Dyson type formula for pure jump Lévy processes with some applications to finance

Sixian Jin[a,1], Henry Schellhorn[a,2], Josep Vives[b,3]

[a] *Institute of Mathematical Sciences, Claremont Graduate University, 710 N. College Ave., Claremont, CA 91711, USA*
[b] *Faculty of Mathematics, Universitat de Barcelona, Gran Via 585, 08007 Barcelona (Catalunya), Spain*

Abstract

In this paper we obtain a Dyson type formula for integrable functionals of a pure jump Lévy process. We represent the conditional expectation of a $\mathcal{F}_T$-measurable random variable F at a time $t \leq T$ as an exponential formula involving 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 the following examples: the first one is the pricing of options in the Poisson-Black-Scholes model; the second one is the pricing of discount bonds in the Lévy quadratic model. We also obtain, for the conditional expectation of a function of a finite number of the process values, a backward Taylor expansion, that turns out to be useful for numerical calculations.

Keywords: Lévy processes, Malliavin calculus, Clark-Ocone formula, Dyson type formula, Backward Taylor expansion.

2010 MSC: 60H07, 91G80

1. Introduction

In [5], a representation theorem for smooth Brownian martingales was obtained. It consists of a series reminiscent of the Dyson series in quantum mechanics. A similar representation was obtained in [6] for functionals of the fractional Brownian motion for $H > 1/2$. In both cases, the representation involves the Malliavin derivative. Here we obtain an analogous formula for functionals of pure jump Lévy processes. Our work is based on Malliavin-Skorohod calculus

[1] E-mail address: sixian.jin@cgu.edu.
[2] Corresponding author. E-mail address: henry.schellhorn@cgu.edu.
[3] E-mail address: josep.vives@ub.edu. Supported by Spanish grants MTM 2015-65092-P-MINECO (2016-2018) and MEC MTM 2016-76420-P (2017-2020).

techniques for Lévy processes. As general references for this calculus we refer the reader to books [3] and [9].

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 Malliavin derivative 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 [8] or [15] for a discussion about the role of the Fock space structure in Malliavin-Skorohod calculus. Also, the method of proof is significantly different.

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 to the risk neutral probability, the obtained representation formula is potentially useful in pricing and hedging, as two examples in the paper show. One of them, the quadratic Lévy model, is, as far as we know, an original result.

The paper is organized as follows. Section 2 is devoted to preliminaries about Malliavin-Skorohod type calculus for Lévy processes and Poisson random measures. For a given conditional expectation of a certain functional, in Section 3 we prove the Dyson type formula and in Section 4 we prove its backward Taylor expansion, which is of independent interest for numerical calculations, see [5]. Finally, in Section 5, two examples of financial application are analyzed in detail.

2. Lévy processes and Malliavin-Skorohod calculus

2.1. Lévy processes and Poisson random measures

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 $\mathbb{P}$ and by $L^2(\Omega)$ the space of the square integrable random variables. 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 [14] 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 on $\mathbb{R}$, that is, a $\sigma-$finite positive measure, null at the origin and such that $\int_{\mathbb{R}_0}(1 \wedge x^2)\nu(\mathrm{d}x) < \infty$.

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 $[0,T]$.

For simplicity, throughout the paper, we denote by $(m(\mathrm{d}s, \mathrm{d}x))^{\otimes n}$ the tensor product of a non-random measure m, that is,

$$(m(\mathrm{d}s, \mathrm{d}x))^{\otimes n} := m(\mathrm{d}s_n, \mathrm{d}x_n) \cdots m(\mathrm{d}s_1, \mathrm{d}x_1).$$

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

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

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

Recall that J is a Poisson random measure on $\mathcal{G}$ with intensity $\ell \otimes \nu$. Let $\mathcal{G}^*$ be the family of Borel sets G such that $(\ell \otimes \nu)(G) < \infty$. Then, for any $G \in \mathcal{G}^*$, $J(G)$ is a Poisson random variable with $E[J(G)] = (\ell \otimes \nu)(G)$. As usual, from now on, we write $\ell(\mathrm{d}t) = \mathrm{d}t$. We can consider the compensated random measure $\widetilde{J}(\mathrm{d}t, \mathrm{d}x) := J(\mathrm{d}t, \mathrm{d}x) - \mathrm{d}t\nu(\mathrm{d}x)$, that is a square integrable centered random measure such that for any G_1 and G_2, subsets of $\mathcal{G}^*$, we have

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

Taking into account that any Lévy process can be decomposed into two independent parts, the Brownian one and the so-called pure jump one, and since the Brownian case was treated in [5], we consider in this paper Lévy processes of the type

$$X_t := \gamma t + \int_0^t \int_{|x|>1} xJ(\mathrm{d}s, \mathrm{d}x) + \int_0^t \int_{|x|\le 1} x\widetilde{J}(\mathrm{d}s, \mathrm{d}x).$$

Recall that the selection of sets $\{|x| > 1\}$ and $\{|x| \le 1\}$ is arbitrary and the border at 1 can be changed by any fixed $\delta > 0$. If for a certain fixed $\delta > 0$,

$$c(\delta) := \int_{|x|\le\delta} x\nu(dx) \tag{2.1}$$

is well defined, we can write

$$\begin{aligned} X_t &= \gamma t + \int_0^t \int_{\mathbb{R}_0} xJ(\mathrm{d}s, \mathrm{d}x) - \int_0^t \int_{|x|\le\delta} x\nu(\mathrm{d}x)\mathrm{d}s \\ &= \int_0^t \int_{\mathbb{R}_0} xJ(\mathrm{d}s, \mathrm{d}x) + (\gamma - c(\delta))t. \end{aligned}$$

In this case, the Lévy process X is of finite variation and we can restrict our analysis to the case

$$X_t = \int_0^t \int_{\mathbb{R}_0} xJ(\mathrm{d}s, \mathrm{d}x),$$

that, for any t, is a.s. a well defined random variable.

If ν has finite expectation, that is, $\int_{\mathbb{R}_0} |x|\nu(dx) < \infty$, condition (2.1) is satisfied and the variable X_t belongs to $L^1(\Omega)$ for any t.

Another relevant particular case is the case when ν is a finite measure, that is $\lambda := \nu(\mathbb{R}_0) < \infty$. In this case condition (2.1) is also satisfied and we have a so-called Compound Poisson process. It is well known that in this case we can associate to X the process

$$N_t := \int_0^t \int_{\mathbb{R}_0} J(ds, dx)$$

that is a standard Poisson process with intensity λ and rewrite $X_t = \sum_{j=0}^{N_t} Y_j$ where Y_j are independent and identically distributed random variables with probability law $Q := \nu/\lambda$. See for example [2].

Two particular cases of Compound Poisson processes are of particular interest, the standard Poisson process that corresponds with the case $Q = \delta_1$, the Dirac-delta at 1, and the so-called simple Lévy processes that correspond to the case Q is concentrated in a finite set of values. In this last case, process X can be rewritten as a sum of different standard Poisson processes with different intensities, see for example [12].

2.2. Chaotic representation property and Malliavin-Skorohod operators

Following [3], we can consider the spaces

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

and define the Itô multiple stochastic integrals $I_n(f)$ with respect to $\tilde{J}$ in the usual way. Then we have the so-called chaos representation property, that is, for any functional $F \in L^2(\Omega)$ we have

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

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

This chaos representation shows that $L^2(\Omega)$ has a Fock space structure. Thus it is possible to apply the machinery related with the annihilation operator (Malliavin derivative) and the creation operator (Skorohod integral) as it is exposed, for example, in [8] and [15].

