisson process. Suppose F is $\mathcal{F}_T$-measurable, then

$$F = E[F] + \int_0^T E[D_tF|\mathcal{F}_t]\mathrm{d}\tilde{N}_t.$$

Moreover, we introduce the compound Poisson process with a finite number of jump sizes $\{z_1, ..., z_J\}$ as $Y_t = \sum_{k=1}^{N_t} Z_k$ where $\{Z_k\}_{k \geq 0}$ is a sequence of i.i.d discrete random variables.

The value of Z represent the jump size at each jump and before time T, we set $p(z_j) = P(Z = z_j)$ for $j = 1, ..., J$, and then we can rewrite $Y_t = \sum_{j=1}^{J} z_j N_t^j$ where $\left\{N_t^j\right\}_j$ is a sequence of independent classical Poisson process with intensities $\{\lambda_j = \lambda p(z_j)\}_j$. Let $\tilde{Y}_t$ denote the compensated compound Poisson process:

$$\tilde{Y}_t = Y_t - \lambda t \sum_{j=1}^{J} z_j p(z_j) = \sum_{j=1}^{J} z_j \left(N_t^j - \lambda p(z_j) t\right) = \sum_{j=1}^{J} z_j \tilde{N}_t^j.$$

Thus the Itô integral for $\tilde{Y}_t$ can be defined as

$$\int_0^T X_t \mathrm{d}\tilde{Y}_t = \sum_{j=1}^{J} \int_0^T z_j X_t \mathrm{d}\tilde{N}_t^j.$$

According to [3], the multivariable Clark-Ocone formula is:

$$F = E[F] + \int_0^T \sum_{j=1}^{J} E[D_s^{(j)} F | \mathcal{F}_s] \mathrm{d}\tilde{N}_s^j$$

where $D_s^{(j)}$ is the Malliavin derivative with respect to $\tilde{N}_s^j$. For simplicity, we introduce the tensor product of a Lebesgue measure l as:

$$(l(\mathrm{d}s))^{\otimes n} := l(\mathrm{d}s_1)...l(\mathrm{d}s_n).$$

3 Main Results

3.1 Backward Taylor expansion

We first define the n-th Charlier polynomial $C_n(x_1, x_2)$ by

$$\exp\left(z x_1 - \frac{z^2}{2} x_2\right) = \sum_{n=0}^{\infty} z^n C_n(x_1, x_2)$$

or the following recurrence formula:

$$\begin{aligned} C_0(x_1, x_2) &= 1, \ C_1(x_1, x_2) = x_1 - x_2; \\ C_{n+1}(x_1, x_2) &= (x_1 - x_2 - n) C_n(x_1, x_2) - nt C_{n-1}(x_1, x_2), \quad n \geq 1. \end{aligned}$$

Then according to Theorem 3.2 and relation (2.7) in [28], we have the following relation:

$$C_n(X_T - X_t, \mu(T - t)) = \int_{t \leq s_1 ... \leq s_n \leq T} \int_{\mathbb{R}_0^n} x_1 \widetilde{N}(\mathrm{d}s_1, \mathrm{d}x_1) ... x_n \widetilde{N}(\mathrm{d}s_n, \mathrm{d}x_n). \tag{3.1}$$

Then with these polynomials, we have our first main result, the Backward Taylor Expansion.

Theorem 3.1 *Assume $F \in L^2(\mathbb{P})$ can be written as $F = F(X_{t_1}, ..., X_{t_K})$ with $t_1 \leq ... \leq t_K$. Then if F satisfies $\sum_{i=0}^{\infty} \int_{\mathbb{R}_0^i} E\left[\left(D_{t_{k+1}, x_1} ... D_{t_{k+1}, x_i} F\right)^2\right] (x^2 \nu(\mathrm{d}x))^{\otimes i} < \infty$ as well as the following condition: for any $1 \leq k \leq K - 1$,*

$$\sum_{i=0}^{L} \int_{\mathbb{R}_0^{2L-i}} E\left[\left(D_{t_{k+1}, x_1} ... D_{t_{k+1}, x_{2L-i}} F\right)^2\right] \left(x^2 \nu(\mathrm{d}x)\right)^{\otimes 2L-i} \binom{L}{i}^4 \frac{i!}{(L!)^2} (t_{k+1} - t_k)^{2L-i} \xrightarrow{L \to \infty} 0. \tag{3.2}$$

Then we have for any $1 \leq k \leq K-1$:

$$E[F|\mathcal{F}_{t_k}] = E[F|\mathcal{F}_{t_{k+1}}] + \sum_{l=1}^{\infty} \gamma(l,k) \int_{\mathbb{R}_0^l} E[D_{t_{k+1},x_1} ... D_{t_{k+1},x_l} F|\mathcal{F}_{t_{k+1}}] \left(x^2 \nu(\mathrm{d}x)\right)^{\otimes l} \tag{3.3}$$

where $\gamma(l,k)$ has the following representation: for $l \geq 1$

$$\gamma(l,k) = (-1)^l \sum_{\substack{i_1+...+i_n+j+n=l \\ j \leq n}} C_{n-j}(X_{t_{k+1}} - X_{t_k}, \mu(t_{k+1}-t_k)) \frac{\mu^j (t_{k+1}-t_k)^j}{(n-j)!j!}. \tag{3.4}$$

As a particular case of the generalized polynomial P_n, and according to 6.2.16 in [22], the n-th Charlier polynomial satisfies:

$$C_n(N_T - N_t, \lambda(T-t)) = n! \int_{t \leq s_1 ... \leq s_n \leq T} (\mathrm{d}\tilde{N}_s)^{\otimes n}. \tag{3.5}$$

Therefore as a corollary of Theorem 3.1, we obtain the backward Taylor expansion for the Poisson and compound Poisson processes.

Corollary 3.2 *Suppose $F \in L^2(\mathbb{P})$ and $M\Delta = T$.*

1. *Poisson process: If F can be represented as $F = F(\tilde{N}_\Delta, ..., \tilde{N}_{M\Delta})$ and satisfies $\sum_{i=0}^{\infty} E\left[(D_T^i F)^2\right] < \infty$ as well as the following condition: for any $m = 1, ..., M-1$*

$$\sum_{i=0}^{L} E\left[\left(D_{(m+1)\Delta}^{2L-i} F\right)^2\right] \lambda^{2L-i} \binom{L}{i}^4 \frac{i!}{(L!)^2} \Delta^{2L-i} \xrightarrow{L\to\infty} 0. \tag{3.6}$$

Then

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{l=0}^{\infty} \gamma(m,l) E[D_{(m+1)\Delta}^l F|\mathcal{F}_{((m+1)\Delta)}] \tag{3.7}$$

where $\gamma(m,l)$ has the following representation:

$$\gamma(m,l) = (-1)^l \sum_{\substack{i_1+...+i_n+j+n=l \\ j \leq n}} C_{n-j}(N_{(m+1)\Delta} - N_{m\Delta}, \lambda\Delta) \frac{\lambda^j \Delta^j}{(n-j)!j!}. \tag{3.8}$$

2. *Compound Poisson process: If F can be represented as $F = F\left(\tilde{Y}_\Delta, ..., \tilde{Y}_{M\Delta}\right)$ and satisfies $\sum_{i=0}^{\infty} E\left[\left(D_T^{i,(j)} F\right)^2\right] < \infty$ for all j as well as the following condition:*

$$\sum_{n_1+...+n_J=L} \left(\sum_{i=0}^{n_j} E\left[\left(D_{(m+1)\Delta}^{2n_j-i,(j)} F\right)^2\right] \lambda_i^{2n_j-i} \binom{n_j}{i}^4 \frac{i!\Delta^{2n_j-i}}{(n_j!)^2} \right) \xrightarrow{L\to\infty} 0. \tag{3.9}$$

for each $m = 1, ..., M$, then we have:

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{l_1,...,l_J=0}^{\infty} \gamma(m; l_1, ..., l_J) D_T^{l_1,(1)} ... D_T^{l_J,(J)} E[F|\mathcal{F}_{(m+1)\Delta}]$$

where

$$\gamma(m;l_1,...,l_J)=(-1)^{l_1+...+l_J}\prod_{j=1}^{J}\sum_{\substack{i_1+...+i_{n_j}+r_j+n_j=l_j\\ r_j\le n_j}}\lambda_j^{r_J}\frac{C_{n_j-r_j}(N_T^j-N_t^j,\lambda_j\Delta)\Delta^{r_j}}{(n_j-r_j)!r_j!}.$$

Remark:

1. A large range of random variables can fit (3.6). By Stirling approximation for the factorial, we can prove that:
$$\sum_{i=0}^{L}\binom{L}{i}^4\frac{i!}{(L!)^2}\lambda^{2L-i}(t_{k+1}-t_k)^{2L-i}\le\frac{C^L}{L^L}$$
for some fixed constant $C>0$. Thus for those random variables F such that
$$E\left[\left(D_T^LF\right)^2\right]\le c^L \tag{3.10}$$
for some constant c, (3.6) always holds. A simple example is $F=e^{\alpha N_T}$ for any constant α.Then we have
$$\begin{aligned}D_T^LF &= \sum_{l=0}^{L}(-1)^l\binom{L}{l}e^{\alpha N_T+\alpha l}=e^{\alpha N_T}\sum_{l=0}^{L}(-1)^l\binom{L}{l}e^{\alpha l}\\ &= e^{\alpha N_T}\left(e^{\alpha}-1\right)^L.\end{aligned}$$
Then,
$$E\left[\left(D_T^LF\right)^2\right]=E[e^{2\alpha N_T}]\left(e^{\alpha}-1\right)^{2L}\le c^L$$
for some fixed c which does not depend on L and (3.6) holds. Therefore, we can regard (3.10) as an alternative condition making the series (3.7) convergent, which is easy to check in practical examples. A similar result holds for compound Poisson processes.

2. About Lévy process's condition (3.2), we can similarly use an alternative condition:
$$\int_{\mathbb{R}_0^L}E\left[\left(D_{t_{k+1},x_1}...D_{t_{k+1},x_L}F\right)^2\right]\left(x^2\nu(\mathrm{d}x)\right)^{\otimes L}\le c^L.$$
Let's choose the same example $F=e^{\alpha X_T}$ for any constant α and take $t_k\le T$ to check it. By induction, it's not hard to obtain
$$x_1...x_LD_{t_{k+1},x_1}...D_{t_{k+1},x_L}F=e^{\alpha X_T}\prod_{l=1}^{L}\left(e^{\alpha x_l}-1\right).$$

Thus

$$\int_{\mathbb{R}_0^L}E\left[\left(D_{t_{k+1},x_1}...D_{t_{k+1},x_L}F\right)^2\right]\left(x^2\nu(\mathrm{d}x)\right)^{\otimes L}=E[e^{2\alpha X_T}]\left(\int_{\mathbb{R}_0}\left(e^{\alpha x}-1\right)x^2\nu(\mathrm{d}x)\right)^L\le c^L$$

for some constant c which does not depend on L. Thus (3.2) holds for $F=e^{\alpha X_T}$.

3.2 Time evolution equation

In order to generalize the series representations to the continuous case, we need first to introduce a new operator, called the "freezing operator", as well as a related time evolution equation.

We denote the compensated Lévy measure by $\tilde{N} \in (L^2([0,T] \times \mathbb{R}_0))'$, the space of distributions. We then have the natural pairing $\langle \cdot, \cdot \rangle$ defined as: for any $f \in L^2([0,T] \times \mathbb{R}_0)$, $\left\langle f, \tilde{N} \right\rangle := \int_{[0,T]\times\mathbb{R}_0} f(s,x) x\tilde{N}(\mathrm{d}s, \mathrm{d}x)$.

Definition 3.3 (Freezing operator) *Given $\omega \in \Omega$, $t \in [0,T]$ and the compensated Lévy measure $\tilde{N}$, we define the freezing path operator ω^t by*

$$\omega^t \circ \left\langle f, \tilde{N} \right\rangle := \int_{[0,T]\times\mathbb{R}_0} f(s,x) x N(\chi_{[0,t]}(s)\mathrm{d}s, \mathrm{d}x) - \int_{[0,T]\times\{|x|\geq 1\}} f(s,x) x\nu(\mathrm{d}x)\mathrm{d}s.$$

Remark 1: Several properties follow directly from the definition.

1. $\omega^t \circ X_u = \omega^t \circ \langle \chi_{[0,u]}(s,x), N \rangle = \int_{[0,T]\times\mathbb{R}_0} \chi_{[0,u]}(s,x) N(\chi_{[0,t]}\mathrm{d}s, \mathrm{d}x) = X_{u\wedge t}$.
2. For any n-variable deterministic function $g(x_1, ..., x_n)$, we have
$$\omega^t \circ g(X_{t_1}, ..., X_{t_n}) = g(X_{t_1 \wedge t}, ..., X_{t_n \wedge t}).$$

For the Itô multiple stochastic integrals $I_n(f_n)$, the following proposition shows how a freezing path operator acts on it.

Proposition 3.4

$$\begin{aligned} & \omega^t \circ I_n(f_n) \\ = & \sum_{k=0}^{n} (-1)^k \binom{n}{k} \int_{[t,T]^k \times \mathbb{R}_0^k} \int_{[0,t]^{n-k}\times\mathbb{R}_0^{n-k}} f_n(s_1, x_1, ..., s_{n-k}, x_{n-k}, u_1, y_1, ..., u_k, y_k) \\ & \times (x\tilde{N}(\mathrm{d}s, \mathrm{d}x))^{\otimes(n-k)} (x\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes k}. \end{aligned} \tag{3.11}$$

Remark 2:

1. If F is given by its chaos decomposition, we can define $\omega^t \circ F := \sum_{n=0}^{\infty} \omega^t \circ I_n(f_n)$ when the series on the right hand side converges. From the denseness of the Charlier polynomials, (3.1), (3.11) and the following relationship of Charlier polynomial
$$C_n(x, y_1) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (y_1 - y_2)^k C_{n-k}(x, y_2),$$
it is not hard to see if $F = g(X_s)$ for a deterministic function g, then $\omega^t \circ F = g(X_{s\wedge t})$, which matches with the second property in Remark 1.
2. This proposition shows the difference between the freezing path operator and the conditional expectation, since $E[I_n(f_n)|\mathcal{F}_t] = I_n(f_n \chi_{[0,t]})$.

In relation with this operator ω^t and based on the above proposition, we can prove the following theorem.

Theorem 3.5 *Let $F \in L^2(\mathbb{P})$. Then for any fixed time t and $t \le s < T$, there exists a sequence $\{F^N\}_{N\ge 0}$ that satisfies:*
(i) $F^N \to F$ in $L^2(\mathbb{P})$
(ii) $D_{u,x}F^N = D_{s+1/N,x}F^N$ for any $u \in (s, s+1/N]$
(iii) there exists $\varepsilon \in (0,1)$ and a constant C which does not depend on N such that

$$E[(\omega^t \circ (F^N - F))^2] \le \frac{C}{N^{2+\varepsilon}}.$$

We define the conditional expectation operator as: for any $s \in [0,T]$,

$$E_s : F \in L^2(\mathbb{P}) \longrightarrow E_s F \in L^2(\mathbb{P})$$

such that

$$E_s F := E[F|\mathcal{F}_s].$$

We also introduce the derivative d in $L^2(\mathbb{P})$ as: for any processes F_s and G_s,

$$\frac{\mathrm{d}F_s}{\mathrm{d}s} = G_s \text{ is defined by } \lim_{\varepsilon\to 0} E\left[\left(\frac{F_{s+\varepsilon} - F_s}{\varepsilon} - G_s\right)^2\right] = 0.$$

Then we can set up an operator differential equation for E_s.

Theorem 3.6 *For $0 \le t \le s \le T$, assume that $F \in L^2(\mathbb{P})$ and satisfies $E\left[\int_{\mathbb{R}_0} (D_s F)^2 x^2 \nu(\mathrm{d}x)\right] < \infty$ we have*

$$\frac{\mathrm{d}\omega^t \circ E_s F}{\mathrm{d}s} = -\omega^t \circ \int_{\mathbb{R}_0} D_{s,x} E_s F x \nu(\mathrm{d}x).$$

Remark: the method of proof of this theorem can be applied to Brownian functionals (under technical conditions). It will then generalize Theorem 2.3 in [11], which was true only for cylindrical functionals, to more general functionals.

Similarly, we have the related equation for the Poisson and compound Poisson processes.

Corollary 3.7 *Suppose $F \in L^2(\mathbb{P})$.*

1. *Poisson process: For any $s \in [t,T]$, if F satisfies $E\left[(D_s F)^2\right] < \infty$, then we have:*

$$\omega^t \circ \left(\frac{\mathrm{d}E_s F}{\mathrm{d}s}\right) \xrightarrow[\epsilon\to 0]{L^2} -\omega^t \circ (\lambda D_s E_s F). \tag{3.12}$$

2. *Compound Poisson process: For any $s \in [t,T]$, if F satisfies $E\left[\left(D_s^{(j)} F\right)^2\right] < \infty$, for each j, then we have:*

$$\omega^t \circ \left(\frac{\mathrm{d}E_s F}{\mathrm{d}s}\right) = -\omega^t \circ \left(\lambda \sum_{j=1}^{J} D_s^{(j)} E_s F\right).$$

Now we are ready to introduce our third main result, the Dyson type formula.

3.3 A Dyson type formula

Definition 3.8 *As in [11] and [29] we introduce the chronological operator $\mathcal{T}$. Let $\{H(t), t \geq 0\}$ be a collection of operators. The chronological operator $\mathcal{T}$ is defined by:*

$$\mathcal{T}(H(t_1) \cdot H(t_2) \cdots H(t_n)) := H(s_1) \cdots H(s_n)$$

where $s_1, \dots, s_n$ is a permutation of $t_1, \dots, t_n$ such that $s_1 \geq s_2 \geq \cdots \geq s_n$.

Then the chronological operator applied to the exponential operator of a time-dependent generator H is given by:

$$\mathcal{T} \exp\left(\int_t^T H(s) \mathrm{d}s \right) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{[t,T]^n} H(t_k) \cdots H(t_1) \, (\mathrm{d}t)^{\otimes n} .$$

As a consequence of the time evolution equation we can prove the following theorem:

Theorem 3.9 *For any $F \in L^2(\mathbb{P})$ that satisfies:*

$$\frac{(T-t)^{2n}}{(2^n n!)^2} E\left[\left(\sup_{u_1,\dots,u_n \in [0,T]} \left| \int_{\mathbb{R}_0^n} \omega^t \circ (D_{u_n,x_n} \dots D_{u_1,x_1} F) \, (x\nu(\mathrm{d}x))^{\otimes n} \right| \right)^2 \right] \xrightarrow[n\to\infty]{} 0$$

we have

$$E[F|\mathcal{F}_t] = \mathcal{T} \exp\left(\int_t^T \int_{\mathbb{R}_0} \omega^t \circ (D_{s,x} F) \, x\nu(\mathrm{d}x) \mathrm{d}s \right) \tag{3.13}$$

provided the right hand side of the formula is well defined in L^2.

Corollary 3.10 *Let $F \in L^2(\mathbb{P})$.*

1. *Poisson process: if F satisfies:*

$$\frac{(T-t)^{2n} \lambda^{2n}}{(2^n n!)^2} E\left[\left(\sup_{u_1,\dots,u_n \in [0,T]} \left| \omega^t \circ (D_{u_n} \dots D_{u_1} F) \right| \right)^2 \right] \xrightarrow[n\to\infty]{} 0. \tag{3.14}$$

Then the following sequence converges in L^2 sense:

$$E[F|\mathcal{F}_t] = \sum_{n=0}^{\infty} \lambda^n \int_{t \leq s_1 \leq \dots \leq s_n \leq T} (\omega^t \circ D_{s_n} \dots D_{s_1} F) \, (\mathrm{d}s)^{\otimes n} . \tag{3.15}$$

2. *Compound Poisson process: we define the operator $A_s := \lambda \sum_{j=1}^{J} D_s^{(j)}$.Then if F satisfies:*

$$\frac{(T-t)^{2n} \lambda^{2n}}{(2^n n!)^2} E\left[\left(\sup_{u_1,\dots,u_n \in [0,T]} \left| \omega^t \circ (A_{u_n} \dots A_{u_1} F) \right| \right)^2 \right] \xrightarrow[n\to\infty]{} 0.$$

then the following sequence converges in L^2 sense:

$$E[F|\mathcal{F}_t] = \sum_{n=0}^{\infty} \int_{t \leq s_1 \leq \dots \leq s_n \leq T} (\omega^t \circ A_{s_n} \circ \dots \circ A_{s_1} F) \, (\mathrm{d}s)^{\otimes n} .$$

Here are two simple examples to illustrate how the exponential formula works for Poisson process.

Example 1: Let $s \le T$ and $F = e^{u\tilde{N}_T}$. Then $D_s^n F = F(e^u - 1)^n$ and $\omega^t \circ F = e^{u\tilde{N}_t - u\lambda(T-t)}$. By applying Theorem 3.10, we have

$$
\begin{aligned}
E[e^{u\tilde{N}_T}|\mathcal{F}_t] &= e^{u\tilde{N}_t - u\lambda(T-t)}(1 + \lambda \int_t^T (e^u - 1)\mathrm{d}s + \lambda^2 \int_t^T \int_{s_1}^T (e^u - 1)^2 \mathrm{d}s_1 \mathrm{d}s_2 + ...) \\
&= e^{u\tilde{N}_t + \lambda(e^u - 1)(T-t) - u\lambda(T-t)}.
\end{aligned}
$$

Example 2: Let $s \le T$ and $F = \exp(\int_0^T u(s)\mathrm{d}\tilde{N}_s)$. Then $D_s^n F = F(e^{u(s)} - 1)^n$ and $\omega^t \circ F = \exp(\int_0^t u(s)\mathrm{d}\tilde{N}_s - \lambda \int_t^T u(s)\mathrm{d}s)$. Similar computations show that

$$
\begin{aligned}
& E\left[\exp(\int_0^T u(s)\mathrm{d}\tilde{N}_s)|\mathcal{F}_t\right] \\
= \ & \exp\left(\int_0^t u(s)\mathrm{d}\tilde{N}_s - \lambda \int_t^T u(s)\mathrm{d}s\right) \\
& \left(1 + \lambda \int_t^T (e^{u(s)} - 1)ds + \lambda^2 \int_t^T \int_{s_1}^T (e^{u(s_1)} - 1)^2 (e^{u(s_2)} - 1)^2 \mathrm{d}s_1 \mathrm{d}s_2 + ...\right) \\
= \ & \exp\left(\int_0^t u(s)\mathrm{d}\tilde{N}_s - \lambda \int_t^T u(s)\mathrm{d}s\right) \exp\left(\lambda \int_t^T (e^{u(s)} - 1)\mathrm{d}s\right) \\
= \ & \exp\left(\int_0^t u(s)\mathrm{d}\tilde{N}_s + \lambda \int_t^T (e^{u(s)} - 1 - u(s))\mathrm{d}s\right).
\end{aligned}
$$

4 Connections with other Approaches

In this section, we show two different approaches which lead to the exponential formula. We deliberately use an informal style in this section, as neither of these approaches should be construed as rigorous proofs of the exponential formula. Indeed, we have not been able to make these ideas rigorous. Nevertheless, we believe it is fruitful to connect the Dyson series with other concepts. The first approach is sometimes called the frequency approach by physicists, and uses the denseness of exponentials.

The second approach shows that we can recover the Dyson series by conditioning on the number of jumps, provided we can permute terms in the series. Another application of that idea is a new formula for the expectation of F conditional on the number of jumps. As far as we know, this is a new or independently rediscovered result.

4.1 Frequency Approach

We justify Theorem 3.9 by denseness of the set of exponentials. Take $h \in \mathbb{L}_1^2$ such that $xh(s,x)$ is bounded. Define $G(h) := e^{I_1(h)}$. Its iterated Malliavin derivative is

$$
D^n_{s_1,x_1,\dots,s_n,x_n} G(h) = G(h) \prod_{i=1}^n \frac{e^{h(s_i,x_i)x_i} - 1}{x_i}
$$

Freezing the path we have

$$\omega^t \circ \left(D^n_{s_1,x_1,\dots,s_n,x_n} G(h)\right) = G(h_t) \exp\left(-\int_t^T \int_{\mathbb{R}_0} h(s,x)x^2\nu(\mathrm{d}x)\mathrm{d}s\right) \prod_{i=1}^n \frac{e^{h(s_i,x_i)x_i}-1}{x_i}$$

where $h_t(s,x) := h(s,x)\chi_{[0,t]}(s)$.