Define $\mathbb{D} := \{F : F \in L^2(\Omega) \text{ with } \sum_{n=1}^{\infty} n n! ||f_n||^2_{\mathbb{L}_n^2} < \infty\}$. The Malliavin derivative of $F \in \mathbb{D}$ is an element of $L^2([0,T] \times \mathbb{R}_0 \times \Omega, \mathcal{G} \otimes \mathcal{F}, \ell \otimes \nu \otimes \mathbb{P})$, defined as

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

Of course the Malliavin derivative can be iterated in the usual way, written as $D^n_{t_1,x_1,\ldots,t_n,x_n} F$. The domain of the nth iterated operator is denoted by $\mathbb{D}^n$.

On the other hand, let $u \in L^2([0,T] \times \mathbb{R}_0 \times \Omega, \mathcal{G} \otimes \mathcal{F}, \ell \otimes \nu \otimes \mathbb{P})$. 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. Denote by $\tilde{f}_n$ the symmetrization in all $n+1$ pairs of variables. Then, we define the Skorohod integral of u by

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

in $L^2(\Omega)$, provided $u \in \operatorname{Dom}\delta$ that means $\sum_{n=0}^{\infty}(n+1)!\,\|\tilde{f}_n\|^2_{\mathbb{L}^2_{n+1}} < \infty$. This integral turns to be an extension of Itô type and pathwise integrals in the sense that for a predictable process $u \in L^2([0,T]\times\mathbb{R}_0\times\Omega)$, we have

$$\delta(u) = \int_{[0,T]\times\mathbb{R}_0} u(s,x)\widetilde{J}(\mathrm{d}s,\mathrm{d}x).$$

Moreover, according to Proposition 5.4 in [16], we have the following integration by parts formula:

Lemma 2.1. *Assume $u \in \operatorname{Dom}\delta$ and $F \in \mathbb{D}$ such that*

$$E\left[\int_{[0,T]\times\mathbb{R}_0} u_{t,x}^2(F^2 + (D_{t,x}F)^2)\,\mathrm{d}t\nu(\mathrm{d}x)\right] < \infty.$$

Then, the following relation holds:

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

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

We define the space $\mathbb{L}^{1,2}$ as the class of processes of Dom δ such that $u(t,x) \in \mathbb{D}$ for almost all (t,x), satisfying

$$E\left[\int_{[0,T]^2\times\mathbb{R}_0^2} (D_{t,x}u(s,y))^2\mathrm{d}t\nu(\mathrm{d}x)\mathrm{d}s\nu(\mathrm{d}y)\right] < \infty.$$

Then, we have the following lemma.

Lemma 2.2. *If both u and $v \in \mathbb{L}^{1,2} \subset \operatorname{Dom}\delta$, we have*

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

2.3. Clark-Hausmann-Ocone formula

Finally, 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 $\{\widetilde{J}(A') : A' \in \mathcal{G}^*, A' \subseteq A\}$. We have, see [8], that F is $\mathcal{F}_A-$measurable if for any $n \geq 1$, $f_n(t_1,x_1,\dots,t_n,x_n) = 0$, $(\ell\otimes\nu)^{\otimes n}$-a.e. unless $(t_i,x_i) \in A$ for all $i = 1,\dots,n$.

In particular, we are interested in the case $A := [0,t)\times\mathbb{R}_0$, let us call the corresponding $\sigma-$algebra as $\mathcal{F}_{t-}$. Obviously, if $F \in \mathbb{D}$ and it is $\mathcal{F}_{t-}-$measurable then $D_{s,x}F = 0$ a.e. for $s \geq t$ and any $x \in \mathbb{R}_0$.

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

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

(see e.g. [8]). 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 [3] section 12.6, the CHO formula. It says that if $F \in \mathbb{D}$ we have

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

where $E[D_{t,x} F|\mathcal{F}_{t-}]$ is the predictable projection of $D_{t,x} F$.

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

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

2.4. 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 a Poisson random measure. It is well known, see for example [13], that this space 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, we define, see [4] or [13], the operator $\Psi_{t,x}$ by

$$\Psi_{t,x} F = F(\omega + e_{t,x}) - F(\omega)$$

where $\omega + e_{t,x}$ denotes to add to ω a jump in time t with amplitude x. This operator is linear, closed and well defined from $L^0(\Omega)$ to $L^0(\Omega \times [0, T] \times \mathbb{R}_0)$. We say that it is a probabilistic interpretation of operator D in the sense that $F \in \mathbb{D}$ if and only if $\Psi F \in L^2([0, T] \times \mathbb{R}_0 \times \Omega)$ and in this case

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

From now on we will write with a certain abuse of notation D instead of Ψ, that is, we will use D as an operator defined on $L^0(\Omega)$ taking values in $L^0(\Omega \times [0, T] \times \mathbb{R}_0)$.

It is immediate to obtain the chain rule: for any t and x and any measurable function f,

$$D_{t,x} f(G) = f(G + D_{t,x} G) - f(G). \tag{2.2}$$

Moreover, by iteration, we have the following useful formula

$$D_{s_1,x_1} \cdots D_{s_n,x_n} F(\omega) = \sum_{k=0}^{n} \sum_{\{j_1 \le \cdots \le j_k\} \subset \{1,\ldots,n\}} (-1)^{n-k} F(\omega + e_{s_{j_1},x_{j_1}} + \ldots + e_{s_{j_k},x_{j_k}}). \tag{2.3}$$

2.5. Particular cases: standard Poisson process and simple Lévy process

In the particular case of the standard Poisson process $\{N_t, t \geq 0\}$ with intensity λ, we have the compensated Poisson process $\{\tilde{N}_t := N_t - \lambda t, t \geq 0\}$. In this case, iterated stochastic integrals are written as

$$I_n(f_n) = n! \int_t^T \int_t^{s_n -} \cdots \int_t^{s_2 -} f_n(s_1, ..., s_n) \mathrm{d}\tilde{N}_{s_1} \cdots \mathrm{d}\tilde{N}_{s_n}$$

and the Malliavin derivative satisfies: $D_t f(G) = f(G + D_t G) - f(G)$.

Accordingly, the CHO formula is

$$F = E[F] + \int_0^T E[D_t F | \mathcal{F}_{t-}] \mathrm{d}\tilde{N}_t. \tag{2.4}$$

for any $F \in \mathbb{D}$ and $\mathcal{F}_T-$measurable.

The simple Lévy process corresponds to the case where Q is concentrated in a finite number of jump sizes $\{z_1, ..., z_J\}$ and $Y_t = \sum_{k=1}^{N_t} Z_k$ where $\{Z_k\}_{k \geq 0}$ is a sequence of independent and identically distributed discrete random variables taking values in $\{z_1, ..., z_J\}$. The values of Z represent the jump size at each jump 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 $\left\{N^j, 1 \leq j \leq J\right\}$ are independent standard Poisson processes with intensities $\{\lambda_j = \lambda p(z_j), 1 \leq j \leq J\}$.

Let $\tilde{Y}_t$ denote the compensated process. Then

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

Thus, the Itô integral for $\tilde{Y}_t$ can be defined as $\int_0^T X_t \mathrm{d}\tilde{Y}_t = \sum_{j=1}^{J} \int_0^T z_j X_t \mathrm{d}\tilde{N}_t^j$ and the CHO formula becomes:

$$F = E[F] + \int_0^T \sum_{j=1}^{J} E[D_s^{(j)} F | \mathcal{F}_{s-}] \mathrm{d}\tilde{N}_s^j \tag{2.5}$$

where $D_s^{(j)} F := D_{s, z_j} F$.

3. A Dyson type formula

We introduce first of all the so-called *freezing path operator*.

Definition 3.1. *Given a functional $F \in L^0(\Omega)$ and given $\omega = ((s_1, x_1), (s_2, x_2), \dots)$, we define the freezing path operator ω^t as the operator*

$$(\omega^t \circ F)(\omega) := F(\omega^t(\omega))$$

where $\omega^t(\omega)$ is the sequence of pairs (s_i, x_i) of ω such that $s_i \leq t$.