Applying Lemma 2.1 and using conditions on function h we have

$$\begin{aligned}
E[G(h)|\mathcal{F}_t] &= G(h_t)E\left[\exp\left(\int_t^T \int_{\mathbb{R}_0} h(s,x)M(\mathrm{d}s,\mathrm{d}x)\right)|\mathcal{F}_t\right] \\
&= G(h_t)\exp\left(\int_t^T \int_{\mathbb{R}_0} (e^{h(s,x)x} - 1 - h(s,x)x)\nu(\mathrm{d}x)\mathrm{d}s\right) \\
&= G(h_t)\exp\left(-\int_t^T \int_{\mathbb{R}_0} h(s,x)x\nu(\mathrm{d}x)\mathrm{d}s\right)\sum_{n=0}^\infty \frac{1}{n!}\prod_{i=1}^n \left(\int_t^T \int_{\mathbb{R}_0} (e^{h(s_i,x_i)x_i} - 1)x_i\nu(\mathrm{d}x_i)\mathrm{d}s_i\right) \\
&= G(h_t)\exp\left(\int_t^T \int_{\mathbb{R}_0} h(s,x)x\nu(\mathrm{d}x)\mathrm{d}s\right) \\
&\quad \times \sum_{n=0}^\infty \int_{t\le s_1\le\dots\le s_n\le T}\int_{\mathbb{R}_0^n} \prod_{i=1}^n \frac{e^{h(s_i,x_i)x_i}-1}{x_i}(x\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n} \\
&= \sum_{n=0}^\infty \int_{t\le s_1\le\dots\le s_n\le T}\int_{\mathbb{R}_0^n} \omega^t \circ D^n_{s_1,x_1,\dots,s_n,x_n} G(h)\,(x\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}
\end{aligned}$$

Thus the exponential formula is correct for $G(h)$. One would like to use the totality of $G(h)$ to obtain the exponential formula. The missing point in the argument is to prove commutativity between the freezing path operator and infinite summation of exponentials.

4.2 Conditioning Approach

There is apparently another approach to obtain the exponential formula. The idea is to evaluate the expectation by conditioning on the number of jumps, and then regroup the terms of the series. We did not follow this approach because of the delicate problems involved in regrouping infinite series. Rather than trying to prove the exponential formula by the conditioning approach, we use the exponential formula to characterize the conditional expectation of F(conditional of the number of jumps) as a series of Malliavin derivatives. As far as we know, this is a new or independently rediscovered the result.

We first consider the case where F depends on a Poisson process. By conditioning, we have:

$$E[F] = \sum_{n=0}^\infty E[F|N_T = n]P(N_T = n). \tag{4.1}$$

Now we define a "jump path" e_s which means that the process incurs a jump at time s. We use $\hat{e}_s$ to denote no such term, then according to the definition of Malliavin derivative, we obtain the following relation:

$$D_{s_1}\dots D_{s_n}F(\omega) = F(\omega + e_{s_1} + \dots + e_{s_n}) - \sum_{i=1}^n F(\omega + e_{s_1} + \dots + \hat{e}_{s_i} + \dots + e_{s_n}) + \dots \tag{4.2}$$

Then the Dyson series becomes:

$$\sum_{n=0}^{\infty}\lambda^n\int_{0\le s_1\le ... s_n\le T}\omega^0\circ D_{s_1}...D_{s_n}F\,(\mathrm{d}s)^{\otimes n}$$
$$= \sum_{n=0}^{\infty}\lambda^n\int_{0\le s_1\le ... s_n\le T}\omega^0\circ\left(F(\omega+e_{s_1}+...+e_{s_n})-\sum_{i=1}^{n}F(\omega+e_{s_1}+...+\hat{e}_{s_i}+...+e_{s_n})+...\right)$$
$$(\mathrm{d}s)^{\otimes n}. \tag{4.3}$$

Define

$$A_n(s_1,...,s_n):=\omega^0\circ F(\omega+e_{s_1}+...+e_{s_n}).$$

Then from the relation (4.2), we can represent A_n by iterated Malliavin derivatives as:

$$A_n(s_1,...,s_n)=\omega^0\circ\left(D_{s_1}...D_{s_n}F+\sum_{j=1}^{n}(-1)^{j-1}\sum_{1\le i_1\le...\le i_j\le n}D_{s_1}...\hat{D}_{s_{i_1}}...\hat{D}_{s_{i_j}}...D_{s_n}F\right). \tag{4.4}$$

Now we regroup the Dyson series (4.3) according to ascending order of A_n, for simplicity, we get rid of the variables in A_n and obtain:

$$\begin{aligned}
(4.3) &= \sum_{n=0}^{\infty}\left(\lambda^n\frac{1}{n!}\int_{[0,T]^n}A_n\,(\mathrm{d}s)^{\otimes n}-\lambda^{n+1}\frac{n+1}{(n+1)!}\int_{[0,T]^{n+1}}A_n\,(\mathrm{d}s)^{\otimes n}\,\mathrm{d}s_{n+1}+...\right)\\
&= \sum_{n=0}^{\infty}\lambda^n\frac{1}{n!}\int_{[0,T]^n}A_n\,(\mathrm{d}s)^{\otimes n}\left(1-\lambda T\frac{n+1}{n+1}+\frac{(\lambda T)^2}{2}\frac{(n+2)(n-1)}{(n+2)(n-1)}+...\right)\\
&= \sum_{n=0}^{\infty}\lambda^n\frac{1}{n!}\int_{[0,T]^n}A_n\,(\mathrm{d}s)^{\otimes n}\left(\sum_{k=0}^{\infty}(-1)^k\frac{(\lambda T)^k}{k!}\right)\\
&= \sum_{n=0}^{\infty}\frac{1}{T^n}\int_{[0,T]^n}A_n\,(\mathrm{d}s)^{\otimes n}\frac{(\lambda T)^n}{n!}e^{-\lambda T}\\
&= \sum_{n=0}^{\infty}\frac{1}{T^n}\int_{[0,T]^n}A_n\,(\mathrm{d}s)^{\otimes n}P(N(T)=n).
\end{aligned}$$

Observe that conditional on $N_T=n$ the arrival times are uniformly distributed, thus $\mathrm{d}s_1...\mathrm{d}s_n/T^n$ is the measure of the arrival times $s_1,...,s_n$. Finally,

$$\frac{1}{T^n}\int_{[0,T]^n}A_n\,(\mathrm{d}s)^{\otimes n}=\frac{1}{T^n}\int_{[0,T]^n}F(e_{s_1}+...+e_{s_n})\,(\mathrm{d}s)^{\otimes n}=E[F|N_T=n].$$

And we arrive at:

$$(4.3)=\sum_{n=0}^{\infty}E[F|N_T=n]P(N_T=n)=E[F].$$

Similar calculation can be applied on Lévy process by setting $e_{s,x}$ as a size x jump occurs at time s. Then the Dyson series becomes:

$$\begin{aligned}
&\textstyle\sum_{n=0}^{\infty}\int_{0\le s_1\le...s_n\le T}\int_{\mathbb{R}_0^n}\omega^0\circ(F(\omega+e_{s_1,x_1}+...+e_{s_n,x_n})\\
&\textstyle-\sum_{i=1}^{n}F(\omega+e_{s_1,x_n}+...+\hat{e}_{s_i,x_i}+...+e_{s_n,x_n})+...)\quad(\mathrm{d}s)^{\otimes n}\,(x\nu(\mathrm{d}x))^{\otimes n}.
\end{aligned} \tag{4.5}$$

Define

$$B_n(s_1, x_1, ..., s_n, x_n) := \omega^0 \circ F(\omega + e_{s_1,x_1} + ... + e_{s_n,x_n})$$

and similarly to (4.4), we can represent B_n also by Malliavin derivatives as:

$$\begin{aligned} & B_n(s_1, x_1, ..., s_n, x_n) \qquad (4.6) \\ = \quad & \omega^0 \circ \left(D_{s_1,x_1} ... D_{s_n,x_n} F + \sum_{j=1}^{n} \sum_{1 \le i_1 \le ... \le i_j \le n} D_{s_1,x_1} ... \hat{D}_{s_{i_1},x_{i_1}} ... \hat{D}_{s_{i_j},x_{i_j}} ... D_{s_n,x_n} F \right). \end{aligned}$$

We regroup the Dyson series according to ascending order of B_n and a similar calculation shows that:

$$(4.5) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{[0,T]^n} \int_{\mathbb{R}_0^n} B_n \left(x\nu(\mathrm{d}x) \right)^{\otimes n} (\mathrm{d}s)^{\otimes n} \left(e^{-T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x)} \right).$$

Now we denote J_T as the jump times in $[0, T]$. According to the infinite divisibility of Lévy process, the arrival times are uniformly distributed. Therefore, similarly to the Poisson process, $\left(\left(x\nu(\mathrm{d}x)\mathrm{d}s \right) / T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x) \right)^{\otimes n}$ is the measure of the arrival times $s_1, ..., s_n$ and

$$P(J_T = n) = \frac{\left(T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x) \right)^n}{n!} \left(e^{-T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x)} \right). \qquad (4.7)$$

We represent the expectation by conditioning on the jump times:

$$\begin{aligned} E[F] \quad &= \quad \sum_{n=0}^{\infty} E[F | J_T = n] P(J_T = n) \\ &= \quad \sum_{n=0}^{\infty} \frac{1}{\left(T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x) \right)^n} \int_{[0,T]^n} \int_{\mathbb{R}_0^n} B_n \left(x\nu(\mathrm{d}x) \right)^{\otimes n} (\mathrm{d}s)^{\otimes n} \frac{\left(T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x) \right)^n}{n!} \left(e^{-T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x)} \right) \qquad (4.8) \end{aligned}$$

Comparing (4.7) and (4.8), we find the relation:

$$E[F | J_T = n] = \frac{1}{\left(T \int_{\mathbb{R}_0} x\nu(\mathrm{d}x) \right)^n} \int_{[0,T]^n} \int_{\mathbb{R}_0^n} B_n \left(x\nu(\mathrm{d}x) \right)^{\otimes n} (\mathrm{d}s)^{\otimes n}. \qquad (4.9)$$

Numerically, relation (4.9) is probably most interesting when combined with (4.6). Indeed, if one can reorganize (4.9) in ascending order of Malliavin differentiation, one can obtain a series that converges faster. Indeed, in most tractable problems, the impact of higher order derivatives is smaller than the impact of lower order derivatives.

5 Applications

5.1 Poisson Black-Scholes Model

In this section, we recover the price of the call option in Poisson Black-Scholes Model with Theorem 3.10. We notice that our Dyson series approach is simpler and more direct than the classical PDE approach (see [23] Chapter 11). Also, we note that the corresponding Black-Scholes result cannot be obtained in the Brownian case, since the maximum function is not infinitely Malliavin differentiable.

Let $F = \exp\left(\log(\sigma+1)N_T - \lambda T\sigma\right)$ for and $\sigma \geq 0$, we want to evaluate $E[(F-K)^+|\mathcal{F}_t]$ for some fixed positive K. Let $G = (F-K)^+$, then for $s_1, ..., s_n \in (t, T]$, using the chain rule of Malliavin differentiation and induction, we obtain:

$$D_{s_1} ... D_{s_n} G = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \left(e^{\log(\sigma+1)N_T - \lambda T\sigma + k\log(\sigma+1)} - K\right)^+$$

and

$$\omega^t \circ D_{s_1} ... D_{s_n} G = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \left(e^{\log(\sigma+1)N_t - \lambda T\sigma + k\log(\sigma+1)} - K\right)^+.$$

Thus by exponential formula (3.15),

$$\begin{aligned}
E[G|\mathcal{F}_t] &= \sum_{n=0}^{\infty} \lambda^n \int_{t \leq s_1 \leq ... \leq s_n \leq T} \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \left(e^{\log(\sigma+1)N_t - \lambda T\sigma + k\log(\sigma+1)} - K\right)^+ (\mathrm{d}s)^{\otimes n} \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \left(e^{\log(\sigma+1)N_t - \lambda T\sigma + k\log(\sigma+1)} - K\right)^+ \frac{\lambda^n (T-t)^n}{n!} \\
&= \sum_{k=0}^{\infty} (-1)^k \left(e^{\log(\sigma+1)N_t - \lambda T\sigma + k\log(\sigma+1)} - K\right)^+ \sum_{n=k}^{\infty} \binom{n}{k} \frac{(-1)^n \lambda^n (T-t)^n}{n!} \\
&= e^{-\lambda(T-t)} \sum_{k=0}^{\infty} \frac{\lambda^k (T-t)^k}{k!} \left(e^{\log(\sigma+1)N_t - \lambda T\sigma + k\log(\sigma+1)} - K\right)^+
\end{aligned}$$

which matches the classical result shown in [23] chapter 11.

5.2 Lévy Quadratic Model

In this section, we use our Dyson type series to evaluate the bond price in the Lévy quadratic model of interest rates. In [11] section 3.4, we applied the Dyson series representation of Brownian motion to evaluate the bond price in the extended Cox-Ingersoll-Ross model, in which the interest rate is given by a summation of the square of Gaussian Ornstein-Uhlenbeck processes. In this section, we extend this model to Lévy processes. In this model, the interest rate is the sum of square of non-Gaussian Ornstein-Uhlenbeck processes, i.e. $r_s := \sum_{i=1}^{d} \left(U_s^{(i)}\right)^2$ such that for all $i = 1, ..., d$,

$$U_s^{(i)} = U_0^{(i)} + \int_0^s \sigma_i(u) \mathrm{d}X_u^{(i)}$$

where for each $i = 1, ..., d$, $\{\sigma_i\}_{i=1,...,d}$ are deterministic volatility functions and $U_0^{(i)}$ is a constant. $\left\{X_u^{(i)}\right\}_{i=1,...,d}$ are d i.i.d. pure jump Lévy processes with their Lévy measure υ. This model can be regarded as a special case of a general Lévy quadratic model. The bond prices are given by (1.2) and (1.3) in [5], as a solution of a system of Riccati equations. For this particular case, we use the Dyson series shown in Theorem 3.9 to give an explicit representation for its bond price $E\left[F|\mathcal{F}_t\right]$ with $F = \exp\left(-\int_t^T r_s \mathrm{d}s\right)$.

In the following, we will denote $s_1 \vee s_2 := \max\{s_1, s_2\}$. We first provide the detailed calculations for the simple case when $d = 1$, $U_0^{(1)} = 0$, $\sigma_1(u) = 1$, i.e. $r_s = X_s^2$. Then by the chain rule, we have

$$x_1 D_{s_1,x_1} F = \exp\left(-\int_t^T X_s^2 \mathrm{d}s - x_1 D_{s_1,x_1}\int_t^T X_s^2 \mathrm{d}s\right) - \exp\left(-\int_t^T X_s^2 \mathrm{d}s\right).$$

By induction as well as the fact $D_{u,x}^n \int_t^T X_s^2 \mathrm{d}s = 0$ for any u and x when $n > 2$, we obtain: for $s_1, ..., s_n \in [t, T]$,

$$\begin{aligned}
& x_1 ... x_n D_{s_1,x_1} ... D_{s_n,x_n} F \\
= \ & \sum_{i=2}^n (-1)^{n-i} \sum_{\{j_1 \leq ... \leq j_i\} \subset \{1,...,n\}} \\
& \exp\left(-\int_t^T X_s^2 \mathrm{d}s - \sum_{r=1}^i x_{j_r} D_{s_{j_r},x_{j_r}} \int_t^T X_s^2 \mathrm{d}s - \sum_{1 \leq r_1 < r_2 \leq i} x_{j_{r_1}} x_{j_{r_2}} D_{s_{j_{r_1}},x_{j_{r_1}}} D_{s_{j_{r_2}},x_{j_{r_2}}} \int_t^T X_s^2 \mathrm{d}s\right) \\
& + (-1)^{n-1} \sum_{j=1}^n \exp\left(-\int_t^T X_s^2 \mathrm{d}s - x_j D_{s_j,x_j} \int_t^T X_s^2 \mathrm{d}s\right) + (-1)^n \exp\left(-\int_t^T X_s^2 \mathrm{d}s\right) \\
= \ & \sum_{i=2}^n (-1)^{n-i} \sum_{\{j_1 \leq ... \leq j_i\} \subset \{1,...,n\}} \\
& \exp\left(-\int_t^T X_s^2 \mathrm{d}s - \sum_{r=1}^i \left(2x_{j_r} \int_{s_{j_r}}^T X_s \mathrm{d}s + (T - s_{j_r}) x_{j_r}^2\right) - \sum_{1 \leq r_1 < r_2 \leq i} 2(T - \max\{s_{j_{r_1}}, s_{j_{r_2}}\}) x_{j_{r_1}} x_{j_{r_2}}\right) \\
& + (-1)^{n-1} \sum_{j=1}^n \exp\left(-\int_t^T X_s^2 \mathrm{d}s - 2x_j \int_{s_j}^T X_s \mathrm{d}s - (T - s_j) x_j^2\right) + (-1)^n \exp\left(-\int_t^T X_s^2 \mathrm{d}s\right).
\end{aligned}$$

Then, by applying freezing path operator ω^t, we obtain:

$$\begin{aligned}
& x_1 ... x_n \omega^t \circ D_{s_1,x_1} ... D_{s_n,x_n} F \\
= \ & \sum_{i=2}^n (-1)^{n-i} \sum_{\{j_1 \leq ... \leq j_i\} \subset \{1,...,n\}} \\
& \exp\left(-(T - t) X_t^2 - \sum_{r=1}^i (2X_t x_{j_r} + x_{j_r}^2)(T - s_{j_r}) - \sum_{1 \leq r_1 < r_2 \leq i} 2(T - s_{j_{r_1}} \vee s_{j_{r_2}}) x_{j_{r_1}} x_{j_{r_2}}\right) \\
& + (-1)^{n-1} \sum_{j=1}^n \exp\left(-(T - t) X_t^2 - (2X_t x_j + x_j^2)(T - s_j)\right) + (-1)^n \exp\left(-X_t^2 (T - t)\right).
\end{aligned}$$

With the help of symmetricity, we can evaluate

$$
\begin{aligned}
&\int_{[t,T]^n}\int_{\mathbb{R}_0^n} \omega^t \circ D_{s_1,x_1}...D_{s_n,x_n}F\,(x\nu(\mathrm{d}x))^{\otimes n}\,(\mathrm{d}s)^{\otimes n} \\
= \;& [\sum_{i=2}^{n}(-1)^{n-i}\binom{n}{i}\int_{[t,T]^n}\int_{\mathbb{R}_0^n} \\
& \exp\left(-(T-t)X_t^2-\sum_{r=1}^{i}(2X_tx_r+x_r^2)(T-s_r)-\sum_{1\le r_1<r_2\le i}2(T-s_{r_1}\vee s_{r_2})x_{r_1}x_{r_2}\right)(\nu(\mathrm{d}x))^{\otimes n}\,(\mathrm{d}s)^{\otimes n}] \\
& +\int_{[t,T]^n}\int_{\mathbb{R}_0^n}(-1)^{n-1}n\exp\left(-(T-t)X_t^2-(2X_tx_1+x_1^2)(T-s_1)\right) \\
& +(-1)^n\exp\left(-X_t^2(T-t)\right)(\nu(\mathrm{d}x))^{\otimes n}\,(\mathrm{d}s)^{\otimes n} \\
= \;& \sum_{i=2}^{n}(-1)^{n-i}\binom{n}{i}\exp\left(-(T-t)X_t^2\right)(T-t)^{n-i}\int_{\mathbb{R}_0^n}\int_{[t,T]^i} \\
& \exp\left(-\sum_{r=1}^{i}(2X_tx_r+x_r^2)(T-s_r)-\sum_{1\le r_1<r_2\le i}2(T-s_{r_1}\vee s_{r_2})x_{r_1}x_{r_2}\right)(\mathrm{d}s)^{\otimes i}\,(\nu(\mathrm{d}x))^{\otimes n} \\
& +(-1)^{n-1}n\exp\left(-(T-t)X_t^2\right)(T-t)^{n-1}\int_{\mathbb{R}_0^n}\frac{\exp\left(-(2X_tx_1+x_1^2)(T-t)\right)}{-(2X_tx_1+x_1^2)}(\nu(\mathrm{d}x))^{\otimes n} \\
& +(-1)^n\exp\left(-X_t^2(T-t)\right)(T-t)^n\int_{\mathbb{R}_0^n}(\nu(\mathrm{d}x))^{\otimes n}. \qquad (5.1)
\end{aligned}
$$

Now the technical part is to evaluate the following integral:

$$\int_{[t,T]^i}\exp\left(-\sum_{r=1}^{i}(2X_tx_r+x_r^2)(T-s_r)-\sum_{1\le r_1<r_2\le i}2(T-s_{r_1}\vee s_{r_2})x_{r_1}x_{r_2}\right)(\mathrm{d}s)^{\otimes i}.$$

We define a sequence of coefficients $\{a_j\}_{j=1,...,i}$ as

$$a_j := 2X_tx_j + x_j^2 + 2x_j(x_{j-1}+...+x_1) \text{ for } j\ge 2;\quad a_1 := 2X_tx_1+x_1^2.$$

Then by simplifying the subscripts of the variables and restricting the domain to the order $t\le s_1\le ...\le s_i\le T$, we define

$$I_i(x_1,...,x_i) := \int_{t\le s_1\le...\le s_i\le T}\exp\left(-a_1(T-s_1)-...-a_i(T-s_i)\right)(\mathrm{d}s)^{\otimes i}.$$

To evaluate this integral, we first set up the following recurrence formula: for simplicity, we get rid of the parameters and set $I_0 := 1$

$$I_i = \frac{(-1)^i}{a_i...(a_i+...+a_1)}e^{-(a_i+...+a_1)(T-t)}+\frac{(-1)^{i+1}}{a_i...(a_i+...+a_1)}I_0+...+\frac{(-1)^3}{a_i(a_i+a_{i-1})}I_{i-2}+\frac{(-1)^2}{a_i}I_{i-1}.$$

Then by induction, we obtain

$$I_i(x_1,...,x_i) = B_i + \sum_{k=0}^{i-1}\frac{B_k}{a_{k+1}(a_{k+1}+a_{k+2})...(a_{k+1}+...+a_i)} \qquad (5.2)$$

where

$$B_k := \frac{(-1)^k \exp\left(-(a_1 + ... + a_k)(T-t)\right)}{a_k(a_k + a_{k-1})...(a_k + ... + a_1)} \text{ for } k \geq 1 \text{ and } B_0 := 1.$$

For example,

$$\begin{aligned}
I_2(x_1, x_2) &= \int_{t \leq s_1 \leq s_2 \leq T} \exp\left(-a_1(T-s_1) - a_2(T-s_2)\right) \mathrm{d}s_2 \mathrm{d}s_1 \\
&= \frac{1}{a_2} \int_t^T \exp(-a_1(T-s_1)) - \exp(-(a_2+a_1)(T-s_1)) \mathrm{d}s_1 \\
&= \frac{1 - \exp(-a_1(T-t))}{a_2 a_1} - \frac{1 - \exp(-(a_2+a_1)(T-t))}{a_2(a_2+a_1)} \\
&= \frac{\exp(-(a_2+a_1)(T-t))}{a_2(a_2+a_1)} - \frac{\exp(-a_1(T-t))}{a_2 a_1} + \frac{1}{a_1(a_1+a_2)} \\
&= B_2 + \frac{B_1}{a_2} + \frac{B_0}{a_1(a_1+a_2)}.
\end{aligned}$$

Then by replacing (5.2) into (5.1), we finally obtain:

$$\begin{aligned}
&\int_{[t,T]^n} \int_{\mathbb{R}_0^n} \omega^t \circ D_{s_1,x_1} ... D_{s_n,x_n} F \left(x\nu(\mathrm{d}x)\right)^{\otimes n} (\mathrm{d}s)^{\otimes n} \\
= &\sum_{i=1}^n (-1)^{n-i} \binom{n}{i} \exp\left(-(T-t)X_t^2\right) (T-t)^{n-i} \int_{\mathbb{R}_0^n} i! I_i(x_1, ..., x_i) \left(\nu(\mathrm{d}x)\right)^{\otimes n} \\
&+ (-1)^n \exp\left(-X_t^2(T-t)\right) (T-t)^n \int_{\mathbb{R}_0^n} \left(\nu(\mathrm{d}x)\right)^{\otimes n}.
\end{aligned}$$