Remark 3.1. *Assume the finite variation case. Consider*

$$X_t := \int_{[0,t]\times\mathbb{R}_0} xJ(\mathrm{d}s, \mathrm{d}x).$$

Based on the definition, we have the following properties of the freezing path operator:

1. $\omega^t \circ X_u = X_{u\wedge t}$.
2. *For any measurable function* $g(x_1, ..., x_n)$*, we have*

$$\omega^t \circ g(X_{t_1}, ..., X_{t_n}) = g(X_{t_1\wedge t}, ..., X_{t_n\wedge t}).$$

3. *Given* $t \in [0, T]$ *and* $f \in L^1([0, T] \times \mathbb{R}_0)$*, we have*

$$\begin{aligned} & \omega^t \circ \int_{[0,T]\times\mathbb{R}_0} f(s, x)\tilde{J}(\mathrm{d}s, \mathrm{d}x) \\ = & \int_{[0,t]\times\mathbb{R}_0} f(s, x)J(\mathrm{d}s, \mathrm{d}x) - \int_{[0,T]\times\mathbb{R}_0} f(s, x)\nu(\mathrm{d}x)\mathrm{d}s. \end{aligned}$$

Now we can establish the main theorems of the paper, the so-called Dyson type formula, for the Compound Poisson case in Theorem 3.1 and for the general pure jump Lévy case in Theorem 3.2.

Theorem 3.1. *Let* $F \in L^1(\Omega)$*. Assume* $\nu(\mathbb{R}_0) = \lambda < \infty$*. Then, a.s.,*

$$E[F|\mathcal{F}_t] = \omega^t \circ F + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{[t,T]^n\times\mathbb{R}_0^n} \omega^t \circ (D_{s_1,x_1} \cdots D_{s_n,x_n} F)(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}. \quad (3.1)$$

Remark 3.2. *Note that the previous equality can be naturally written as*

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

and note also that the iteration of operator $D_{s,x}$ *is symmetric in* $(s_1, x_1), ..., (s_n, x_n)$*.*

Proof. If F is integrable and thanks to Jensen's inequality, the left hand side of the equality (3.1) is, for any $t \in [0, T]$, a well defined random variable that belongs to $L^1(\Omega)$.

If $\nu(\mathbb{R}_0) = \lambda < \infty$, the underlying pure jump Lévy process X associated to ν is a compound Poisson process, that is, there exists a standard Poisson process N with intensity λ that determines the jump instants and $\nu = \lambda Q$ where Q is the probability law of the jump amplitudes.

First of all, we show that the right hand side element of (3.1) is bounded by $e^{2\lambda(T-t)}E[|F||\mathcal{F}_t]$, so it is a random variable in $L^1(\Omega)$. It is enough to show that

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{[t,T]^n\times\mathbb{R}_0^n} |\omega^t \circ (D_{s_1,x_1} \cdots D_{s_n,x_n} F)|(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n} \leq e^{2\lambda(T-t)}E[|F||\mathcal{F}_t].$$

Using formula (2.3) and the fact that the underlying process is a compound Poisson process, the left hand side term in the previous inequality is equivalent to

$$\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\sum_{\{j_1\leq\cdots\leq j_k\}\subset\{1,\ldots,n\}}\int_{[t,T]^n\times\mathbb{R}_0^n}|F(\omega^t(\omega+e_{s_{j_1},x_{j_1}}+...+e_{s_{j_k},x_{j_k}}))|(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}.$$

Note that integrals

$$\int_{[t,T]^n\times\mathbb{R}_0^n}|F(\omega^t(\omega+e_{s_{j_1},x_{j_1}}+...+e_{s_{j_k},x_{j_k}}))|(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}$$

are the same for any selection of $j_1,\ldots,j_k$, fixed k. So the previous series is equal to

$$\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\int_{[t,T]^n\times\mathbb{R}_0^n}|F(\omega^t(\omega+e_{s_1,x_1}+...+e_{s_k,x_k}))|(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}.$$

Interchanging sums and computing the $n-k$ integrals not affected by s and x, we have

$$\sum_{k=0}^{\infty}\frac{1}{k!}\sum_{n=k}^{\infty}\frac{\lambda^{n-k}(T-t)^{n-k}}{(n-k)!}\int_{[t,T]^k\times\mathbb{R}_0^k}|F(\omega^t(\omega+e_{s_1,x_1}+...+e_{s_k,x_k}))|(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes k}$$

that is equal to

$$\begin{aligned}
& e^{\lambda(T-t)}\sum_{k=0}^{\infty}\frac{1}{k!}\int_{[t,T]^k\times\mathbb{R}_0^k}|F(\omega^t(\omega+e_{s_1,x_1}+\cdots+e_{s_k,x_k}))|(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes k}\\
= \; & e^{\lambda(T-t)}\sum_{k=0}^{\infty}\frac{\lambda^k(T-t)^k}{k!}\int_{[t,T]^k\times\mathbb{R}_0^k}|F(\omega^t(\omega+e_{s_1,x_1}+\cdots+e_{s_k,x_k}))|\left(\frac{\nu(\mathrm{d}x)\mathrm{d}s}{\lambda(T-t)}\right)^{\otimes k}\\
= \; & e^{2\lambda(T-t)}\sum_{k=0}^{\infty}P(N_T-N_t=k)E\left[E[|F||N_T-N_t=k]\Big|\mathcal{F}_t\right]\\
= \; & e^{2\lambda(T-t)}E[|F||\mathcal{F}_t].
\end{aligned}$$

Doing similar computations we prove the equality:

$$
\begin{aligned}
& \sum_{n=0}^{\infty} \frac{1}{n!} \int_{[t,T]^n \times \mathbb{R}_0^n} \omega^t \circ (D_{s_1,x_1} \cdots D_{s_n,x_n} F)(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}) \\
= & \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \sum_{\{j_1 \le \cdots \le j_k\} \subset \{1,\ldots,n\}} (-1)^{n-k} \\
& \times \int_{[t,T]^n \times \mathbb{R}_0^n} F(\omega^t(\omega + e_{s_{j_1},x_{j_1}} + \ldots + e_{s_{j_k},x_{j_k}}))(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n} \\
= & \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} \\
& \times \int_{[t,T]^n \times \mathbb{R}_0^n} F(\omega^t(\omega + e_{s_1,x_1} + \ldots + e_{s_k,x_k}))(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n} \\
= & \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n=k}^{\infty} \frac{(-1)^{n-k} \lambda^{n-k} (T-t)^{n-k}}{(n-k)!} \\
& \times \int_{[t,T]^k \times \mathbb{R}_0^k} F(\omega^t(\omega + e_{s_1,x_1} + \ldots + e_{s_k,x_k}))(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes k} \\
= & e^{-\lambda(T-t)} \sum_{k=0}^{\infty} \frac{1}{k!} \int_{[t,T]^k \times \mathbb{R}_0^k} F(\omega^t(\omega + e_{s_1,x_1} + \ldots + e_{s_k,x_k}))(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes k} \\
= & e^{-\lambda(T-t)} \sum_{k=0}^{\infty} \frac{\lambda^k (T-t)^k}{k!} \int_{[t,T]^k \times \mathbb{R}_0^k} F(\omega^t(\omega + e_{s_1,x_1} + \ldots + e_{s_k,x_k})) \left(\frac{\nu(\mathrm{d}x)\mathrm{d}s}{\lambda(T-t)} \right)^{\otimes k} \\
= & \sum_{k=0}^{\infty} P(N_T - N_t = k) E\left[E[F | N_T - N_t = k] \Big| \mathcal{F}_t \right] \\
= & E[F | \mathcal{F}_t].
\end{aligned}
$$

□

The main idea of this proof is to use the binomial recombination. Similar idea for related combinatorics can be found in the proofs of Proposition 4 and 5 of [10].

Theorem 3.2. *Assume $F \in L^1(\Omega)$, and*

$$
\omega^t \circ F + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{[t,T]^n \times \mathbb{R}_0^n} |\omega^t \circ (D_{s_1,x_1} \cdots D_{s_n,x_n} F)| (\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n} < \infty, \ a.s..
$$

Then, a.s.,

$$
E[F|\mathcal{F}_t] = \omega^t \circ F + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{[t,T]^n \times \mathbb{R}_0^n} \omega^t \circ (D_{s_1,x_1} \cdots D_{s_n,x_n} F)(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}.
$$

Proof. If $\Omega_{T,\epsilon}$ is the canonical Poisson space when ν is concentrated in $\{|x| > \epsilon\}$, from the finite activity case (Theorem 3.1), we have

$$E[F\mathbb{1}_{\Omega_{T,\epsilon}}|\mathcal{F}_t] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{[t,T]^n \times \{|x|>\epsilon\}^n} \omega^t \circ (D_{s_1,x_1} \cdots D_{s_n,x_n} F)(\nu(\mathrm{d}x)\mathrm{d}s)^{\otimes n}.$$

Using the hypothesis and dominated convergence, the equality is valid a.s. for a general ν. □

Remark 3.3. *Following Section 2.5, we directly obtain the Dyson type formula for Poisson process and simple Lévy process as corollaries of Theorem 3.1.*

In the standard Poisson case, if $F \in L^1(\Omega)$, we have

$$E[F|\mathcal{F}_t] = \omega^t \circ F + \sum_{n=1}^{\infty} \lambda^n \int_{t \le s_1 \le \ldots \le s_n \le T} \omega^t \circ (D_{s_n} \cdots D_{s_1} F)\, (\mathrm{d}s)^{\otimes n}. \quad (3.2)$$

In the simple Lévy process case, we define the operator $A_s := \sum_{j=1}^{J} \lambda_j D_s^{(j)}$. Then,

$$E[F|\mathcal{F}_t] = \omega^t \circ F + \sum_{n=1}^{\infty} \int_{t \le s_1 \le \ldots \le s_n \le T} \omega^t \circ (A_{s_n} \cdots A_{s_1} F)\, (\mathrm{d}s)^{\otimes n}.$$

4. Backward Taylor expansion

In this section, we provide the backward Taylor expansion (BTE) for a function of discrete pure jump Lévy processes. As was showed in [5] for the Brownian case, the BTE is useful for numerical applications. We provide an example of an application in the next section. One can also use, as we did in [5], the BTE to prove the Dyson type formula by an approximation argument. However, the conditions for convergence of the Dyson series that we give in Theorem 3.1 are simpler than what we obtain following the BTE approach, so we do not show this method of proof here.

We first consider the case of the Poisson process. The case of the compound Poisson process will be a corollary. Recall that the Charlier polynomials $C_n(x, y)$ can be defined by the formula $(1+z)^x e^{-yz} = \sum_{n=0}^{\infty} \frac{z^n}{n!} C_n(x, y)$ or by recurrence as $C_0(x, y) = 1$, $C_1(x, y) = x - y$ and

$$C_{n+1}(x, y) = (x - y - n)C_n(x, y) - nyC_{n-1}(x, y).$$

A well-known relation between Charlier polynomials and multiple stochastic integrals (see [16] section 3) is given by the formula

$$C_n(N_T - N_t, \lambda(T - t)) = n! \int_t^T \int_t^{s_n-} \cdots \int_t^{s_2-} \mathrm{d}\tilde{N}_{s_1} \cdots \mathrm{d}\tilde{N}_{s_n}.$$

By the integration by parts formula given in Lemma 2.1 and induction, we can prove the following lemma, which is the counterpart of Lemma 5.3 in [5].

Lemma 4.1. *Let $F = f(\tilde{N}_T) \in L^2(\Omega)$ be $\mathcal{F}_T$-measurable and $\sum_{i=0}^{\infty} E\left[\left(D_T^i F\right)^2\right] < \infty$. Then for $n \geq 1$,*

$$\int_t^T \int_t^{s_n-} \cdots \int_t^{s_2-} F\delta\tilde{N}_{s_1} \cdots \delta\tilde{N}_{s_n} = \sum_{j=0}^{n} \sum_{i_1,\ldots,i_n=0}^{\infty} (-1)^{i_1+\ldots+i_n+j} D_T^{i_1+\ldots+i_n+j} F \times C_{n-j}\left(N_T - N_t, \lambda(T-t)\right) \frac{(\lambda(T-t))^j}{(n-j)!j!}.$$

Proof. As a consequence of Lemma 2.1 (see also Proposition 6.5.1 in [9]), the integration by parts formula for the Skorohod integration with respect to the compensated Poisson process is:

$$\int_0^T F\delta\tilde{N}_s = F\int_0^T \mathrm{d}\tilde{N}_s - \lambda\int_0^T D_s F \mathrm{d}s - \int_0^T D_s F\delta\tilde{N}_s. \tag{4.1}$$

Since F depends only on $\tilde{N}_T$, we have $D_s F = D_T F$ when $s \leq T$. Therefore, we apply (4.1) repeatedly to obtain

$$\begin{aligned}
\int_0^T F\delta\tilde{N}_s &= F\int_0^T \mathrm{d}\tilde{N}_s - \lambda D_T F \int_0^T \mathrm{d}s \\
&\quad - \left(D_T F \int_0^T \mathrm{d}\tilde{N}_s - \lambda D_T^2 F \int_0^T \mathrm{d}s - \int_0^T D_T^2 F\delta\tilde{N}_s\right) \\
&= G\int_0^T \mathrm{d}\tilde{N}_s - \lambda D_T G \int_0^T \mathrm{d}s
\end{aligned} \tag{4.2}$$

where $G := \sum_{i=0}^{\infty} (-1)^i D_T^i F$.

In general, repeatedly applying (4.2), by induction:

$$\begin{aligned}
&\int_t^T \int_t^{s_n-} \cdots \int_t^{s_2-} F\delta\tilde{N}_{s_1} \cdots \delta\tilde{N}_{s_n} \\
&= \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \ldots \sum_{i_n=0}^{\infty} \Big\{ (-1)^{i_1+..+i_n} D_T^{i_1+..+i_n} F \int_t^T \int_t^{s_n-} \cdots \int_t^{s_2-} \mathrm{d}\tilde{N}_{s_1} \cdots \mathrm{d}\tilde{N}_{s_n} \\
&\quad + \lambda(-1)^{i_1+..+i_n+1} D_T^{i_1+..+i_n+1} F \int_t^T \int_t^{s_{n-1}-} \cdots \int_t^{s_2-} \mathrm{d}\tilde{N}_{s_1} \cdots \mathrm{d}\tilde{N}_{s_{n-1}} \int_t^T ds_1 \\
&\quad + \ldots + \lambda^n (-1)^{i_1+..+i_n+n} D_T^{i_1+..+i_n+n} F \int_{t \leq s_1 \leq \ldots \leq s_n \leq T} (\mathrm{d}s)^{\otimes n} \Big\} \\
&= \sum_{j=0}^{n} \sum_{i_1,\ldots,i_n=0}^{\infty} (-1)^{i_1+\ldots+i_n+j} D_T^{i_1+..+i_n+j} F C_{n-j}(N_T - N_t, \lambda(T-t)) \frac{\lambda^j (T-t)^j}{(n-j)!j!}.
\end{aligned}$$

□

Note that any Lévy process can be assumed right continuous with left limits without losing generality. Being T is a prefixed point, the probability of a jump on T is null. Therefore N_T is almost surely equal to N_{T-} and of course these two random variables are equal as elements of $L^2(\Omega)$. Similar argument holds also in the following proof of Theorem 4.1.