Therefore by Theorem 3.9 and changing summation between n and i, we get

$$\begin{aligned}
&E\left[F|\mathcal{F}_t\right] \\
= &\sum_{n=0}^\infty \frac{1}{n!} \int_{[t,T]^n} \int_{\mathbb{R}_0^n} \omega^t \circ D_{s_1,x_1} ... D_{s_n,x_n} F \left(x\nu(\mathrm{d}x)\right)^{\otimes n} (\mathrm{d}s)^{\otimes n} \\
= &\exp\left(-X_t^2(T-t) + (T-t)\int_{\mathbb{R}_0} \nu(\mathrm{d}x)\right) \left(1 + \sum_{i=1}^\infty (-1)^i \int_{\mathbb{R}_0^i} I_i(x_1, ..., x_i) \left(\nu(\mathrm{d}x)\right)^{\otimes i}\right).
\end{aligned}$$

Now we provide the bond pricing formula for the case when $d > 1$ and $\{\sigma_i\}_{i=1,...,d}$ are deterministic volatility functions. We calculate $E\left[F|\mathcal{F}_t\right]$ by decomposing the filtration into $\left(\mathcal{F}_t^{(i)}\right)^{\otimes d}$ where for each i, $\mathcal{F}_t^{(i)}$ is the natural filtration generated by $X^{(i)}$. Then by the independence of $\left\{X^{(i)}\right\}_{i=1,...,d}$ we have

$$E\left[F|\mathcal{F}_t\right] = \prod_{i=1}^d E\left[F^{(i)}|\mathcal{F}_t^{(i)}\right] := \prod_{i=1}^d E\left[\exp\left(-\int_t^T \left(U_0^{(i)} + \int_0^s \sigma_i(u)\mathrm{d}X_u^{(i)}\right)^2 \mathrm{d}s\right)|\mathcal{F}_t^{(i)}\right].$$

If we denote $D^{(i)}$ as the Malliavin derivative for $X^{(i)}$, then similar calculations show that

$$
\begin{aligned}
& x_1 ... x_n D^{(1)}_{s_1,x_1} ... D^{(i)}_{s_n,x_n} F^{(i)} \\
= \; & \sum_{l=2}^{n} (-1)^{n-l} \sum_{\{j_1 \leq ... \leq j_l\} \subset \{1,...,n\}} \\
& \exp\Big(- \int_t^T \left(U_s^{(i)}\right)^2 \mathrm{d}s - \sum_{r=1}^{l} \left(2x_{j_r} \int_{s_{j_r}}^T U_s^{(i)} \sigma_i(s) \mathrm{d}s + x_{j_r}^2 \int_{s_{j_r}}^T \sigma_i^2(s) \mathrm{d}s \right) \\
& - \sum_{1 \leq r_1 < r_2 \leq l} 2x_{j_{r_1}} x_{j_{r_2}} \int_{s_{j_{r_1}} \vee s_{j_{r_2}}}^T \sigma_i^2(s) \mathrm{d}s \Big) \\
& + (-1)^{n-1} \sum_{j=1}^{n} \exp\left(- \int_t^T \left(U_s^{(i)}\right)^2 \mathrm{d}s - \sum_{r=1}^{n} \left(2x_j \int_{s_j}^T U_s^{(i)} \sigma_i(s) \mathrm{d}s + x_j^2 \int_{s_j}^T \sigma_i^2(s) \mathrm{d}s \right) \right) \\
& + (-1)^n \exp\left(- \int_t^T \left(U_s^{(i)}\right)^2 \mathrm{d}s \right).
\end{aligned}
$$

Now if we define

$$
\begin{aligned}
b_l^{(i)}(x_1, ..., x_l) \; : \; &= \int_{[t,T]^l} \exp\Big(- \sum_{r=1}^{l} \left(2U_t^{(i)} x_r \int_{s_r}^T \sigma_i(s) \mathrm{d}s + x_r^2 \int_{s_r}^T \sigma_i^2(s) \mathrm{d}s \right) \\
& \quad - \sum_{1 \leq r_1 < r_2 \leq l} 2x_{r_1} x_{r_2} \int_{s_{r_1} \vee s_{r_2}}^T \sigma_i^2(s) \mathrm{d}s \Big) (\mathrm{d}s)^{\otimes i}; \\
b_1^{(i)}(x_1) \; : \; &= \int_t^T \exp\left(-2U_t^{(i)} x_1 \int_{s_1}^T \sigma_i(s) \mathrm{d}s - x_1^2 \int_{s_1}^T \sigma_i^2(s) \mathrm{d}s \right) \mathrm{d}s_1, \quad b_0^{(i)} = 1,
\end{aligned}
$$

we finally obtain

$$
E\left[F^{(i)} | \mathcal{F}_t^{(i)} \right] = \exp\left(- \left(U_t^{(i)}\right)^2 (T-t) + (T-t) \int_{\mathbb{R}_0} \nu(\mathrm{d}x) \right) \left(\sum_{l=0}^{\infty} \int_{\mathbb{R}_0^l} b_l^{(i)} \left(\nu(\mathrm{d}x)\right)^{\otimes l} \right)
$$

and

$$
E\left[F | \mathcal{F}_t \right] = \prod_{i=1}^{d} E\left[F^{(i)} | \mathcal{F}_t^{(i)} \right] = \exp\left(-r_t(T-t) + d(T-t) \int_{\mathbb{R}_0} \nu(\mathrm{d}x) \right) \prod_{i=1}^{d} \left(\sum_{l=0}^{\infty} \int_{\mathbb{R}_0^l} b_l^{(i)} \left(\nu(\mathrm{d}x)\right)^{\otimes l} \right),
$$

which, to the best of our knowledge, is an original result.

6 Appendix

6.1 Proof of Theorem 3.1

The proof of this theorem is similar with the proof of Theorem 2.1 for the Backward Taylor Expansion of Brownian motion in [11]. Here we skip the details but only list several important lemmas and steps.

First for all $k \leq K-1$, similar to (5.4) in [11], by iterating applying the Clark-Haussmann-Ocone formula (2.6) we obtain

$$E[F|\mathcal{F}_{t_k}] = \sum_{l=0}^{L-1} \int_{t_k \leq s_1 \leq \ldots \leq s_l \leq t_{k+1}} \int_{\mathbb{R}_0^l} E[D_{t_{k+1},x_1} \ldots D_{t_{k+1},x_l} F|\mathcal{F}_{t_{k+1}}](x\widetilde{N}(\mathrm{d}s, \mathrm{d}x))^{\otimes l} + R^L_{[t_k,t_{k+1}]} \tag{6.1}$$

where the remainder is defined as

$$R^L_{[t_k,t_{k+1}]} := \int_{t_k \leq s_1 \leq \ldots \leq s_L \leq t_{k+1}} \int_{\mathbb{R}_0^L} E[D_{t_{k+1},x_1} \ldots D_{t_{k+1},x_L} F|\mathcal{F}_{s_L}](x\widetilde{N}(\mathrm{d}s, \mathrm{d}x))^{\otimes L}.$$

Filling the same role as Lemmata 5.1 and 5.2 in [11], we have the following lemma and relation (6.3) to guarantee the series (3.3)'s convergence.

By Proposition 5.4 and 5.7 in [28], (2.12) in [15] and induction, we have

Lemma 6.1 *If both u and $v \in \mathbb{D} \subset \mathrm{Dom}\,\delta$, we have*

$$\begin{aligned} E[\delta(u)\delta(u)] &= \int_{[0,T]\times\mathbb{R}_0} E[u_{t,x}v_{t,x}]\mu(\mathrm{d}t, \mathrm{d}x) \\ &\quad - \int_{[0,T]\times\mathbb{R}_0} \int_{[0,T]\times\mathbb{R}_0} E[D_{s,y}u_{t,x}D_{t,x}y_{s,y}] \left(x^2\nu(\mathrm{d}x)\mathrm{d}t\right) \left(y^2\nu(\mathrm{d}y)\mathrm{d}s\right). \end{aligned}$$

Then by induction, we can show that:

$$E\left[\left(\int_{[0,T]^L\times\mathbb{R}_0^L} X(s_1, x_1, \ldots, s_L, x_L)(\widetilde{N}(\mathrm{d}s, \mathrm{d}x))^{\otimes L}\right)^2\right] = \sum_{i=0}^{L} \binom{L}{i}^2 i! E\left[\left\|D^{L-i}X\right\|^2_{H^{\otimes(2L-i)}}\right], \tag{6.2}$$

where

$$\begin{aligned} &\left\|D^{L-i}X\right\|^2_{H^{\otimes(2L-i)}} := \int_{[0,T]^{2L}\times\mathbb{R}_0^{2L}} D_{t_1,x_1} \ldots D_{t_{L-i},x_{L-i}} X(s_1, y_1, \ldots, s_L, y_L)(x^2\nu(\mathrm{d}x)\mathrm{d}t)^{\otimes(L-i)} \\ &\times D_{s_1,y_1} \ldots D_{s_{L-i},y_{L-i}} D^{L-i}_{s_1,\ldots,s_{L-i}} X(t_1, x_1, \ldots, t_L, x_L)(y^2\nu(\mathrm{d}y)\mathrm{d}s)^{\otimes(L-i)} \\ &\times \left(x^2_{L-i+1}\nu(\mathrm{d}x_{L-i+1})\mathrm{d}t_{L-i+1} \ldots x_L^2\nu(\mathrm{d}x_L)\mathrm{d}t_L\right) \left(y^2_{L-i+1}\nu(\mathrm{d}y_{L-i+1})\mathrm{d}s_{L-i+1} \ldots y_L^2\nu(\mathrm{d}y_L)\mathrm{d}s_L\right). \end{aligned}$$

Applying this lemma to the remainder $R^L_{[t_k,t_{k+1}]}$ and with the help of condition (3.2), similarly to the proof of Lemma 5.2 in [11], we obtain:

$$\begin{aligned} E\left[\left(R^L_{[t_k,t_{k+1}]}\right)^2\right] &\leq \sum_{i=0}^{L} \int_{\mathbb{R}_0^{2L-i}} E\left[\left(D_{t_{k+1},x_1} \ldots D_{t_{k+1},x_{2L-i}} F\right)^2\right] \left(x^2\nu(\mathrm{d}x)\right)^{\otimes 2L-i} \binom{L}{i}^4 \frac{i!}{(L!)^2}(t_{k+1}-t_k)^{2L-i} \\ &\xrightarrow{L\to\infty} 0. \end{aligned} \tag{6.3}$$

Now according to the integration by parts formula given in Lemma 2.2 and induction, we can prove the following lemma, which is the counterpart of Lemma 5.3 in [11].

Lemma 6.2 *Let F be defined as in Theorem 3.1, then for any $k \leq K-1$ and $l \geq 1$*

$$\begin{aligned} \int_{t_k \leq s_1 \leq \ldots \leq s_l \leq t_{k+1}} \int_{\mathbb{R}_0^l} F(x\widetilde{N}(\mathrm{d}s, \mathrm{d}x))^{\otimes l} &= \sum_{j=0}^{l} \sum_{i_1,\ldots,i_l=0}^{\infty} (-1)^{i_1+\ldots+i_l+j} \int_{\mathbb{R}_0^l} D_{t_{k+1},x_1} \ldots D_{t_{k+1},x_l} F(x^2\nu(\mathrm{d}x))^{\otimes l} \\ &\quad \times C_{n-j}(X_{t_{k+1}} - X_{t_k}, \mu(t_{k+1}-t_k)) \frac{\mu^j(t_{k+1}-t_k)^j}{(n-j)!j!}. \end{aligned}$$

Then applying this lemma to each term (6.1) with the remainder tends to 0, we finish the proof of Theorem 3.1.