According to Lemma 2.2, Proposition 5.4 and 5.7 in [16], and (2.12) in [7], we have the following lemma. It is the counterpart of Lemma 5.1 in [5].

Lemma 4.2. *For the multiple Skorohod integral δ^L with respect to the compensated Poisson process, if $u \in \mathbb{L}^{L,2} \subset \mathrm{Dom}\delta^L$ we have:*

$$E\left[\left(\delta^L(u)\right)^2\right] = \sum_{i=0}^{L} \binom{L}{i}^2 i! E\left[\left\|D^{L-i}u\right\|^2_{H^{\otimes(2L-i)}}\right],$$

where

$$\left\|D^{L-i}u\right\|^2_{H^{\otimes(2L-i)}} := \int_{[0,T]^{2L}} D_{t_1}\cdots D_{t_{L-i}}u(s_1,...,s_L)(\mathrm{d}t)^{\otimes(L-i)}$$
$$\times D_{s_1}\cdots D_{s_{L-i}}u(t_1,...,t_L)(\mathrm{d}s)^{\otimes(L-i)}\mathrm{d}t_{L-i+1}\cdots \mathrm{d}t_L \mathrm{d}s_{L-i+1}\cdots \mathrm{d}s_L.$$

Recall that the space $\mathbb{L}^{L,2}$ is the generalization of $\mathbb{L}^{1,2}$ changing $\mathbb{D}$ by $\mathbb{D}^L$ in the definition, see Lemma 2.2 above (or Definition 5.5 in [16]).

Based on these two lemmas, we can prove the Backward Taylor Expansion. For simplicity, we introduce a notation $p(m,n)$ to denote the number of partitions of a positive integer n with exactly m parts, allowing 0.

Theorem 4.1. *Suppose $F \in L^2(\Omega)$ and $0 \le t_1 \le ... \le t_M$.*

1. Poisson process: Assume $F = F(\tilde{N}_{t_1},...,\tilde{N}_{t_M})$ such that for any $m = 1,\dots,M-1$, $\sum_{i=0}^{\infty} E\left[\left(D^i_{t_m}F\right)^2\right] < \infty$ and

$$\sum_{i=0}^{L} E\left[\left(D^{2L-i}_{t_{m+1}}F\right)^2\right]\lambda^{2L-i}\binom{L}{i}^4 \frac{i!}{(L!)^2}(t_{m+1}-t_m)^{2L-i} \xrightarrow{L\to\infty} 0. \tag{4.3}$$

Then we have

$$E[F|\mathcal{F}_{t_m}] = \sum_{l=0}^{\infty} \gamma_l(m,\Delta N_{t_m})E[D^l_{t_{m+1}}F|\mathcal{F}_{t_{m+1}}] \tag{4.4}$$

where $\Delta N_{t_m} := N_{t_{m+1}} - N_{t_m}$ and the random coefficient $\gamma_l(m,\Delta N_{t_m})$ has the following representation for $l \ge 0$:

$$\gamma_l(m,\Delta N_{t_m}) = (-1)^l \sum_{\substack{j+n\le l\\ j\le n}} p(n,l-j-n)C_{n-j}(\Delta N_{t_m},\lambda(t_{m+1}-t_m))\frac{\lambda^j(t_{m+1}-t_m)^j}{(n-j)!j!}.$$

2. Simple Lévy process: Assume $F = F\left(\tilde{Y}_{t_1},...,\tilde{Y}_{t_M}\right)$ such that for any $m = 1,\dots,M-1$ and all j, $\sum_{i=0}^{\infty} E\left[\left(D^{i,(j)}_{t_m}F\right)^2\right] < \infty$; as well as

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

Then we have

$$E[F|\mathcal{F}_{t_m}] = \sum_{l_1,\dots,l_J=0}^{\infty} \gamma_{l_1,\dots,l_J}(m, \Delta N_{t_m}^1, \dots, \Delta N_{t_m}^J) E[D_{t_{m+1}}^{l_1,(1)} \cdots D_{t_{m+1}}^{l_J,(J)} F|\mathcal{F}_{t_{m+1}}]$$

where

$$\gamma_{l_1,\dots,l_J}(m, \Delta N_{t_m}^1, \dots, \Delta N_{t_m}^J) = (-1)^{l_1+\dots+l_J}$$
$$\prod_{j=1}^{J} \sum_{\substack{r_j+n_j \le l_j \\ r_j \le n_j}} \frac{p(n_j, l_j - n_j - r_j) C_{n_j - r_j}(\Delta N_{t_m}^j, \lambda_j(t_{m+1} - t_m)) \lambda_j^{r_j}(t_{m+1} - t_m)^{r_j}}{(n_j - r_j)! r_j!}.$$

Proof. The proof of this theorem is similar to the proof of Theorem 2.1 for the BTE of Brownian motion in [5]. First, for any $m \le M-1$, we apply the CHO formula (2.4) to obtain:

$$\begin{aligned} E[F|\mathcal{F}_{t_m}] &= E[F] + \int_0^{t_m} E[D_{s_1} F|\mathcal{F}_{s_1-}] \mathrm{d}\tilde{N}_{s_1} \\ &= E[F|\mathcal{F}_{t_{m+1}}] - \int_{t_m}^{t_{m+1}} E[D_{t_{m+1}} F|\mathcal{F}_{s_1-}] \mathrm{d}\tilde{N}_{s_1}. \end{aligned} \tag{4.5}$$

The equality $D_{s_1} F = D_{t_{m+1}} F$ in the previous line comes from the assumption that $F = F(\tilde{N}_{t_1}, ..., \tilde{N}_{t_M})$. Applying the CHO formula iteratively, we obtain:

$$\begin{aligned} E[F|\mathcal{F}_{t_m}] &= \sum_{l=0}^{L-1} (-1)^l \int_{t_m}^{t_{m+1}} \int_{t_m}^{s_l-} \cdots \int_{t_m}^{s_2-} E[D_{t_{m+1}}^l F|\mathcal{F}_{t_{m+1}-}] \delta\tilde{N}_{s_1} \cdots \delta\tilde{N}_{s_l} \\ &\quad + R_{[t_m, t_{m+1}]}^L \end{aligned} \tag{4.6}$$

where the remainder is defined as

$$R_{[t_m, t_{m+1}]}^L := \int_{t_m}^{t_{m+1}} \int_{t_m}^{s_L-} \cdots \int_{t_m}^{s_2-} E[D_{t_{m+1}}^L F|\mathcal{F}_{s_L-}] \mathrm{d}\tilde{N}_{s_1} \cdots \mathrm{d}\tilde{N}_{s_L}.$$

Applying Lemma 4.2 to the remainder $R_{[t_m, t_{m+1}]}^L$ and with the help of condition (4.3), similarly to the proof of Lemma 5.2 in [5], we obtain:

$$E\left[\left(R_{[t_m, t_{m+1}]}^L\right)^2\right] \le \sum_{i=0}^{L} E\left[\left(D_{t_{m+1}, x_1} \cdots D_{t_{m+1}, x_{2L-i}} F\right)^2\right] \binom{L}{i}^4 \frac{i!}{(L!)^2} (t_{m+1} - t_m)^{2L-i}$$

and this converges to 0, which guarantees that the series (4.4) converge. Now we define, for any $\varepsilon > 0$,

$$G(L, m, \varepsilon) := \int_{t_m}^{t_{m+1-\varepsilon}} \int_{t_m}^{s_L-} \cdots \int_{t_m}^{s_2-} E[D_{t_{m+1}}^L F|\mathcal{F}_{t_{m+1}-\varepsilon}] \delta\tilde{N}_{s_1} \cdots \delta\tilde{N}_{s_L}$$

where $G(L,m,\varepsilon)$ is a $\mathcal{F}_{t_{m+1}-\varepsilon}$ measurable random variable. Notice that $D_sF = D_{t_{m+1}}F$ if $s \in (t_m, t_{m+1}]$, by taking $\varepsilon < t_{m+1} - t_m$ and applying Lemma 4.1 to $G(L,m,\varepsilon)$, we obtain

$$G(L,m,\varepsilon) = \sum_{j=0}^{n} \sum_{i_1,\ldots,i_n=0}^{\infty} (-1)^{i_1+\ldots+i_n+j} E[D_{t_{m+1}}^{i_1+\ldots+i_n+j+L} F|\mathcal{F}_{t_{m+1}-\varepsilon}]$$
$$\times C_{n-j}\left(N_{t_{m+1}-\varepsilon} - N_{t_m}, \lambda(t_{m+1} - \varepsilon - t)\right) \frac{(\lambda(t_{m+1} - \varepsilon - t_m))^j}{(n-j)!j!}.$$

Let $\mathcal{T}(L,m,\varepsilon)$ be the set

$$\mathcal{T}(L,m,\varepsilon) = \{t_m \leq s_1 \leq \ldots \leq s_L \leq t_{m+1}\} - \{t_m \leq s_1 \leq \ldots \leq s_L \leq t_{m+1-\varepsilon}\}.$$

Then $E\left[\left(\int_{\mathcal{T}(L,m,\varepsilon)} E[D_{t_{m+1}}^L F|\mathcal{F}_{t_{m+1}-\varepsilon}]\delta\tilde{N}_{s_1}\cdots\delta\tilde{N}_{s_L}\right)^2\right] \xrightarrow{\varepsilon\to 0} 0.$

On the other hand, $E[D_{t_{m+1}}^{i_1+\ldots+i_n+j+L}F|\mathcal{F}_{t_{m+1}-\varepsilon}] \to E[D_{t_{m+1}}^{i_1+\ldots+i_n+j+L}F|\mathcal{F}_{t_{m+1}}]$ almost surely as $\varepsilon \to 0$ because $\tilde{N}$ has a finite number of jumps in any finite interval. Therefore, with the decomposition of the Skorohod integral

$$\begin{aligned}
&\int_{t_m}^{t_{m+1}} \int_{t_m}^{s_L-} \cdots \int_{t_m}^{s_2-} E[D_{t_{m+1}}^L F|\mathcal{F}_{t_{m+1}-}]\delta\tilde{N}_{s_1}\cdots\delta\tilde{N}_{s_L} \\
= \;& G(L,m,\varepsilon) + \int_{\mathcal{T}(L,m,\varepsilon)} E[D_{t_{m+1}}^L F|\mathcal{F}_{t_{m+1}-\varepsilon}]\delta\tilde{N}_{s_1}\cdots\delta\tilde{N}_{s_L}
\end{aligned}$$

and Cauchy-Schwarz inequality, in $L^2(\Omega)$,

$$\begin{aligned}
&\int_{t_m}^{t_{m+1}} \int_{t_m}^{s_L-} \cdots \int_{t_m}^{s_2-} E[D_{t_{m+1}}^L F|\mathcal{F}_{t_{m+1}-}]\delta\tilde{N}_{s_1}\cdots\delta\tilde{N}_{s_L} \\
= \;& \sum_{j=0}^{L} \sum_{i_1,\ldots,i_L=0}^{\infty} (-1)^{i_1+\ldots+i_L+j} E[D_{t_{m+1}}^{i_1+\ldots+i_L+j+L}F|\mathcal{F}_{t_{m+1}}] \\
& \times C_{L-j}\left(N_{t_{m+1}} - N_{t_m}, \lambda(t_{m+1} - t_m)\right) \frac{(\lambda(t_{m+1} - t_m))^j}{(L-j)!j!}.
\end{aligned}$$

We notice that the remainder tends to 0. This completes the proof of the backward Taylor expansion for Poisson processes.

Since the simple Lévy processes can be regarded as a finite sum of Poisson processes, similar to (4.5), we apply the CHO formula (2.5) and obtain:

$$E[F|\mathcal{F}_{t_m}] = E[F|\mathcal{F}_{t_{m+1}}] - \int_{t_m}^{t_{m+1}} \sum_{j=1}^{J} E[D_{t_{m+1}}^{(j)}F|\mathcal{F}_{s_1-}]\mathrm{d}\tilde{N}_{s_1}^j.$$

Keep iterating the above formula using (2.5), with the similar discussions as what we did for (4.6), the backward Taylor expansion for simple Lévy processes can be obtained.

□

As a corollary of Theorem 4.1, we can similarly prove the backward Taylor expansion for the compound Poisson process. We write $X_t = \int_{[0,t]\times\mathbb{R}_0} xJ(\mathrm{d}s,\mathrm{d}x)$.

Corollary 4.1. *Assume $F \in L^2(\Omega)$ such that $F = f(X_{t_1},...,X_{t_M})$ with $0 \leq t_1 \leq ... \leq t_M$. Then if $\sum_{i=0}^{\infty}\int_{\mathbb{R}_0^i} E\left[\left(D_{t_{m+1},x_1}\cdots D_{t_{m+1},x_i}F\right)^2\right](\nu(\mathrm{d}x))^{\otimes i} < \infty$, $\nu(\mathbb{R}_0) = \lambda < \infty$, and the following condition holds: for any $1 \leq m \leq M-1$,*

$$\sum_{i=0}^{L}\int_{\mathbb{R}_0^{2L-i}} E\left[\left(D^{2L-i}_{t_{m+1},x_1,...,t_{m+1},x_{2L-i}}F\right)^2\right](\nu(\mathrm{d}x))^{\otimes 2L-i}\binom{L}{i}^4\frac{i!}{(L!)^2}(t_{m+1}-t_m)^{2L-i} \xrightarrow{L\to\infty} 0, \tag{4.7}$$

we have, for any $1 \leq m \leq M-1$,

$$\begin{aligned} & E[F|\mathcal{F}_{t_m}] \\ = \; & E[F|\mathcal{F}_{t_{m+1}}] + \sum_{l=1}^{\infty}\gamma_l(m,\Delta X_m)\int_{\mathbb{R}_0^l} E[D^l_{t_{m+1},x_1,...,t_{m+1},x_l}F|\mathcal{F}_{t_{m+1}}]\,(\nu(\mathrm{d}x))^{\otimes l} \end{aligned}$$

where $\gamma_l(m,\Delta X_m)$ has the following representation: for $l \geq 0$

$$\gamma_l(m,\Delta X_m) = (-1)^l \sum_{\substack{i_1+\cdots+i_n+j+n=l \\ j\leq n}} C_{n-j}\left(\Delta X_{t_m},\lambda(t_{m+1}-t_m)\right)\frac{(\lambda(t_{m+1}-t_m))^j}{(n-j)!j!}.$$

Remark 4.1. *A large range of random variables can fit (4.3). Using the Stirling approximation, we can prove that:*

$$\sum_{i=0}^{L}\binom{L}{i}^4\frac{i!}{(L!)^2}\lambda^{2L-i}(t_{m+1}-t_m)^{2L-i} \leq \frac{C^L}{L^L}$$

for some fixed constant $C > 0$. Thus, for those random variables F such that for any l and t,

$$E\left[\left(D^l_tF\right)^2\right] \leq c^l \tag{4.8}$$

for some constant c, (4.3) always holds. A simple example is $F = e^{\alpha N_T}$ for any constant α. We have

$$D^L_TF = \sum_{l=0}^{L}(-1)^{L-l}\binom{L}{l}e^{\alpha N_T+\alpha l} = e^{\alpha N_T}\sum_{l=0}^{L}(-1)^{L-l}\binom{L}{l}e^{\alpha l} = e^{\alpha N_T}\left(e^{\alpha}-1\right)^L.$$

Then, $E\left[\left(D^L_TF\right)^2\right] = E[e^{2\alpha N_T}]\left(e^{\alpha}-1\right)^{2L} \leq c^L$ for some fixed c which does not depend on L and (4.3) holds. Therefore, we can regard (4.8) as an alternative condition of (4.3), which is easy to check in practical examples. Similar discussion also holds for condition (4.7). We can use the alternative condition:

$$\int_{\mathbb{R}_0^L} E\left[\left(D_{t_{m+1},x_1}\cdots D_{t_{m+1},x_L}F\right)^2\right](\nu(\mathrm{d}x))^{\otimes L} \leq c^L.$$

to simplify the checking process in practical calculations.