6.2 Proof of Proposition 3.4

It is well-known (see for instance theorem 12.51 in [8]) that smooth functions with compact support are dense in $L^2([0,T]^n)$. Thus we assume that the Wiener chaos decomposition holds with differentiable kernels. Namely, in $L^2(\mathbb{P})$:

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

where f_n are infinitely differentiable. We will use this important fact in the proof of this proposition as well as in the proof of the next one. To simplify notation, we only prove the lemma for the Poisson process. Without loss of generality, we suppose the Poisson intensity $\lambda = 1$. Then we want to show

$$\omega^t \circ I_n(f_n) = \sum_{k=0}^{n}(-1)^k \binom{n}{k} \int_{[t,T]^k} I_{n-k}(f_n \chi_{[0,t]^{n-k}})(u_1, ..., u_k)(\mathrm{d}u)^{\otimes k}, \tag{6.4}$$

where $I_{n-k}(f_n)(u_1, ..., u_k)$ is defined as:

$$I_{n-k}(f_n)(u_1, ..., u_k) := \int_{[0,T]^{n-k}} f_n(s_1, ..., s_{n-k}, u_1, ..., u_k)(\mathrm{d}\tilde{N}_s)^{\otimes(n-k)}(\mathrm{d}u)^{\otimes k}.$$

The proof for pure jump Lévy processes follows exactly the same procedure.

Because that we define the freezing path operator ω^t for Poisson process, i.e. $\omega^t \circ N_s = N_s \chi_{[0,t]}(s) + N_t \chi_{[t,T]}(s)$. We first present the multiple stochastic integral $I_n(f_n)$ using Poisson processes instead of compensated Poisson processes. We emphasize the symmetricity of f_n here, then

$$\begin{aligned} I_n(f_n) &= \int_{[0,T]^n} f_n(s_1, ..., s_n) \left(\mathrm{d}\tilde{N}_s\right)^{\otimes n} \\ &= \sum_{k=0}^{n}(-1)^k \binom{n}{k} \int_{[0,T]^k} \int_{[0,T]^{n-k}} f_n(s_1, ..., s_{n-k}, u_1, ..., u_k) \left(\mathrm{d}N_s\right)^{\otimes(n-k)} \left(\mathrm{d}u\right)^{\otimes k} \\ &= \sum_{k=0}^{n}(-1)^k \binom{n}{k} \int_{[0,T]^k} J_{n-k}(f_n)(u_1, ..., u_k) \left(\mathrm{d}u\right)^{\otimes k} \end{aligned} \tag{6.5}$$

with

$$J_{n-k}(f_n)(u_1, ..., u_k) := \int_{[0,T]^{n-k}} f_n(s_1, ..., s_{n-k}, u_1, ..., u_k) \left(\mathrm{d}N_s\right)^{\otimes n-k}.$$

We then establish the integration by parts formula for Poisson process. According to (2.5), we have

$$\int_0^T F u_s \mathrm{d}\tilde{N}_s = F \int_0^T u_s \mathrm{d}\tilde{N}_s - \int_0^T D_s F u_s \mathrm{d}s - \int_0^T D_s F u_s \mathrm{d}\tilde{N}_s.$$

We can rewrite as:

$$\begin{aligned} &\int_0^T F u_s \mathrm{d}N_s - \int_0^T F u_s \mathrm{d}s \\ = \;& F \int_0^T u_s \mathrm{d}N_s - F \int_0^T u_s \mathrm{d}s - \int_0^T D_s F u_s \mathrm{d}s - \int_0^T D_s F u_s \mathrm{d}N_s + \int_0^T D_s F u_s \mathrm{d}s \end{aligned}$$

which implies

$$\int_0^T F u_s \mathrm{d}N_s = F\int_0^T u_s \mathrm{d}N_s - \int_0^T D_s F u_s \mathrm{d}N_s. \tag{6.6}$$

Now we begin to prove the lemma. We split it into two steps.

Step 1. We first prove that

$$\omega^t \circ J_n(f_n) = J_n(f_n \chi_{[0,t]^n}). \tag{6.7}$$

By induction, we assume that for $k \leq n-1$,

$$\omega^t \circ J_k(f_n)(s_1, ..., s_{n-k}) = J_k(f_n \chi_{[0,t]^k})(s_1, ..., s_{n-k}) \tag{6.8}$$

the variables $s_1, ..., s_{n-k}$ can be regarded as constants since f_n is a n-variable function and J_k contains k integrals. Now we are going to set up the recurrence formula. Firstly, we apply the Itô's lemma:

$$\begin{aligned}
&\int_{[0,T]^n} f_n(s_1, ..., s_n)\,(\mathrm{d}N_s)^{\otimes n} \\
= &\int_{[0,T]^{n-1}} N_T f_n(s_1, ..., s_{n-1}, T)\,(\mathrm{d}N_s)^{\otimes(n-1)} - \int_{[0,T]^{n-1}} \int_0^T \frac{\partial f_n(s_1, ..., s_{n-1}, s_n)}{\partial s_n} N_{s_n} \mathrm{d}s_n\,(\mathrm{d}N_s)^{\otimes(n-1)}
\end{aligned} \tag{6.9}$$

Then we consider the two terms separately and apply the integration by parts formula (6.6): the first term of (6.9) can be simplified as

$$\begin{aligned}
&\int_{[0,T]^{n-1}} N_T f_n(s_1, ..., s_{n-1}, T)\,(\mathrm{d}N_s)^{\otimes(n-1)} \\
= &\int_{[0,T]^{n-2}} (N_T - 1) \int_0^T f_n(s_1, ..., s_{n-1}, T) \mathrm{d}N_{s_{n-1}}\,(\mathrm{d}N_s)^{\otimes(n-2)} \\
= &\ \dots \\
= &\ (N_T - n + 1) \int_{[0,T]^{n-1}} f_n(s_1, ..., s_{n-1}, T)\,(\mathrm{d}N_s)^{\otimes(n-1)}.
\end{aligned} \tag{6.10}$$

The second term of (6.9) is more complicated. By (6.6) as well as the symmetricity of f_n, we obtain:

$$
\begin{aligned}
& \int_{[0,T]^{n-1}} \int_0^T \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} N_{s_n} \mathrm{d}s_n \, (\mathrm{d}N_s)^{\otimes(n-1)} \\
= & \int_0^T \int_{[0,T]^{n-2}} N_{s_n} \int_0^T \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \mathrm{d}N_{s_{n-1}} \, (\mathrm{d}N_s)^{\otimes(n-2)} \, \mathrm{d}s_n \\
& - \int_0^T \int_{[0,T]^{n-2}} \int_0^{s_n} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \mathrm{d}N_{s_{n-1}} \, (\mathrm{d}N_s)^{\otimes(n-2)} \, \mathrm{d}s_n \\
= & \ ... \\
= & \int_0^T N_{s_n} \int_{[0,T]^{n-1}} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \, (\mathrm{d}N_s)^{\otimes(n-1)} \, \mathrm{d}s_n \\
& -(n-1) \int_0^T \int_{[0,T]^{n-2}} \int_0^{s_n} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \mathrm{d}N_{s_{n-1}} \, (\mathrm{d}N_s)^{\otimes(n-2)} \, \mathrm{d}s_n \\
= & \int_0^T N_{s_n} \int_{[0,T]^{n-1}} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \, (\mathrm{d}N_s)^{\otimes(n-1)} \, \mathrm{d}s_n \\
& -(n-1) \int_{[0,T]^{n-1}} \left(f_n(s_1,...,s_{n-1},T) - f_n(s_1,...,s_{n-1},s_{n-1}) \right) (\mathrm{d}N_s)^{\otimes(n-1)} . \qquad (6.11)
\end{aligned}
$$

Combining (6.10) and (6.11), we obtain

$$
\begin{aligned}
& \int_{[0,T]^n} f_n(s_1,...,s_n) \, (\mathrm{d}N_s)^{\otimes n} \\
= & \ N_T \int_{[0,T]^{n-1}} f_n(s_1,...,s_{n-1},T) \, (\mathrm{d}N_s)^{\otimes(n-1)} - \int_0^T N_{s_n} \int_{[0,T]^{n-1}} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \, (\mathrm{d}N_s)^{\otimes(n-1)} \, \mathrm{d}s_n \\
& -(n-1) \int_{[0,T]^{n-1}} f_n(s_1,...,s_{n-1},s_{n-1}) \, (\mathrm{d}N_s)^{\otimes(n-1)} . \qquad (6.12)
\end{aligned}
$$

Notice that when acted by the freezing path operator, the first and second term in (6.12) can be applied by our induction assumption (6.8) but the third term can not, because of the non-symmetricity on the variable s_{n-1}. So we have

$$
\begin{aligned}
& \omega^t \circ \int_{[0,T]^n} f_n(s_1,...,s_n) \, (\mathrm{d}N_s)^{\otimes n} \\
= & \ N_t \int_{[0,t]^{n-1}} f_n(s_1,...,s_{n-1},T) \, (\mathrm{d}N_s)^{\otimes(n-1)} - \int_0^t N_{s_n} \int_{[0,t]^{n-1}} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \, (\mathrm{d}N_s)^{\otimes(n-1)} \, \mathrm{d}s_n \\
& -N_t \int_t^T \int_{[0,t]^{n-1}} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \, (\mathrm{d}N_s)^{\otimes(n-1)} \, \mathrm{d}s_n \\
& -(n-1) \left(\omega^t \circ \int_{[0,T]^{n-1}} f_n(s_1,...,s_{n-1},s_{n-1}) \, (\mathrm{d}N_s)^{\otimes(n-1)} \right) \\
= & \ N_t \int_{[0,t]^{n-1}} f_n(s_1,...,s_{n-1},t) \, (\mathrm{d}N_s)^{\otimes(n-1)} - \int_0^t N_{s_n} \int_{[0,t]^{n-1}} \frac{\partial f_n(s_1,...,s_{n-1},s_n)}{\partial s_n} \, (\mathrm{d}N_s)^{\otimes(n-1)} \, \mathrm{d}s_n \\
& -(n-1) \left(\omega^t \circ \int_{[0,T]^{n-1}} f_n(s_1,...,s_{n-1},s_{n-1}) \, (\mathrm{d}N_s)^{\otimes(n-1)} \right) . \qquad (6.13)
\end{aligned}
$$

In order to finish the induction, we need to prove

$$
\omega^t \circ \int_{[0,T]^{n-1}} f_n(s_1,...,s_{n-1},s_{n-1}) \, (\mathrm{d}N_s)^{\otimes(n-1)} = \int_{[0,t]^{n-1}} f_n(s_1,...,s_{n-1},s_{n-1}) \, (\mathrm{d}N_s)^{\otimes(n-1)} . \quad (6.14)
$$

We apply induction again. Assume that for some $n-k \leq n-2$ we have:

$$\begin{aligned}
& \omega^t \circ \int_{[0,T]^{n-k}} f(s_1, ..., s_{n-k-1}, s_{n-k}, ..., s_{n-k}) \left(\mathrm{d}N_s\right)^{\otimes(n-k-1)} \mathrm{d}N_{s_{n-k}} \\
= & \int_{[0,t]^{n-k}} f(s_1, ..., s_{n-k-1}, s_{n-k}, ..., s_{n-k}) \left(\mathrm{d}N_s\right)^{\otimes(n-k-1)} \mathrm{d}N_{s_{n-k}}. \qquad (6.15)
\end{aligned}$$

We consider the $n-k+1$ case. As in the calculation of (6.12), we obtain

$$\begin{aligned}
& \int_{[0,T]^{n-k}} \int_0^T f(s_1, ..., s_{n-k}, s_{n-k+1}, ..., s_{n-k+1}) \mathrm{d}N_{s_{n-k+1}} \mathrm{d}N_{s_{n-k}} \left(\mathrm{d}N_s\right)^{\otimes(n-k-1)} \\
= & \ N_T \int_{[0,T]^{n-k}} f(s_1, ..., s_{n-k}, T, ..., T) \left(\mathrm{d}N_s\right)^{\otimes(n-k)} \\
& - \int_0^T N_{s_{n-k+1}} \int_{[0,T]^{n-k}} \frac{\partial f(s_1, ..., s_{n-k}, s_{n-k+1}, ..., s_{n-k+1})}{\partial s_{n-k+1}} \left(\mathrm{d}N_s\right)^{\otimes(n-k)} \mathrm{d}s_{n-k+1} \\
& -(n-k-1) \int_{[0,T]^{n-k}} f(s_1, ..., s_{n-k-1}, s_{n-k}, ... s_{n-k}) \left(\mathrm{d}N_s\right)^{\otimes(n-k-1)} \mathrm{d}N_{s_{n-k}}. \qquad (6.16)
\end{aligned}$$

It is clear to see that the first and second term in (6.16) can apply our original assumption for J_n, i.e. (6.8); for the third term we can apply our second assumption (6.15). Thus, when applying the freezing path operator, the variable T becomes t, i.e.

$$\begin{aligned}
& \omega^t \circ \int_{[0,T]^{n-k+1}} f(s_1, ..., s_{n-k}, s_{n-k+1}, ..., s_{n-k+1}) \left(\mathrm{d}N_s\right)^{\otimes(n-k+1)} \\
= & \int_{[0,t]^{n-k+1}} f(s_1, ..., s_{n-k}, s_{n-k+1}, ..., s_{n-k+1}) \left(\mathrm{d}N_s\right)^{\otimes(n-k+1)}.
\end{aligned}$$

Finally we proved (6.14) and therefore from (6.13), we proved (6.7).

Step 2. We now prove (6.4). We use (6.5) to develop the right hand side of (6.4) as:

$$\begin{aligned}
& \sum_{k=0}^{n} (-1)^k \binom{n}{k} \int_{[t,T]^k} I_{n-k}(f_n \chi_{[0,t]^{n-k}})(u_1, ..., u_k)(\mathrm{d}u)^{\otimes k} \\
= & \sum_{k=0}^{n} \sum_{k_1=0}^{n-k} (-1)^k \binom{n}{k} (-1)^{k_1} \binom{n-k}{k_1} \int_{[t,T]^k} \int_{[0,t]^{k_1}} \\
& J_{n-k-k_1}(f_n \chi_{[0,t]^{n-k-k_1}})(u_1, ..., u_k, v_1, ..., v_{k_1}) \left(\mathrm{d}u\right)^{\otimes k} \left(\mathrm{d}v\right)^{\otimes k_1} \\
= & \sum_{m=0}^{n} (-1)^m \binom{n}{m} \sum_{k=0}^{m} \binom{m}{k} \int_{[t,T]^k} \int_{[0,t]^{m-k}} \\
& J_{n-m}(f_n \chi_{[0,t]^{n-m}})(u_1, ..., u_k, v_1, ..., v_{m-k}) \left(\mathrm{d}u\right)^{\otimes k} \left(\mathrm{d}v\right)^{\otimes m-k}
\end{aligned}$$

by supposing again that $m = k + k_1$ and sum over it first. We then need another basic rule of integration with symmetric function g:

$$\begin{aligned}
& \int_{[0,T]^n} g(s_1, ..., s_n)(\mathrm{d}s)^{\otimes n} = \int_{([0,t] \cup [t,T])^n} g(s_1, ..., s_n)(\mathrm{d}s)^{\otimes n} \\
= & \sum_{k=0}^{n} \binom{n}{k} \int_{[0,t]^{n-k}} \int_{[t,T]^k} g(u_1, ..., u_k, v_1, ..., v_{n-k})(\mathrm{d}u)^{\otimes k}(\mathrm{d}v)^{\otimes(n-k)}.
\end{aligned}$$

Therefore according to (6.5) we finally obtain:

$$\begin{aligned}
&\int_{[t,T]^k} I_{n-k}(f_n\chi_{[0,t]^{n-k}})(u_1,...,u_k)(\mathrm{d}u)^{\otimes k}\\
=&\ \sum_{m=0}^{n}(-1)^m\binom{n}{m}\int_{[0,T]^{m-k}} J_{n-m}(f_n\chi_{[0,t]^{n-m}})(u_1,...,u_m)\,(\mathrm{d}u)^{\otimes m}\\
=&\ \omega^t\circ I_n(f_n).
\end{aligned}$$

We finished the proof of the lemma. □

6.3 Proof of Theorem 3.5

The proof is constructive. For any fixed $t\in[0,T]$, if F has its chaos decomposition $\sum_{n=0}^{\infty} I_n(f_n)$, then for fixed N (depending on M), we will study $F^{M,N} := \sum_{n=0}^{M} I_n(f_n^N)$, where:

$$\begin{aligned}
&f_n^N(s_1,x_1,...,s_n,x_n) := \qquad (6.17)\\
&f_n(t\chi_{[s,s+1/N]}(s_1)+s_1\chi_{[0,T]/[s,s+1/N]}(s_1),x_1,...,t\chi_{[s,s+1/N]}(s_n)+s_n\chi_{[0,T]/[s,s+1/N]}(s_n),x_n).
\end{aligned}$$

In words, the kernel f_n^N is constant when its arguments lie between s and $s+1/N$. Then we have the following lemma.

Lemma 6.3 $\omega^t\circ I_n(f_n^N)\xrightarrow[N\to\infty]{L^2(\mathbb{P})}\omega^t\circ I_n(f_n)$ *and in particular:*

$$E\left[(\omega^t\circ\left(I_n(f_n^N)-I_n(f_n)\right)^2\right]\le\frac{C(n!)^2n^7}{N^3}$$

where C is a constant which does not depend on N and n.

Proof. For any fixed n, we define a sequence of sets $\{A_{k_1,k_2}\}_{k_1+k_2\le n}$ as

$$A_{k_1,k_2}:=\left\{s_1,...,s_n: 0\le s_1\le...\le s_{k_1}\le t\le s_{k_1+1}\le...\le s_{k_1+k_2}\le t+\frac{1}{N}\le s_{k_1+k_2+1}\le...\le s_n\le T\right\}.$$

Observe that on $A_{k_1,0}$ the kernels f_n and f_n^N coincide. According to (6.17), we obtain:

$$\begin{aligned}
&\omega^t\circ I_n(f_n)-\omega^t\circ I_n(f_n^N)\\
=&\ n!\sum_{k_1+k_2\le n,\ k_2\neq 0}\omega^t\circ\int_{A_{k_1,k_2}}\int_{\mathbb{R}_0^n}\left(f_n-f_n^N\right)(s_1,...,s_n,x^{\otimes n})(x\tilde{N}(\mathrm{d}s,\mathrm{d}x))^{\otimes n}. \qquad (6.18)
\end{aligned}$$

To bound (6.18), we apply Proposition 3.4 to obtain:

$$\begin{aligned}
&E\left[(\omega^t\circ I_n(f_n))^2\right]\\
=&\ (n!)^2\sum_{k=0}^{n}\frac{1}{(k!)^2}\int_{[t,T]^k\times\mathbb{R}_0^k}\int_{\{0\le s_1\le...\le s_{n-k}\le t\}\times\mathbb{R}_0^{n-k}} f_n(s_1,x_1...,s_{n-k},x_{n-k},u_1,y_1,...,u_k,y_k)^2\\
&(x^2\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n-k}(y^2\nu(\mathrm{d}y)\mathrm{d}u)^{\otimes k}\\
<&\ \infty. \qquad (6.19)
\end{aligned}$$

Now we apply (6.19) on (6.18) and by Cauchy-Schwartz inequality, we have:

$$
\begin{aligned}
&E[\big(\omega^t \circ I_n(f_n^N) - \omega^t \circ I_n(f_n)\big)^2] \\
\leq\ & (nn!)^2 \sum_{\substack{k_1+k_2\leq n \\ k_2\neq 0}} E\left[\left(\omega^t \circ \int_{A_{k_1,k_2}} \int_{\mathbb{R}_0^n} \left(f_n - f_n^N\right)(s_1,...,s_n,x^{\otimes n})(x\tilde{N}(\mathrm{d}s,\mathrm{d}x))^{\otimes n}\right)^2\right] \\
=\ & (nn!)^2 \sum_{\substack{k_1+k_2\leq n \\ k_2\neq 0}} \sum_{k=n-k_1}^{n} \frac{1}{(k!)^2} \int_{[t,T]^k\times\mathbb{R}_0^k} \int_{\{0\leq s_1\leq ...\leq s_{n-k}\leq t\}\times\mathbb{R}_0^{n-k}} \\
& \left(f_n - f_n^N\right)(s_1,...,s_{n-k},x^{\otimes n-k},u_1,y_1,...,u_k,y_k)^2 \chi_{A_{k_1,k_2}}(s_1,...,s_n)(x^2\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n-k}(y^2\nu(\mathrm{d}y)\mathrm{d}u)^{\otimes k}
\end{aligned} \tag{6.20}
$$

Since f_n is differentiable with respect to $s_1,...,s_n$, there exists a constant C_n such that

$$
|f_n(s_1,x_1,...,s_n,x_n) - f_n(t,x_1,...,t,x_n)| \leq C_n n(\sup_{s_1,...,s_n}(s_i - t)).
$$

Therefore following (6.20), we obtain

$$
E[\big(\omega^t \circ I_n(f_n^N) - \omega^t \circ I_n(f_n)\big)^2] \leq \frac{C(nn!)^2 n^5}{N^3}
$$

where C is a constant which does not depend on n and N. ■

Now we construct F^N by $\sum_{n=0}^{\infty} I_n(f_n^N)$. To prove the theorem, we introduce two sub-series $F^{M,N}$ and F^M by

$$
F^{M,N} := \sum_{n=0}^{M} I_n(f_n^N) \xrightarrow[L^2(\mathbb{P})]{M\to\infty} F^N; \quad F^M := \sum_{n=0}^{M} I_n(f_n) \xrightarrow[L^2(\mathbb{P})]{M\to\infty} F.
$$

For enough large N, we choose M such that $\left(M^7(M!)^2\right)^{1/3} M \leq N$. Then by Lemma 6.3 and Cauchy-Schwarz inequality, there exists a constant $\varepsilon \in (0,1)$ such that

$$
E[\big(\omega^t \circ \left(F^{M,N} - F^M\right)\big)^2] = E\left[\left(\sum_{n=0}^{M}\left(\omega^t \circ I_n(f_n) - \omega^t \circ I_n(f_n^N)\right)\right)^2\right] \leq CM\left(\sum_{n=0}^{M}\frac{(nn!)^2 n^5}{N^3}\right) \leq \frac{C}{N^{2+\varepsilon}}.
$$

Then using triangle inequality, we proved the theorem. □

6.4 Proof of Theorem 3.6

For any $F \in L^2(\mathbb{P})$, $s \in [t,T]$, we choose the sequence $\left\{F^N\right\}_{N\geq 0}$ constructed in Theorem 3.5. Then by the Clark-Ocone formula, we obtain

$$
\begin{aligned}
E[F^N|F_{s-1/N}] &= E[F^N|\mathcal{F}_s] - \int_{s-1/N}^{s}\int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x) \\
&\quad + \int_{s-1/N}^{s}\int_{s_1}^{s}\int_{\mathbb{R}_0^2} E[D_{s_2,x_2}D_{s,x_1}F^N|\mathcal{F}_{s_2}]x_2\widetilde{N}(\mathrm{d}s_2,\mathrm{d}x_2)x_1\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x_1).
\end{aligned}
$$

Since

$$\begin{aligned}
&\int_{s-1/N}^{s}\int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x)\\
=\;& \int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\int_{s-1/N}^{s}\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x) - \int_{\mathbb{R}_0}\int_{s-1/N}^{s} E[D^2_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\mathrm{d}s_1\\
&-\int_{s-1/N}^{s}\int_{\mathbb{R}_0} E[D^2_{s,x}F^N|\mathcal{F}_s]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x).
\end{aligned}$$

Therefore with $\omega^t \circ x\widetilde{N}(\mathrm{d}s,\mathrm{d}x) = -x^2\upsilon(\mathrm{d}x)$ for any $t \leq s$, we obtain

$$\begin{aligned}
&\omega^t \circ \left(\int_{s-1/N}^{s}\int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x)\right) = -\omega^t \circ \int_{\mathbb{R}_0}\frac{1}{N}E[D_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\\
&-\omega^t \circ \left(\int_{\mathbb{R}_0}\int_{s-1/N}^{s} E[D^2_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\mathrm{d}s_1 + \int_{s-1/N}^{s}\int_{\mathbb{R}_0} E[D^2_{s,x}F^N|\mathcal{F}_s]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x)\right).
\end{aligned}$$