5. Applications

5.1. Poisson Black-Scholes Model

In this section, we recover the price of the call option under the Poisson Black-Scholes Model. We notice that the Dyson series approach is simpler and more direct than the classical PIDE approach, see for example [11], Chapter 11. Also, we note that we can not use the Dyson series in the Brownian case, since the maximum function is not differentiable.

Let $F = \exp(N_T \log(\sigma+1) - \lambda T\sigma)$ with $\sigma \geq 0$. 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 derivative and induction, we obtain:

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

Thus by the Dyson type formula for Poisson process (3.2),

$$\begin{aligned} & E[G|\mathcal{F}_t] \\ = \quad & \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \left(e^{N_t \log(\sigma+1) - \lambda T\sigma + k\log(\sigma+1)} - K \right)^+ \frac{\lambda^n (T-t)^n}{n!} \end{aligned} \tag{5.1}$$

$$= \quad e^{-\lambda(T-t)} \sum_{k=0}^{\infty} \frac{\lambda^k (T-t)^k}{k!} \left(e^{N_t \log(\sigma+1) - \lambda T\sigma + k\log(\sigma+1)} - K \right)^+ \tag{5.2}$$

which matches the classical result shown in [11], Chapter 11.

Rather than these two approaches, we can also obtain a series of the pricing formula with BTE using Theorem 4.1:

$$\begin{aligned} \sum_{n=0}^{\infty} \Bigg\{ & \sum_{\substack{i_1+\ldots+i_l+j+l=n \\ j \leq l}} C_{l-j}(N_T - N_t, \lambda(T-t)) \frac{\lambda^j (T-t)^j}{(l-j)!j!} \\ & \times \sum_{k=0}^{n} (-1)^k \binom{n}{k} \left(e^{N_T \log(\sigma+1) - \lambda T\sigma + k\log(\sigma+1)} - K \right)^+ \Bigg\}. \end{aligned} \tag{5.3}$$

In the following, we numerically compare these three different approaches: by Dyson series (5.1), BTE (5.3) and regular series (5.2). In each case, we truncate the series along the first summation symbol up to N terms.

We choose $T = \lambda = 1,\ t = 0$ and the strike $K = 1$, and approximate these three series by taking $N = 30$ and we show two cases with different values of σ. From Figure 1 and 2, we can easily observe that both backward Taylor expansion and Dyson type series converge faster than the regular approach. Moreover, this comparison also implies that the Dyson series matches the BTE numerically for this particular example.

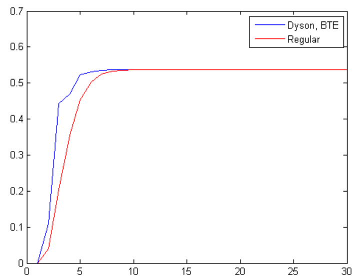

Figure 1: Option price with $\sigma = 2$ as a function of the number of terms in the series

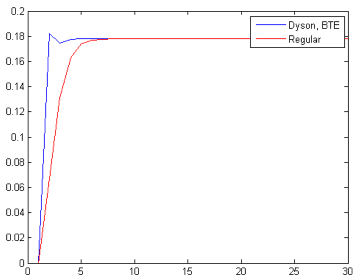

Figure 2: Option price with $\sigma = 0.5$ as a function of the number of terms in the series

5.2. Lévy Quadratic Model

In this section, we use the Dyson type series to evaluate the bond price in the Lévy quadratic model of interest rates. In [5], 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, that is, $r_s := \sum_{i=1}^d \left(U_s^{(i)}\right)^2$ with $U_s^{(i)} = U_0^{(i)} + \int_0^s \sigma_i(u) \mathrm{d}X_u^{(i)}$ for each $i = 1, ..., d$. $\{\sigma_i\}_{i=1,\ldots,d}$ are deterministic volatility functions, $U_0^{(i)}$ are constants and X_u^i are independent and identically distributed pure jump Lévy processes with finite Lévy measure ν and intensity $\lambda := \nu(\mathbb{R}_0) < \infty$.

This model can be regarded as a special case of a general Lévy quadratic model. The bond price of the general case is given by (1.2) and (1.3) in [1] as a solution of a system of Riccati equations. For this particular case, we use the Dyson series shown in Theorem 3.1 to give an explicit representation for the bond price $E[F|\mathcal{F}_t]$ with $F = \exp\left(-\int_t^T r_s \mathrm{d}s\right)$. We have the following proposition, which to the best of our knowledge, is an original result.

Proposition 5.1. *1. If $d = 1$, $U_0^{(1)} = 0$ and $\sigma_1(u) = 1$, we have:*

$$E[F|\mathcal{F}_t] = \exp\left(-r_t(T-t) - \lambda(T-t)\right)\left(1 + \sum_{i=1}^{\infty} \int_{\mathbb{R}_0^i} I_i(x_1, ..., x_i)\left(\nu(\mathrm{d}x)\right)^{\otimes i}\right)$$

where I_i is defined in (5.5) below.

2. If $d > 1$ and $\{\sigma_i\}_{i=1,\ldots,d}$ are deterministic volatility functions, we have:

$$E[F|\mathcal{F}_t] = \exp\left(-r_t(T-t) - d\lambda(T-t)\right)\prod_{i=1}^{d}\left(\sum_{l=0}^{\infty} \int_{\mathbb{R}_0^l} b_l^{(i)}(x_1, ..., x_i)\left(\nu(\mathrm{d}x)\right)^{\otimes l}\right)$$

where $b_l^{(i)}$ is defined in (5.6) below.

Before proving the proposition, we provide two numerical simulations. Let X be a standard Poisson process with $\lambda = 1$, take $T = 1$ and $t = 0$. We approximate the bond prices using two series in Proposition 5.1. For case 1, we simulate $e^{-1}\left(1 + \sum_{i=1}^{N} I_i(1, \cdots, 1)\right)$ up to $N = 10$, see Figure 3. For case 2, we assume $d = 2$, $U_0^{(1)} = U_0^{(2)} = 0$ and $\sigma_1(u) = \sigma_2(u) = e^{-(1-u)}$ and simulate

$$e^{-2}\left(1 + \sum_{l=1}^{N} b_l^{(1)}(1, ..., 1)\right)\left(1 + \sum_{l=1}^{N} b_l^{(2)}(1, ..., 1)\right)$$

up to $N = 10$ also, see Figure 4. From these two figures, we can observe a quite stable series converging to the target estimation quickly in four terms. And the limit values match with the values obtained by Monte Carlo simulation generating the Poisson variable 2^{20} times. Rather than the large amount of the data necessary to generate Poisson variables in order to simulate the integral in $E\left[\exp\left(-\int_0^1 N_s^2)\mathrm{d}s\right)\right]$, using Riemann sum, the Dyson series has an obvious simpler way to reach a good approximation with less request of time and data.

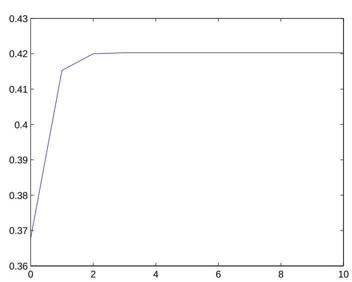

Figure 3: The first 10 partial sums of Dyson series with $d = \sigma = 1$.

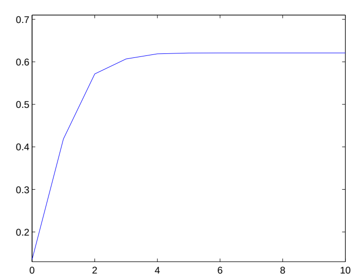

Figure 4: The first 10 partial sums of Dyson series with $d = 2$ and $\sigma_1(u) = \sigma_2(u) = e^{-(1-u)}$.

Proof of Proposition 5.1. In the following, we will denote $s_1 \vee s_2 := \max\{s_1, s_2\}$. For the first case when $d = 1$, $U_0^{(1)} = 0$ and $\sigma_1(u) = 1$, i.e. $r_s = X_s^2$ and $F = \exp\left(-\int_t^T X_s^2 \mathrm{d}s\right)$. By the chain rule (2.2) and the fact that $D_{u,x}^n \int_t^T X_s^2 \mathrm{d}s = 0$ for any u and x when $n > 2$, basic calculations show that

$$\begin{aligned} D_{s_1,x_1} \cdots D_{s_n,x_n} F = &\sum_{i=1}^{n} (-1)^{n-i} \sum_{\{j_1 \leq \ldots \leq j_i\} \subset \{1,\ldots,n\}} \\ &\exp\left\{ -\int_t^T X_s^2 \mathrm{d}s - \sum_{r=1}^{i} \left(2x_{j_r} \int_{s_{j_r}}^T X_s \mathrm{d}s + (T - s_{j_r}) x_{j_r}^2 \right) \right. \\ &\left. - \sum_{1 \leq r_1 < r_2 \leq i} 2(T - s_{j_{r_1}} \vee s_{j_{r_2}}) x_{j_{r_1}} x_{j_{r_2}} \right\} + (-1)^n \exp\left(-\int_t^T X_s^2 \mathrm{d}s \right), \end{aligned} \tag{5.4}$$

where by convention, the sum over $1 \leq r_1 < r_2 \leq i$ disappears when $i = 1$.

With the help of symmetry and Remark 3.1, we obtain

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

where

$$\begin{aligned}
I_i(x_1, ..., x_i) &:= \int_{t \leq s_1 \leq ... \leq s_i \leq T} \exp\left(-a_1(X_t)(T-s_1) - ... - a_i(X_t)(T-s_i)\right) (\mathrm{d}s)^{\otimes i}; \\
a_j(X_t) &:= 2X_t x_j + x_j^2 + 2x_j(x_{j-1} + ... + x_1), j \geq 2; \quad a_1(X_t) := 2X_t x_1 + x_1^2.
\end{aligned}$$

Since X_t can be regarded as a constant in the integral, I_i can be evaluated explicitly. For simplicity, we write a_j instead of $a_j(X_t)$. By setting up a recurrence formula for I_i and using induction, we obtain

$$I_i(x_1, ..., x_i) = B_i + \sum_{k=0}^{i-1} \frac{B_k}{a_{k+1}(a_{k+1} + a_{k+2}) \cdots (a_{k+1} + \cdots + a_i)} \tag{5.5}$$

where

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

Therefore by Theorem 3.1 and changing summation between n and i, we get

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

Now we provide the bond pricing formula for the case when $d > 1$ and $\{\sigma_i\}_{i=1,\ldots,d}$ 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 processes $X^{(i)}$ we have

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

If we denote $D^{(i)}$ as the Malliavin derivative for $X^{(i)}$, then similar calculations as in (5.4) give us

$$D^{(i)}_{s_1,x_1}\cdots D^{(i)}_{s_n,x_n}F^{(i)} = \sum_{l=1}^{n}(-1)^{n-l}\sum_{\{j_1\leq\ldots\leq j_l\}\subset\{1,\ldots,n\}}$$
$$\exp\Bigg(-\int_t^T\left(U_s^{(i)}\right)^2\mathrm{d}s-\sum_{r=1}^{l}\left(2x_{j_r}\int_{s_{j_r}}^T U_s^{(i)}\sigma_i(s)\mathrm{d}s+x_{j_r}^2\int_{s_{j_r}}^T\sigma_i^2(s)\mathrm{d}s\right)$$
$$-\sum_{1\leq r_1<r_2\leq l}2x_{j_{r_1}}x_{j_{r_2}}\int_{s_{j_{r_1}}\vee s_{j_{r_2}}}^T\sigma_i^2(s)\mathrm{d}s\Bigg)+(-1)^n\exp\left(-\int_t^T\left(U_s^{(i)}\right)^2\mathrm{d}s\right),$$

where by convention, the sum over $1\leq r_1<r_2\leq l$ disappears when $l=1$.

Now if we define

$$b_l^{(i)}(x_1,...,x_l) := \int_{[t,T]^l}\exp(-\sum_{r=1}^{l}\left(2U_t^{(i)}x_r\int_{s_r}^T\sigma_i(s)\mathrm{d}s+x_r^2\int_{s_r}^T\sigma_i^2(s)\mathrm{d}s\right)$$
$$-\sum_{1\leq r_1<r_2\leq l}2x_{r_1}x_{r_2}\int_{s_{r_1}\vee s_{r_2}}^T\sigma_i^2(s)\mathrm{d}s)(\mathrm{d}s)^{\otimes i};$$
$$b_1^{(i)}(x_1) := \int_t^T\exp\left(-2U_t^{(i)}x_1\int_{s_1}^T\sigma_i(s)\mathrm{d}s-x_1^2\int_{s_1}^T\sigma_i^2(s)\mathrm{d}s\right)\mathrm{d}s_1,\quad b_0^{(i)}=1, \tag{5.6}$$

with the simple notation $b_l^{(i)}$ instead of $b_l^{(i)}(x_1,...,x_l)$, we finally obtain

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

which completes the proof. □

Acknowledgments

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 and a short visit of professor Josep Vives at the Université François Rabelais, Tours (France). The authors thank IMUB and Université François Rabelais for the support and the kind hospitality. The authors are also thankful to the anonymous referee for his valuable comments and suggestions towards the improvement of this paper.

References

[1] L. Chen, D. Filipović and H. V. Poor (2004): *Quadratic term structure models for risk-free and defaultable rates.* Mathematical Finance 14 (4): 515–536.

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

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

[4] G. Di Nunno and J. Vives (2017): *A Malliavin-Skorohod calculus in L^0 and L^1 for additive and Volterra-type processes.* Stochastics 89 (1): 142-170.

[5] S. Jin, Q. Peng and H. Schellhorn (2016): *A representation theorem for smooth Brownian martingales.* Stochastics 88 (5): 651-679.

[6] S. Jin, Q. Peng and H. Schellhorn (2015): *Fractional Hida-Malliavin Derivatives and Series Representations of Fractional Conditional Expectations.* Communications on Stochastic Analysis 9 (2): 213-238.

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

[8] 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.

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

[10] N. Privault (2016): *Combinatorics of Poisson stochastic integrals with random integrands.* Stochastic analysis for Poisson point processes, Peccati and Reitzner Eds., Bocconi & Springer 37-80.

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

[12] J. León, J. L. Solé, F. Utzet and J. Vives (2002): *On Lévy processes, Malliavin calculus and market models with jumps.* Finance and Stochastics 6 (2): 197-225.

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

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

[15] 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.

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