Then we can establish the equation as:

$$\begin{aligned}
&E\left[\left(\omega^t \circ N\left(E[F^N|\mathcal{F}_{s-1/N}] - E[F^N|\mathcal{F}_s]\right) + \omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\right)^2\right]\\
\leq\;& 2E\left[\left(\omega^t \circ \int_{\mathbb{R}_0}\int_{s-1/N}^{s} E[D^2_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\mathrm{d}s_1\right)^2\right]\\
&+8E\left[\left(\omega^t \circ \int_{\mathbb{R}_0}\int_{s-1/N}^{s} E[D^2_{s,x}F^N|\mathcal{F}_s]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x)\right)^2\right] = O\left(\frac{1}{N^2}\right)
\end{aligned}\tag{6.21}$$

where the last equality follows from lemma 6.1, (6.19), and the boundedness of $E\left[\left(\int_{s-1/N}^{s}\int_{\mathbb{R}_0} E[D^2_{s,x}F^N]x\widetilde{N}(\mathrm{d}s_1,\mathrm{d}x)\right)^2\right]$. Then, from Theorem 3.5, for some constant $\varepsilon < 1$,

$$\left\|\omega^t \circ N\left(E[F|\mathcal{F}_{s-1/N}] - E[F^N|\mathcal{F}_{s-1/N}]\right)\right\|^2_{L^2(\mathbb{P})} \leq \frac{C}{N^{\varepsilon}};\tag{6.22}$$

Here, for simplicity, we define the L^2 norm $\|\cdot\|^2_{L^2(\mathbb{P})} := E[(\cdot)^2]$. From the closability of the Malliavin derivative operator and Theorem 3.5

$$\left\|\omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x) - \omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\right\|^2_{L^2(\mathbb{P})} \leq \frac{C}{N^{2+\varepsilon}}\tag{6.23}$$

With triangle inequality and Cauchy-Schwartz inequality, we finally have, using (6.21), (6.22), and (6.23):

$$
\begin{aligned}
&\left\|\omega^t \circ N\left(E[F|\mathcal{F}_{s-1/N}] - E[F|\mathcal{F}_s]\right) - \omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\right\|^2_{L^2(\mathbb{P})} \\
\leq\ & \left\|\omega^t \circ N\left(E[F|\mathcal{F}_{s-1/N}] - E[F^N|\mathcal{F}_{s-1/N}]\right)\right\|^2_{L^2(\mathbb{P})} \\
&+ \left\|\omega^t \circ N\left(E[F|\mathcal{F}_s] - E[F^N|\mathcal{F}_s]\right)\right\|^2_{L^2(\mathbb{P})} \\
&+ \left\|\omega^t \circ \left(N(E[F^N|\mathcal{F}_s] - E[F^N|\mathcal{F}_{s-1/N}]) - \omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x)\right)\right\|^2_{L^2(\mathbb{P})} \\
&+ \left\|\omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F^N|\mathcal{F}_s]x\upsilon(\mathrm{d}x) - \omega^t \circ \int_{\mathbb{R}_0} E[D_{s,x}F|\mathcal{F}_s]x\upsilon(\mathrm{d}x)|\right\|^2_{L^2(\mathbb{P})} \\
\leq\ & \frac{C}{N^\varepsilon} \qquad (6.24)
\end{aligned}
$$

or in other words:

$$\left(\frac{\mathrm{d}\omega^t \circ E_sF}{\mathrm{d}s}\right) = -\omega^t \circ \int_{\mathbb{R}_0} D_{s,x}E_sFx\nu(\mathrm{d}x). \qquad (6.25)$$

□

6.5 Proof of Theorem 3.9

Define for $i = 1, .., N(T-s)$:

$$x_i^N := N\omega^t \circ \left(E[F|\mathcal{F}_{s+(i-1)/N}] - E[F|\mathcal{F}_{s+i/N}] - \frac{1}{N}\int_{\mathbb{R}_0} E[D_{s+i/N,x}F|\mathcal{F}_{s+i/N}]x\nu(\mathrm{d}x)\right)$$

We rewrite (6.24) as:

$$E\left[\left(\frac{x_i^N}{N}\right)^2\right] \leq \frac{C}{N^{2+\varepsilon}}$$

Jensen's inequality states that:

$$E\left[\left(\frac{x_i^N}{N}\right)^2\right] \leq \frac{\sum_{i=1}^N E[(x_i^N)^2]}{N} \leq \frac{C}{N^\varepsilon} \qquad (6.26)$$

Since $\int_s^T \int_{\mathbb{R}_0} E[D_{s+i/N,x}F|\mathcal{F}_{s+i/N}]x\nu(\mathrm{d}x))$ is bounded in $L^2(\mathbb{P})$,then

$$\sum_{i=1}^N \frac{x_i^N}{N} \xrightarrow[L^2(\mathbb{P})]{N\to\infty} \omega^t \circ \left(E[F|\mathcal{F}_s] - F - \int_s^T \int_{\mathbb{R}_0} E[D_{s+i/N,x}F|\mathcal{F}_{s+i/N}]x\nu(\mathrm{d}x)\right)$$

Using (6.26), We thus proved that, in $L^2(\mathbb{P})$:

$$\omega^t \circ (E_sF) = \omega^t \circ F + \int_s^T \int_{\mathbb{R}_0} \omega^t \circ D_{s,x} \circ E_s \circ Fx\nu(\mathrm{d}x)\,\mathrm{d}u. \qquad (6.27)$$

Then for positive integer n we define the operator $T_s^{(n)}$ by:

$$T_s^{(n)}F := \sum_{i=0}^n \mathcal{A}_{i,s}F,$$

where

$$\mathcal{A}_{i,s}F := \int_{s\le s_1\le\ldots\le s_i\le T}\int_{\mathbb{R}_0^i} D_{s_1,x_1}\ldots D_{s_i,x_i}F\ (x\nu(\mathrm{d}x))^{\otimes i}\,(\mathrm{d}s)^{\otimes i}.$$

Then by iterating (6.27) we obtain: for $n>0$

$$\begin{aligned}\omega^t\circ(E_sF) &= \omega^t\circ(T_s^{(n-1)}F)\\ &\quad+\int_{s\le u_1\le\ldots\le u_n\le T}\int_{\mathbb{R}_0^n}\omega^t\circ(D_{u_n,x_n}\ldots D_{u_1,x_1}\circ E_{u_n}F)\ (x\nu(\mathrm{d}x))^{\otimes n}\,(\mathrm{d}u)^{\otimes n}.\end{aligned}$$

Thus according to assumption (3.14):

$$\begin{aligned}&E\left[\left(\omega^t\circ\left((E_s-T_s^{(n-1)})F\right)\right)^2\right]\\ =\ & E\left[\left(\int_{s\le u_1\le\ldots\le u_n\le T}\int_{\mathbb{R}_0^n}\omega^t\circ(D_{u_n,x_n}\ldots D_{u_1,x_1}\circ E_{u_n}F)\ (x\nu(\mathrm{d}x))^{\otimes n}\,(\mathrm{d}u)^{\otimes n}\right)^2\right]\\ \le\ & \frac{(T-s)^{2n}}{(2^n n!)^2}E\left[\left(\sup_{u_1,\ldots,u_n\in[0,T]}\left|\int_{\mathbb{R}_0^n}\omega^t\circ(D_{u_n,x_n}\ldots D_{u_1,x_1}F)\,(x\nu(\mathrm{d}x))^{\otimes n}\right|\right)^2\right]\\ &\xrightarrow[n\to\infty]{} 0.\end{aligned}$$

We now take $s=t$ and obtain:

$$\begin{aligned}E[F|\mathcal{F}_t] &= E_tF=\omega^t\circ(T_t^{(\infty)}F)\\ &= \sum_{n=0}^{\infty}\int_{t\le u_1\le\ldots\le u_n\le T}\int_{\mathbb{R}_0^n}\omega^t\circ(D_{u_n,x_n}\ldots D_{u_1,x_1}F)\ (x\nu(\mathrm{d}x))^{\otimes n}\,(\mathrm{d}u)^{\otimes n}.\end{aligned}$$

□

Acknowledgements

This work has been partially developed during two short visits of professor Henry Schellhorn at the Institute of Mathematics of the University of Barcelona (IMUB) in Summer 2014 and Summer 2015. The authors thank IMUB for the support and the kind hospitality.

References

[1] E. Alòs, J. A. León, and J. Vives (2008): *An anticipating Itô formula for Lévy processes.* ALEA Latin American Journal of Probability and Mathematical Statistics 4: 285-305.

[2] E.Alòs, J. León, M. Pontier and J. Vives (2008): *A Hull and White formula for a general stochastic volatility jump-diffusion model with applications to the study of the short-time behavior of the implied volatility.* Journal of Applied Mathematics and Stochastic Analysis.

[3] K. Aase, B. Øksendal, N. Privault and J. Uboe (2000): *White noise generalizations of the Clark-Haussmann-Ocone theorem, with application to mathematical finance.*

[4] F. E. Benth, G. Di Nunno, A. Løkka, B. Øksendal, and F. Proske (2003): *Explicit representations of the minimal variance portfolio in markets driven by Lévy processes*. Mathematical Finance 13: 55-72.

[5] L. Chen, D. Filipović, and H. Poor (2004): *Quadratic term structure models for risk-free and defaultable rates*. Math. Finance 14 no. 4, 515–536.

[6] R. Cont and P. Tankov (2003): *Financial Modelling with Jump Processes*. Chapman-Hall / CRC.

[7] G. Di Nunno, B. Øksendal, and F. Proske (2009): *Malliavin calculus for Lévy processes and Applications to Finance*. Springer.

[8] J. Hunter, and B. Nachtergaele (2001): *Applied Analysis*. World Scientific.

[9] Y. Ishikawa (2013): *Stochastic calculus of variations for jump processes*. Walter de Gruyter.

[10] K. Itô (1956): *Spectral type of shift transformations of differential processes with stationary increments*. Transactions of the American Mathematical Society 81: 252-263.

[11] S. Jin, Q. Peng and H. Schellhorn (2016): *A representation theorem for smooth Brownian martingales*. Stochastics. Vol. 88, No. 5, 651-679.

[12] S. Jin, Q. Peng and H. Schellhorn (2015): Fractional Hida-Malliavin Derivatives and Series Representations of Fractional Conditional Expectations. Communications on Stochastic Analysis. Vol. 9, No. 2, 213-238.

[13] A. Løkka (2004): *Martingale representation of functionals of Lévy processes.* Stochastic Analysis and Applications 22 (4): 867-892.

[14] J. Neveu (1976): *Processus Ponctuels*. Lecture Notes in Mathematics 598. Springer.

[15] I.Nourdin and D.Nualart (2010): *Central Limit Theorem for Multiple Skorohod Integrals*. Journal of Theoretical Probability 23: 39-64.

[16] D. Nualart (2006): *The Malliavin Calculus and Related Topics*. Second edition. Springer.

[17] D. Nualart and J. Vives (1990): *Anticipative calculus for the Poisson process based on the Fock space*. Séminaire des Probabilités XXIV. Lectures Notes in Mathematics 1426: 154-165.

[18] D. Nualart and J. Vives (1995): *A duality formula on the Poisson space and some applications*. Proceedings of the Ascona Conference on Stochastic Analysis. Progress in Probability. Birkhauser.

[19] E. Petrou (2008): *Malliavin calculus in Lévy spaces and Applications in Finance*. Electronic Journal of Probability 13: 852-879

[20] J. Picard (1996): *Formules de dualité sur l'espace de Poisson*. Annales de l'IHP, section B, 32 (4): 509-548.

[21] J. Picard (1996): *On the existence of smooth densities for jump processes*. Probability Theory and Related Fields 105: 481-511.

[22] N. Privault (2009): Stochastic Analysis in Discrete and Continuous Settings. Springer.

[23] S. Shreve (2004): *Stochastic Calculus for Finance II: Continuous-Time Models.* Springer.

[24] J. L. Solé, F. Utzet, and J. Vives (2007): *Chaos expansions and Malliavin calculus for Lévy processes.* Proceedings of the Abel Symposium 2005: 595-612. Springer.

[25] J. L. Solé, F. Utzet, and J. Vives (2007): *Canonical Lévy processes and Malliavin calculus.* Stochastic Processes and their Applicacions 117: 165-187.

[26] K. I. Sato (1999): *Lévy processes and Infinitely Divisible Distributions.* Cambridge.

[27] J. Vives (2013): *Malliavin calculus for Lévy processes: a survey.* Proceedings of the 8th Conference of the ISAAC-2011. Rendiconti del Seminario Matematico, Università e Politecnico di Torino 71 (2): 261-272.

[28] A. Yablonski (2008): *The calculus of variations for processes with independent increments.* Rocky Mountain Journal of Mathematics 38 (2): 669-701.

[29] E. Zeidler (2006): *Quantum Field Theory 1: Basics in Mathematics and Physics*. Springer.